JOURNAL OF ASSOCIATED MEDICAL SCIENCES

JAMS

JOURNAL OF ASSOCIATED MEDICAL SCIENCES

THE OFFICIAL PEER-REVIEWED ONLINE JOURNAL

Volume 59 Number 2 May-August 2026 E-ISSN: 2539-6056

Journal of Associated Medical Sciences

Aims and scope

The Journal of Associated Medical Sciences belongs to the Faculty of Associated Medical Sciences (AMS), Chiang Mai University, Thailand. The journal aims explicitly to provide a platform for medical technologists, radiologic technologists, occupational therapists, physical therapists, speech-language pathologists, and other related professionals to distribute, share, and discuss their research findings, inventions, and innovations in the areas of:

1. Medical Technology
2. Radiologic Technology
3. Occupational Therapy
4. Physical Therapy
5. Communication Disorders
6. Other related fields

Submitted manuscripts within the journal's scope will be processed strictly following the journal's double-blinded peer review process. Therefore, the final decision can be completed in 1-3 months on average, depending on the number of rounds of revision.

Objectives

The Journal of Associated Medical Sciences aims to publish integrated research papers in medical technology, physical therapy, occupational therapy, radiologic technology, and related research under peer review via a double-blinded process by at least two
Peer review process

Peer review process

By submitting a manuscript to the *Journal of Associated Medical Sciences*, the authors agree to subject it to the confidential double-blinded peer-review process. Editors and reviewers are informed that the manuscripts must be considered confidential. After a manuscript is received, it is assigned by a specific Associate Editor. The Associate Editor prepares a list of expert reviewers, which may include some suggestions by the Editor-in-Chief. Authors can indicate individuals whom they would like to have excluded as reviewers. Generally, requests to exclude certain potential reviewers will be honored except in fields with a limited number of experts.

All potential reviewers are contacted individually to determine availability. Manuscript files are sent to at least two expert reviewers. Reviewers are asked to complete the review of the manuscripts within 2 weeks and to return a short review form. Based on the reviewers' comments, the Associate Editor recommends a course of action and communicates the reviews and recommendations to the Editor-in-Chief for a final decision.

The Associate Editor considers the comments made by the reviewers and the recommendation of the Editor-in-Chief, selects those comments to be shared with the authors, makes a final decision concerning the manuscripts, and prepares the decision letter for signature by the Editor-in-Chief. If revisions of the manuscripts are suggested, the Associate Editor also recommends who should review the revised paper when resubmitted. Authors are informed of the decision by e-mail, appropriate comments from reviewers and editors.

Publication schedule

The *Journal of Associated Medical Sciences* is published three issues a year

Issue 1: January-April
Issue 2: May-August
Issue 3: September-December

Access Policy and Archiving

The *Journal of Associated Medical Sciences* publishes three issues a year. The articles are available online to the public, supporting and exchanging knowledge. They can be accessed free of charge through DOI at https://he01.tci-thaijo.org/index.php/bulletinAMS/index based on the Thailand Citation Index (TCI) online service.

Types of Manuscripts

Manuscripts may be submitted in the form of review articles, original articles, or short communications as an approximate guide to length:

- **Review articles** must not exceed 20 journal pages (not more than 5,000 words), including 6 tables/figures and references (maximum 75, recent and relevant).
- **Research articles** must not exceed 15 journal pages (not more than 3,500 words), including 6 tables/figures and references (maximum 40, recent and relevant).
- **Short communications**, including technical reports, notes, and letters to the editor, must not exceed 5 journal pages (not more than 1,500 words), including 2 tables/figures and references (maximum 10, recent and relevant).

Ownership and Revenue Sources

The *Journal of Associated Medical Sciences* is supported by the Faculty of Associated Medical Sciences, Chiang Mai University, Thailand.

Journal History

The Journal of Associated Medical Sciences was established in 1968 as the *Bulletin of Chiang Mai Associated Medical Sciences*. For over 48 years, printed versions from Vol. 1, No. 1 to Vol. 49 were published. To advance internationally and strengthen academic quality, the journal title was changed to *Journal of Associated Medical Sciences* (*J Assoc Med Sci, JAMS*) in 2017 with Vol. 50, No. 1 and onwards, with online access only.

Language

Manuscripts must be prepared in English, and a native speaker must proofread and edit. American or British usage is accepted, but not a mixture of these.

Publication Fee

The Journal of Associated Medical Sciences does not charge for submitting and publishing an article. All articles published in the journal are open-access and freely and widely available to all readers via the journal website.

Informed Consent Policy

All manuscripts based on research involving human participants or using human blood, tissues, other samples, and medical records must be submitted with a clearance certificate with an approval number from the Institutional Ethics Committee. Informed consent must be signed to participate in the study after proper and sufficient explanation from the researchers. Subjects must be allowed to ask if they have questions, and they must be answered truthfully. The subject, not the researcher, decides whether to participate in the study. Adults have the legal capacity to sign the informed consent. If the subjects are children, their parents or guardians must sign the informed consent. All consent must be voluntarily given without any influence

Advertising and Direct Marketing

The Journal of Associated Medical Sciences does not require advertising and direct marketing activities, including soliciting manuscripts conducted on the journal's behalf.

Editorial Board

Editor-in-Chief

Preeyanat Vongchan — Chiang Mai University, Thailand.
preeyanat.v@cmu.ac.th

Associate Editor

Araya Yankai — Chiang Mai University, Thailand.
araya.yankai@cmu.ac.th

Khanittha Punturee — Chiang Mai University, Thailand.
khanittha.taneyhill@cmu.ac.th

Montree Tungjai — Chiang Mai University, Thailand.
montree.t@cmu.ac.th

Singkome Tima — Chiang Mai University, Thailand.
singkome.tima@cmu.ac.th

Suchart Kothan — Chiang Mai University, Thailand.
suchart.kothan@cmu.ac.th

Supaporn Chinchai — Chiang Mai University, Thailand.
supaporn.c@cmu.ac.th

Thanusak Tatu — Nation University, Thailand.
thanusak_tat@nation.ac.th

Editorial Board

Cecilia Li-Tsang — Hong Kong Polytechnic University, Hong Kong.
cecilia.li@polyu.edu.hk

Christopher Lai — School of Medical and Health Sciences, Hong Kong.
christopherlai@twc.edu.hk

Clare Hocking — Auckland University of Technology, New Zealand.
clare.hocking@aut.ac.nz

Darawan Rinchai — The Rockefeller University, United States of America.
drinchai@gmail.com

David Man — Hong Kong Poly Technic University, Hong Kong.
David.Man@polyu.edu.hk

Elizabeth Wellington — University of Warwick, United Kingdom.
Wellington@warwick.ac.uk

Fuming Zhang — Renssellaer Polytechnic Institute, United State of America.
ZHANGF2@rpi.edu

Ganjana Lertmemongkolchai — Chiang Mai University, Thailand.
ganja_le@kku.ac.th

Hans Bäumler — Universitätsmedizin Berlin, German.
hans.baeumler@charite.de

Hong Joo Kim — Kyungpook National University, South Korea.
hongjoo@knu.ac.kr

Isamu Shibamoto — Seirei Christopher University, Japan.
isamu-s@seirei.ac.jp

Jourdain Gonzague — French National Research Institute for Sustainable Development (IRD), France.
gonzague.jourdain@phpt.org

Kesara Na Bangchang — Thammasart University, Thailand.
kesaratmu@yahoo.com

Leonard Henry Joseph — University of Brighton, United Kingdom.
L.HenryJoseph@brighton.ac.uk

Marc Lallemant *marclallemant@gmail.com*	Drugs for Neglected Diseases Initiative (DNDi), Switzerland.
Masahiro Hosoda *m_hosoda@hirosaki-u.ac.jp*	Hirosaki University, Japan.
Mohamad Warda *mohamad.warda@atauni.edu.tr*	Ataturk University, Türkiye.
Montree Tungjai *montree.t@cmu.ac.th*	Chiang Mai University, Thailand.
Nicole Ngo Glang Huang *nicole.ngo-giang-huong@phpt.org*	French National Research Institute for Sustainable Development (IRD), France.
Prawit Janwantanakul *prawit.j@chula.ac.th*	Chulalongkorn University, Thailand.
Roongtiwa Vachalathiti *roongtiwa.vac@mahidol.ac.th*	Mahidol University, Thailand.
Rumpa Boonsinsukh *rumpa@g.swu.ac.th*	Srinakharinwirot University, Thailand.
Sakorn Pornprasert *sakorn.pornprasert@cmu.ac.th*	Chiang Mai University, Thailand.
Sophie Le Coeur *sophielecoeur0@gmail.com*	French Institute for Demographic Studies (INED), France.
Srijit Das *s.das@squ.edu.om*	Sultan Qaboos University, Oman.
Sugalya Amatachaya *samata@kku.ac.th*	Khon Kaen University, Thailand.
Supan Fucharoen *supan@kku.ac.th*	Khon Kaen University, Thailand.
Tengku Shahrul Anuar *tengku9235@uitm.edu.my*	Universiti Teknologi MARA, Malaysia.
Thanaporn Tunprasert *T.Tunprasert2@brighton.ac.uk*	University of Brighton, United Kingdom.
Timothy R. Cressey *Tim.cressey@phpt.org*	French National Research Institute for Sustainable Development (IRD), France.
Valerie Wright-St Clair *vwright@aut.ac.nz*	Auckland University of Technology, New Zealand.
Witaya Mathiyakom *mcnitt@usc.edu*	University of Southern California, United States of America.
Zhirong Zhong *zhongzhirong@126.com*	Southwest Medical University, Luzhou, China.

Business manager
Kantaphon Promkam

Treasurer
Angsumalee Srithiruen

Webpage Administrative Staff
Kantaphon Promkam
Tippawan Sookruay

Journal website
https://www.tci-thaijo.org/index.php/bulletinAMS/index (Homepage)

Journal E-ISSN:
2539-6056

Editorial Office
Faculty of Associated Medical Sciences, Chiang Mai University
110 Inthawaroros Road, Suthep, Muang, Chiang Mai, 50200
Phone 053 935072 Facsimile 053 936042

Disclaimer

Every effort is made by the Editor-in-Chief and the Editorial Board of the *Journal of Associated Medical Sciences* to see that no inaccurate or misleading data, opinion, or statement appears in the Journal. However, they wish to clarify that the data and opinions appearing in the articles and advertisements herein are the responsibility of the contributor, sponsor, or advertiser concerned. Accordingly, the Editor-in-Chief and the Editorial Board accept no liability whatsoever for the consequences of any such inaccurate or misleading data, opinion, or statement.

Journal of Associated Medical Sciences

Vol. 59 No. 2 : May - August 2026

Content

1 Cut-off determination of HLA crossmatching by flow cytometry

Puntharawan Sitthipornworakul[1], Chonticha Sirikul[1], Sucheela Chomsook[2], Praphaphorn Daoram[2], Preeyarat Kamsupa[2], Praijitr Tanun[2], Thitiporn Dispao[1], Nipapan Leetrakool[2,3], Nampeung Anukul[1]*

9 Case report: Two cases of Hb Liuzhou-Yufeng (HBA1:c.334G>T) found in Northern Thailand

Yasmina Azfa Putri Alif[1], Maulida Syifa Kamila Makarim[1], Chedtapak Ruengdit[2], Yona Mimanda[1], Chris Adhiyanto[1], Laifa Annisa Hendarmin[1], Sakorn Pornprasert[2]*

15 Prevalence and risk factors of elevated urinary albumin-to-creatinine ratio in a Thai academic population

Suwatsin Kittikunnathum, Pharisa Nanthawong, Piyawan Bunpo[*]

24 Elevated Sall4 expression correlates with prostate cancer Gleason score and metastasis using immunohistochemistry and RNAscope®

Dhafer A. Alghezi[1] and Ali Harb[2]*

34 Genomic insights into colistin-resistant Escherichia coli from clinical isolates in Thailand: Diversity of sequence types, plasmid-borne *mcr* variants, and One Health implications

Kwanchon Jearakitiwanich[1], Anusak Kerdsin[2], Monchai Siribamrungwong[3], Panan Ratthawongjirakul[4]*

44 Comparative evaluation and interpretability analysis of modern CNN architectures for brain tumor MRI classification

Nitipon Pongphaw[*] and Prommin Buaphan

63 Effects of core strengthening exercises on inspiratory muscle strength and functional capacity among obese college students

Kaviya Muralidharan[*] and Santhosh Kumar Sankar

72 Comprehensive benchmarking of machine learning models for blood glucose classification and prediction: new approach for improved hyperglycemia and hypoglycemia detection and prediction

Houda El Bouhissi[1], Tatiana Ermakova[2], Rabie A. Ramadan[3], Djamila Ouaret[1]*

84 Radiation protection in dental radiology: Compliance, continuing education and equipment audit in imaging clinics

Emmanuel O. Esien-umo[1], Valentine C. Ikamaise[2], Blessing S. Ibe[1], Akwa E. Erim[1], Christopher A. Ishiekwen[3], Kingsley C. Omeke[1], Ndubuisi O. Chiaghanam[1]*

91 Blood donation and blood components utilization in Suratthani Hospital

Weerasak Jornpakdee[1], Buchita Swatwan[2], Suwannee Jitueakul[2], Rungkarn Sangkaruk[1]*

Content

97 Association of Intron 22 inversion and HLA-DRB1*15:01 allele with inhibitor development in hemophilia A: A genetic analysis of risk factors

Chitpon Seeljunya[1], Chollanot Kaset[1,2], Sirinart Chomean[1,2]*

105 Fall risk, fear of falling and their association with depressive symptoms among institutionalized older adults: A cross-sectional study

Bowas S, Loganathan S, Raghuram P, Jaivignesh S*

113 Factors increasing telemedicine adoption across adult non-healthcare professionals: A systematic review of observational studies published between 2021 and 2025

Trisakti Halimah Delimasari, Elsa Tursina, Annisa' Arifatul Hikmah, Riska Fajar Fatony*

125 The effects of motor-cognitive training on balance and working memory in young adult women with low physical activity: A randomized controlled trial

Supanee Chounchay, Pinpapha Retsuk, Parintorn Mantam, Kannika Seetha, Warangkana Arsadee, Aparat Chanchuphon, Nattida Pantub, Supitcha Wongphasuksathaporn*

134 The study of maximum electromyogram of masseter muscle before and during modified electroconvulsive therapy in psychiatric patients

*Namsukh Kongkalai**

142 Fabrication and evaluation of 3D-printed customizable boluses for optimized radiotherapy applications

Chitchaya Suwanraksa[1], Athitan Khiewded[2], Gadesuda Jankaew[1], Thirawut Rojchanaumpawan[1], Surapong Chatpun[3], Pornchai Phukpattaranont[2]*

153 Effectiveness of guided imagery to improve sleep quality among young adults with post-traumatic stress disorder (PTSD)

Vikraman G. and Benitta Paulin Giftlin R.*

163 Assessment of imaging radiation dose from tomotherapy systems

Pratchayakan Hompeng[1,2], Wannapha Nobnop[3], Tinnagorn Donmoon[4], Komkrit Krongkietlearts[5], Kananan Utitsarn[5], Sumalee Yabsantia[1,6], Thunyarat Chusin[1,6]*

178 Serum Klotho as a feasible diagnostic biomarker for metabolic syndrome in Iraqi adults aged over 50 years

Haitham Ibrahim Azrak and Entedhar Rifaat Sarhat*

Content

187 Towards precision hematology: Machine learning-driven comparative feature importance analysis of hematologic parameters in newly diagnosed chronic myeloid leukemia vs healthy controls

Hossein Ayatollahi, Arefeh Mazhari, Seyed Mohammad Tabatabaei, Samaneh Talebi, Mahnaz Dehghan, Maryam Sheikhi, Yasamin Ayatollahi, Zahra Darvish Khalilabadi

196 The usefulness of endometrial receptivity biomarkers as predictors of IVF outcome

Zainab A. Aljarah, Rana M. Hameed, Wasan G. Al Safi

208 Examination of the validity and reliability of the Perceive, Recall, Plan, and Perform-long COVID (PRPP-Long COVID) self-report assessment for Thai adults with long COVID

Sutinun Juntorn[1], Peeradech Thichanpiang[1], Pawared Ontong[2], Thitiya Wangkawan[1], Peeraya Munkhetvit[3]*

220 Effect of acceptance and commitment therapy on psychological well-being and quality of life among individuals living with HIV

*Punitha P and Santhiya M**

226 Effectiveness of reminiscence therapy to improve sleep quality among older adults with hypertension

J. Roselin Jaba Sheela[1] and J. Abiya Evangelin Jayapriya[2]*

234 Interprofessional education and collaboration in health professions: A narrative review of global evidence and lessons from the Philippines

Paolo Miguel P. Bulan[1], Nikki Y. Pestaño[1], Romel V. Cabazor[1], Marelen H. Amanuence[2], Lois Vanjie G. Napalit[2], Niño Paolo M. Tan[3], Catherine S. Malaca[2], Eubina C. Estoy[4]*

250 Advanced YOLOv8 architecture for multi-class brain tumor detection

Kavita Singh, Rajesh Kumar, Satyendra Singh*

259 Combined neutrophil-to-lymphocyte, monocyte-to-lymphocyte, and platelet- to-lymphocyte ratios as inflammatory biomarkers in tuberculosis: A retrospective study from Lampang, Thailand

Sittichok Chaiwong, Niphawan Jettavan, Yuttana Mundee, Pawaret Panyajai, Wuttichote Jansaento, Watcharapong Jugnam-ang

268 Effect of botulinum toxin-A phonophoresis on spasticity in children with hemiplegia

Ahmed F Attia[1], Amira M El-Tohamy[1], Gehan H El-Meniawy[1], Walaa ElNaggar[2]*

Content

280 Association of aerobic and anaerobic capacities with body composition in basketball players aged 13-16 years in Chiang Mai Province

Panida Prasertpong[1], Sothida Nantakool[2], Busaba Chuatrakoon[2] A*

286 Impact of ergonomic workstation interventions and scapular bracing on neck muscle performance, postural alignment, neck disability and work productivity in bankers with non-specific neck pain

Muhammad Awais Tassawar, Soyba Nazir, Montakarn Chaikumarn

298 Comparison of carotid intima media thickness and high sensitive c-reactive protein with hemostasis model assessment for insulin resistance in prediabetes and normoglycemic patients: A case-control study

Kamani Bharath Teja, R Prasanna Venkatesh, Vignessh Raveekumaran

307 Dual-scale evaluation of pain after hypofractionated post-mastectomy radiotherapy

Pratibha Prashar, Brish Bhanu Bhardwaj, Vandana Thakur, Poorva Vias, Ankita Pandey, Hardik Sharma, Kulbir Singh, Vivek Kumar

316 Development of a YOLOv8-based deep learning framework for detection and classification of periapical radiolucencies in dental radiographs using the periapical index

Natdanai Hirata[1], Wannakamon Panyarak[2], Kittipit Klanliang[3], Arnon Charuakkra[2], Kittichai Wantanajittikul[1]*

327 Brief review: Effects of textural properties of lamellar silica in drug delivery applications

Dyah Ellyawati Kusumaningtyas Maharani[1], Fajar Rakhman Wibowo[1], Rino Rakhmata Mukti[2]*

339 Long-term effects of a structured training program on physical fitness in male football players aged 18-22 years

Peerawats Dejprom[1], Sothida Nantakool[2], Busaba Chuatrakoon[2]*

348 Development and preliminary evaluation of a dynamic cognitive training program for healthy older adults

*Warunya Sutthikampa, Suchitporn Lersilp, Autchariya Punyakaew, Supawadee Putthinoi**

364 Evaluating the integration of assistive devices in occupational therapy across North India

Shivani Bhardwaj, Bilal Ahmed, Saba Irem, Saqueba Shahi, Nikhat Sultana*

372 Role of vitamin C or vitamin E on platelet levels in dengue fever: A randomized controlled trial

Akshay Basireddy, Vignessh Raveekumaran, Mohamed Hanifah*

Content

382 Radioprotective effect and antioxidant activities of Thai Gac fruit (Momordica cochinchinensis Spreng)

Anont Laorngnual, Sareeya Reungpatthanaphong, Paiboon Reungpatthanaphong

393 Obesity risk and body composition among Bangkok adolescents with autism spectrum disorder: Stratification by sensory and cognitive severity using data mining

Sarun Kunwittaya[1], Apilak Worachartcheewan[2], Chattip Silarangsri[3], Khanittha Punturee[4], Peeradech Thichanpiang[5], Kannika Permpoonputtana[1]*

421 Efficacy of the modified carbapenem inactivation method (mCIM) for detecting carbapenemase-producing Pseudomonas aeruginosain a 1,000-bed tertiary hospital in Thailand

Jakkrit Khampakul[1,2], Noppadon Jumroon[1], Narong Chaihongsa[2], Pitak Santanirand[2]*

428 Prevalence and characterization of β-thalassemia minor in a premarital screening at a hereditary blood diseases center in Babylon, Iraq

Shekair WH[1], Al-Maaroof ZW[2], Aljawazery SAA[3], Haddawi KH[4], Al-Khyate SAM[5]*

438 Efficacy of digital gaming vs traditional Indian gaming on memory and attention in patients with persistent post-concussion syndrome: A feasibility study

J. Mohanasundaram, Vignesh Srinivasan, Prathap Suganthirababu, Harishraj A., Madeshwaran S.*

Journal of Associated Medical Sciences 2026; 59(2): 1-8

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Cut-off determination of HLA crossmatching by flow cytometry

Puntharawan Sitthipornworakul[1], Chonticha Sirikul[1], Sucheela Chomsook[2], Praphaphorn Daoram[2], Preeyarat Kamsupa[2], Praijitr Tanun[2], Thitiporn Dispao[1], Nipapan Leetrakool[2,3], Nampeung Anukul[1*]

[1]Division of Transfusion Science, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Blood Bank Section, Maharaj Nakorn Chiang Mai Hospital, Faculty of Medicine, Chiang Mai University, Chiang Mai Province, Thailand.
[3]Faculty of Medical Technology, Nation University, Lampang Province, Thailand.

ARTICLE INFO

Article history:
Received 26 May 2025
Accepted as revised 16 December 2025
Available online 8 January 2026

Keywords:
HLA crossmatching, flow cytometry, cut-off, kidney transplantation.

ABSTRACT

Background: Microlymphocytotoxicity test is a gold standard for HLA crossmatching but provides low sensitivity. Emerging flow cytometry crossmatching (FCXM) provides more sensitivity, however, suitable cut-off values for FCXM must still be established for crossmatching result interpretation.

Objectives: This study aimed to determine the appropriate cut-off value for FCXM in kidney transplant patients and donors.

Materials and methods: A total of 30 sample pairs with known HLA typing and known antibody specificity were analyzed. Median channel fluorescence (MCF) and median channel fluorescence shift (MFS) were used to determine cut-off value using three different approaches and % accuracy was determined.

Results: The result showed that the cut-off values from first approach using lowest MCF of positive control and highest MCF of negative control were 1655.99 MCF for T-cells (yielding 100% accuracy) and 43668.06 MCF for B-cells (yielding 36.67% accuracy). The cut-off values from second approach which calculated from standard deviation (SD) of negative control MCF showed SD of T-cells MCF as 203.81 (100% accuracy) and SD for B-cells MCF as 11109.30 (60% accuracy). The cut-off values from third approach derived from the lowest MFS of the expected positive results were 1753.50 for T-cells (100% accuracy) and 593.50 for B-cells (73.33% accuracy).

Conclusion: The cut-off values from the third approach exhibited the highest accuracy. However, FCXM still had false positives and false negatives, underscoring the necessity for continuous refinement and adjustment of cut-off values based on the data accumulated within the laboratory to be used further for routine laboratory practice.

Introduction

For patients with end-stage organ failure, transplantation has emerged as a crucial therapeutic strategy that offers a chance for survival and an enhanced quality of life. However, immunological compatibility between the donor and the recipient has a significant impact on the outcome of a successful transplant, which is not exclusively determined by the surgical procedure.[1] Human leukocyte antigen system located on the short arm of chromosome no. 6 is a key component of the immune response and being one of the most important factors affecting transplant outcome. Two major classes of *HLA* gene system which

* **Corresponding contributor.**
Author's Address: *Division of Transfusion Science, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: nampeung.a@cmu.ac.th
doi: 10.12982/JAMS.2026.031
E-ISSN: 2539-6056

are class I include HLA-A, HLA-B, HLA-C and class II including HLA-DR, HLA-DQ, HLA-DP are responsible for antigen presentation. The class I molecules primarily present endogenous peptides to cytotoxic T cells, while class II molecules present exogenous peptides to helper T cells. When donor and recipient HLA molecules are not compatible, the immune system may recognize and react, which could result in immune-mediated graft rejection and transplant failure.[2] Not only HLA matching is important in selecting compatible donors for recipients, but preventing HLA-mediated rejection is also paramount for successful transplantation outcomes. Central to this process is HLA compatibility testing or HLA crossmatching which is the detection of pre-existing HLA antibodies in patient serum which can prevent the risk of hyperacute rejection from donor specific antibody (DSA).[1]

A serological method, microlymphocytotoxicity testing or complement-dependent cytotoxicity crossmatch (CDCXM), has been traditionally utilized in HLA crossmatching to detect these antibodies against donor HLA antigens. Although it is the gold standard method for HLA compatibility testing, its limitations including false negative results due to low sensitivity and the inability to detect low levels of antibodies and false positive results from autoantibodies could limit the successful transplantation outcome and may still cause graft rejection.[2,3] Studies have shown that patients with negative CDCXM but positive flow cytometry crossmatch (FCXM) results have an increased risk of organ rejection.[4] As a result, flow cytometry crossmatching, which offers improved sensitivity and precision in identifying donor-specific antibodies at lower levels, has been widely adopted to mitigate this risk.[2,5]

FCXM provides a more sensitive and accurate result for identifying antibody-mediated rejection that might be missed by the CDC approach. Despite its benefits, to guarantee its best application in organ e.g. kidney transplantation, exact cut-off values for FCXM must still be established. These parameters aid in reducing false positives, ensuring donor-recipient compatibility, and striking a balance between sensitivity and specificity. However, a significant challenge remains in determining the optimal cut-off value for defining positive and negative FCXM results which can vary among laboratories.[6] The lack of standardized cut-off values across laboratories necessitates the establishment of laboratory-specific thresholds based on robust data analysis. Therefore, this study aimed to determine the appropriate cut-off value for HLA crossmatching by flow cytometry in kidney transplant patients and donors, to be used further for routine laboratory practice in the HLA laboratory. The findings of this study will provide a basis for optimizing the pre-transplant assessment of kidney transplant candidates, thereby improving patient outcomes.

Materials and methods

Preparation of lymphocyte cells

Flow cytometry crossmatching (FCXM) experiments were performed using cell preparations from 29 donor cell samples and 7 recipient sera, resulting in a total of 30 FCXMs reactions. All samples of this study were the residual blood samples from HLA routine laboratory, Blood Bank Section, Maharaj Nakorn Chiang Mai Hospital, Faculty of Medicine, Chiang Mai University, Thailand. Donor lymphocytes were isolated from residual kidney donor blood samples after routinely tested with microlymphocytotoxicity test. Acid Citrate Dextrose (ACD) anticoagulated whole blood was used in a volume of 50 mL. Blood was divided into two equal portions and 5 mL of 5% dextran was added to facilitate red blood cell agglutination. The samples were then incubated at 37 °C with a 45-degree tilt. Lymphocytes were separated using Ficoll-Hypaque solution with a density of 1.077 (Robbins Scientific Corporation, Norway) in a volume of 10 mL. The samples were centrifuged at 2000 rpm for 20 minutes to obtain a lymphocyte ring layer. This layer was carefully extracted and diluted to 50 mL with normal saline solution. The cells were washed three times by centrifugation at 1400 rpm for 10 minutes and 1000 rpm for 10 minutes. The resulting cell pellet was resuspended in 2 mL of RPMI 1640 medium. Cell concentration was determined using a hemocytometer after mixing the cell suspension with 0.2% Trypan blue dye in a 1:1 ratio. Based on the cell count, the appropriate volume of cold flow wash buffer was calculated to achieve a final cell concentration of 10,000 cells/µL. After centrifugation and removal of the RPMI 1640 medium, the calculated volume of cold flow wash buffer was added to the cell pellet.

Preparation of patient serum

Seven of patient sera were prepared from residual blood samples collected after routine serological testing in the HLA laboratory, Blood Bank Section, Maharaj Nakorn Chiang Mai Hospital, Faculty of Medicine, Chiang Mai University, Thailand. The blood samples were centrifuged at 1000 rpm for 10 minutes to separate red blood cells. One milliliter of serum was then collected and transferred into a 1.5 mL microcentrifuge tube. All patient sera were characterized using LABScreen™ PRA Class I (LS1PRA) and LABScreen™ PRA Class II (LS2PRA) (Luminex®, One Lambda, USA), which reliably detect the presence of anti-HLA antibodies at the antigen group level. Positive control sera were pooled serum collected from 4-6 sensitized patients. These sera were heat-inactivated at 58 °C for 30 minutes to inactivate complement and then characterized by anti-HLA antibody screening using routine lymphocytotoxicity test (LCT). The positive control sera demonstrated greater than 75% panel reactive antibody (PRA) with whole lymphocytes of 14

random whole blood donors. Negative control sera were pooled from 4-6 male, blood group AB, apheresis healthy donors without a history of transfusion. These sera underwent the same heat inactivation procedure and were confirmed negative by LCT using cells from 10 different donors. Both positive and negative control sera were not tested by Luminex PRA assay but tested by LCT only as they were routinely used in LCT crossmatching. Their consistency in this assay was validated.

Flow cytometric crossmatch (FCXM) procedure

Flow cytometric crossmatch (FCXM) procedure followed Chamsai *et al.*,[7] 20 µL of prepared cells were pipetted into test tubes and mixed with 25 µL of patient serum. The mixture was incubated at room temperature for 30 minutes. Following incubation, the cells were washed three times with cold flow wash buffer, centrifuging at 500 ×g at 4 °C for 10 minutes per wash. Subsequently, the following antibodies were added: 5 µL of anti-CD19-PE, 5 µL of anti-CD3-PC5, and 10 µL of anti-IgG-FITC (diluted 1:100 with cold flow wash buffer). The samples were then incubated in the dark at 4 °C for 20 minutes. After incubation, the cells were washed twice with cold flow wash buffer, centrifuging at 500 ×g at 4 °C for 10 minutes per wash. The reaction was stopped by adding 250 µL of 1% cold paraformaldehyde. Finally, the samples were analyzed using a Flow cytometer (Beckman Coulter, DxFLEX) with CytExpert for DxFLEX software.

Determination of cut-off values for HLA crossmatching using flow cytometry

Three different approaches were used to calculate the cut-off values for HLA crossmatching to determine the best suitable thresholds. Across all three analytical approaches, cut-offs were determined using results from 30 independent FCXM reactions. The first approach involved recording the lowest mean channel fluorescence (MCF) value of the positive control and the highest MCF value of the negative control for both T and B cells and selected the most appropriate cut-off based on the observed range.[8] The second approach employed the mean and standard deviation (SD) obtained from the negative control's MCF for both T and B cells.[7,9] The third approach analyzed using the median channel fluorescence shift (MFS) to determine the lowest MFS value that reliably indicated a positive result and highest MFS value that showed negative result in both T and B cells.[10] All the cut-off values calculated from these three methods were compared to HLA class I and II typing of donor lymphocytes using polymerase chain reaction-sequence specific oligonucleotides (PCR-SSO) (LABType™, One Lambda, USA) with antibody specificity in patient sera to determine the possibility of positive or negative crossmatching result. Results obtained from three cut-off methodologies were subsequently calculated for the percentage of accuracy as the sum of true positives (TP) and true negatives (TN), divided by the total number of samples tested. The most appropriate cut-off values for flow cytometry crossmatch (FCXM) were then identified.

Results

HLA typing and antibody matching were performed on 30 cell-serum combinations, with 20 and 10 intended to be true positive and true negative for T-cells, respectively. Similarly, for B-cells, 21 were intended to be true positive, and 9 were intended to be true negative.

In the first approach, the lowest MCF value of the positive control for T-cells was 37817.28 with 40% accuracy, while the highest MCF value of the negative control was 1655.99 with 100% accuracy. For B-cells, the lowest MCF value of the positive control was 59115.80 with 40% accuracy, whereas the highest MCF value of the negative control was 43668.06 with 36.67% accuracy (Figure 1). The second approach showed that the T-cells negative control serum had a mean MCF+SD of 1103.59+203.81. The B-cells negative control serum had a mean MCF+SD of 12027.66+11109.30. The accuracy of the T-cells means MCF+SD and mean MCF+2SD were 100%. Meanwhile, the accuracy of B-cells means MCS+SD and mean MCF+2SD were 60% and 56.67%, respectively (Figure 2). In the third approach, the lowest MFS value that reliably indicated a T-cells positive result was 1753.50 with 100% accuracy, while the highest MFS value for a T-cells negative result was 991.50 with 100% accuracy. For B-cells, the lowest MFS value that reliably indicated a positive result was 593.50 with 73.33% accuracy. Whereas the highest MFS value for a B-cells negative result was 30996.64 with 40% accuracy (Figure 3). All results are represented in Table 1, which compares the cut-off values for T and B cells FCXM.

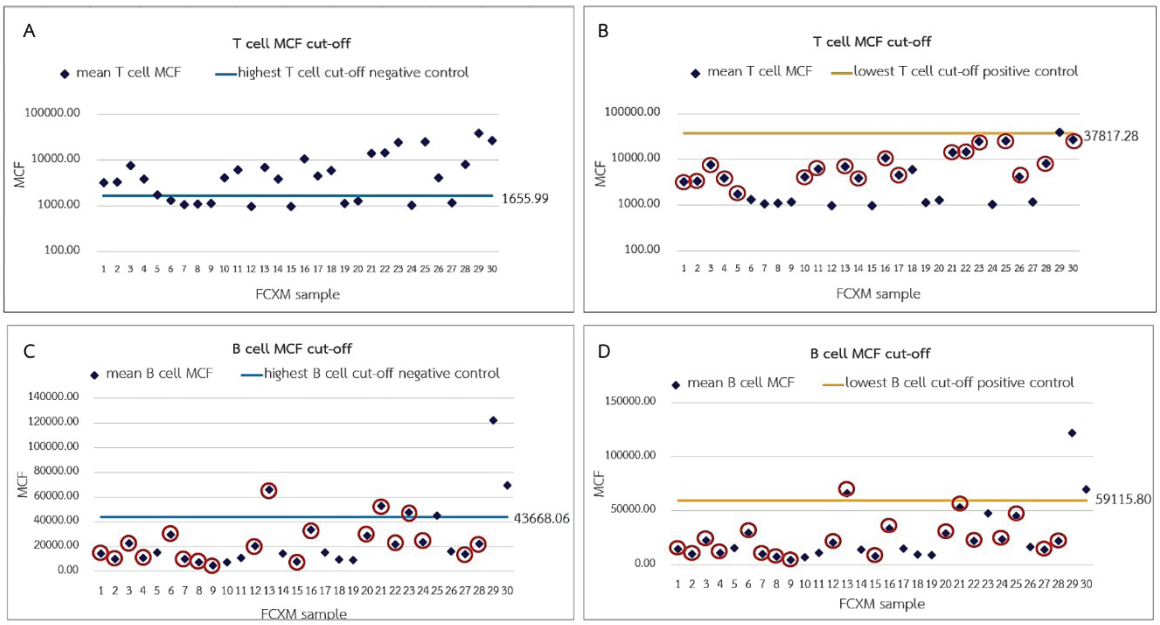


Figure 1. *Results from the first approach illustrating mean MCF across 30 FCXMs compared to each cut-off value. Red circles indicate discordant results between FCXM and expected outcomes. A: highest MCF T-cell negative control as cut-off, B: lowest MCF T-cell positive control as cut-off, C: highest MCF B-cell negative control as cut-off, D: lowest MCF B-cell positive control as cut-off.*

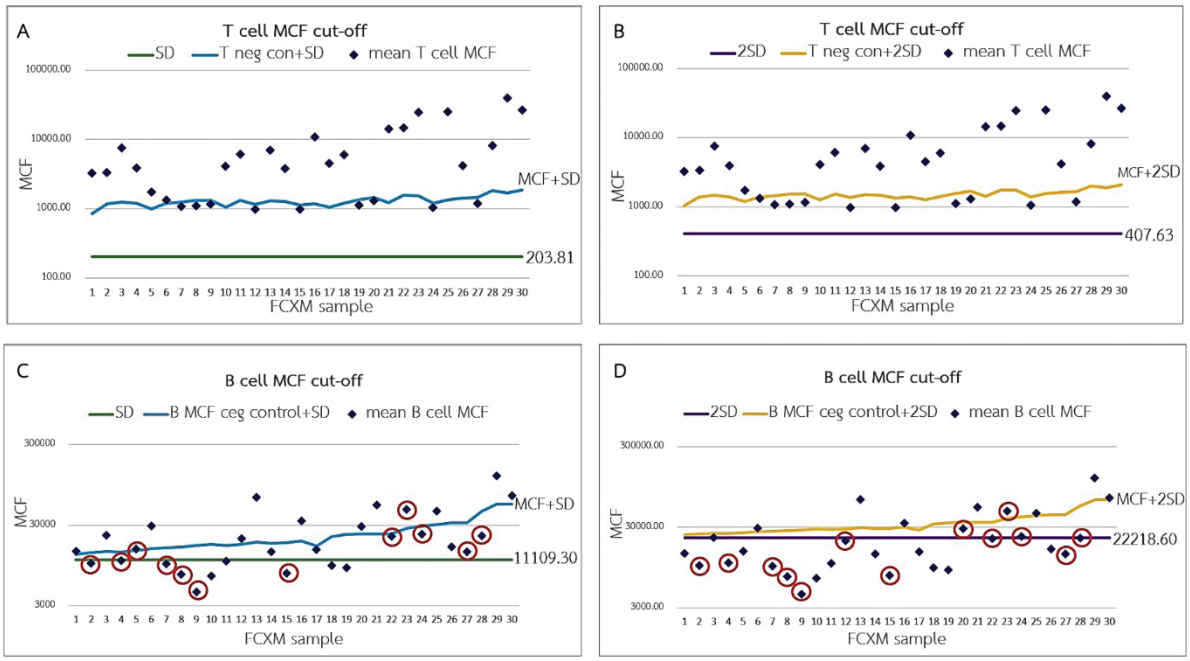


Figure 2. *Results from the second approach using statistical cut-offs based on mean MCF and standard deviations (SD) of negative control sera in 30 FCXMs. Red circles represent discrepant outcomes. A: Mean MCF+1 SD cut-off for T-cells, B: Mean MCF+2 SD cut-off for T-cells, C: Mean MCF+1 SD cut-off for B-cells, D: Mean MCF+2 SD cut-off for B-cells.*

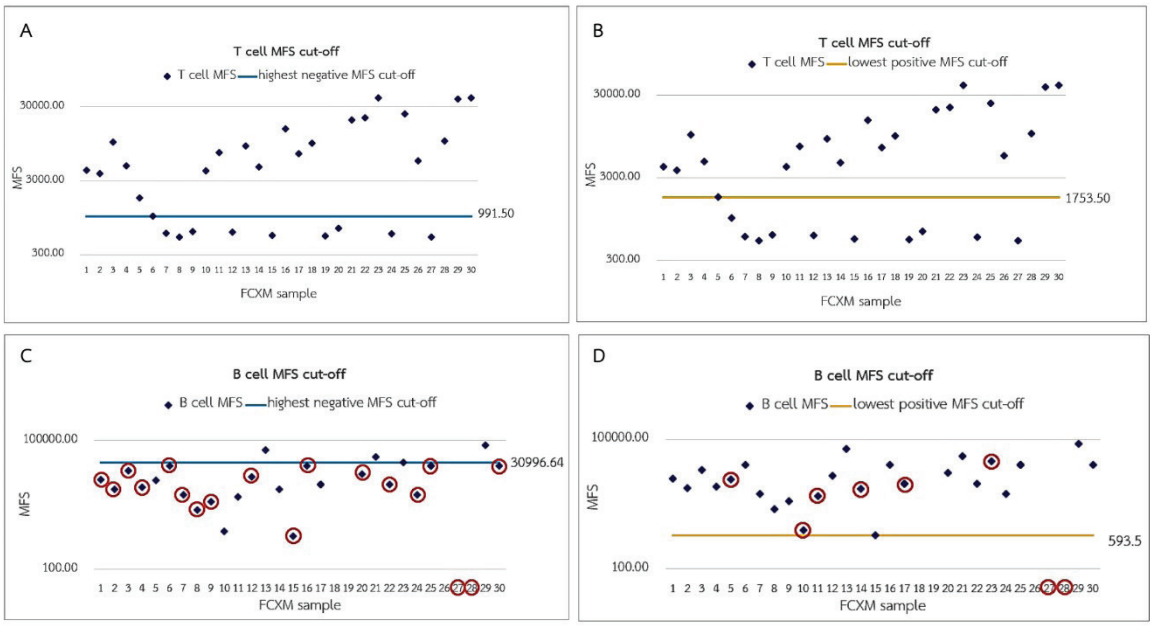


Figure 3. *Results from the third approach illustrating MFS across 30 FCXMs compared to defined cut-off values. Red circles indicate discordant results between observed and expected crossmatch outcomes. A: highest MFS value for T-cell predicting a negative result used as a cut-off, B: lowest MFS value for T-cell predicting a positive result used as a cut-off, C: highest MFS value for B-cell predicting a negative result used as a cut-off, D: lowest MFS value for B-cell predicting a positive result used as a cut-off.*

Table 1. Comparison of the cut-off values for T- and B-cell FCXM.

Parameter	Cut-off	T-cell FCXM				B-cell FCXM			
		Cut-off	% Accuracy	FP	FN	Cut-off	% Accuracy	FP	FN
1st Approach	Lowest MCF positive control	37817.28	40.00	0/10	18/20	59115.80	40.00	0/9	19/21
	Highest MCF negative control	1655.99	100.00	0/10	0/20	43668.06	36.67	1/9	18/21
2nd Approach	SD cut-off value, MCS	203.81	100.00	0/10	0/20	11109.30	60.00	2/9	10/21
	2SD cut-off value, MCS	407.63	100.00	0/10	0/20	22218.60	56.67	1/9	12/21
3rd Approach	Lowest MFS positive	1753.50	100.00	0/10	0/20	593.50	73.33	6/9	2/21
	Highest MFS negative	991.50	100.00	0/10	0/20	30996.64	40.00	0/9	18/21

Note: FP: false positive, FN: false negative.

Discussion

Flow cytometric crossmatch (FCXM) plays a critical role in preventing acute antibody-mediated rejection. The establishment of appropriate laboratory-specific cut-off values is essential to minimize false positive and false negative results. In this study, three different methodologies were evaluated for determining FCXM cut-off values using median channel fluorescence (MCF) and median channel fluorescence shift (MFS) obtained from 30 donor-recipient pairs. These approaches were compared to assess their accuracy based on the results of HLA typing and LABScreen™ PRA Class I and II for antibody typing to identify the most suitable cut-off values for routine use. The selected three cut-off methodologies were based on previously described approaches in the literature. The first approach, based on the lowest MCF value of the positive control and the highest MCF value of the negative control, was adopted from a prior undergraduate term paper conducted in our laboratory.[8] This method was selected because it demonstrated the highest level of agreement between flow cytometric crossmatch and complement-dependent cytotoxicity crossmatch results in that study. However, low accuracy with 80% specificity in T-cells and 60% specificity in B-cells was observed, therefore, alternative approaches for calculating the cut-off values were considered to achieve improved performance. The second approach with using the mean and SD of the negative control's MCF applied in this study is consistent with the methodology previously reported in Chamsai *et al.*[7] and Chulalongkorn University thesis.[9] The third approach with using median fluorescence shift (MFS) was consistent with published in de Moraes *et al.* [10]

From the results in this study, the first approach resulted in 18 false negatives out of 20 for T-cells. For B-cells, the lowest MCF value of the positive control led to 19 false negatives out of 21. Additionally, the highest MCF value of the negative control produced 1 false positive out of 9 and 18 false negatives out of 21 for B-cells, while no false positive and negative was observed for T-cells (Table 1). Comparison of sample MCF with the cut-off revealed a high rate of false negative results, likely due to the elevated background in both positive and negative control sera. Therefore, this cut-off value is not suitable for reliable use. The cut-off value determined by the second approach demonstrated a 100% match for T-cells. However, for B-cells, using MCF negative control+SD resulted in 2 false positives out of 9 and 10 false negatives out of 21, whereas MCF negative control+2SD yielded 1 false positive out of 9 and 12 false negatives out of 21 (Table 1). Comparing sample MCF with these cut-off values suggests that MCF negative control+SD is the preferable choice for this method, as it provides higher overall accuracy. All false-negative results had MCF values lower than the cut-off due to the high and variable background of the negative control. These false negative results may have occurred for several reasons, including poor washing technique leading to incomplete removal of serum immunoglobulin, poor lymphocyte purity, high background in the negative control serum, and a low serum-to-target cell ratio.[2,11,12] The third approach achieved a 100% match for T-cells. For B-cells, the lowest MFS value that reliably indicated a positive result resulted in 6 false positives out of 9 and 2 false negatives out of 21. In contrast, the highest MFS value for a negative B-cell result led to 18 false negatives out of 21 (Table 1). By comparing sample MFS with the cut-off and selecting the value with the highest accuracy, the results support using the lowest MFS that indicates a positive result as the cut-off. These two false negative results may have been influenced by a low antibody titer in the serum and low expression of loci on donor cells.[2,11] The low-expression loci present on these two donor cells included HLA-DR51, DR52, and DR53.[13,14] However, the selected cut-off value also resulted in false positive reactions for B-cells FCXM likely reflects the challenges in B-cells crossmatching sensitivity and specificity. B-cells express Fc receptors that can bind nonspecifically to immunoglobulins, which can lead to elevated background fluorescence and false positive results.[15] Additionally, the lowest MFS positive cut-off may be too low or less stringent, making it more prone to detecting weak or nonspecific binding signals that do not represent true donor-specific antibody reactions. Also, the presence of therapeutic antibodies (such as Rituximab) in the recipient's serum, autoantibodies, and the low background of the negative control serum can be the causes.[2,11,12]

In summary, the appropriate cut-off values are the lowest MFS indicating a positive result, set at 1753.50 for T-cells and 593.50 for B-cells. There should be some exceptions, though, because of false negative results, and samples must be collected continuously so that a clinically significant cut-off value can be found based on the data gathered in each lab.[6,16] The implementation of these cut-off values in a clinical setting should be carefully evaluated in collaboration with physicians. Given its direct impact on transplant rejection and patient outcomes, the importance of setting accurate cut-off values in FCXM has grown. Inappropriate threshold settings can either increase the risk of graft loss by failing to detect clinically relevant donor-specific antibodies or unnecessarily exclude potential donors by producing false positive results. Previous study has shown that recipients with positive FCXM results, particularly for HLA class I, including those with low-level antibodies, are more likely to increase risk of acute rejection and decrease graft survival rate.[4] The correlation between DSA and FCXM can vary significantly due to the difference in the cut-off value used which FCMX failed to detect DSA in samples with lower MFI (less than 2,000).[4] Accordingly, continual adjustment and validation of these cut-off values in the clinical context are essential to enhance transplant success and improve long-term graft function. However, as FCMX found to be 8-32 times more sensitive than CDC and 4-16 times more sensitive than AHG-CDC,[17] it was reported a high rate of false positive FCXM results as high as 35.9%.[4]

Therefore, appropriate conditions should be carefully evaluated to minimize false positive and false negative results. Although our FCXM protocol is based on the three-colored assay reported by Chamsai *et al.*,[7] additional optimizations, for example, the serum-cell ratio and titrated the fluorochrome-conjugated antibodies should be made to improve assay specificity and efficiency when different flow cytometer machine and settings are applied. Notably, the Halifax and Halifaster protocols employ optimized lymphocyte isolation techniques, shortened incubation times, and incorporate pronase treatment to reduce nonspecific binding in B-cells FCXM, increasing accuracy and decreasing false positive rates.[12,18-20] Additionally, some laboratories recommend using $F(ab')_2$ fragments IgG as secondary antibodies to further improve specificity.[20] False negative results can be reduced by using an alternative negative control serum, enhancing the washing technique, and adjusting the serum-to-cell volume ratio.[12] One limitation of this study is the source and selection of negative control sera. In the current study, negative control sera were pooled from 4-6 male, AB blood type, apheresis healthy donors without a history of transfusion. These sera underwent the heat inactivation procedure and were confirmed negative by LCT using cells from 10 different donors. As we used both positive and negative control in routine LCT crossmatching, we used the same control in FCXM. We realized that at least 30 different negative control sera is an ideal approach to minimize background and improve specificity. Due to limitations in sample availability, we

were unable to recruit enough of such negative control sera, which may affect the robustness of the negative control group. This limitation is acknowledged and should be addressed in future studies, where we plan to collect a larger number of negative control sera from AB blood type, non-sensitized donors confirmed negative for HLA antibodies. Notably, negative control sera should be confirmed negative, at minimum by solid phase antibody screening. Additionally, a comparative analysis of MFI values between donor-derived and patient-derived negative sera should be performed to better define baseline fluorescence and improve specificity in FCXM cut-off determination. Similarly, the positive control sera should be also confirmed broad HLA specificity and exhibit a high anti-HLA IgG titer with specific MFI range from solid phase immunoassay.

One more issue that might affect false positive and false negative in our study is selection of the positive samples. Although single antigen bead (SAB) assay provide higher resolution allele-specific antibody identification, they were not available during our study. Our study utilized LABScreen™ PRA Class I and PRA Class II for antibody detection in positive sera. These panel assays are proven methods for screening for class I, class II, or both anti-HLA antibodies. Future research would further improve antibody specificity assessment and cut-off accuracy by incorporating SAB assay and external quality assessment.

With some limitations of FCXM, alternatively, some laboratories have adopted virtual crossmatching, an immunological compatibility assessment that compares a patient's alloantibody profile with a donor's histocompatibility antigens.[2] This approach enhances antibody detection sensitivity and enables a virtual evaluation of compatibility. Additionally, it serves as a tool for deceased donor allocation, aiming to balance the likelihood of transplantation with the risk of a positive crossmatch result.[21]

Conclusion

The established cut-off values, representing the lowest MFS indicating a positive result, are 1753.50 for T-cells and 593.50 for B-cells. Ongoing sample collection and protocol optimization are essential to determine clinically significant cut-off values tailored to each laboratory's data especially applied flow cytometer machine and settings are different to the reference protocol. Implementing these cut-off values in clinical practice should be carefully evaluated in collaboration with physicians.

Conflict of interest

The authors declare no conflict of interest.

Ethical approval

This study was approved by the Research Ethics Committee, Faculty of Medicine, Chiang Mai University, Thailand (approval code: NONE-2567-0674; date of approval: 7 February 2025).

References

[1] Kasiske BL, Zeier MG, Chapman JR, Craig JC, Ekberg H, Garvey CA, et al. KDIGO clinical practice guideline for the care of kidney transplant recipients: a summary. Kidney Int. 2010; 77(4): 299-311.

[2] Paul A, Dileep RCA, Gracious N. Crossmatch and Histocompatibility. Int J Open Inf Technol. 2024. doi: 10.5772/intechopen.1006348.

[3] Fuggle SV, Taylor CJ. Chapter 10 - Histocompatibility in renal transplantation. In: Morris PJ, Knechtle SJ, Editors. Kidney transplantation. 6th Ed. Philadelphia: WB Saunders; 2008. pp: 140-57. doi: 10.1016/B978-1-4160-3343-1.50014-1.

[4] Kupatawintu P, Tatawatorn A, Premasathian N, Avihingsanon Y, Leelahavanichkul A, Hirankarn N. Association between flow cytometric cross-matching and graft survival in Thai cadaveric-donor kidney transplantation. Asian Pac J Allergy Immunol. 2016; 34(1): 86-93. doi: 10.12932/ap0569.34.1.2016.

[5] Lao-araya M. Flow cytometry in medical research and clinical applications. Biomed Sci Clin Med. 2014; 53(2): 99-109.

[6] Park Y, Kim HS. P140 Proper setting of cut-off value for the lymphocyte crossmatch by flow cytometry in kidney transplantation. Hum Immunol. 2018; 79-165. doi: 10.1016/j.humimm.2018.07.197.

[7] Chamsai S, Kittisares K, Thanaketpaisarn Y, Ngamthawornwong S, Koktathong K. Evaluation of the three-colored flow cytometric crossmatch assay in living-related kidney transplantation. J Hematol Transfus Med. 2018; 28: 239-48.

[8] Dispao T. Comparison of HLA crossmatching results between microlymphocytotoxicity test and flow cytometry in kidney transplant patients [Term paper]. Faculty of Associated Medical Science: Chiang Mai University; 2023. [in Thai].

[9] Choktaweesak N. Flow cytometric crossmatch: an alternative technique of HLA compatibility testing in deceased donor kidney transplant waited-list [Thesis]. Faculty of Allied Health Sciences: Chulalongkorn University; 2020 [in Thai].

[10] de Moraes P, Fagundes I, Cardone JM, Gil BC, Kulzer ASS, Hadi RA, et al. Accuracy of the median channel shift in the flow cytometry for predicting complement dependent cytotoxicity crossmatching in kidney transplant candidates. Transpl Immunol. 2019; 52: 27-31. doi: 10.1016/j.trim.2018.10.005.

[11] Guillaume N. Improved flow cytometry cross-matching in kidney transplantation. HLA. 2018; 92(6): 375-83. doi: 10.1111/tan.13403.

[12] Liwski RS, Gebel HM. Of cells and microparticles: assets and liabilities of HLA antibody detection. Transplantation. 2018; 102(1S): S1-S6. doi: 10.1097/tp.0000000000001818.

[13] Holdsworth R, Hurley CK, Marsh SGE, Lau M, Noreen HJ, Kempenich JH, et al. The HLA dictionary 2008: a summary of HLA-A, -B, -C, -DRB1/3/4/5, and -DQB1 alleles and their association with serologically defined HLA-A, -B, -C, -DR, and -DQ antigens. Tissue

Antigens. 2009; 73(2):95-170. doi: 10.1111/j.1399-0039.2008.01183.x.

[14] Park BG, Park Y, Joo DJ, Huh KH, Kim MS, Kim SI, et al. Clinical significance of donor-specific anti-HLA-DR51/52/53 antibodies for antibody-mediated rejection in kidney transplant recipients. Korean J Transplant. 2019; 33(3): 47-54. doi: 10.4285/jkstn.2019.33.3.47.

[15] Brian JF, Rosanne P, Cynthia DF, Tricia C, Shari S, Candy V, et al. Abrogating biologics interference in flow cytometric crossmatching. Hum Immunol. 2021; 82(8): 574-80. doi: 10.1016/j.humimm.2021.01.016.

[16] Downing J. The lymphocyte crossmatch by flow cytometry for kidney transplantation. In: Christiansen FT, Tait BD, editors. Immunogenetics: methods and applications in clinical practice. Totowa, NJ: Humana Press; 2012. p. 379-90. doi: 10.1007/ 978-1-61779-842-9_22.

[17] Koktathong K, Vejbaesya S, Bejrachandra S, Pattanapanyasat K. Flow cytometric crossmatch for kidney transplantation. J Med Assoc Thai. 2005; 88: 769-74.

[18] Sarah JH, Karl PS, Andrea AZ, Annette MJ. Impact of pronase on flow cytometric crossmatch outcome. Hum Immunol. 2011; 74(4): 330-6. doi: 10.1016/j.humimm.2011.01.005.

[19] Cláudia S, Paula A, Paula S, Paula X. Low expression loci and the use of pronase in flow cytometry crossmatch. Transplant Proc. 2023; 55(6): 1389-9.

[20] Liwski RS, Greenshields AL, Conrad DM, Murphey C, Bray RA, Neumann J, Gebel HM. Rapid optimized flow cytometric crossmatch (FCXM) assays: the Halifax and Halifaster protocols. Hum Immunol. 2018; 79(1): 28-38. doi: 10.1016/j.humimm. 2017.10.020.

[21] Moszkowska G, Zieliński M, Zielińska H, Dukat-Mazurek A, Dębska-Zielkowska J, Dębska-Ślizień A, *et al*. Evaluation of pretransplant donor-specific alloantibodies with different crossmatch techniques. Transplant Proc. 2018: 50(6): 1625-30. doi: 10.1016/j.transproceed.2018.02.165.

Journal of Associated Medical Sciences 2026; 59(2): 9-14

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Case report: Two cases of Hb Liuzhou-Yufeng (*HBA1*:c.334G>T) found in Northern Thailand

Yasmina Azfa Putri Alif[1], Maulida Syifa Kamila Makarim[1], Chedtapak Ruengdit[2], Yona Mimanda[1], Chris Adhiyanto[1], Laifa Annisa Hendarmin[1], Sakorn Pornprasert[2*]

[1]Faculty of Medicine, Syarif Hidayatullah Jakarta Islamic State University, Indonesia.
[2]Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 25 October 2025
Accepted as revised 22 December 2025
Available online 9 January 2026

Keywords:
Hemoglobin Liuzhou-Yufeng; α-globin variant; thalassemia; next-generation sequencing; Northern Thailand.

ABSTRACT

Background: Hemoglobin (Hb) Liuzhou-Yufeng is an α-globin variant caused by a heterozygous *HBA1* mutation (c.334G>T). However, its clinical significance is not well-defined. Moreover, it is hard to detect using standard capillary electrophoresis (CE) or high-performance liquid chromatography (HPLC) methods and often need molecular techniques like next-generation sequencing (NGS).

Objective: This study aims to describe genotype and hematological features of the first two cases with heterozygosity for Hb Liuzhou-Yufeng found in Northern Thailand and the diagnostic challenges faced.

Materials and methods: Two patients with anemia, a 51-year-old female and a 32-year-old male, underwent hematological evaluation, Hb analysis using CE for case 1 and HPLC for case 2, iron studies, and red cell morphology tests. The dichlorophenolindophenol (DCIP) and osmotic fragility (OF) tests were also conducted. The molecular work-up included multiplex real-time PCR for α-thalassemia deletions, followed by NGS.

Results: Both patients were confirmed by the NGS to carry Hb Liuzhou-Yufeng ($\alpha\alpha^{LY}/\alpha\alpha$). Their total Hb levels were found within the ranges of mild anemia (10-12 g/dL). However, they had normocytic and normochromic red cells, and the morphology of the red blood cells appeared to be normal. In addition, the Hb analysis by HPLC and CE methods revealed normal peaks and levels of HbA, HbA_2, and HbF.

Conclusion: Hb Liuzhou-Yufeng seems clinically harmless but is challenging to detect using conventional methods. This study highlights the importance of NGS in finding rare Hb variants. It improves diagnostic accuracy, genetic counseling, and personalized management of hemoglobinopathy.

Introduction

Inherited hemoglobin (Hb) disorders rank among the most prevalent genetic conditions globally and are generally divided into thalassemia (deficiencies in globin chain synthesis) and structural hemoglobinopathies (defects caused by abnormal globin chains).[1,2] In Southeast Asia, α-thalassemia is notably common, with carrier rates reaching 30-40% in Thailand, particularly in the northern areas. Most cases result from α-globin gene deletions like the $--^{SEA}$ and $-\alpha^{3.7}$ deletions, but non-deletional variants such as Hb

*** Corresponding contributor.**
Author's Address: *Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: sakornmi001@gmail.com
doi: 10.12982/JAMS.2026.032
E-ISSN: 2539-6056

Constant Spring (HbCS, *HBA2*:c.427T>C) and Hb Paksé (HbPS, *HBA2*:429A>T) are also prevalent.[3–5] Although regular hematological assessments and Hb analysis by capillary electrophoresis (CE) or high-performance liquid chromatography (HPLC) detect common variants, they often miss rare mutations. Consequently, molecular techniques, such as next-generation sequencing (NGS), are necessary for precise diagnosis, particularly in unusual cases.[4]

A new α-globin variant, Hb Liuzhou, was initially reported in China and results from a point mutation in *HBA1* (c.182A>G), causing a lysine-to-arginine at codon 60. It was named according to the birthplace of the patient.[6] Carriers usually show no symptoms and have normal blood indices.[6,7] Later, a different yet associated variant, Hb Liuzhou-Yufeng (*HBA1*:c.334G>T; Ala→Ser), was discovered. The "Yufeng" is one of four districts in Liuzhou, a prefecture-level city within the Guangxi Zhuang Autonomous Region of China; hence, the Hb Liuzhou-Yufeng was named according to the geographic origin of the initial case. Although Hb Liuzhou is recognized as clinically silent, the clinical significance of Hb Liuzhou-Yufeng remains uncertain due to the scant literature available.[8] Thus, we present the initial two cases of Hb Liuzhou-Yufeng identified in Chiang Mai, Northern Thailand. These instances highlight the diagnostic difficulties presented by this uncommon variant and stress the significance of molecular testing in screening for thalassemia and hemoglobinopathy.

Case Presentation

Case 1: A 51-year-old female patient presented with vertigo persisting for four days, accompanied by generalized weakness. Complete blood count revealed Hb concentration of 11.4 gm/dL, mean corpuscular volume (MCV) of 86.0 fL, and mean corpuscular Hb (MCH) of 28.0 pg, indicating mild normocytic normochromic anemia. The morphology of the red blood cells appeared to be normal, and the findings from both the DCIP and OF tests were negative. Further Hb analysis through CE (Capillarys 2 Flex piercing, Sebia, Evry, France) was performed, yielding results within the normal range (Figure 1A and Table 1), thereby excluding a β-thalassemia trait. However, a potential α-thalassemia syndrome was still suspected. Subsequent to these findings, molecular testing was conducted using a single-tube multiplex real-time PCR with Evagreen and high-resolution melting (HRM) analysis to identify common α^0-thalassemia deletions, specifically $--^{SEA}$, $--^{THAI}$, and $--^{CR}$ deletions.[7] In addition, the common α^+-thalassemia $-\alpha^{3.7}$ and $-\alpha^{4.2}$ deletions were investigated through a multiplex Gap-PCR technique.[8] The patient tested negative for these mutations. Following the negative PCR results, further genetic analysis employing NGS was performed. The NGS panel designed to target the coding regions of the *HBA1*, *HBA2*, and *HBB* genes. This assay utilized the Thalassemia Gene Detection Kit provided by BGI Group (Shenzhen, China). Subsequently, sequence reads were mapped against the human genome reference hg19, and detected genetic variants were analyzed using THACARE HALOS software (BGI Group). It revealed a mutation from GCC to TCC at codon 111 of the α1-globin gene, identified as Hb Liuzhou-Yufeng variant (Figure 2).

Case 2: A 32-year-old HIV infected male came in with weakness and nausea for 24 hours. His Hb level was 10.7 gm/dL, showing mild anemia. The MCV was 94.0 fL and MCH was 31.9 pg, indicating normocytic normochromic anemia. The red cell distribution width (RDW) was 16.9%. Red cell appearance was normal, and both DCIP and OF tests were negative. Iron studies indicated elevated ferritin levels, decreased serum iron and total iron-binding capacity (TIBC), and normal transferrin saturation (Table 1), so refusing iron deficiency anemia. The Hb analysis by HPLC (VARIANT β-thalassemia Short Program; Bio-Rad Laboratories, Hercules, CA, USA) was normal (Figure 1B and Table 1), thereby excluding a β-thalassemia trait. However, a potential α-thalassemia syndrome was still suspected. Subsequent to these findings, molecular testing was conducted using a single-tube multiplex real-time PCR with Evagreen and HRM analysis to identify common α^0-thalassemia deletions, specifically $--^{SEA}$, $--^{THAI}$, and $--^{CR}$ deletions.[7] In addition, the common α^+-thalassemia $-\alpha^{3.7}$ and $-\alpha^{4.2}$ deletions were investigated through a multiplex Gap-PCR technique.[8] The patient tested negative for these mutations. Later, the NGS revealed a mutation from GCC to TCC at codon 111 of the α1-globin gene, identified as Hb Liuzhou-Yufeng variant.

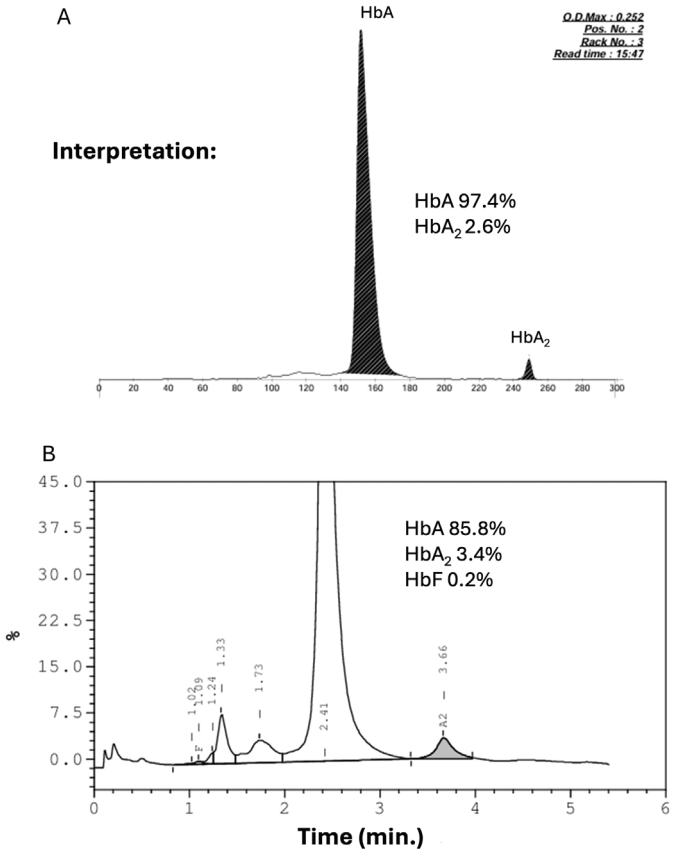


Figure 1. *CE electropherogram and HPLC chromatogram without seeing the separate peak of Hb Liuzhou-Yufeng. A: case 1, B: case 2.*

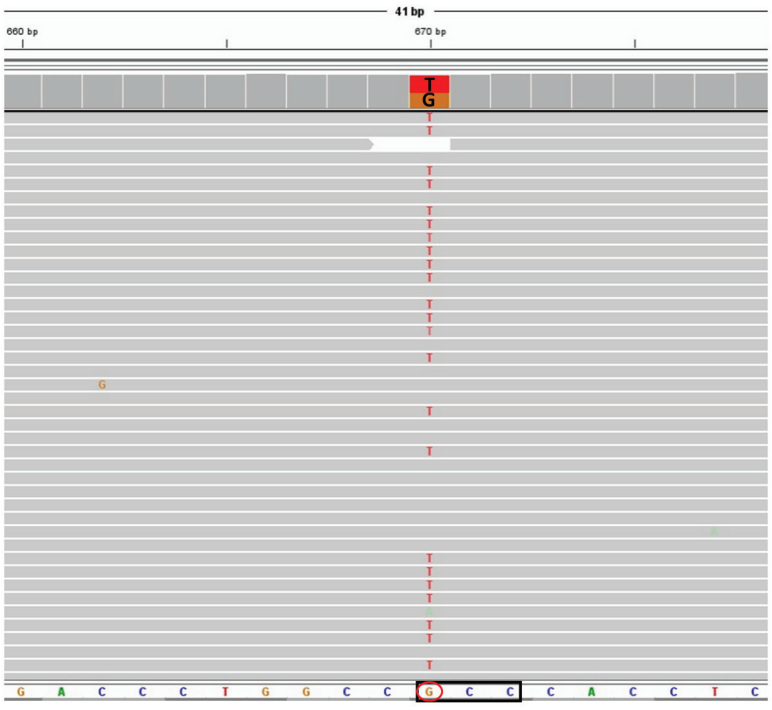


Figure 2. *Representative of NGS results exported from the integrative genomics viewer (IGV) of Hb Liuzhou-Yufeng (HBA1:c.334G>T).*

Table 1. The characteristics and hematological parameters of cases

Characteristics and hematological parameters	Case No. 1	Case No. 2
Age (Years)	51	32
Gender	Female	Male
α-globin genotype	$\alpha\alpha^{LY}/\alpha\alpha$	$\alpha\alpha^{LY}/\alpha\alpha$
β-globin genotype	β^A/β^A	β^A/β^A
RBCs ($x10^{12}$/L)	4.13	3.4
Total Hb (gm/dL)	11.4	10.7
PCV (L/L)	0.35	0.32
MCV (fL)	86.0	94.0
MCH (pg)	28.0	31.0
MCHC (gm/L)	325	334.4
RDW (%)	15.3	16.9%
Ferritin (ng/mL)	N/A	2,020
Serum iron (µg/dL)	N/A	45
TIBC (µg/dL)	N/A	175
Transferrin saturation (%)	N/A	25.7
HbA (%)	97.4†	85.8*
HbA_2 (%)	2.6†	3.4*
HbF (%)	0.0†	0.2*

Note: *Normal range of adults: red blood cell counts (RBCs) 4.2-6.1 x 10^{12}/L, total Hb 12.0-18.0 gm/dL, packed cell volume (PCV) 0.37-0.52 L/L, mean corpuscular volume (MCV) 80-100 fL, mean corpuscular Hb (MCH) 27.0-31.0 pg, mean corpuscular Hb concentration (MCHC) 320-360 gm/L, red cell distribution width (RDW) 11.5-15.0%, HbA 95.0-98.0%, HbA2 1.5-3.5%, HbF 0.0-1.0%, Ferritin 30-300 ng/mL, Serum iron 50-150 µg/dL, total iron-binding capacity (TIBC) 240-450 µg/dL, Transferrin saturation 20-55%. †Hb analysis by capillary electrophoresis (CE), *Hb analysis by high-performance liquid chromatography (HPLC)*

Discussion

The α-globin genes, *HBA1* and *HBA2*, are located on chromosome 16p13.3. They encode nearly identical α-globin chains. *HBA2* produces about three times more transcripts than *HBA1* in red blood cells. Changes in these genes cause α-thalassemia, with deletions being more common. Non-deletional mutations in *HBA2* typically lead to more severe symptoms. This is because disrupting the higher-expressing gene significantly reduces α-globin production. On the other hand, changes in *HBA1* usually result in milder effects due to its lower baseline expression. [9,10]

In both individuals in the current study, mild to moderate Hb reduction was noted along with normocytic normochromic anemia. The significantly elevated ferritin level (2,020 ng/mL) in Case 2 may be attributed to HIV infection, as HIV induces chronic inflammation, leading to an increase in ferritin levels. Hb analysis through CE (case 1) and HPLC (case 2) revealed no Hb variants. Molecular testing excluded significant α-thalassemia deletions ($--^{SEA}$, $--^{THAI}$, and $--^{CR}$). The NGS revealed an identical point mutation in the *HBA1* gene (c.334G>T), leading to the Hb Liuzhou-Yufeng variant. The substitution from alanine (nonpolar) to serine (polar) at codon 111 may subtly affect Hb structure and oxygen affinity, potentially accounting for mild symptoms.[11,12] While this mutation doesn't alter Hb levels identifiable through Hb typing and often results in relatively mild anemia, the subtle clinical symptoms noted may be attributed to the variant's influence on the oxygen affinity of Hb. Its placement in *HBA1*, a gene with reduced expression relative to *HBA2*, also contributes to the attenuated severity.

Research from Northern Thailand has shown that combined α-thalassemia mutations can produce a wide range of clinical and hematological phenotypes, often influenced by coexisting variants such as HbE or α-thalassemia deletions.[13] For example, the silent Hb

Hekinan II variant (*HBA1*:c.84G>T) remains clinically harmless unless co-inherited with other α-thalassemia mutations.[14] Currently, there are no published reports of combined mutations involving Hb Liuzhou-Yufeng. Because this variant locates in the less-expressed *HBA1* gene, its pathogenic effect is usually mild. However, when co-inherited with α-thalassemia deletions (e.g., α^{+}-thalassemia), it could present as a mild phenotype characterized by slight anemia or thalassemia features, reflecting the altered balance between α- and β-globin chain synthesis.

From a diagnostic perspective, these cases emphasize the limitations of Hb analysis methods like CE and HPLC, which showed normal outcomes even with the presence of the Hb Liuzhou-Yufeng variant. This underlines the risk of underdiagnosis if laboratories rely solely on routine Hb analysis. Diagnostic workflows should incorporate the NGS and other molecular tools, especially in areas with a high occurrence of thalassemia and hemoglobinopathies, to guarantee precise identification of rare variants. Therefore, it is recommended that NGS be performed on all anemic samples, even when red blood cells are normocytic and normochromic, and common causes of anemia remain undetected. Accurate diagnosis is crucial for effective genetic counseling, family risk evaluation, and guiding public health initiatives for thalassemia management.

Conclusion

This research demonstrated the effective application of a targeted NGS panel for the identification of rare α1-globin gene mutation, Hb Liuzhou-Yufeng. Clinically, individuals carrying Hb Liuzhou-Yufeng showed few or no symptoms, reflecting a benign clinical course, which highlights the difficulty of diagnosing this variant through clinical evaluation or conventional screening methods alone. Therefore, integrating advanced molecular techniques such as NGS into thalassemia screening programs is essential for the accurate identification of rare Hb variants. This approach not only enhances diagnostic precision but also supports improved genetic counseling and personalized clinical management, ultimately leading to better outcomes for individuals affected by hemoglobinopathies.

Ethical approval

This study was approved by the Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University, Thailand (AMSEC-68EM-014).

Conflict of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

CRediT authorship contribution statement

Yasmina Azfa Putri Alif: conceptualization, investigation, writing original draft; **Maulida Syifa Kamila Makarim:** conceptualization, investigation, writing original draft; **Chedtapak Ruengdit:** blood and data collections, investigation, methodology, review and editing; **Yona Mimanda:** conceptualization, review and editing; **Chris Adhiyanto:** conceptualization, review and editing; **Laifa Annisa Hendarmin:** conceptualization, review and editing; **Sakorn Pornprasert:** conceptualization, project administration, validation, writing, reviewing and editing.

Acknowledgements

The authors thank the technicians at Associated Medical Sciences-Clinical Service Center, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand for their assistance.

References

[1] Leckngam P. Thalassemia and hemoglobinopathies in Thailand: A systematic review. J Health Sci Altern Med. 2023; 5(3): 104-13. doi: 10.14456/jhsam.2023.24.

[2] WHO. Regional desk review of haemoglobinopathies with an emphasis on thalassaemia and accessibility and availability of safe blood and blood products as per these patients' requirement in South-East Asia under universal health coverage [Internet]. 2021. Available from: http://apps.who.int/bookorders.

[3] Jomoui W, Panyasai S, Sripornsawan P, Tepakhan W. Revisiting and updating molecular epidemiology of α-thalassemia mutations in Thailand using MLPA and new multiplex gap-PCR for nine α-thalassemia deletion. Sci Rep. 2023; 13(1): 9850. doi: 10.1038/s41598-023-36840-8.

[4] Musallam KM, Cappellini MD, Coates TD, Kuo KHM, Al-Samkari H, Sheth S, *et al*. Alpha-thalassemia: A practical overview. Blood Rev. 2024; 64: 101165. doi: 10.1016/j.blre.2023.101165.

[5] Paiboonsukwong K, Jopang Y, Winichagoon P, Fucharoen S. Thalassemia in Thailand. Hemoglobin. 2022; 46(1): 53-7. doi:10.1080/03630269.2022.2025824.

[6] Xu A, Chen W, Xie W, Ji L. Identification of a new hemoglobin variant Hb Liuzhou [HBA1:C.182A→G] by MALDI-TOF mass spectrometry during HbA1c measurement. Scand J Clin Lab Invest. 2020; 80(6): 479-83. doi: 10.1080/00365513.2020.1783698.

[7] Ruengdit C, Punyamung M, Intasai N, Pornprasert S. Single-tube multiplex real-time PCR with EvaGreen and high-resolution melting analysis for diagnosis of α0-thalassemia--SEA,--THAI, and--CR type deletions. PLoS One. 2023; 18(11): e0293838. doi: 10.1371/journal.pone.0293838.

[8] Fucharoen S, Fucharoen G, Sanchaisuriya K, Pengjam Y. Molecular analysis of a Thai β-thalassaemia heterozygote with normal haemoglobin A2 level: implication for population screening. Ann Clin Biochem. 2002; 39(1): 44-9.

doi: 10.1258/0004563021901720.

[9] Traeger-Synodinos J, Vrettou C, Sofocleous C, Zurlo M, Finotti A, Gambari R. Impact of α-globin gene expression and α-globin modifiers on the phenotype of β-thalassemia and other hemoglobinopathies: Implications for patient management. Int J Mol Sci. 2024; 25(6): 3400. doi: 10.3390/ijms25063400.

[10] Vijian D, Wan Ab Rahman WS, Ponnuraj KT, Zulkafli Z, Bahar R, Yasin N, et al. Gene mutation spectrum among alpha-Thalassaemia patients in Northeast Peninsular Malaysia. Diagnostics. 2023; 13(5): 894. doi: 10.3390/diagnostics13050894.

[11] Knappenberger JA, Kuriakose SA, Vu BC, Nothnagel HJ, Vuletich DA, Lecomte JTJ. Proximal influences in two-on-two globins: Effect of the Ala69Ser replacement on synechocystis sp. PCC 6803 hemoglobin. Biochemistry. 2006; 45(38): 11401-13. doi: 10.1021/bi060691x.

[12] Narayanan S, Mathew B, Srinivasu BY, Bhat V, Ross C, Mandal AK. Effect of point mutation on structure–function correlation of hemoglobin variants, HbE and HbD Punjab. Amino Acids. 2020; 52(6-7): 893-904. doi: 10.1007/s00726-020-02858-9.

[13] Pornprasert S, Ruengdit C, Punyamung M, Sripichai O. Alpha-Thalassemia caused by ααIVSI-1(AGGT> AGAT) (HBA1: c.95 + 1G > A) Mutation and its combinations with other forms of thalassemia or hemoglobinopathy in Northern Thailand. Indian J Hematol Blood Transfus. 2025; 41(3): 674-9. doi: 10.1007/s12288-024-01895-8.

[14] Panyasai S, Chantanaskulwong P, Permsripong N, Mokmued T. Interactions of electrophoretically silent hemoglobin Hekinan II [HBA1:c.84G>T] with various forms of α-thalassemias and other hemoglobinopathies: novel insights into the molecular and hematological characteristics and genetic origins. Libyan J Med. 2024; 19(1): 2406620. doi: 10.1080/19932820.2024.2406620.

Journal of Associated Medical Sciences 2026; 59(2): 15-23

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Prevalence and risk factors of elevated urinary albumin-to-creatinine ratio in a Thai academic population

Suwatsin Kittikunnathum, Pharisa Nanthawong, Piyawan Bunpo*

Division of Clinical Chemistry, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 17 September 2025
Accepted as revised 16 December 2025
Available online 9 January 2026

Keywords:
Urinary albumin-to-creatinine ratio, microalbuminuria, obesity, diabetes, systemic lupus erythematosus, chronic kidney disease.

ABSTRACT

Background: The urinary albumin–to–creatinine ratio (UACR) is a key marker for early kidney dysfunction. This study assessed UACR levels and their associations with demographic and clinical factors in a Thai university population.

Materials and methods: A cohort of 158 participants (53 males, 105 females) was recruited from Chiang Mai University between January and May 2024. Urinary albumin and creatinine were measured by turbidimetric immunoassay and creatininase method, respectively, and expressed as UACR (mg/gm creatinine). UACR was categorized as normoalbuminuria (<30 mg/gm creatinine), microalbuminuria (30-299 mg/gm creatinine), or clinical albuminuria (≥300 mg/gm creatinine). Data on body mass index (BMI), blood pressure, fasting blood glucose, physical activity, use of supplements or medications, and the presence of chronic conditions were also collected.

Results: Most participants (94.3%) had normoalbuminuria, whereas microalbuminuria and clinical albuminuria were observed in 4.4% and 1.3% of participants, respectively. Overall, 71.5% were normotensive and 28.5% hypertensive. Based on BMI, 41.1% had normal weight, 18.4% were overweight, and 34.8% were obese. Nearly half reported no regular exercise, and the majority were non–smokers (98.1%) and non–drinkers (83.5%). Regular supplement or medication use was reported by 57.0%, mainly vitamins/minerals and fish oil. Chronic diseases occurred in 37.3%, with hypercholesterolemia (10.8%) and hypertension (8.9%) being most common. Hypercholesterolemia, hypertension, allergy, type II diabetes, lifestyle factors, and supplement or medication use were not associated with albuminuria (Chi-square, $p>0.05$), whereas obesity class II, elevated fasting blood glucose (≥126 mg/dL) and systemic lupus erythematosus (SLE) were significantly associated with increased urinary albumin excretion (Chi-square, $p<0.05$).

Conclusion: Elevated blood glucose, obesity, and SLE were significantly linked to increased UACR, emphasizing the importance of early identification and management of these risk factors to prevent kidney disease progression.

Introduction

Chronic kidney disease (CKD) is a significant global health burden, with its prevalence increasing due to the rising incidence of metabolic disorders such as obesity, diabetes mellitus, and hypertension.[1] Early detection of renal dysfunction is crucial for preventing disease progression, and one of the most widely used markers for early kidney damage is the urinary albumin–to–creatinine ratio (UACR).[2-4] Microalbuminuria,

* **Corresponding contributor.**
Author's Address: *Division of Clinical Chemistry, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: piyawan.b@cmu.ac.th
doi: 10.12982/JAMS.2026.033
E-ISSN: 2539-6056

an intermediate stage between normoalbuminuria and clinical albuminuria, serves as an important indicator of endothelial dysfunction and renal impairment, particularly among individuals with metabolic syndrome and cardiovascular risk factors.[5]

Evidence on UACR prevalence in healthy Thai adults is lacking, as existing studies are limited to clinic–based or high–risk populations. In diabetic cohorts, microalbuminuria prevalence ranged from 26-27% in hospital and multi–center studies.[6,7] Among nondiabetic hypertensive patients, 16.6% had microalbuminuria.[8] A large community hospital review (N=36,172) applying KDIGO criteria reported CKD in 8.2% of adults, with UACR identifying early CKD cases beyond eGFR detection.[9] In community screening of high risk adults, albuminuria or reduced eGFR was found in 20.3% of participants.[10] These findings highlight the absence of data on UACR in healthy Thai populations. In Thailand, based on patient reports from hospitals under the Ministry of Public Health, there were a total of 1,007,251 patients diagnosed with CKD across stages 1 to 5 in 2021.[11] This highlights the urgent need for early detection and intervention to mitigate disease progression and reduce healthcare burdens. In response to this growing concern, the present research was conducted from January to May 2024 at Chiang Mai University's Faculty of Associated Medical Sciences to investigate key metabolic risk factors associated with CKD.

Obesity has been strongly linked to kidney dysfunction, with evidence suggesting that excess adiposity contributes to glomerular hyperfiltration, inflammation, and renal fibrosis.[12] Furthermore, elevated fasting blood sugar levels and type II diabetes mellitus are well–documented risk factors for nephropathy, leading to progressive albuminuria and eventual decline in renal function.[13,14] While hypercholesterolemia is a recognized contributor to cardiovascular disease, its direct impact on renal impairment remains controversial, with some studies failing to establish a significant correlation between dyslipidemia and microalbuminuria.[15]

The present study investigates the relationship between metabolic risk factors and renal dysfunction, as measured by UACR, in a sample population. Specifically, it examines the associations between obesity, diabetes, hypercholesterolemia, and other lifestyle factors with albuminuria status. By identifying key predictors of renal dysfunction, the study aims to provide insight into the early detection and prevention of CKD.

Materials and methods

Study participants

This study, conducted from January to May 2024 at the Faculty of Associated Medical Sciences, Chiang Mai University, included 158 participants (both male and female) aged 20 years and older. The inclusion criteria required participants to be in stable health without any acute illness at the time of enrollment. Prior to participation, all individuals received a comprehensive briefing regarding the study's objectives and methodology. Informed consent was obtained from each participant. Data collection included a range of biometric parameters, such as gender, age, height, weight, body mass index (BMI), blood glucose levels, physical activity, smoking and alcohol consumption habits, medication adherence, blood pressure, and any existing congenital conditions. Participants were instructed to avoid excessive fluid intake within 12 hours before sample collection and to refrain from physical exercise immediately prior to providing urine samples to minimize potential effects on urinary measurements. The urine collection process involved discarding the initial portion of the sample and retaining 10-15 milliliters of the midstream sample for analysis. Samples were stored and analyzed within 1-2 hours of collection.

Ethical considerations

The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Research Committee from the Faculty of Associated Medical Sciences, Chiang Mai University (AMSEC-66EX-046).

Laboratory measurements

Urinary albumin concentrations were measured using a turbidimetric immunoassay (BA400, BioSystems S.A., Barcelona, Spain). Urinary creatinine was measured using the creatininase method with the BA400 automated system (BioSystems S.A., Spain), and results were expressed as the UACR in mg/gm of creatinine. Microalbuminuria was defined according to the American Diabetes Association and the National Kidney Foundation guidelines, where UACR values ranged from 30 to 299 mg/gm creatinine. Reagents, calibrators, and controls were sourced from BioSystems S.A., Barcelona, Spain.

Statistical analysis

At the end of the study, there were 158 participants. Data were processed using the standard statistical software SPSS 17.0 (SPSS Software, Thailand). Categorical variables were analyzed using the Chi-square test, with statistical significance defined as $p<0.05$. Continuous variables were summarized as means±SD.

Results

Demographic and clinical characteristics of the study population

The study included 158 participants, comprising 53 males (33.5%) and 105 females (66.5%), aged between 20 and 64 years. Most participants were within the age ranges of 25.0-34.9 years (33.5%) and 35.0-44.9 years (32.3%), while smaller proportions were aged

45.0-54.9 years (23.4%), 55.0-64.9 years (9.5%), and 20.0-24.9 years (1.3%). Normotension was observed in 71.5% of individuals, while 28.5% had hypertension (systolic blood pressure >130 mmHg) (Table 1).

Table 1. Demographic and clinical features of the study population (N=158).

Parameter	Frequency	Percentage
Gender		
Male	53	33.5
Female	105	66.5
Age (years)		
20.0-24.9	2	1.3
25.0-34.9	53	33.5
35.0-44.9	51	32.3
45.0-54.9	37	23.4
55.0-64.9	15	9.5
Blood pressure		
Normotension	113	71.5
Hypertension (>130 mmHg systolic)	45	28.5
BMI (kg/m^2)		
Underweight (<18.5)	9	5.7
Normal (18.5-22.9)	65	41.1
Overweight (23.0-24.9)	29	18.4
Obese I (25.0-29.9)	40	25.3
Obese II (≥ 30.0)	15	9.5
Fasting blood sugar (mg/dL)		
<126	76	48.1
≥126	2	1.3
Not Determined	80	50.6
Exercise		
No	78	49.4
1-2 times/week	40	25.3
3-5 times/week	35	22.1
6-7 times/week	5	3.2
Current smoking		
No	155	98.1
Yes	3	1.9
Alcohol consumption		
No	132	83.5
1-2 times/week	18	11.4
3-5 times/week	5	3.2
6-7 times/week	3	1.9

Regarding body mass index (BMI), 5.7% of participants were underweight (<18.5 kg/m^2), 41.1% had normal weight (18.5-22.9 kg/m^2), 18.4% were overweight (23.0-24.9 kg/m^2), 25.3% had obesity class I (25.0-29.9 kg/m^2), and 9.5% had obesity class II (≥30.0 kg/m^2). Among participants, 48.1% had fasting blood glucose levels <126 mg/dL, while 1.3% had levels ≥126 mg/dL. However, 50.6% of participants did not have fasting glucose measurements available, representing a substantial proportion of missing data.

Physical activity levels varied, with 49.4% reporting no exercise, 25.3% exercising 1-2 times per week, 22.1% engaging in physical activity 3-5 times per week, and 3.2% exercising daily. Smoking prevalence was low, with 98.1% being non–smokers. Similarly, alcohol consumption was minimal, with 83.5% reporting no alcohol intake, 11.4% drinking 1-2 times per week, 3.2% consuming alcohol 3-5 times per week, and 1.9% drinking daily.

Medical and supplement use

A total of 57.0% of participants regularly consumed supplements or medications. The most used supplements included vitamins and minerals (35.4%), fish oil (7.0%), and collagen/fiber supplements (3.8%). Other substances, such as anti-hypertensive drugs (7.6%), lipid-reducing medications (2.5%), and diabetes medications (1.3%), were also reported (Table 2).

Table 2. Supplement or regular medicine use among the study population (N=158).

Parameter	Frequency	Percentage
Supplement or regular medicine		
No	68	43.0
Yes	90	57.0
- Vitamin and mineral	56	35.4
- Fish oil	11	7.0
- Collagen and fiber	6	3.8
- L–Carnitine	1	0.6
- Glutathione	1	0.6
- Probiotics and Prebiotics	3	1.9
- Protein supplement	4	2.5
- Glucosamine	2	1.3
- Medicinal plant	8	5.1
- Birth control pill	1	0.6
- Antacids	1	0.6
- Migraine medication	1	0.6
- Anti-inflammatory drug	1	0.6
- Anti-hypertensive drug	11	7.0
- Anti-allergic drug	1	0.6
- Thyroid drug	4	2.5
- Lipid reducing drug	12	7.6
- Uric acid reducing drug	1	0.6
- Anti-viral drug	1	0.6
- Adrenal grand drug	1	0.6
- Steroid drug	2	1.3
- Diabetes medication	1	0.6

Prevalence of chronic diseases

Among the participants, 37.3% had at least one chronic disease, with hypercholesterolemia (10.8%) and hypertension (8.9%) being the most prevalent. Other reported conditions included type II diabetes (2.5%), hyperuricemia (1.9%), hepatitis B virus infection (1.9%), and autoimmune diseases such as systemic lupus erythematosus (SLE) (0.6%) (Table 3).

Table 3. Prevalence of chronic diseases in the study population (N=158).

Parameter	Frequency	Percentage
Chronic disease		
No	99	62.7
Yes	59	37.3
- Diabetes type II	4	2.5
- Hypertension	14	8.9
- Hypercholesterolemia	17	10.8
- G6PD deficiency	3	1.9
- Allergy	14	8.9
- Kidney tumor	1	0.6
- Hyperuricemia	3	1.9
- Asthma	2	1.3
- Migraine	2	1.3
- Hepatitis B virus infection	3	1.9
- Cervical spondylosis	1	0.6
- Liver fibrosis	1	0.6
- Adrenal grand enlargement	1	0.6
- Systematic lupus erythematosus (SLE)	2	1.3
- Gastric ulcer	1	0.6
- Autoimmune disease	1	0.6
- Tonsillectomy	1	0.6
- Beta-thalassemia trait	1	0.6
- Irritable bowel syndrome	1	0.6

Table 3. Prevalence of chronic diseases in the study population (N=158) (continued).

Parameter	Frequency	Percentage
- Acid reflux	1	0.6
- Hyperthyroidism	1	0.6
- Hypothyroidism	2	1.3

Urinary albumin-to-creatinine ratio and kidney function

Analysis of UACR levels revealed that 94.3% of participants had normoalbuminuria (0.3-28.5 mg/gm creatinine), while 4.4% exhibited microalbuminuria (35.7-190.6 mg/gm creatinine), and 1.3% had clinical albuminuria (576.8-670.0 mg/gm creatinine) (Table 4).

Table 4. Distribution of albuminuria categories based on UACR levels in spot urine samples (N=158).

Category	UACR ratio (mg albumin/gm creatinine)	Mean mg albumin/gm creatinine (range)	Spot urine N (%)
Normoalbuminuria	<30	6 (0.3-28.5)	149 (94.3)
Microalbuminuria	30-299	113 (35.7-190.6)	7 (4.4)
Clinical albuminuria	≥300	623 (576.8-670.0)	2 (1.3)

Comparison of normoalbuminuria and microalbuminuria/ clinical albuminuria groups

Participants with microalbuminuria or clinical albuminuria (N=9, 5.7%) were compared to those with normoalbuminuria (N=149, 94.3%). The mean age was slightly higher in the microalbuminuria/clinical albuminuria group (mean 42.5 years, range 30.2-60.9) compared with the normoalbuminuria group (mean 40.1 years, range 22.8-59.6), however, this difference was not statistically significant (p>0.05) (Table 5).

Table 5. Comparison of demographic and clinical parameters between normoalbuminuria and elevated albuminuria groups (N=158).

Parameter	Normoalbuminuria	Microalbuminuria and Clinical albuminuria	*p* value
Number	149 (94.3)	9 (5.7)	
Age (years) (range)	Mean = 40.1 (22.8-59.6)	Mean = 42.5 (30.2-60.9)	>0.05
Sex			
Male (%)	51 (32.3)	2 (1.3)	>0.05
Female (%)	98 (62.0)	7 (4.4)	>0.05
BMI (kg/m²)			
Underweight (%)	9 (5.7)	0	>0.05
Normal (%)	62 (39.2)	3 (1.9)	>0.05
Overweight (%)	29 (18.4)	0	>0.05
Obese I (%)	37 (23.4)	3 (1.9)	>0.05
Obese II (%)	12 (7.6)	3 (1.9)	<0.05
Blood pressure			
Normotension (%)	107 (67.7)	6 (3.8)	>0.05
Hypertension (%)	42 (26.6)	3 (1.9)	>0.05
Fasting blood sugar (mg/dL)			
<126 (%)	70 (44.3)	6 (3.8)	>0.05
≥126 (%)	1 (0.6)	1 (0.6)	<0.05
Not Determined (%)	78 (49.4)	2 (1.3)	>0.05
Exercise			
No (%)	71 (44.9)	7 (4.4)	>0.05
1-2 times/week (%)	40 (25.3)	0	>0.05
3-5 times/week (%)	33 (20.9)	2 (1.3)	>0.05
6-7 times/week (%)	5 (3.2)	0	>0.05
Current smoking			
No (%)	146 (92.4)	9 (5.7)	>0.05
Yes (%)	3 (1.9)	0	>0.05

Table 5. Comparison of demographic and clinical parameters between normoalbuminuria and elevated albuminuria groups (N=158) (continued).

Parameter	Normoalbuminuria	Microalbuminuria and Clinical albuminuria	*p* value
Alcohol consumption			
No (%)	124 (78.5)	8 (5.1)	>0.05
1-2 times/week (%)	18 (11.4)	0	>0.05
3-5 times/week (%)	4 (2.5)	1 (0.6)	>0.05
6-7 times/week (%)	3 (1.9)	0	>0.05
Supplement or regular medicine			
No (%)	64 (40.5)	4 (2.5)	>0.05
Yes (%)	85 (53.8)	5 (3.2)	>0.05
Chronic disease			
No (%)	94 (59.5)	5 (3.2)	>0.05
Yes (%)	55 (34.8)	4 (2.5)	<0.05
- Diabetes type II (%)	3 (1.9)	1 (0.6)	>0.05
- Hypercholesterolemia (%)	16 (10.1)	1 (0.6)	>0.05
- Allergy (%)	13 (8.2)	1 (0.6)	>0.05
- SLE (%)	2 (1.3)	1 (0.6)	<0.05

Sex distribution did not differ significantly between the groups (p>0.05). However, a significantly higher proportion of individuals with microalbuminuria/clinical albuminuria were classified as obese class II (p<0.05). No significant associations were found between hypertension, smoking, alcohol consumption, or physical activity and albuminuria status (p>0.05). Participants with fasting blood glucose ≥126 mg/dL were significantly more likely to present with microalbuminuria or clinical albuminuria compared to those with lower glucose levels (p<0.05).

Participants with chronic diseases such as hypercholesterolemia, type II diabetes, allergy, and other conditions showed no significant association with albuminuria (p>0.05). In contrast, systemic lupus erythematosus (SLE) was significantly associated with increased urinary albumin excretion (p<0.05).

Discussion

The findings of this study provide important insights into urinary albumin excretion patterns within a cohort of 158 participants from the Faculty of Associated Medical Sciences, Chiang Mai University. The measurement of UACR serves as a key biomarker in assessing kidney function and detecting early signs of kidney damage, particularly in individuals with risk factors for chronic kidney disease.[2-4] The findings revealed that most participants (94.3%) had normoalbuminuria, with a mean age of 40.1 years (range 22.8-59.6), indicating preserved renal function in most of this cohort. In contrast, a small proportion (5.7%) exhibited microalbuminuria or clinical albuminuria, with a mean age of 42.5 years (range 30.2-60.9), suggesting early kidney dysfunction that may require closer clinical monitoring to prevent further progression.

Microalbuminuria, defined as UACR levels between 30 and 299 mg albumin/g creatinine, is a recognized marker for renal injury and an early indicator of CKD. Evidence on UACR prevalence in healthy Thai adults is limited, as most studies focus on clinic-based or high risk populations. Microalbuminuria was reported in 26-27% of diabetic.[6,7] and 16.6% of hypertensive patients.[8] A community hospital review found CKD in 8.2% of adults, and high risk community screening detected albuminuria or reduced eGFR in 20.3%.[10] These findings emphasize the lack of representative data on UACR in healthy Thais. In U.S. adults without diabetes (NHANES 1999-2020), the prevalence of albuminuria was 9.7%,[16] whereas the present study observed a lower prevalence of 4.4%, consistent with findings in lower-risk populations compared to those with diabetes or hypertension.[17] Clinical albuminuria (UACR ≥300 mg albumin/g creatinine), indicative of more advanced kidney damage, was observed in 1.3% of participants. This finding highlights the importance of early detection and intervention to prevent progression to more severe forms of CKD.

In terms of factors associated with albuminuria, the study did not find significant correlations between hypertension, smoking, alcohol consumption, or physical activity and the presence of microalbuminuria or clinical albuminuria. The lack of association with hypertension, smoking, and physical inactivity was unexpected given their known impact on kidney disease. This may be explained by the relatively young age of participants, the high proportion of normotensive individuals (71.5%), and the low prevalence of smoking (1.9%) and alcohol consumption (16.5%), which likely reduced statistical power and exposure variability to detect significant associations.[18,19]

Interestingly, a significant association was observed between elevated fasting blood glucose levels (≥126 mg/dL) and the presence of microalbuminuria/ clinical albuminuria. Elevated blood glucose is a well-established risk factor for the development of diabetic nephropathy, a leading cause of CKD globally.[20] Impaired fasting glucose (IFG) was independently associated with the development of albuminuria in a Chinese community–based population. The risk of albuminuria increased progressively with higher fasting blood glucose levels, suggesting that individuals with IFG should be prioritized for albuminuria screening to enable early detection and intervention.[21] The association between hyperglycemia and increased urinary albumin excretion emphasizes the need for diligent management of blood glucose levels in preventing kidney damage, especially in individuals at risk of diabetes or those already diagnosed with the condition.

In this study, among chronic diseases, only SLE was significantly associated with increased urinary albumin excretion. This finding aligns with previous literature demonstrating that SLE predisposes individuals to kidney damage. Although diabetes mellitus (DM) is known to cause renal injury through mechanisms such as hyperglycemia-induced glomerular hypertension,[20] no significant association with albuminuria was observed in this cohort. This may be due to the fact that two participants with type 2 DM had normal blood glucose levels (data not shown), one had a glucose level of 127 mg/dL (data not shown), and all three exhibited normoalbuminuria, whereas the single participant with type 2 DM and clinical albuminuria had a fasting blood glucose level of 300 mg/dL (data not shown). These findings suggest that individuals with type 2 DM who maintain good glycemic control may have a lower risk of developing albuminuria. In contrast, SLE, an autoimmune disorder that frequently affects the kidneys and can lead to glomerulonephritis,[22] remains a strong determinant of albuminuria. The finding underscores the importance of monitoring kidney function in individuals with SLE to prevent further renal deterioration. A nationwide study (2017-2020) reported an SLE prevalence of 85.8 per 100,000 persons in Thailand, with a female-to-male ratio of 9.6:1 and the highest rates in adults aged 40-49 years and in the southern region (178.5/100,000).[23] Thai lupus nephritis studies consistently show high rates of proteinuria, typically assessed by 24-hour urine protein or urine protein creatinine ratio, while UACR measurements are rarely reported in published Thai cohorts.[24,25]

Moreover, while obesity is a known risk factor for the development of kidney disease,[26] a significantly higher proportion of participants with microalbuminuria or clinical albuminuria were classified as obese class II. This is an important observation, as the relationship between obesity and kidney function is well established. Obesity can lead to glomerular hyperfiltration, increased intraglomerular pressure, and ultimately kidney damage over time.[27] These results highlight those higher degrees of obesity, particularly class II obesity, may exert an independent and measurable effect on early kidney dysfunction, reinforcing the need for targeted weight management strategies.

Limitations

Several limitations should be noted. First, the study was conducted at a single academic institution with a relatively small sample size (N=158), which may limit the generalizability of the findings to the broader Thai population. A second limitation was the absence of fasting glucose data in 80 participants, of whom only two had albuminuria; this missing information may have introduced bias. Third, the small number of individuals with elevated UACR (N=9) limited the statistical power to detect associations with some risk factors. Finally, only univariate analyses were performed; future studies should employ multivariate models to adjust for potential confounding variables.

Conclusion

In conclusion, the findings highlight the significance of urinary albumin as an indicator of early kidney dysfunction and reinforce the need for regular monitoring, especially among individuals with risk factors such as elevated blood glucose, obesity, and SLE. Although the overall prevalence of albuminuria was low in this cohort, the associations observed with certain demographic and clinical factors suggest that targeted interventions could help mitigate the risk of progression to more severe kidney disease. Further research is needed to explore the long–term implications of these early signs of kidney dysfunction and to identify effective strategies for prevention and management.

Ethical approval

This study was reviewed and approved by the Ethics Research Committee from Faculty of Associated Medical Sciences, Chiang Mai University, (AMSEC-66EX-046).

Conflict of interests

The authors declare no conflict of interest.

CRediT authorship contribution statement

Piyawan Bunpo: conceived and designed the study, carried out all the experimental work and statistical analysis, interpreted and prepared the manuscript; **Suwatsin Kittikunnathum**: carried out all the experimental work and statistical analysis, interpreted and prepared the manuscript; **Pharisa Nanthawong**: carried out all the experimental work and statistical analysis. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by the Faculty of Associated Medical Sciences, Chiang Mai University. The funders had no role in study design, data collection and analysis, interpretation of data and preparation of the manuscript.

References

[1] Mahmoud T, Borgi L. The interplay between nutrition, metabolic, and endocrine disorders in chronic kidney disease. Semin Nephrol. 2021; 41(2): 180-8. doi: 10.1016/j.semnephrol.2021.03.012.

[2] American Diabetes Association. Standards of medical care in diabetes-2021. Diabetes Care. 2021; 44: S1-S232. doi: 10.2337/dc21-S002.

[3] Alicic RZ, Rooney MT, Tuttle KR. Diabetic kidney disease: Challenges, progress, and possibilities. Clin J Am Soc Nephrol. 2017; 12(12): 2032-45. doi: 10.2215/CJN.11491116.

[4] Roscioni SS, Lambers Heerspink HJ, de Zeeuw D. Microalbuminuria: Target for renoprotective therapy pro. Kidney international. 2014; 86(1): 40-9. doi: 10.1038/ki.2013.490.

[5] Zarshenas F, Dehghan A, Mirzaei M. Association between chronic kidney disease and cardiovascular disease risk factors in elderly: Results from the first phase of fasa and shahedieh cohort studies. BMC Nephrol. 2024; 25(1): 413. doi: 10.1186/s12882-024-03566-2.

[6] Kang SH, Jung DJ, Choi EW, Cho KH, Park JW, Do JY. Hba1c levels are associated with chronic kidney disease in a non-diabetic adult population: A nationwide survey (knhanes 2011-2013). PloS one. 2015; 10(12): e0145827. doi: 10.1371/journal.pone.0145827.

[7] Pranart Chiowanich TK, Srivadee Oravivattanakul, Viroj Wiwanitkit. Prevalence and risk factors of microalbuminuria in patient with diabetic mellitus at northern part referral hospital in thailand. Diabetes and Metabolic Syndrome: Clinical Research and Reviews. 2009; 3(3): 152-4. doi: 10.1016/J.DSX.2009.02.010.

[8] Pongsathorn Gojaseni AP, Worawon Chailimpamontree, Thaweepong Pajareya, Anutra Chittinandana. Prevalence and risk factors of microalbuminuria in thai nondiabetic hypertensive patients. Vascular Health and Risk Management. 2010. doi: 10.2147/VHRM.S9739.

[9] Lekskulchai V. Use of estimated glomerular filtration rate and urine albumin-to-creatinine ratio based on kdigo 2012 guideline in a thai community hospital: Prevalence of chronic kidney disease and its risk factors. Med Sci Monit Basic Res. 2022; 28: e938176. doi: 10.12659/MSMBR.938176.

[10] Udom Krairittichai SP, Amporn Jongsareejit, Charnvate Sattaputh. Prevalence and risk factors of diabetic nephropathy among thai patients with type 2 diabetes mellitus. Journal of the Medical Association of Thailand. 2011. doi:

[11] Tsai MC, Lojanapiwat B, Chang CC, Noppakun K, Khumrin P, Li SH, et al. Risk prediction model for chronic kidney disease in thailand using artificial intelligence and shap. Diagnostics (Basel). 2023; 13(23). doi: 10.3390/diagnostics13233548.

[12] Garofalo C, Borrelli S, Minutolo R, Chiodini P, De Nicola L, Conte G. A systematic review and meta-analysis suggests obesity predicts onset of chronic kidney disease in the general population. Kidney international. 2017; 91(5): 1224-35. doi: 10.1016/j.kint.2016.12.013.

[13] Ahmed MA, Ferede YM, Takele WW. Incidence and predictors of chronic kidney disease in type-ii diabetes mellitus patients attending at the amhara region referral hospitals, ethiopia: A follow-up study. PloS one. 2022; 17(1): e0263138. doi: 10.1371/journal.pone.0263138.

[14] Pugliese G, Penno G, Natali A, Barutta F, Di Paolo S, Reboldi G, et al. Diabetic kidney disease: New clinical and therapeutic issues. Joint position statement of the italian diabetes society and the italian society of nephrology on "the natural history of diabetic kidney disease and treatment of hyperglycemia in patients with type 2 diabetes and impaired renal function". J Nephrol. 2020; 33(1): 9-35. doi: 10.1007/s40620-019-00650-x.

[15] Moorhead JF, Chan MK, El-Nahas M, Varghese Z. Lipid nephrotoxicity in chronic progressive glomerular and tubulo-interstitial disease. Lancet. 1982; 2(8311): 1309-11. doi: 10.1016/s0140-6736(82)91513-6.

[16] Bragg-Gresham JL, Annadanam S, Gillespie B, Li Y, Powe NR, Saran R. Using risk assessment to improve screening for albuminuria among us adults without diabetes. J Gen Intern Med. 2025; 40(13): 3159-69. doi: 10.1007/s11606-024-09185-9.

[17] Hillege HL, Fidler V, Diercks GF, van Gilst WH, de Zeeuw D, van Veldhuisen DJ, et al. Urinary albumin excretion predicts cardiovascular and noncardiovascular mortality in general population. Circulation. 2002; 106(14): 1777-82. doi: 10.1161/01.cir.0000031732.78052.81.

[18] Yang HY, Lu KC, Fang WH, Lee HS, Wu CC, Huang YH, et al. Impact of interaction of cigarette smoking with angiotensin-converting enzyme polymorphisms on end-stage renal disease risk in a han chinese population. J Renin Angiotensin Aldosterone Syst. 2015; 16(1): 203-10. doi: 10.1177/1470320313481837.

[19] Yang L, Chu TK, Lian J, Lo CW, Lau PK, Nan H, et al. Risk factors of chronic kidney diseases in chinese adults with type 2 diabetes. Sci Rep. 2018; 8(1): 14686. doi: 10.1038/s41598-018-32983-1.

[20] Tziomalos K, Athyros VG. Diabetic nephropathy: New risk factors and improvements in diagnosis. Rev Diabet Stud. 2015; 12(1-2): 110-8. doi: 10.1900/RDS.2015.12.110.

[21] Jiang Y, Jia J, Li J, Huo Y, Fan F, Zhang Y. Impaired fasting blood glucose is associated with incident albuminuria: Data from a chinese community-based cohort. J Diabetes Complications. 2022; 36(2): 108125. doi: 10.1016/j.jdiacomp.2022.108125.

[22] Lichtnekert J, Anders HJ. Lupus nephritis-related chronic kidney disease. Nat Rev Rheumatol. 2024; 20(11): 699-711. doi: 10.1038/s41584-024-01158-w.

[23] Pongkulkiat P, Foocharoen C, Onchan T, Suwannaroj S, Mahakkanukrauh A. Prevalence and incidence of systemic lupus erythematosus in thailand based on national health data. Lupus Sci Med. 2025; 12(2). doi: 10.1136/lupus-2025-001621.

[24] Amorn Sankhaanuruk NK, Worawit Louthrenoo. Incidence of lupus nephritis flares after complete response in a thai population. Medical Research Archives. 2022; 10(5). doi: 10.18103/mra.v10i5.2831.

[25] Nun Singpan RC, Boonyarit Cheunsuchon. Clinicopathological characteristics of lupus nephritis in thai males. Journal of Nephropathology. 2021; 10(2): 1-5. doi: 10.34172/jnp.2021.19.

[26] Kim TB, Ahn SY, Oh J, Bae EH, Chin HJ, Kim MG, et al. The impact of obesity on kidney disease: Observational cohort study analyzing 14,492 kidney biopsy cases. J Korean Med Sci. 2024; 39(3): e12. doi: 10.3346/jkms.2024.39.e12.

[27] Song SH, Oh TR, Suh SH, Choi HS, Kim CS, Ma SK, et al. Obesity is associated with incident chronic kidney disease in individuals with normal renal function. Korean J Intern Med. 2024; 39(5): 813-22. doi: 10.3904/kjim.2023.491.

Journal of Associated Medical Sciences 2026; 59(2): 24-33

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Elevated Sall4 expression correlates with prostate cancer Gleason score and metastasis using immunohistochemistry and RNAscope®

Dhafer A. Alghezi[1*] and Ali Harb[2]

[1]*Department of Microbiology, College of Medicine, University of Thi-Qar, Thi-Qar, Iraq.*
[2]*College of Nursing, National University of Science and Technology, Thi-Qar, Iraq.*

ARTICLE INFO

Article history:
Received 30 May 2025
Accepted as revised 21 December 2025
Available online 13 January 2026

Keywords:
Immunohistochemistry, metastasis, PCa Gleason score, RNAscope®, Sall4.

ABSTRACT

Background: Prostate cancer lacks reliable biomarkers to distinguish between indolent and aggressive forms, posing diagnostic and prognostic challenges. Sall4, primarily found in embryonic stem cells, is reactivated in various cancers, but its role in prostate cancer remains unclear.

Objectives: This study aims to evaluate Sall4 protein and mRNA levels in malignant and normal prostate tissues and explore their association with clinical data.

Materials and methods: This study was conducted from December 2022 to April 2024 at Al-Hussein Teaching Hospital, Thi-Qar, Iraq. Sall4 protein and mRNA expression levels were assessed in 40 normal tissues and 194 malignant prostate tissues using immunohistochemistry and RNAscope® methods. The data were analyzed using unpaired t-tests.

Results: The study identified a significant increase in nuclear Sall4 protein expression, assessed by immunohistochemistry, in prostate cancer tissues compared to normal tissues (p=0.001). Similarly, Sall4 mRNA levels, measured using RNAscope®, were significantly higher in malignant tissues (p<0.001). Increased Sall4 expression at both protein and mRNA levels was significantly associated with higher Gleason scores (protein: p=0.003; mRNA: p=0.009), lymph node involvement (protein: p=0.002; mRNA: p=0.006), and metastasis (protein: p=0.001; mRNA: p=0.017). However, no significant correlation was found between Sall4 expression and tumor size.

Conclusion: Elevated Sall4 expression may be associated with prostate tumorigenesis and aggressiveness. Further studies are needed to clarify its role and evaluate its potential as a prognostic biomarker for prostate cancer.

Introduction

Prostate cancer (PCa) is a significant global health concern, characterized by a high fatality rate and widespread impact on men's health.[1,2] It accounts for approximately 300,000 new cases and 41,000 deaths annually in the United States[3] and ranks as the tenth most predominant malignancy in Iraq, accounting for 1,224 newly diagnosed cases in 2019.[4] The vast majority of PCa cases, approximately 95%, are known to be acinar adenocarcinomas, which develop in glandular areas of the prostate.[5,6] In contrast, only around 5% of these cases have ductal adenocarcinomas, which grow from the glandular

*** Corresponding contributor.**
Author's Address: *Department of Microbiology, College of Medicine, University of Thi-Qar, Thi-Qar, Iraq.*
E-mail address: Dhafer.a.f.alghezi@bath.edu
doi: 10.12982/JAMS.2026.034
E-ISSN: 2539-6056

epithelial cells and are often identified through histopathological analysis.[6]

Two key systems assess PCa progression: the Gleason grade and Tumor-Node-Metastasis (TNM) systems.[7] The Gleason grading system classifies PCa based on microscopic tissue architecture, ranging from 1 to 5, indicating how closely the cancer resembles normal tissue.[8] Less aggressive cancers resemble healthy tissue, while more aggressive ones show abnormal features and higher metastatic potential.[9] PCa is heterogeneous, with varying patterns in the same tumor, so the Gleason score combines the two most common grades in biopsy or prostatectomy samples, with the highest score being 10. A Gleason score of 7 (e.g., 3+4 or 4+3) reflects different proportions of grade 3 and 4 components.[9,10] The TNM system stages cancer based on tumor size (T), lymph node involvement (N), and metastasis (M), providing key prognostic information and guiding treatment.[11] However, neither system can reliably differentiate between aggressive and non-aggressive tumors or predict outcomes after initial therapy, such as relapse or sustained remission.[8]

Spalt-like transcription factor 4 (Sall4), a zinc finger transcription factor, is crucial in maintaining pluripotency and promoting oncogenesis, with its expression linked to aggressive cancer behaviors. Sall4 is predominantly found in embryonic stem cells and during embryonic development, with its expression markedly absent or reduced in the majority of healthy adult human tissues. However, it has been reactivated in various cancers, including approximately 30% of solid tumors, such as PCa.[12-14]

Its role in regulating key pathways associated with cell survival and proliferation makes it a compelling candidate for further investigation in PCa. Sall4 may contribute to the progression of PCa primarily by acting as a transcriptional activator of the Bmi-1 gene, which is well-known for its function in controlling stem cell self-renewal and differentiation as well as preventing cells from dying.[15-17] Another study on transgenic mice with the Sall4 B mutation showed lower levels of PTEN expression than wild-type controls, indicating that Sall4 may inhibit PTEN transcription. Elevated Sall4 levels could, therefore, contribute to PCa development by inhibiting PTEN and fostering tumorigenic processes, emphasizing the need to explore further Sall4's regulatory impact on PTEN in PCa.[18] Yang and colleagues have demonstrated that Sall4 upregulates several anti-apoptotic genes, including Bmi-1, Bcl2, DAD1, TEGT, BIRC7, and BIRC4, while negatively regulating pro-apoptotic genes such as TNF, TP53, and PTEN.[18]

These findings encourage us to assess the expression level of Sall4 in normal and malignant prostate tissues and to investigate its correlation with PCa clinical data using immunohistochemistry (IHC) and RNAscope® methods.

Materials and methods

Patients and control samples

This retrospective study was approved by the Ethics Board of Al Hussein Teaching Hospital, Thi-Qar, Iraq (Approval No. 2021159, dated 7/12/2022). A total of 234 prostate samples were analyzed, including 194 formalin-fixed, paraffin-embedded malignant specimens and 40 normal prostate (NP) tissues as controls. Gleason scores and pathological stages were assessed by histopathologists. Patients who received chemotherapy or hormonal therapy were excluded. Normal testis tissue served as a positive control in both IHC and RNAscope® assays, while negative controls omitted the primary antibody. Clinical data related to prostate tissues are summarized in Table 1.

Table 1. The distribution of normal and malignant prostate cases according to clinical data.

Prostate clinical data		Number	%	*p* value
Number of samples	Normal	40	17.1	<0.0001
	Malignant	194	82.9	
Age range	Normal	21-50		<0.0001
	Malignant	20-85		
Gleason score	Low Gleason score	30	15.5	
	High Gleason score	160	82.5	
	N/A	4	2	
T category	T1	18	9.3	
	T2	103	53.1	
	T3	48	24.7	
	T4	17	8.8	
	N/A	8	4.1	

Table 1. The distribution of normal and malignant prostate cases according to clinical data (continued).

Prostate clinical data		Number	%	*p* value
N category	N0	154	81	
	N1	31	8.8	
	N/A	9	4.6	
M category	M0	148	76.3	
	M1	33	17	
	N/A	13	6.7	

Immunohistochemistry

IHC was performed on prostate tissue sections using two anti-Sall4 antibodies: a rabbit polyclonal (1:100, Abcam, Ab29112, UK) and a mouse monoclonal (1:100, Novus Bio, H00057167-M03, UK). Detection used the DAKO Envision™+HRP (DAB) kit (K4010, DAKO, UK). This method was carried out according to the previously protocol described by Alalwany, *et al.*[19] Sections were deparaffinized with Histoclear (H5-200, National Diagnostics, UK), rehydrated through graded ethanol (100%, 95%, 70%) (20821 330, VWR, UK), and permeabilized with 0.5% Triton X-100 (T8787-100ML, Sigma, UK) in PBS (BR00140, OXOID, UK).

Heat-induced epitope retrieval was done at 90 °C for 30 minutes in Tris/EDTA buffer (pH 9) (T/p90/630, Fisher Scientific, UK), followed by cooling. Endogenous peroxidase activity was blocked with %3 hydrogen peroxide, and non-specific binding was reduced using %10 normal goat serum (G9023, Sigma, UK) and %0.05 BSA (A4503, Sigma, UK) in PBS.

Primary antibodies were applied overnight at 4 °C. After PBS washes, sections were incubated with secondary antibody for 30 minutes, developed with DAB for 5 minutes, counterstained with hematoxylin (H-3401, Vector, UK), and mounted using DPX (06522, Sigma-Aldrich, UK). Images were captured using a Nikon Eclipse E800 microscope.

IHC quantification

The H-score approach was used to evaluate the Sall4 immunostaining in prostate tissue samples by histopathologists. This method quantifies protein expression in IHC by evaluating the percentage and intensity of stained cells. The H-score range is 0-300.[25] Nuclear staining was shown to have four intensity levels: 0, indicating no staining, 1, indicating mild staining, 2, indicating moderate staining, and 3, indicating severe staining. The formula used to determine the H-score was as follows:

"H-score = (3 × % strongly positive nuclei) + (2 × % moderately positive nuclei) + (1 × % weakly positive nuclei)". The total score ranges from 0 to 300.

RNAscope® protocol

RNAscope® was used to detect mRNA molecules in tissue samples via light microscopy, following the manufacturer's instructions over two days using RNAscope® 2.5. HD detection reagent –Brown (Cat. No. 322310, ACD, Italy). First, tissue sections were incubated in a dry oven at 60°C for an hour to ensure adherence, followed by deparaffinization and rehydration as described in the IHC protocol. Endogenous peroxidase activity was blocked by applying H_20_2 for 10 minutes, followed by washing with deionized water (ddH_2O). Target retrieval (cat. no 322000, ACD, Italy) was performed by heating the slides at 98 °C in 1x target retrieval solution (cat. no 310091, ACD, Italy) for 15 minutes, then rinsing with ddH_2O and 100% ethanol.

On the second day, RNAscope® amplifier reagents were brought to room temperature, and probes were pre-warmed to 40 °C in a HybEZ II oven to dissolve any precipitated salts. Tissue sections were treated with protease pulse (Cat. No. 322330, ACD, Italy), incubated for 30 minutes at 40 °C in a humidified tray, and washed with ddH_2O to remove RNA-binding proteins, permeabilize tissue, and enhance probe penetration.

Next, mRNA probes were added, and the sections were incubated for two hours at 40 °C. After a two-minute wash with 1x wash buffer, six amplification stages (AMP 6-1) were carried out, with incubation times of 30, 15, 30, 30, and 15 minutes, respectively, at 40 °C. Each amplification step was followed by two washes in buffer for three minutes. DAB was applied to the sections for 10 minutes at room temperature, followed by ddH_2O washing. Nuclear staining was done with 50% hematoxylin for two minutes, rinsed under tap water, and briefly washed with 0.02% ammonia water. Finally, dehydration, clearing, mounting, and photography were performed as per the IHC protocol.

RNAscope® quantification

Using the methods advised by the manufacturer, the RNAscope® score was employed to assess the expression of mRNA Sall4 in tissue samples. Briefly, the number of brown spots inside the prostate cells were counted and scored after five random fields were photographed at a 20x magnification. The positive cells were scored using the following scoring system: 0 represented no staining, one represented 1-3 dots/cell, two represented 4-9 dots/cell, three represented 10-15 dots/ cell, and <10% dots are in clusters, and four represented >15 dots/cell and >10% dots are in clusters.

Statistical analysis

The data of this study was analyzed by GraphPad Prism version 8.00 for Windows (GraphPad Software, La Jolla, California, USA, www.graphpad.com). Data were first assessed for normality using the Shapiro–Wilk test. Variables that followed a normal distribution ($p>0.05$) were analyzed using unpaired t-tests to compare means between groups. For non-normally distributed data, non-parametric tests were considered. For statistical significance, a significance level of $p<0.05$ was used.

Results

Clinical data of study population

The current study used 194 PCa (82.9%) and 40 NP tissue samples (17.1%). In PCa, the age range was between 20 and 80, while in NP, it was between 21 and 50. In addition, the majority of PCa cases were classified as high Gleason scores (N=160, 82.9%), whereas only 30 PCa cases (15.5%) were classified as low Gleason scores. Of the 121 PCa patients (62.4%) with T1-2 malignancies, 65 cases (33.5%) had T3-4 malignancies. The majority of PCa cases (N=154, 81%) had no lymph node invasion, whereas the remaining cases (N=31, 8.8%) had lymph node invasion. A total of 33 cases (17%) of PCa were identified as metastatic, while the majority of cases (N=148, 76.3%) were non-metastatic, as detailed in Table 1.

Sall4 expression in prostate samples using IHC

IHC was performed on normal and malignant prostate tissues to assess the immunostaining of anti-Sall4. The results indicated that normal and malignant prostate tissue showed nuclear Sall4 immunostaining, with signal intensities varying from strong to moderate. Variable intensities of nuclear Sall4 were observed in NP tissue. Staining intensity ranged from strong (Figure 1B, arrow) to negative (Figure 1A, arrow). PCa tissues also exhibited nuclear Sall4 staining, with signal strength varying significantly from strong (Figure 1C, arrow) to moderate (Figure 1D, arrow), weak (Figure 1E, arrow), and negative (Figure 1F, arrow). Sall4 has been previously reported in spermatogonia cells of the testis.[27] In this study, testis was used as a positive control for Sall4 expression, and IHC analysis revealed strong nuclear staining in the spermatogonia cells (Figure 1G, arrow). Negative control, no primary antibody added, showed free of Sall4 staining (Figure 1H, arrow).

Sall4 immunostaining validation using testis and postate tissues

To validate Sall4 immunostaining in both prostate and positive control, Testis tissues, IHC was conducted on sections to ensure that the two independent Sall4 antibodies (rabbit polyclonal and mouse monoclonal) exhibited a consistent staining pattern. The rabbit polyclonal Sall4 antibody result showed that strong nuclear Sall4 was seen in both the Testis (Figure 2A, arrow) and Prostate (Figure 2C, arrow). Sall4 mouse monoclonal also showed a similar expression pattern in the Testis (Figure 2B, arrow) and Prostate (Figure 2D, arrow) tissues.

Sall4 expression in prostate samples using RNA-scope®

RNAscope® analysis was performed on normal and malignant prostate tissues. The findings demonstrated that Sall4 mRNA stained both groups. The glandular epithelial cells of NP had Sall4 mRNA staining (Figure A, arrow). Malignant tissues were found to have varied degrees of Sall4 mRNA staining; these included widespread (Figure 3B, arrows), scattered (Figure 3C, arrow), and negative (Figure 3D, arrow) staining. Furthermore, it was shown that the spermatogonia cells in the testis tissue expressed Sall4 mRNA (Figure 2E, arrow). The negative control's prostate sections had no staining (Figure 3F, arrow).

Quantification of Sall4 staining in prostate tissue samples

The measurement of Sall4 expression in prostate samples was investigated using both methods. Quantitative analysis of IHC data demonstrated a significantly higher nuclear Sall4 immunostaining in PCa compared to NP tissues ($p=0.001$) (Table 2 and Figure 4A). This elevation of Sall4 expression exhibited a positive correlation with a high Gleason score ($p=0.003$) (Table 2 and Figure 4B). Furthermore, it was demonstrated that Sall4 nuclear expression was markedly elevated in N1 patients compared to those with (N0) ($p=0.002$) (Table 2 and Figure 4C). Nuclear Sall4 staining and metastasis showed a strong correlation ($p=0.001$) (Table 2 and Figure 4D). In contrast, Sall4 expression did not show a significant association with tumor size ($p=0.728$) (Table 2).

The quantification of Sall4 mRNA revealed higher mRNA levels of Sall4 in PCa than in normal individuals ($p<0.001$) (Table 2 and Figure 5A). Increased Sall4 at the mRNA level was seen in the high Gleason score compared to the low Gleason score ($p=0.009$) (Table 2 and Figure 5B). A significant association was seen between high mRNA levels of Sall4 and the presence of lymph node involvement (N1) ($p=0.006$) (Table 2 and Figure 5C) as well as metastatic PCa patients ($p=0.017$) (Table 2 and Figure 5D). However, this study did not show a significant association between Sall4 mRNA levels and tumor size ($p=0.852$) (Table 2).

Table 2. Sall4 quantification results in Prostate cancer clinical data.

Comparison	The main findings	Nuclear Sall4 result		Sall4 mRNA result	
		Mean±SD	*p* value	Mean±SD	*p* value
NP vs PCa	Increased in PCa	0.650±1.988	0.001	0.003±0.025	<0.001
		12.24±19.93		0.488±0.826	
Gleason scores (LGS vs HGS)	Increased in high Gleason score	8.743±16.31	0.003	0.960±0.838	0.009
		22.88±26.17		1.543±1.275	
Stage T (T1-2 vs T3-4)	No difference	11.26±18.57	0.728	0.498±0.88	0852
		12.39±20.57		0.489±0.746	
Stage N (NO vs N1)	Increased in N1	0.464±18.30	0.002	0.516±0.853	0.006
		21.73±20.36		1.061±1.254	
Stage M (M0 vs M1)	Increased in M1	9.345±17.71	0.001	0.484±0.813	0.017
		22.02±22.13		0.931±1.254	

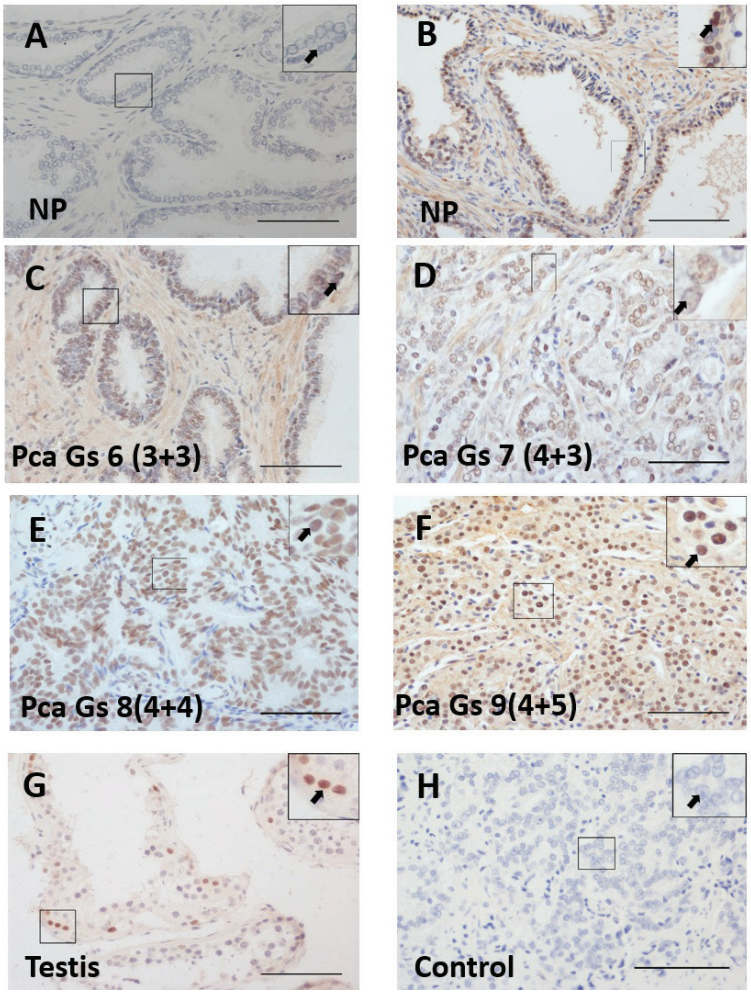


Figure 1. *The expression of Sall4 in normal and malignant tissues. A: no Sall4 expression was seen in NP (arrow), B: strong nuclear Sall4 staining was found in NP (arrow). C: PCa Gleascon score (Gs) 6 showed weak nuclear Sall4 staining (arrow). D: PCa Gs 8 tissue has weak nuclear Sall4 staining (arrow). E: PCa Gs 8 tissues showed moderate nuclear Sall4 staining (arrow). F: strong nuclear Sall4 staining was seen in PCa high Gs 8 (arrow), G: positive control, testis tissue, had strong nuclear Sall4 (arrow), H: no background staining is observed in prostate tissue, confirming the specificity of the assay. Scale bars = 100 µm; insets show 2x magnification.*

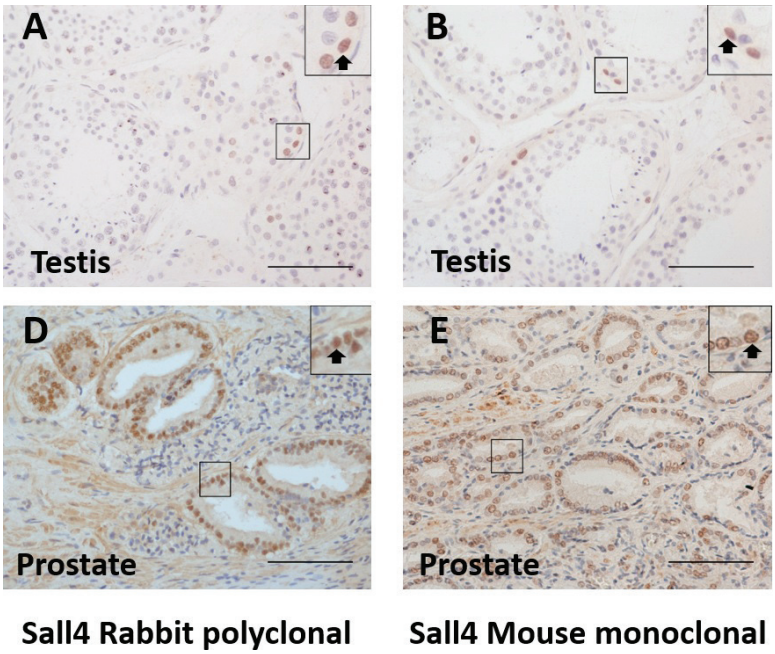


Figure 2. *Validation of Sall4 staining using two antibodies. A and C: nuclear Sall4 staining (rabbit polyclonal) was observed in both tissues of the Testis and Prostate (arrows), B and D: intense nuclear Sall4 staining (mouse monoclonal) detected in tissues of the testis and prostate (arrows).*

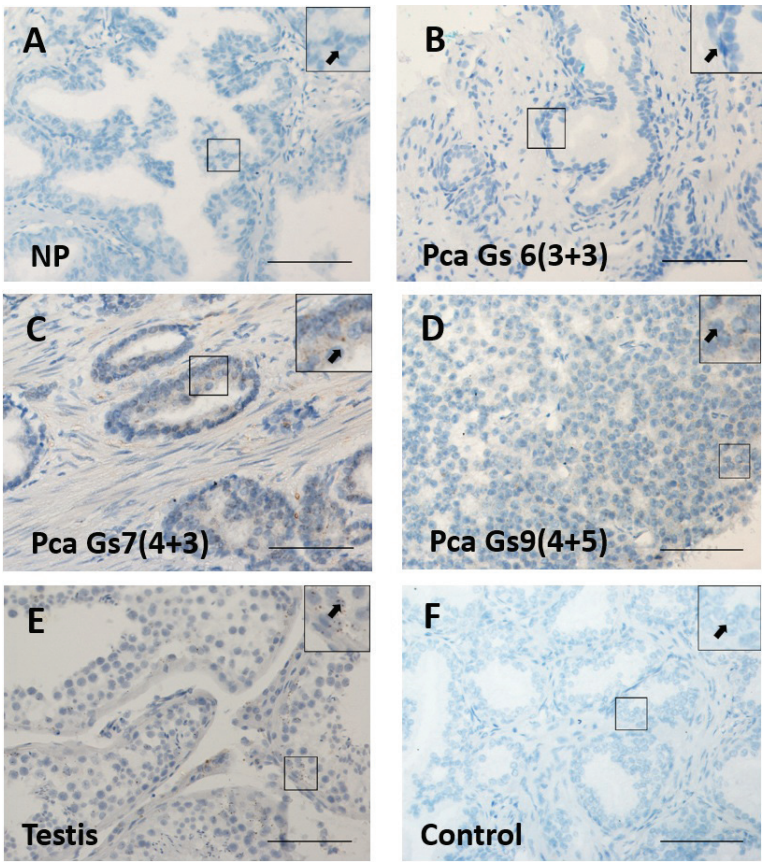


Figure 3. *Sall4 mRNA in prostate samples. A: mRNA staining (arrow) of Sall4 was detected in a few epithelial cells of NP, B: no Sall4 mRNA dots were observed in PCa Gs 6 (arrow), C: mRNA dots of Sall4 were scattered in PCa Gs 7 (arrow), D: Sall4 mRNA dots were found in PCa Gs 9 (arrow), E: Sall4 mRNA dots were observed in the spermatogonia cells of the testis (arrow), F: negative control had no Sall4 mRNA dots (arrow). Scale bars = 100 µm; insets show 2x magnification.*

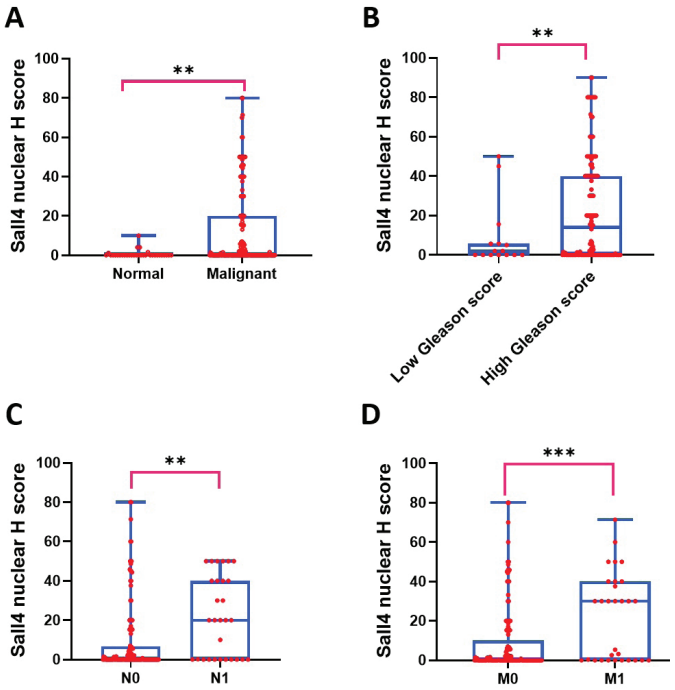


Figure 4. *Sall4 expression quantification in prostate samples using IHC. A: PCa had more nuclear expression of Sall4 than NP (P 0.001), B: there was a noticeable increase in nuclear Sall4 expression from high Gleason scores to low scores (p=0.003), C: Sall4 expression was increased significantly in N1 compared to N0 (p=0.002), D: increased Sall4 expression was associated with metastasis (p=0.001). Each case's five randomly selected photographs make up the data. Unpaired t-tests were used in this study. NP (N=40), PCa (N=194), Low Gleason score (N=30), High Gleason score (N=160), T1-2 (N=121), T3-4 (N=65), N0 (N=154), N1(N=31), M0(N=148), M1(N=33).*

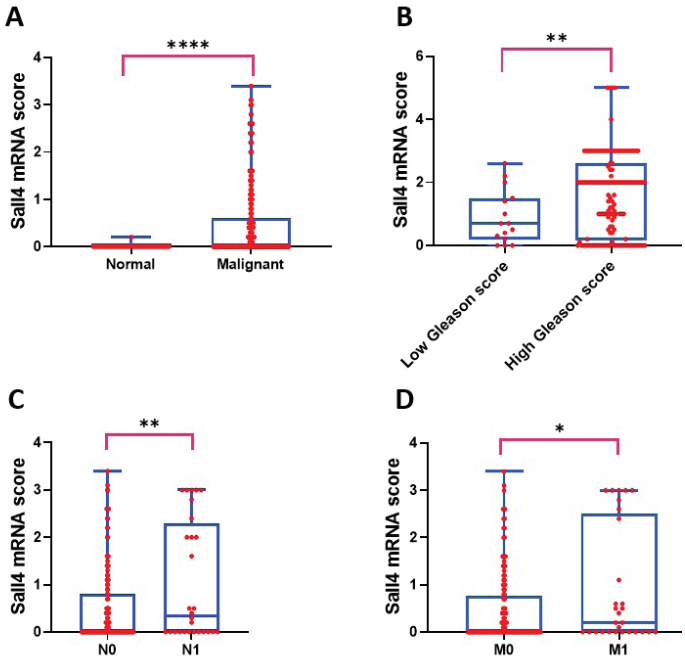


Figure 5. *Sall4 mRNA quantification in Prostate samples using RNAscope®. A: PCa had more Sall4 mRNA than NP (p<0.0001), B: higher Sall4 mRNA was significantly related to Gleason Score (p=0.009), C: increased Sall4 mRNA level in N1 compared to N0 (p=0.006), D: a significant association was detected between Sall4 mRNA level and metastasis (p=0.002). Each case's five randomly selected photographs make up the data. Unpaired t-tests were used in this study. NP (N=40), PCa (N=194), Low Gleason score (N=30), High Gleason score (N=160), T1-2 (N=121), T3-4 (N=65), N0(N=154), N1(N=31), M0(N=148), M1(N=33).*

Discussion

Sall4, a key regulator of pluripotency and differentiation, is linked to oncogenesis and tumor progression.[12] While its role in various cancers is documented, its involvement in PCa remainsunder-explored. This study used RNAscope® and IHC assays to investigate Sall4 levels in prostate samples. Sall4 is significantly up-regulated at protein and mRNAlevels in PCa compared to NP, correlating positively with Gleason score. Its expression is significantly associated with PCa pathological stage, including N and M, but not T, suggesting Sall4's role in tumor grade and stage.

Since there are few studies on Sall4 in PCa, this study may represent a pioneering effort to thoroughly investigate Sall4 expression in PCa. It evaluates both protein and mRNA levels and explores their associations with the Gleason score and pathological stage. SALL4 is primarily known as an embryonic stem cell marker and is typically low or absent in most adult tissues; occasional low-level expression has been reported in regenerative or proliferative epithelial compartments.[12] However, the SALL4 staining observed in a few NP tissues, such as in Figure 1B, may be attributed to non-specific background staining or focal expression in certain basal or progenitor-like epithelial cells. Nonetheless, this finding was not consistent across all NP samples examined, and overall, SALL4 expression was significantly higher in Pca tissues compared to NP. A Chinese study found that increased Sall4 staining in cancer patients with a high Gleason score (≥) compared to those with a low Gleason score (<7), as well as in advanced vs early stages.20 This study agreed with the current data. In addition to this Chinese study, previous studies identified Sall4 levels in various cancer types, including PCa samples, but they did not particularly investigate the relationship between PCa clinical characteristics and Sall4 expression.

For example, modest Sall4 expression was seen in NP tissues (57%) in research examining non-small cell lung carcinomas using a tissue microarray with 112 benign tissue controls, including 7 NP samples.[29] This observation aligns with our research results.[21] Conversely, another investigation revealed that none of the 62 PCa samples in a large cohort had nuclear Sall4 staining.[22] This discrepancy may be due to differences in sample size, the type of antibody, or the scoring system used. To the best of our knowledge, no previous study has utilized RNAscope® technology to evaluate Sall4 mRNA levels in prostate tissue, whether malignant or non-malignant, or to explore the correlation between Sall4 expression and clinical data related to malignancy. In this study, RNAscope in situ hybridization was used rather than quantitative RT-PCR and RNA-sequencing to assess mRNA expression, due to the former's superior ability to preserve tissue architecture and provide spatial resolution at the single-cell level. RNAscope® allows for direct visualization of transcript localization within intact tissue, making it especially valuable in heterogeneous tissues such as prostate cancer. This technique also minimizes issues related to RNA degradation commonly encountered in FFPE samples and allows for precise localization of transcripts, offering both morphological and molecular insights.[23]

This research addresses this significant gap by utilizing advanced RNAscope® assays to elucidate the differential expression patterns of Sall4 and their potential correlation with PCa clinical parameters, thereby contributing novel insights into its role and relevance in the pathology of PCa.

To provide context for our findings, we compared Sall4 expression in PCa with other cancer types. This analysis enhances our understanding of Sall4's role across different malignancies, aligning with research on Sall4 in colon and lung carcinomas, further supporting its involvement in various cancers.[21,24] The anti-Sall4 and mRNA data from this investigation, however, differ from those of Hao *et al.*, who found that colorectal carcinoma (CRC) had less Sall4 staining than either atypical hyperplasia or normal colon tissues.[25] Discrepancies in Sall4 expression observed across different studies may be attributed to variations in sample types, sample sizes, or the antibodies used. The IHC data for Sall4 expression in malignant versus normal tissues across several cancer types are complicated. Further research, including large cohorts of various cancer kinds and monitoring both Sall4 expression, is required, and somewhat inconsistent, suggesting that tumor type may affect these results.

In addition, our data about the correlation between Sall4 level and cancer grade partially agreed with the previous study, which found a low level of Sall4 mRNA in moderately differentiated CRC tissues compared to well-differentiated and poorly differentiated tissues, as measured by RT-PCR.24 These observations suggest that Sall4 could be used as a candidate biomarker for PCa prognosis, offering insights into its role in disease progression and differentiation. This current study, however, disagreed with another CRC study.[25] This discrepancy could be due to differences in cancer type and diagnostic methodologies, as RT-PCR was used in that study to quantify Sall4 mRNA. Furthermore, other research has reported no statistical association between Sall4 level and cancer grades in various malignancies.[21,26,27] These inconsistencies suggest that Sall4 may have varied roles in different cancers, potentially influenced by the specific cancer type and molecular environment.

The current results show a positive association between Sall4 expression and pathological stages, such as lymph node status and metastasis, suggesting its link to PCa progression. This finding suggests that SALL4 may influence tumor aggressiveness and metastatic potential more than primary tumor volume. SALL4 may play a greater role in promoting processes such as invasion, migration, and angiogenesis, which are more reflective of metastatic behavior than of

localized tumor growth. This finding is supported by other study which found that SALL4 knockdown impaired cell proliferation and migration but had limited effect on overall cell mass or apoptosis.[20] Thus, SALL4 may contribute to disease progression through mechanisms that are not directly dependent on tumor size, highlighting the importance of evaluating multiple tumor characteristics when assessing oncogenic drivers. This aligns with a previous CRC study, which found a correlation between Sall4 mRNA and lymph node metastases, but not clinical stage T.[24] Our findings were inconsistent with previous lung cancer[21,27],esophageal cancer[22] and hepatocellular carcinoma studies.[28] The absence of a correlation between SALL4 expression and tumor size suggests that Sall4 may be more relevant to tumor aggressiveness and metastatic potential rather than proliferative volume, aligning with findings in other cancers where it influences invasiveness independently of size.

Clinical approaches for diagnosing and prognosing PCa include IHC-based tumor biomarkers, clinical staging, Gleason scores, and PSA levels. IHC is widely used to identify key proteins in cells and tissues.[23,29] However,somevariables,suchastheselectionofantibodies, antibody concentrations, and antigen retrieval techniques, might affect IHC findings and cause variations in staining results.[30] The lack of a standardized IHC protocol further complicates its use, and non-specific staining can result in false-positive findings.30 Despite its limitations, IHC remains a clinically relevant and widely accessible method for protein localization. Despite its variability, IHC remains IHC remains a clinically relevant and widely accessible method for protein localization. We used RNAscope® to complement IHC and overcome its limitations.

Sall4 may play a role in PCa development and progression, as shown by consistent results from two antibodies and a Sall4 probe using high-specificity approaches. Combining RNAscope® and IHC provides more accurate expression data than IHC alone.

Conclusion

This study is the first to show a positive correlation between elevated Sall4 levels and Gleason grades and metastasis in PCa, suggesting Sall4's role in disease initiation and progression. Sall4 emerges as a potential biomarker for PCa, with diagnostic and prognostic value. Further research using additional prostate cancer cell lines, larger cohorts including more normal prostate tissue samples, and complementary molecular techniques such as RNA-sequencing and multivariate analyses will be essential to validate and extend our findings.

Ethical approval

The approval for conducting this retrospective study was obtained from the Ethics Board of Al Hussein Teaching Hospital in Thi-Qar, Iraq (approval number: 2021159, date: 7/12/2022). The study was conducted in accordance with applicable ethical guidelines, and patient confidentiality was strictly maintained throughout the research process.

Funding

This study did not receive any financial support from public, commercial, or non-profit funding bodies and was conducted exclusively using the authors' own financial resources.

Conflict of interest

The authors declare no conflicts of interest.

CRediT authorship contribution statement

Both Authors designed the study, collected data, conducted experiments and statistical analysis, and drafted the report, reviewed and edited the final text.

Acknowledgements

The authors thank the College of Medicine, University of Thi-Qar, for imaging facilities and Al-Hussein Teaching Hospital staff for help with sample collection.

References

[1] Lang SH, Frame FM, Collins AT. Prostate cancer stem cells. J Pathol. 2009; 217(2): 299-306. doi: 10.1002/path.2478.

[2] Alghezi DA. Identifying potential new stem cell biomarkers for prostate cancer [Doctoral Thesis]: University of Bath; 2019. doi: 10.13140/RG.2.2.15345.08802.

[3] Swami U, McFarland TR, Nussenzveig R, Agarwal N. Advanced prostate cancer: Treatment advances and future directions. Trends Cancer. 2020; 6(8): 702-15. doi: 10.1016/j.trecan.2020.04.010.

[4] Al Alwan N A . Cancer Control and Oncology Care in Iraq. J Contemp Med Sci. 2022; 8(1): 82-5. doi. org/10.22317/jcms.v8i1.1154.

[5] Dunn MW, Kazer MW. Prostate cancer overview. Semin Oncol Nurs. 2011; 27(4): 241-50. doi: 10.1016/j.soncn.2011.07.002.

[6] Andriole GL, Catalona WJ. The diagnosis and treatment of prostate cancer. Annu Rev Med. 1991; 42:9-15. doi:10.1146/annurev.me.42.020191.000301.

[7] Alghezi DA, Whitley P, Beresford M, Bowen R, Mitchard J, Chalmers AD. Decreased ABCG2 expression in prostate cancer and negatively associated with poorly differentiated grade and biochemical recurrence. Thi-Qar Med J. 2020; 20(2): 242-56. doi.org/10.32792/jmed.v20i2.144.

[8] Alghezi D, Aljawher R, Musawi S. Increased CD73 expression is associated with poorly differentiated Gle ason score and tumor size in prostate cancer. J Adv Biotechnol Exp Ther. 6(1): 161. doi: 10.5455/jabet.2023.d115.

[9] Matoso A, Epstein JI. Grading of prostate cancer: Past, present, and future. Curr Urol Rep. 2016; 17(3): 25. doi: 10.1007/s11934-016-0576-4.

[10] Sharpe B, Alghezi DA, Cattermole C, Beresford M, Bowen R, Mitchard J, et al. A subset of high Gleason grade prostate carcinomas contain a large burden of prostate cancer syndecan-1 positive stromal cells. The Prostate. 2017; 77(13): 1312-24. doi: 10.1002/pros.23391.

[11] Goldstein AS, Huang J, Guo C, Garraway IP, Witte ON. Identification of a cell of origin for human prostate cancer. Science. 2010; 329(5991): 568-71. doi: 10.1126/science.1189992.

[12] Zhang X, Yuan X, Zhu W, Qian H, Xu W. SALL4: an emerging cancer biomarker and target. Cancer Lett. 2015; 357(1): 55-62. doi: 10.1016/j.canlet.2014.11.037.

[13] Tatetsu H, Kong NR, Chong G, Amabile G, Tenen DG, Chai L. SALL4, the missing link between stem cells, development and cancer. Gene. 2016; 584(2): 111-9. doi: 10.1016/j.gene.2016.02.019.

[14] Abouelnazar FA, Zhang X, Wang M, Zhang J, Yu D, Zang X, et al. The new advance of SALL4 in cancer: Function, regulation, and implication. J Clin Lab Anal. 2023; 37(9-10): e24927. doi: 10.1002/jcla.24927.

[15] Jiang L, Li J, Song L. Bmi-1, stem cells and cancer. Acta Biochim Biophys Sin (Shanghai). 2009; 41(7): 527-34. doi: 10.1093/abbs/gmp040.

[16] Wang F, Zhao W, Kong N, Cui W, Chai L. The next new target in leukemia: The embryonic stem cell gene SALL4. Mol Cell Oncol. 2014; 1(4): e969169. doi: 10.4161/23723548.2014.969169.

[17] Álvarez C, Quiroz A, Benítez-Riquelme D, Riffo E, Castro AF, Pincheira R. SALL Proteins: Common and antagonistic roles in cancer. Cancers (Basel). 2021; 13(24). doi: 10.3390/cancers13246292.

[18] Yang J, Chai L, Gao C, Fowles TC, Alipio Z, Dang H, et al. SALL4 is a key regulator of survival and apoptosis in human leukemic cells. Blood. 2008; 112(3): 805-13. doi: 10.1182/blood-2007-11-126326.

[19] Alalwany O, Alghezi DA, Aljawher RQ, Harb A. Increased CD3 immunostaining associated with high grade and tumor size in colorectal carcinoma. Egypt J Med Microbiol. 2025; 34(1): 205-12. doi: 10.21608/ejmm.2024.331363.1364.

[20] Zhou J, Peng S, Fan H, Li J, Li Z, Wang G, et al. SALL4 correlates with proliferation, metastasis, and poor prognosis in prostate cancer by affecting MAPK pathway. Cancer Med. 2023; 12(12): 13471-85. doi: 10.1002/cam4.5998.

[21] Rodriguez E, Chen L, Ao MH, Geddes S, Gabrielson E, Askin F, et al. Expression of transcript factors SALL4 and OCT4 in a subset of non-small cell lung carcinomas (NSCLC). Transl Respir Med. 2014; 2(1): 10. doi: 10.1186/s40247-014-0010-7.

[22] Kilic E, Tennstedt P, Högner A, Lebok P, Sauter G, Bokemeyer C, et al. The zinc-finger transcription factor SALL4 is frequently expressed in human cancers: association with clinical outcome in squamous cell carcinoma but not in adenocarcinoma of the esophagus. Virchows Arch. 2016; 468(4): 483-92. doi: 10.1007/s00428-016-1908-y.

[23] Wang F, Flanagan J, Su N, Wang LC, Bui S, Nielson A, et al. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. J Mol Diagn. 2012; 14(1): 22-9. doi: 10.1016/j.jmoldx.2011.08.002.

[24] Forghanifard MM, Moghbeli M, Raeisossadati R, Tavassoli A, Mallak AJ, Boroumand-Noughabi S, et al. Role of SALL4 in the progression and metastasis of colorectal cancer. J Biomed Sci. 2013; 20(1): 6. doi: 10.1186/1423-0127-20-6.

[25] Hao L, Zhao Y, Wang Z, Yin H, Zhang X, He T, et al. Expression and clinical significance of SALL4 and β-catenin in colorectal cancer. J Mol Histol. 2016; 47(2): 117-28. doi: 10.1007/s10735-016-9656-5.

[26] Han LY, Fletcher MS, Urbauer DL, Mueller P, Landen CN, Kamat AA, et Doi: al. HLA class I antigen processing machinery component expression and intratumoral T-Cell infiltrate as independent prognostic markers in ovarian carcinoma. Clin Cancer Res. 2008; 14(11): 3372-9. doi: 10.1158/1078-0432.CCR-07-4433.

[27] Gautam AK, Wang C, Zeng J, Wang J, Lu J, Wei J, et al. Expression and clinical significance of SALL4 and LGR5 in patients with lung cancer. Oncol Lett. 2015; 10(6): 3629-34. doi: 10.3892/ol.2015.3772.

[28] Han SX, Wang JL, Guo XJ, He CC, Ying X, Ma JL, et al. Serum SALL4 is a novel prognosis biomarker with tumor recurrence and poor survival of patients in hepatocellular carcinoma. J Immunol Res. 2014; 2014: 262385. doi: 10.1155/2014/262385.

[29] Oliver C. Jamur MC. Immunocytochemical methods and protocols. Totowa N, editor: Humana Press; 2010. doi: 10.1007/978-1-59745-324-0

[30] Whitaker HC, Girling J, Warren AY, Leung H, Mills IG, Neal DE. Alterations in beta-catenin expression and localization in prostate cancer. Prostate. 2008; 68(11): 1196-205. doi: 10.1002/pros.20780.

Journal of Associated Medical Sciences 2026; 59(2): 34-43

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Genomic insights into colistin-resistant *Escherichia coli* from clinical isolates in Thailand: Diversity of sequence types, plasmid-borne *mcr* variants, and One Health implications

Kwanchon Jearakitiwanich[1*], Anusak Kerdsin[2], Monchai Siribamrungwong[3], Panan Ratthawongjirakul[4]

[1]*Department of Clinical Pathology and Medical Technology, Lerdsin General Hospital, Bangkok, Thailand.*
[2]*Faculty of Public Health, Kasetsart University, Chalermphrakiat Sakon Nakhon Province Campus, Sakon Nakhon, Thailand.*
[3]*Department of Internal Medicine, Lerdsin General Hospital, College of Medicine, Rangsit University, Bangkok, Thailand.*
[4]*Center of Excellence for Innovative Diagnosis of Antimicrobial Resistance, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand.*

ARTICLE INFO

Article history:
Received 1 September 2025
Accepted as revised 13 December 2025
Available online 13 January 2026

Keywords:
Colistin resistance, Escherichia coli, mcr genes, plasmids, one health.

ABSTRACT

Background: Colistin is a last-resort antimicrobial agent for treating multidrug-resistant Enterobacterales. However, resistance has emerged globally through plasmid-mediated *mcr* genes and chromosomal mutations. Reports of whole-genome sequencing of colistin-resistant *Escherichia coli* from clinical settings in Thailand remain limited.

Objectives: To characterize the molecular and phenotypic features of colistin-resistant *E. coli* isolated from a tertiary hospital in Thailand using whole-genome sequencing (WGS).

Materials and methods: Fourteen colistin-resistant *E. coli* isolates were collected from clinical specimens between 2021 and 2022. Antimicrobial susceptibility testing was performed using broth microdilution. WGS was applied to identify resistance determinants, plasmid replicons, virulence genes, phylogroups, and sequence types.

Results: The prevalence of colistin-resistant *E. coli* was 1.2% (14/1203 isolates). Most isolates exhibited an ESBL-like multidrug-resistant profile with preserved carbapenem susceptibility. WGS revealed diverse sequence types (including ST131, ST95, ST58, ST69) and phylogroups, indicating polyclonal dissemination. Two *mcr* variants (*mcr-1.1* and *mcr-3.5*) were identified on mobile plasmids (IncX4, IncI2, IncFII, IncHI2), with some isolates carrying both variants. Several isolates without *mcr* carried chromosomal mutations in *mgrB*, *phoPQ*, or *pmrAB*. Virulence genes, particularly adhesins, siderophore systems, and capsule-related determinants, were widely distributed. The detection of *mcr-3.5*, previously reported in livestock, within clinical isolates highlights potential zoonotic or foodborne transmission.

Conclusion: Colistin-resistant *E. coli* in Thailand shows significant genetic diversity and frequent coexistence of *mcr* and ESBL genes, often on plasmids with high potential for horizontal transfer. These findings emphasize the importance of antimicrobial stewardship, stricter control of drug use in animals, and integrated genomic surveillance under a One Health framework to mitigate the dissemination of colistin resistance.

*** Corresponding contributor.**
Author's Address: *Department of Clinical Pathology and Medical Technology, Lerdsin General Hospital, Bangkok, Thailand.*
E-mail address: kwanchon2529@gmail.com
doi: 10.12982/JAMS.2026.035
E-ISSN: 2539-6056

Introduction

An antimicrobial resistance (AMR) has emerged as one of the most global health threats. Among the Enterobacterales, *Escherichia coli* is a pathogen responsible for hospital-acquired infections and is

frequently implicated in multidrug-resistant (MDR) outbreaks. The increasing prevalence of colistin-resistant strains has the widespread use of colistin.[1] a last resort polymyxin antibiotic. However, colistin resistance has rapidly emerged through both chromosomal mechanisms and horizontally transferable plasmid-mediated *mcr* genes.[2]

In Thailand, the emergence of colistin-resistant Enterobacterales is a growing concern for tertiary care hospital[3]. This trend not only compromises empirical options but also a potential in infection control. These aims to prevent the misuse of last-line agents like colistin and strengthen surveillance systems that can detect and respond to emerging resistance trends.[4,5] In Thailand, the emergence of colistin-resistant Enterobacterales is a growing concern for tertiary care hospitals.[6] Specifically, colistin resistance in *Escherichia coli* has been previously documented from various settings, including clinical samples and slaughter pigs in Thailand.[7]

The Whole-genome sequencing (WGS) offers a powerful platform for AMR surveillance, providing detailed insights into resistance mechanisms, clonal relatedness, and gene transmission pathways.[8] Unlike conventional phenotypic tests, WGS can differentiate between plasmid-mediated and chromosomal resistance, identify high-risk sequence types (STs), and trace the distribution of virulence and resistance genes across clinical settings These genomic data not only inform antimicrobial stewardship by guiding empirical treatment guidelines and targeted infection control interventions, but also emphasize a One Health perspective highlighting how resistance genes such as *mcr* can disseminate across humans, animals, and the environment. Integrating genomic surveillance across sectors is therefore crucial to track the evolution and transmission of resistance determinants and to inform coordinated control strategies. In this study, we applied WGS to characterize colistin-resistant *E. coli* isolates obtained from clinical specimens at Lerdsin General Hospital. The findings aim to provide molecular epidemiological evidence that supports local infection control and national AMR policies, and to demonstrate the utility of genomic surveillance in antimicrobial use in Thailand.

Materials and methods

Collection and identification of bacterial isolates

All 1,203 isolates of *Escherichia coli* were collected from clinical specimens. Of these, 14 colistin-resistant *Escherichia coli* isolates were subsequently collected from residual clinical specimens processed at the Clinical Microbiology Laboratory, Lerdsin Hospital, Bangkok, Thailand. between October 2021 and September 2022. Species identification was performed using MALDI-TOF MS (Sirius system; Bruker Daltonik, Germany). All isolates were recoded to ensure patient anonymity and preserved at -80 °C in tryptic soy broth supplemented with 20% glycerol. Prior to testing, isolates were subculture on MacConkey agar containing colistin (4 µg/mL) and incubated at 37 °C for 18 hours.

Antimicrobial susceptibility testing

All isolates were tested for susceptibility to 16 antibiotics. The minimum inhibitory concentrations (MICs) were determined by broth microdilution using the automated Sensititre™ THAN2F system (Thermo Fisher, UK), following the manufacturer's protocol.[9] Quality control was performed using *E. coli* ATCC 25922. The MICs were interpreted according to CLSI guidelines (M100-S31, 2021 Edition)[10] Isolates with colistin MICs ≥4 µg/mL were classified as colistin-resistant and included in the study.

Whole genome sequencing (WGS)

Genomic DNA was extracted using the ZymoBIOMICS™ DNA Miniprep Kit (Zymo Research, USA) and quantified with a NanoDrop spectrophotometer (Biotek, USA). Whole-genome sequencing was performed at MicrobeNG (Birmingham, UK) using paired-end 2 × 250 bp Illumina NextSeq technology with a minimum mean coverage of 30×. Raw reads (FASTQ format) were trimmed using Trimmomatic and assessed for quality using FastQC. Genome assembly was performed using SPAdes v3.15.5. The assembly were checking quality by filter sequence by length minimum 200 bp and QUAST for assembly quality (https://usegalaxy.eu, Galaxy Europe version 25.0, access on 31 Aug 2025). Read and assembly were used to bioinformatics analysis. Antibiotic resistance genes were identified using ResFinder (https://genepi.food.dtu.dk/resfinder), access on 24 Aug 2025[5] and the CARD RGI tool (https://card.mcmaster.ca/analyze/rgi, access on 24 Aug 2025).[11] Mobile genetic elements were identified using MGE in (https://cge.food.dtu.dk/services/MobileElementFinder, access on 25 Aug 2025).[5] Chromosomal mutations associated with colistin resistance, including *mgrB, pmrAB* and *phoPQ* were analyzed using BLASTn and ClustalW, aligned with *E. coli* K-12 MG1655 reference genomes.[5] Amino acid substitutions were predicted using BLASTx, and insertion sequences were identified using ISfinder (https://isfinder.biotoul.fr/blast.php, access on 24 Aug 2025).[12] Virulence genes were identified using VirulenceFinder (https://cge.food.dtu.dk/services/VirulenceFinder, access on 25 Aug 2025).[5] Plasmid replicons were detected using PlasmidFinder (https://cge.food.dtu.dk/services/PlasmidFinder, access on 25 Aug 2025.[5] Phylogrouping of *E. coli* was performed using Clermont Typing (https://ezclermont.hutton.ac.uk, access on 30 Aug 2025),[13] and sequence types (STs) were assigned using MLST based on the (https://pubmlst.org, access on 30 Aug 2025).[5]

Statistical analysis

Descriptive statistics were used to summarize the characteristics of the study isolates. Categorical data,

such as specimen type, department of admission, and gender, were presented as counts and percentages. The association between colistin-resistant and non-colistin-resistant *E. coli* isolates and various categorical variables was tested using the Chi-square test or Fisher's exact test, where appropriate, to determine the *p*-values shown in Table 1. A *p*-value of less than 0.05 was considered statistically significant. All statistical analyses were performed using SPSS version 29.0.1.0.

Data availability statement

The complete genomes were deposited in GenBank under Bioproject Accession No. PRJNA1312245.

Results

Clinical isolates from clinical specimens

A total of 14 colistin-resistant isolates were collected from residual clinical specimens between October 2021 and September 2022. The isolates originated from various specimen types, including urine, blood, sputum, pus, and other sources, and were derived from multiple hospital departments including intensive care unit (ICU) and non-ICU Table 1. Among 1,203 *E. coli* isolates, 14 isolates (1.2%) were colistin-resistant *E. coli*. The distribution of resistant isolates did not significantly differ from non-colistin resistant isolates with respect to specimen type (p=0.318), department of admission (p=0.850), or gender (p=0.925). The median age of patients with colistin-resistant isolates was 70 years (IQR 60-78), compared with 65 years (IQR 55-73) among those with non-resistant isolates.

Table 1. Comparison of specimen sources, departments, and demographic characteristics between colistin-resistant and non-resistant *Escherichia coli* isolates.

	Number of isolates		
	Colistin resistant *E. coli* N=14 (1.2%)	**Non-colistin resistant *E. coli* N=1189 (98.8%)**	***p*-value**
Specimens			0.318
Blood	1 (7.1%)	183 (15.4%)	
Sputum	1 (7.1%)	118 (9.9%)	
Urine	9 (64.3%)	713 (59.9%)	
Pus	3 (21.4%)	95 (8.0%)	
Department			0.850
ICU	0	52 (4.4%)	
Non-ICU	14 (100%)	1,057 (95.6%)	
Gender			0.925
Male	6 (42.9%)	452 (38%)	
Female	8 (57.1%)	737 (62.0%)	
Age (years)			
Median age	70 (60-78)	65 (55-73)	

Phenotypic antimicrobial resistance profiles

Among the 14 colistin-resistant *E.coli* isolates, a high prevalence of multidrug resistance (MDR) was observed. All 14 colistin-resistant *E.coli* isolates exhibited ESBL phenotypes and met the MDR definition. The most frequent resistance pattern included ß-lactams, aminoglycosides, and fluoroquinolones, with six isolates also resistant to trimethoprim–sulfamethoxazole. Resistances to ampicillin and third-generation cephalosporins (including ceftazidime, ceftriaxone, and cefotaxime) were common, indicating the presence of an extended-spectrum ß-lactamase (ESBL)-like phenotype in all isolates. In addition, resistance to fluoroquinolones (ciprofloxacin and levofloxacin) was frequently detected among the same isolates, further supporting their classification as MDR strains. The full antimicrobial susceptibility profile is presented in Table 2.

Table 2. Antibiotic susceptibility pattern of colistin-resistant *E. coli* 14 isolates.

Isolates	Colistin (µg/mL)	Year	Sample	Resistant mechanism	MDR status	AK	CN	AP	AUG	TZP	CAZ	CRO	CTX	FOX	CIP	LEV	ETP	IMP	MEM	SXT
EC1	8	2021	Urine	ESBL	MDR	S	R	R	S	S	R	R	R	S	R	R	S	S	S	S
EC2	8	2021	Sputum	ESBL	MDR	S	R	R	S	S	R	R	R	S	R	R	S	S	S	S
EC3	8	2021	Urine	ESBL	MDR	S	R	R	S	S	R	R	R	S	R	R	S	S	S	S
EC4	8	2022	Urine	ESBL	MDR	S	R	R	S	S	R	R	R	S	R	R	S	S	S	R
EC5	8	2022	Pus	ESBL	MDR	S	R	R	S	S	R	R	R	S	R	R	S	S	S	R
EC6	>8	2022	Urine	ESBL	MDR	S	R	R	S	S	R	R	R	S	R	R	S	S	S	R
EC7	4	2022	Urine	ESBL	MDR	S	R	R	S	S	R	R	R	S	S	S	S	S	S	R
EC8	8	2022	Urine	ESBL	MDR	S	R	R	S	S	R	R	R	S	S	S	S	S	S	S
EC9	4	2022	Urine	ESBL	MDR	S	R	R	S	S	R	R	R	S	R	R	S	S	S	R
EC10	8	2022	Pus	ESBL	MDR	S	R	R	S	S	R	R	R	S	R	R	S	S	S	S
EC11	>8	2022	Urine	ESBL	MDR	S	S	R	S	S	R	R	R	R	R	R	S	S	S	S
EC12	>8	2022	Pus	ESBL	MDR	S	S	R	S	S	R	R	R	S	R	R	S	S	S	R
EC13	4	2022	Blood	ESBL	MDR	S	S	R	S	S	R	R	R	S	R	R	S	S	S	R
EC14	>8	2022	Urine	ESBL	MDR	S	R	R	S	S	R	R	R	S	S	S	S	S	S	R

Note: S: susceptible, I: intermediate, R: resistance, CT: colistin, AK: amikacin, CN: gentamicin, AP: ampicillin, AUG: amoxicillin-clavulanic acid, TZP: piperacillin-tazobactam, CAZ: ceftazidime, CRO: ceftriaxone, CTX: cefotaxime, FOX: cefoxitin, CIP: ciprofloxacin, LEV: levofloxacin, ETP: ertapenem, IMP: imipenem, MEM: meropenem, SXT: trimethoprim-sulfamethoxazole.

By contrast, susceptibility to carbapenems (ertapenem, imipenem, and meropenem) was retained in all isolates, suggesting the absence of carbapenemase producers in this collection. Among aminoglycosides, amikacin susceptibility was common, whereas gentamicin resistance was observed in several isolates.

Notably, three isolates (EC6, EC8, and EC14) exhibited largely susceptible phenotypes, with resistance confined to ampicillin or a limited number of agents, and two isolates EC9 and EC11 demonstrated intermediate susceptibility to cefoxitin or fluoroquinolones. Overall, the resistance pattern highlights the predominance of ESBL-producing MDR *E.coli* with preserved carbapenem susceptibility, while a minority of isolates displayed non-ESBL or susceptible phenotypes. The co-existence of ESBL-mediated multidrug resistance with colistin resistance indicates that these isolates represent high-risk lineages with limited therapeutic options.

Genotypic features of colistin-resistant E. coli isolates

The 14 colistin-resistant *E. coli* isolates represented diverse sequence types (STs) including high-risk extraintestinal *E. coli* lineages ST131 and ST95, as well as ST58, ST69, ST117, ST457. Multiple plasmid replicons were detected, predominantly IncF variants (FIA, FIB, FII) in combination with Col-type replicons. Plasmid backbones such as IncX4, IncI2, IncHI2, and IncFII carried *mcr* determinants. Two *mcr* variants were identified, namely *mcr*-1.1 and *mcr*-3.5. Several isolates carried both *mcr*-variants (EC3, EC8, EC14) with *mcr*-1.1 located on IncX4 or IncI2 plasmids replicon and *mcr*-3.5 on IncFII or IncX1 plasmids replicons. The co-existence of two *mcr* genes within single isolates suggests possible plasmid co-residence of multiple plasmids. In contrast, three isolates (EC9, EC11, EC12, EC13) lacked *mcr* genes, implying colistin resistance was mediated by chromosomal mutations. Point mutations were identified in multiple chromosomal loci implicated in colistin resistance. Variants in *mgrB* (e.g., N36S, N3K, Y12H), *phoPQ* (e.g., I44L, R6H, S138T), and *pmrAB* (e.g., S29G, D283G, H2R, V351I) were observed across isolates. These mutations were especially relevant in isolates without *mcr* genes, supporting their role in chromosomal-mediated colistin resistance.

An acquired resistance determinants were detected. The ESBL-associated genes, including $bla_{\text{CTX-M-14}}$, $bla_{\text{CTX-M-15}}$, $bla_{\text{CTX-M-24}}$, and $bla_{\text{CTX-M-55}}$, were predominant. The other β-lactamases such as $bla_{\text{OXA-1}}$ and $bla_{\text{CMY-42}}$ were also present. Aminoglycoside resistance genes included *aac(3)-IId*, *aac(3)-VIa*, and *aac(6')-Ib-cr*, while fluoroquinolone resistance was shown *to be the qnrS1* gene. The co-carriage of ESBLs and aminoglycoside or fluoroquinolone resistance genes highlights the multidrug-resistant *E. coli* of many isolates. The details of genomic features are in Table 3.

Table 3. Genomic features of colistin-resistant *E. coli* 14 isolates.

Isolates	Sequence types	Colistin (μg/mL)	*mcr* variant (location)	Plasmid replicon type	*mgrB*	*phoP*	*phoQ*	*pmrA*	*pmrB*	Acquired resistance genes
EC1	131	8	*mcr*-3.5 (IncFII)	Col-type, IncF (FIA, FIB, FIC), IncFII	N36S	I44L	WT	S29G, T31S	V351I, H2R, E123D, D283G	*bla* $_{\text{CTX-M-15}}$, *bla* $_{\text{OXA-1}}$, *aac(6)-Ib-cr, aac(3)-IIa*
EC2	648	8	*mcr*-3.5 (IncFII)	Col-type, IncF (FIA, FIB, FIC), IncFII	WT	I44L	L467M	S29G	A360V, H2R, D283G	*bla* $_{\text{TEM-1B}}$, *bla* $_{\text{CTX-M-55}}$, *qnS1, aac(3)-IId*
EC3	4679	8	*mcr*-1.1 (IncX4), *mcr*-3.5 (IncX1)	Col-type, IncF (FIB, FIC), IncI1-I, IncQ1, IncR, IncX1, IncX4	WT	I44L	WT	S29G	Y358N, D283G	*bla* $_{\text{CTX-M-55}}$, *bla* $_{\text{TEM-1B}}$, *aac(3)-IId*
EC4	58	8	*mcr*-1.1 (IncX4)	Col-type, IncF (FIA, FIB, FIC), IncI1-I, IncQ1, IncX1, IncX4, IncY	WT	I44L	WT	S29G	Y358N, D283G	*bla* $_{\text{TEM-1B}}$, *bla* $_{\text{CTX-M-55}}$, *qnS1, aac(3)-IId*
EC5	69	8	*mcr*-3.5 (IncFII)	Col440, ColRNAI, IncFII, IncQ1	WT	I44L	S138T, A482T	S29G	D283G, H2R, S138N	*bla* $_{\text{TEM-1B}}$, *qnS1, aac(3)-VIa*
EC6	117	>8	*mcr*-1.1 (IncHI2)	Col-type, IncB, IncF (FIB, FIC), IncHI2	WT	I44L	WT	S29G	A360V, H2R, D283G	*bla* $_{\text{TEM-1B}}$, *bla* $_{\text{CTX-M-55}}$, *qnS1, aac(3)-IId*
EC7	95	4	*mcr*-3.5 (IncFII)	Col-type, IncFIB, IncFII, IncQ1	WT	I44L	R6H	S29G, G144S, T31S, L128N	V351I, H2R, E123D, D283G	*bla* $_{\text{TEM-1B}}$, *aac(3)-IId*
EC8	58	8	*mcr*-1.1 (IncX4), *mcr*-3.5 (IncFII)	Col-type, IncB, IncF (FIB, FIC), IncHI2, IncFII, IncI1, IncX4	WT	I44L	WT	S29G	Y358N, D283G	*bla* $_{\text{TEM-1B}}$, *qnS1, aac(3)-IId, aac(3)-VIa*
EC9	457	4	-	Col-type, IncF (FIA, FIB, FII), IncX1, IncY	N3K, Y12H, N36S	I44L	L467M	S29G	D283G	*bla* $_{\text{TEM-1B}}$, *bla* $_{\text{CTX-M-55}}$, *qnS1, aac(3)-IId*
EC10	58	8	*mcr*-3.5 (IncFII)	Col-type, IncF (FIA, FIB, FIC), IncFII	WT	I44L	WT	S29G	Y358N, H2R, D283G	*bla* $_{\text{TEM-1B}}$
EC11	11021	>8	-	Col-type, IncF (FIA, FIB, FII), IncI	WT	I44L	WT	S29G, A105P	H2R	*bla* $_{\text{OXA-1}}$, *bla* $_{\text{CMY-42,}}$ *aac(6)-Ib-cr*
EC12	10	>8	-	IncFIB, IncI	WT	WT	WT	S29G	H2R	*bla* $_{\text{TEM-1B,}}$ *aac(3)-IId*
EC13	131	4	-	Col-type, IncF (FIA, FIB, FII)	N36S	I44L	WT	S29G, T31S	V351I, E123D, D283G	*bla* $_{\text{CTX-M-14}}$, *bla* $_{\text{CTX-M-24}}$
EC14	58	>8	*mcr*-1.1 and *mcr*-3.5 (IncI2)	Col-type, IncF (FIA, FIB, FII), IncI2	WT	I44L	WT	S29G	Y358N, H2R, D283G	*bla* $_{\text{TEM-1B}}$, *aac(3)-IId*

Analysis of virulence genes revealed that all 14 colistin-resistant *E. coli* isolates possessed multiple virulence determinants across diverse functional categories shown in Table 4. Adhesin-associated genes were universally present (14/14, 100%), including *fimH*, *pap* family genes *papA, papC, fdeC*, *yeh*, and *yfcV,* indicating a strong potential for host colonization and epithelial attachment. Iron acquisition systems were detected in almost all isolates (13/14, 92.9%), with siderophore-related gene *chuA, fyuA,* and *iutA* frequently observed, reflecting the ability to thrive in iron-limited host environments. Capsule and serum resistance genes *kps, iss, ompT, traT* were also common (12/14, 85.7%), supporting mechanisms for immune evasion and serum survival.

Table 4. Virulence genes distribution in colistin-resistant *E. coli*.

Functional group	Virulence genes	No. positive isolates (%)
Adhesins	*fimH, papA/papC, fdeC, focC, sfaD, hra, iha, lpfA, yeh, yfcV, csgA*	14/14 (100%)
Iron acquisition	*chuA, fyuA, iroN, ireA, irp2, iucC, iutA, sitA, katP*	13/14 (92.9%)
Capsule/Serum resistance	*kpsE, kpsMII_K1/K5/K11, neuC, iss, ompT, traT*	12/14 (85.7%)
Toxins	*hlyA/E/F, cnf1, sat, vat, usp, pic, tsh, cma, cib*	9/14 (64.3%)
Invasins	*tia, eilA, AslA*	5/14 (35.7%)
Other/Regulators	*anr, gad, hha, nlpI, terC, shiA/B, capU, etsC, cvaC, mchF, traJ*	14/14 (100%)

Toxin-associated genes *hlyA, cnf1, sat*, and *vat* showed variable distribution (9/14, 64.3%), indicating heterogeneity in cytotoxic potential among the isolates. Invasion gene *tia, eilA, AslA* were less prevalent (5/14, 35.7%), suggesting that invasive capacity is limited to a subset of strains. Conversely, regulatory and stress-response genes *anr, gad, hha, nlpI, terC, shiA, capU, etsC, cvaC*, *mchF, traT* were consistently present in all isolates (14/14, 100%), underscoring the genomic stability of stress-adaptive and plasmid maintenance systems in this MDR *E. coli* collection.

Discussion

The prevalence of colistin-resistant *E. coli* has been reported worldwide, including in livestock from farms in Japan[14,15] and China.[1] In Thailand, previous studied have identified strains in slaughter pigs from farms[7] and in clinical samples from tertiary hospitals.[8] This study highlights the complex epidemiology of colistin-resistant *E. coli* in Lerdsin General Hospital a tertiary hospital in Thailand. The isolates belonged to diverse sequence types including high-risk extrapathogenic *E. coli* lineages such as ST131[16] and ST648, as well as virulent but emerging high risk lineages like ST95. Community and animal-associated STs, including ST58 and ST69,[17] were also detected. Together, these findings indicate polyclonal dissemination, suggesting multiple introductory events from both clinical and non-clinical reservoirs rather than clonal expansion within the hospital.

Phenotypic testing revealed that most isolates displayed an ESBL-like multidrug-resistant profile with additional fluoroquinolone resistance, although carbapenem susceptibility was preserved. A report in this study is that most colistin-resistant *E. coli* isolates also exhibited an ESBL phenotype. This co-occurrence may be explained by the carriage of both resistance determinants on the same or compatible plasmid backbones. In several isolates, *mcr* genes were located on IncF or IncI-type plasmids, which are also known to frequently harbor ESBL genes such as bla_{CTX-M}. The genetic linkage of *mcr* and ESBL-associated genes within the same plasmid increases the likelihood of co-selection.[18] Although most colistin-resistant *E. coli* isolates in this study exhibited an ESBL phenotype, phenotypic testing showed that several antimicrobial agents remained effective. All isolates were susceptible to carbapenems, and most retained susceptibility to amikacin. The clinicians may still consider carbapenems or selected aminoglycosides as therapeutic options. However, the broad-spectrum activity of these agents carries a risk of further resistance, underscoring the need for antimicrobial stewardship and ongoing genomic surveillance.

The detection of *mcr*-1.1 and *mcr*-3.5 on highly mobile plasmids (IncX4, IncI2, IncFII, IncHI2) underscores their potential for horizontal transfer across *E. coli* lineages and possibly other Enterobacterales.[19] The co-existence of both *mcr* variants in single isolates reflects ongoing plasmid creates opportunities for accelerated dissemination. Notably, both *mcr* variants have been widely reported in livestock, especially swine, raising the likelihood of zoonotic or foodborne transmission.[1,20]

Virulence gene profiles, including adhesins, siderophore systems, and capsule-associated genes were widely distributed, suggesting that these resistant strains also retain significant pathogenic potential. The convergence of virulence and resistance is particularly concerning lineages such as ST131, which are globally linked to invasive infections.[21]

Our findings highlight the connection of patients in hospitals, animals, and the community in the emergence of colistin resistance. The core evidence supporting the one health implication lies in the molecular characteristics of the isolates. Specifically, the detection of *mcr*-3.5 within clinical isolates is a key finding. The *mcr*-3 group, and particularly *mcr*-3.5 has been widely reported in livestock, especially swine,[22,23] in Thailand and globally suggesting a strong non-human reservoir linkage.[24] The identification of this specific variant in human clinical samples strongly emphasizes the potential for zoonotic or foodborne transmission from the animal/agricultural sector or the community into the hospital setting. Furthermore, the colistin-resistant isolates belonged to diverse Sequence Types (STs), including community and animal-associated lineages such as ST58 and ST69. This polyclonal dissemination suggests multiple introduction events from both clinical and non-clinical reservoirs rather than clonal expansion within the hospital. The co-existence of these community/animal-associated STs alongside high-risk clinical lineages (e.g., ST131) further strengthens the argument for cross-sectoral movement of colistin resistance genes, necessitating a One Health approach.[25] The mechanism facilitating this cross-sectoral spread is likely the carriage of *mcr* determinants on highly mobile plasmids (IncX4, IncI2, IncFII, IncHI2).[26] These plasmids can circulate widely and transfer resistance genes across different bacterial hosts and lineages. These observations underscore the urgent need for tighter regulation of antimicrobial use in animals and the integration of genomic surveillance with stewardship programs under a One Health framework to guide preventive interventions across clinical, agricultural, and community settings.

Conclusion

This study demonstrates that colistin-resistant *E. coli* isolated from a tertiary hospital in Thailand are genetically diverse, encompassing high-risk extrapathogenic *E. coli* lineages and community-associated strains. The frequent coexistence of colistin resistance with ESBL phenotypes, often mediated by plasmid-borne determinants, underscores the threat of multidrug resistance. Notably, the detection of *mcr*-3.5, previously reported in livestock, within clinical isolates highlights the possibility of zoonotic or foodborne transmission. Plasmids carrying *mcr* genes possess strong potential for horizontal transfer across bacterial hosts, facilitating rapid dissemination. These findings emphasize the need for rational antimicrobial use, continuous genomic surveillance, and integrated stewardship efforts under a One Health framework to mitigate the spread of colistin resistance across human, animal, and environmental sectors.

Ethical approval

This study was exempted from ethical review for human research by the Ethics Committee for Human Research at Lerdsin Hospital, Department of Medical Services, under project code LH661016.Laboratory work complied with national biosafety and biosecurity regulations under supervision of the institutional Biosafety and Biosecurity Committee.

Funding

Whole genome sequencing charge from Assoc. Prof. Dr. Panan Ratthawongjirakul

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Kwanchon Jearakitiwanich: conceptualization, methodology, data curation, formal analysis, writing: original draft, review and editing, project administration; **Anusak Kerdsin**: data analysis guidance, bioinformatics supervision; **Monchai Siribamrungwong:** clinical supervision, review of antimicrobial resistance aspects; **Panan Ratthawongjirakul**: funding acquisition, resources, project support.

Acknowledgements

None.

References

[1] Li H, Liu Y, Yang L, Wu X, Wu Y, Shao B. Prevalence of *Escherichia coli* and antibiotic resistance in animal-derived food samples - Six districts, Beijing, China, 2020. China CDC Wkly. 2021; 3(47): 999-1004. doi: 10.46234/ccdcw2021.243.

[2] El-Sayed Ahmed MAE-G, Zhong L-L, Shen C, Yang Y, Doi Y, Tian G-B. Colistin and its role in the era of antibiotic resistance: an extended review (2000–2019). Emerg Microbes infect. 2020; 9(1): 868-85. doi: 10.1080/22221751.2020.1754133.

[3] Eiamphungporn W, Yainoy S, Jumderm C, Tan-Arsuwongkul R, Tiengrim S, Thamlikitkul V. Prevalence of the colistin resistance gene *mcr*-1 in colistin-resistant *Escherichia coli* and *Klebsiella pneumoniae* isolated from humans in Thailand. J Glob Antimicrob resist. 2018; 15: 32-5. doi: 10.1016/j.jgar.2018.06.007.

[4] Sumpradit N, Wongkongkathep S, Malathum K, Janejai N, Paveenkittiporn W, Yingyong T, et al. Thailand's national strategic plan on antimicrobial resistance: progress and challenges. Bull World Health Organ. 2021; 99(9): 661-73. doi::10.2471/BLT.20.280644

[5] Paveenkittiporn W, Kamjumphol W, Ungcharoen R, Kerdsin A. Whole-genome sequencing of clinically isolated carbapenem-resistant Enterobacterales harboring *mcr* genes in Thailand, 2016-2019. Front microbiol. 2021; 11: 586368. doi: 10.3389/fmicb.2020.586368

[6] Tangsawad W, Kositamongkol C, Chongtrakool P, Phisalprapa P, Jitmuang A. The burden of carbapenem-resistant Enterobacterales infection

in a large Thai tertiary care hospital. Front Pharmacol. 2022; 13: 972900. doi.org/10.3389/fphar.2022.972900

[7] Khanawapee A, Kerdsin A, Chopjitt P, Boueroy P, Hatrongjit R, Akeda Y, et al. Distribution and molecular characterization of *Escherichia coli* harboring *mcr* genes isolated from slaughtered pigs in Thailand. Microb Drug Resist. 2021; 27(7): 971-9. doi: 10.1089/mdr.2020.0242.

[8] Nobthai P, Ruekit S, Peerapongpaisarn D, Sukhchat P, Swierczewski BE, Ruamsap N, et al. The co-existence of *mcr*-1.1 and *mcr*-3.5 in *Escherichia coli* isolated from clinical samples in Thailand. Antibiotics. 2025; 14(6): 596. doi.org/10.3390/antibiotics14060596

[9] Thermo Scientific. Instructions for use sensititre susceptibility plates. 2021.

[10] Clinical and Laboratory Standards Institute. Performance standards for antimicrobial susceptibility testing; Approved standard document M100S. 2021.

[11] Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res. 2017; 45: D566-D573. doi: 10.1093/nar/gkw1004.

[12] Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res. 2006; 34(Database issue): D32-6. doi: 10.1093/nar/gkj014.

[13] Beghain J, Bridier-Nahmias A, Le Nagard H, Denamur E, Clermont O. ClermonTyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping. Microb Genom. 2018; 4(7): e000192. doi: 10.1099/mgen.0.000192.

[14] Nakano A, Nakano R, Nishisouzu R, Suzuki Y, Horiuchi S, Kikuchi-Ueda T, et al. Prevalence and relatedness of *mcr*-1-mediated colistin-resistant *Escherichia coli* isolated from livestock and farmers in Japan. Front Microbiol. 2021; 12: 664931. doi: 10.3389/fmicb.2021.664931

[15] Kawano K, Masaki T, Kawaguchi T, Kuroda M. Persistence of colistin resistance and *mcr*-1.1-positive *E. coli* in poultry despite colistin ban in Japan. Antibiotics. 2025; 14(4): 360. doi: 10.3390/antibiotics14040360

[16] Cho S-T, Mills EG, Griffith MP, Nordstrom HR, McElheny CL, Harrison LH, et al. Evolution of extended-spectrum β-lactamase-producing ST131 *Escherichia coli* at a single hospital over 15 years. Sci Rep. 2024; 14(1): 19750. doi: 10.1186/s13104-022-06079-z

[17] Benlabidi S, Raddaoui A, Lengliz S, Cheriet S, Hynds P, Achour W, et al. Occurrence of high-risk clonal lineages ST58, ST69, ST224, and ST410 among extended-spectrum β-lactamase-producing *Escherichia coli* isolated from healthy free-range chickens (Gallus gallus domesticus) in a rural region in Tunisia. Genes. 2023; 14(4): 875. doi: 10.3390/genes14040875.

[18] Wranne MS, Karami N, Kk S, Jaén-Luchoro D, Yazdanshenas S, Lin Y-L, et al. Comparison of CTX-M encoding plasmids present during the early phase of the ESBL pandemic in western Sweden. Sci Rep. 2024; 14(1): 11880. doi: 10.1038/s41598-024-62663-2.

[19] Dadashi M, Sameni F, Bostanshirin N, Yaslianifard S, Khosravi-Dehaghi N, Nasiri MJ, et al. Global prevalence and molecular epidemiology of *mcr*-mediated colistin resistance in *Escherichia coli* clinical isolates: a systematic review. J of Glob Antimicrob Resist. 2022; 29: 444-61. doi: 10.1016/j.jgar.2021.10.022

[20] Xu Y, Zhong LL, Srinivas S, Sun J, Huang M, Paterson DL, et al. Spread of *MCR*-3 Colistin Resistance in China: An Epidemiological, Genomic and Mechanistic Study. EBioMedicine. 2018; 34: 139-57. doi: 10.1016/j.ebiom.2018.07.027.

[21] Chopjitt P, Boueroy P, Morita M, Iida T, Akeda Y, Hamada S, et al. Genetic characterization of multidrug-resistant *Escherichia coli* harboring colistin-resistant gene isolated from food animals in food supply chain. Front Cell Infect Microbiol. 2024;14: 1289134. doi::10.3389/fcimb.2024.1289134.

[22] Sudatip D, Mostacci N, Tiengrim S, Thamlikitkul V, Chasiri K, Kritiyakan A, et al. The risk of pig and chicken farming for carriage and transmission of *Escherichia coli* containing extended-spectrum beta-lactamase (ESBL) and mobile colistin resistance (*mcr*) genes in Thailand. Microb Genom. 2023; 9(3): e000951. doi: 10.1099/mgen.0.000951.

[23] Pungpian C, Lee S, Trongjit S, Sinwat N, Angkititrakul S, Prathan R, et al. Colistin resistance and plasmid-mediated *mcr* genes in *Escherichia coli* and *Salmonella* isolated from pigs, pig carcass and pork in Thailand, Lao PDR and Cambodia border provinces. J Vet Sci. 2021; 22(5): e68. doi: 10.4142/jvs.2021.22.e68

[24] Boonyasiri A, Brinkac LM, Jauneikaite E, White RC, Greco C, Seenama C, et al. Characteristics and genomic epidemiology of colistin-resistant Enterobacterales from farmers, swine, and hospitalized patients in Thailand, 2014-2017. BMC Infect Dis. 2023; 23(1): 556. doi: 10.1186/s12879-023-08539-8.

[25] Yamamoto Y, Higashi A, Ikawa K, Hoang HTT, Yamaguchi T, Kawahara R, et al. Horizontal transfer of a plasmid possessing *mcr*-1 marked with a single nucleotide mutation between *Escherichia coli* isolates from community residents. BMC Research Notes. 2022; 15(1):196. doi: 10.1186/s13104-022-06079-z

[26] Leangapichart T, Stosic MS, Hickman RA, Lunha K, Jiwakanon J, Angkititrakul S, et al. Exploring the epidemiology of *mcr* genes, genetic context and plasmids in *Enterobacteriaceae* originating from pigs and humans on farms in Thailand. J Antimicrob Chemother. 2023; 78(6): 1395-405. doi: 10.1093/jac/dkad097

Journal of Associated Medical Sciences 2026; 59(2): 44-62

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Comparative evaluation and interpretability analysis of modern CNN architectures for brain tumor MRI classification

Nitipon Pongphaw* and Prommin Buaphan

Department of Electrical and Computer Engineering, Faculty of Science and Engineering, Kasetsart University Chalermphrakiat Sakon Nakhon Province Campus, Sakon Nakhon, Thailand.

ARTICLE INFO

Article history:
Received 3 November 2025
Accepted as revised 14 January 2026
Available online 23 January 2026

Keywords:
Brain tumor classification, magnetic resonance imaging, comparative study, transfer learning, grad-CAM.

ABSTRACT

Background: Accurate and interpretable brain tumor classification from MRI images remains a key challenge in medical image analysis, particularly when using publicly available datasets of moderate size.

Objective: This study investigates the performance of a ConvNeXt-Tiny based framework for four-class brain tumor classification glioma, meningioma, pituitary tumor, and no tumor and compares it with established convolutional architectures.

Materials and methods: Using transfer learning and identical experimental settings, ConvNeXt-Tiny was evaluated against DenseNet169, Xception, MobileNetV3-Large, CNN+DenseNet169, and ResNet50. Standard evaluation metrics (accuracy, precision, recall, and F1-score) were used, and Grad-CAM was applied to visualize model attention for interpretability. Generalization was further assessed using an independent dataset.

Results: ConvNeXt-Tiny achieved high overall performance (accuracy = 0.9924, F1-score = 0.9918), comparable to DenseNet169 and Xception but with lower computational cost. The model maintained stable learning behavior, minimal overfitting, and consistent accuracy on unseen data. Grad-CAM visualizations confirmed that the network focused on clinically relevant tumor regions, improving transparency and reliability of predictions.

Conclusion: ConvNeXt-Tiny provides a strong and efficient baseline for interpretable brain tumor classification, balancing accuracy and computational efficiency. While the results are promising, differences among recent architectures were modest, and clinical validation using multi-center MRI datasets is necessary to confirm broader applicability.

Introduction

The classification of brain tumors using deep learning techniques has become a transformative approach in medical imaging, particularly in magnetic resonance imaging (MRI). This progress is largely driven by the capability of deep learning models, especially Convolutional Neural Networks (CNNs), to automatically extract features from images and improve diagnostic accuracy.[1,2] Recent studies have demonstrated that integrating advanced architectures, such as residual networks and attention mechanisms, can enhance performance in differentiating various types of brain tumors, including gliomas and meningiomas.[3,4] Transfer learning has been widely

* **Corresponding contributor.**
Author's Address: *Department of Electrical and Computer Engineering, Faculty of Science and Engineering, Kasetsart University Chalermphrakiat Sakon Nakhon Province Campus, Sakon Nakhon, Thailand.*
E-mail address: nitipon.p@ku.th
doi: 10.12982/JAMS.2026.036
E-ISSN: 2539-6056

adopted to further improve classification accuracy, with models pre-trained on large datasets being fine-tuned on specific radiological datasets to better capture the nuances of brain tumor images.[1,5,6] For example, the use of ResNet50 and other architectures has shown significant improvements in metrics such as accuracy and F1-score.[2,7] Moreover, deep learning approaches have been reported to outperform traditional diagnostic methods, which often rely heavily on subjective interpretation.[8,9] Numerous studies have also explored model optimization strategies. Hybrid models combining CNNs with other machine learning classifiers, such as SVMs and decision trees, have shown improvements in both performance and generalizability.[10,11] Architectural innovations, including attention blocks and multi-tiered networks, aim to reduce overfitting while enhancing model interpretability.[3,12] Additionally, metaheuristic optimization techniques have been applied for feature selection to strengthen classification robustness.[13] Context-aware deep learning frameworks have further enriched the capability to segment and classify brain tumors while also predicting patient survival rates.[14] Multi-faceted strategies of this nature have demonstrated significant improvements in clinical outcomes, providing healthcare practitioners with effective tools for brain tumor diagnosis and management.[15] Ongoing research continues to advance deep learning applications in brain tumor classification, highlighting the potential for future clinical adoption. The ability to achieve high accuracy while handling the complex and heterogeneous nature of brain tumors underscores the critical role of deep learning in enhancing diagnostic workflows and patient care. Explainable Artificial Intelligence (XAI) techniques enhance model transparency and interpretability, allowing clinicians to understand the decision-making process behind deep learning predictions. Techniques such as Local Interpretable Model-agnostic Explanations (LIME) and Gradient-weighted Class Activation Mapping (Grad-CAM) have been employed to explain how CNNs classify tumor types and detect tumor presence. For instance, Ullah *et al*. proposed DeepEBTDNet, which effectively utilized LIME to interpret MRI-based predictions, achieving a validation accuracy of 98.96%.[16] Similarly, Esmaeili *et al*. reported that incorporating interpretable methods could improve the accuracy of trained models.[17] Grad-CAM has been applied to visualize image regions most influential to the model's decisions. Goyal and Sharma used Grad-CAM to highlight features driving brain tumor detection, enhancing trust in AI-assisted diagnostics.[18] Tan *et al*. demonstrated the practical application of these visualization techniques in clinical workflows, improving diagnostic reliability and addressing concerns about the "black-box" nature of deep learning models, though reference[19] does not directly support Grad-CAM application in this context. Beyond interpretability, these techniques also enhance model robustness and reliability by showing which features are considered important, enabling practitioners to validate them against established medical knowledge.[20] Furthermore, attention mechanisms incorporated into deep learning frameworks allow models to focus on clinically relevant regions in MRI scans, which is critical for tumor segmentation and classification. Tonni *et al*. introduced a hybrid transfer learning framework incorporating attention-based features to improve both interpretability and performance.[19] Despite significant advances in deep learning-based brain tumor classification, several research gaps remain that limit model robustness and clinical applicability. One major gap is the reliance on large, well-annotated datasets for effective model training. Many studies rely on publicly available datasets that vary in size, balance, and acquisition conditions, which may lead to overfitting and reduced generalization when models are applied to unseen or heterogeneous MRI scans.[2] This highlights the importance of evaluating deep learning frameworks under realistic data availability and heterogeneous imaging conditions, while avoiding over-reliance on excessively large or curated clinical datasets. Another critical gap is the lack of interpretability of deep learning models in clinical practice. Although CNN-based architectures achieve high accuracy, their "black-box" nature makes it difficult for clinicians to understand the rationale behind predictions, reducing trust and hindering adoption in real-world healthcare settings.[20]

Materials and methods

This study proposes a deep learning framework for brain tumor classification from MRI images. The overall workflow is illustrated in Figure 1.

The methodology consists of five main components:

1. Proposed approach: This study presents a fair experimental comparison between ConvNeXt-Tiny as the baseline model and other deep learning architectures for four-class brain tumor classification. Beyond multi-class prediction, the proposed approach provides additional insights through subclass-level tumor analysis. Furthermore, binary classification between tumor and no-tumor cases is investigated to offer a clearer understanding of the model's diagnostic behavior.

2. Model configuration: The model performance was also evaluated in comparison with several CNN architectures, including ResNet50, CNN+DenseNet169, DenseNet169, MobileNetV3-Large, and Xception, using transfer learning.

3. Data preprocessing: The images were organized into training, validation, and test sets. All images were resized to 224×224 pixels prior to model training.

4. Evaluation: Performance of all models was assessed using standard metrics, including accuracy, precision, recall, and F1-score. Confusion matrix was used to visualize class-wise performance. In addition, training efficiency was evaluated by recording the

total training time and the average training time per epoch. These measurements provide insight into the computational cost of each model, allowing for a holistic comparison of predictive performance and training efficiency.

5. Grad-CAM: To provide insight into model decisions, Grad-CAM (Gradient-weighted Class Activation Mapping) was applied to the trained models. Grad-CAM generates heatmaps highlighting the regions of input images that contribute most strongly to the predicted class. This allows visual assessment of whether the model focuses on relevant tumor regions, supporting interpretability and trustworthiness of the predictions.

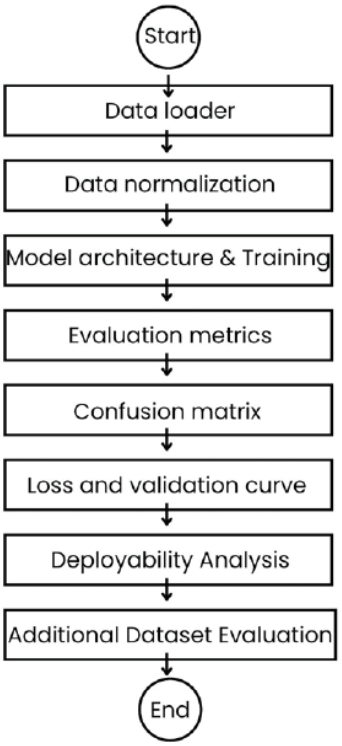


Figure 1. *Flowchart of methodology.*

Contributions

To address these challenges, this study makes two primary contributions. First, a fair and systematic comparison is conducted using ConvNeXt-Tiny as the baseline model to evaluate the performance of other deep learning architectures under identical training and evaluation conditions, thereby ensuring an objective assessment under identical experimental settings using a publicly available benchmark dataset. Second, the framework incorporates interpretability-enhancing mechanisms, including Grad-CAM analysis, to provide clear and clinically meaningful explanations for each classification decision. This integrated approach enhances diagnostic reliability while fostering clinical confidence and practical applicability in real-world brain tumor analysis.

Model architecture

ConvNeXt-Tiny was selected as the baseline model due to its modern convolutional architecture that bridges traditional CNNs with transformer-inspired design principles. By incorporating updated architectural choices, including larger kernel sizes, depthwise separable convolutions, and simplified normalization and activation strategies, ConvNeXt-Tiny enhances representational capacity while maintaining computational efficiency. The model was chosen a priori to balance expressive power and parameter efficiency, thereby reducing the risk of overfitting and remaining compatible with gradient-based interpretability techniques such as Grad-CAM. Larger ConvNeXt variants and transformer-based architectures were not prioritized because of their higher computational requirements and stronger dependence on large-scale training data.

Compared with lightweight models such as MobileNetV3-Large and deeper architectures including DenseNet169 and ResNet50, ConvNeXt-Tiny offers a well-balanced trade-off among model complexity, predictive performance, and generalization capability, making it a robust and reliable baseline for fair comparison. To ensure a comprehensive evaluation, six deep learning models were selected for comparison: CNN+DenseNet169, Xception, DenseNet169, ConvNeXt-Tiny, MobileNetV3-Large, and ResNet50. All models were trained for 20 epochs using the Adam optimizer with a learning rate of 1×10^{-4} and CrossEntropyLoss as the objective function on a four-class dataset. Detailed hyperparameter settings are summarized in Table 1, and schematic architectures of all models are illustrated in Figure 2.

Table 1. Parameter settings of all models.

Model	Parameters
CNN+DenseNet169	CNN(Conv2d(64) → BatchNorm → ReLU → MaxPool → Conv2d(128) → BatchNorm → ReLU → MaxPool) → Adapter (Conv2d 128→3) → Upsample (224×224) → DenseNet169(pretrained) → Dense(num_classes)
Xception	Xception(pretrained) → Dense(num_classes)
DenseNet169	DenseNet169(pretrained) → Dense(num_classes)
ConvNeXt-Tiny	ConvNeXt-Tiny(pretrained) → Dense(num_classes)
MobileNetV3-Large	MobileNetV3-Large(pretrained) → Dense(num_classes)
ResNet50	ResNet50(pretrained) → Dense(num_classes)

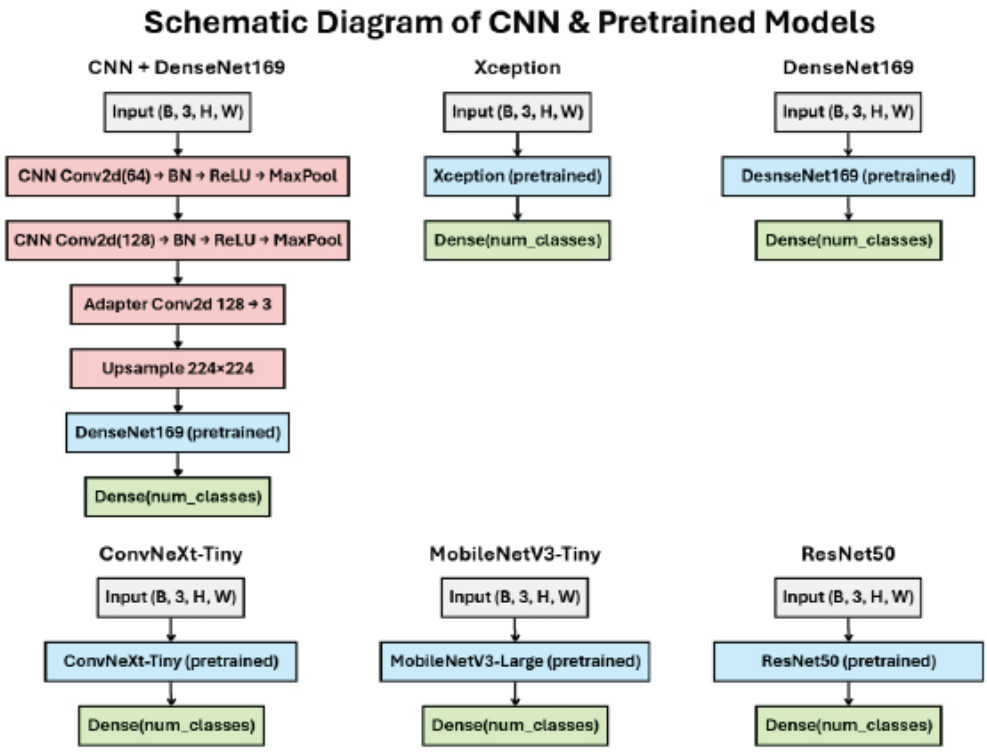


Figure 2. *Schematic diagrams of all models.*

Dataset

In this study, Dataset 1 was adopted from,[21] which is distributed under the CC0 license, and was used as the primary dataset for model training and initial evaluation. This dataset consists of 5,139 training images, 573 validation images, and 1,311 test images. For external validation, Dataset 2 was employed, namely the *Labeled MRI Brain Tumor Dataset Computer Vision Model* curated by Ali Rostami,[22] which is publicly available under the CC BY 4.0 license and contains a total of 2,443 MRI images. Dataset 2 was exclusively used to assess the generalization capability of the proposed model on unseen data. Detailed descriptions of both datasets, including class composition and source attribution, are provided in the Supplementary File. Figure 3 illustrates representative samples from each class in Dataset 1, while Figure 4 presents example images from Dataset 2.

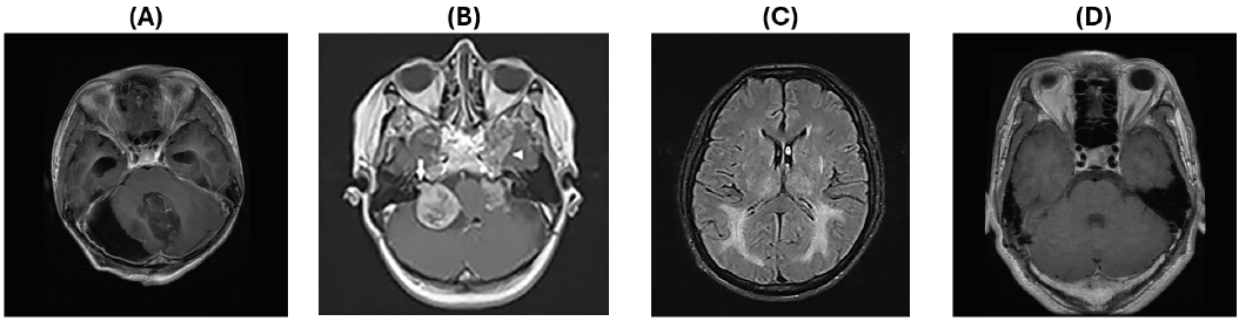


Figure 3. *Sample images from dataset 1. A: glioma, B: meningioma, C: no tumor, D: pituitary tumor.*

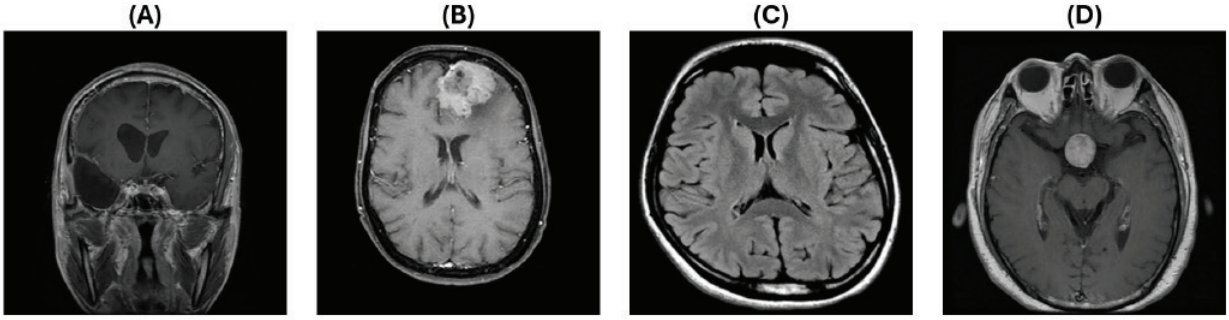


Figure 4. *Sample images from alternative dataset. A: glioma, B: meningioma, C: no tumor, D: pituitary tumor.*

Data preprocessing

In this study, all MRI images from both datasets were preprocessed before being used in the deep learning models. Dataset 1 was divided into training and validation sets using a 90:10 stratified split to preserve class balance, while all images in Dataset 2 were used exclusively for testing. All images were resized to 224×224 pixels, the standard input size for most convolutional neural networks (CNNs), to ensure consistency across architectures, and then converted into PyTorch tensors. The data were loaded using PyTorch's *ImageFolder* and *DataLoader* with a batch size of 32. Training data were shuffled at each epoch, whereas validation and test data were loaded sequentially. The final dataset included 5,139 training images, 573 validation images, and 1,311 test images from Dataset 1, together with all images from Dataset 2 used for external testing.

Evaluation metrics

To assess the performance of the proposed models, the following standard evaluation metrics were used: accuracy, precision, recall, and F1-score. Accuracy measures the proportion of correctly classified images among all images. Precision indicates the proportion of correctly predicted positive samples among all samples predicted as positive. Recall represents the proportion of correctly predicted positive samples among all actual positive samples. F1-score is the harmonic means of precision and recall, providing a balance between the two. All metrics were calculated using the macro-average approach, which treats each class equally by computing the metric independently for each class and then averaging the results, regardless of class imbalance.

Computational resources

All experiments were executed on Google Colab Pro using an NVIDIA H100 GPU. Model training and inference were performed using CUDA acceleration, as indicated by the device configuration (using device: cuda).

Results

In this study, we hypothesized that the ConvNeXt-Tiny model could serve as an effective baseline for comparative evaluation of convolutional architectures in multi-class brain tumor classification. All models were trained and evaluated under identical experimental settings to ensure fair comparison.

Evaluation metrics

The performance of all evaluated models is summarized in Table 2 Overall, the results show that several architectures achieved comparably high performance under identical experimental settings. DenseNet169 yielded the highest accuracy (0.9969), followed closely by MobileNetV3-Large (0.9947) and Xception (0.9939). ConvNeXt-Tiny, used as the main baseline in this study, also demonstrated competitive results (accuracy=0.9924, precision=0.9920, recall=0.9917, F1-score=0.9918), indicating its reliability and efficiency for this task. CNN+DenseNet169 performed slightly lower but remained strong overall, while ResNet50 showed comparatively lower metrics across all measures. These findings suggest that multiple convolutional architectures can achieve similar performance levels in multi-class brain tumor classification, emphasizing the value of comparative evaluation rather than reliance on a single model.

Table 2. Performance of all models.

Model	Accuracy	Precision	Recall	F1-score
CNN+Densenet169	0.9916	0.9914	0.9909	0.9912
Xception	0.9939	0.9936	0.9934	0.9934
Densenet169	0.9969	0.9969	0.9967	0.9968
ConvNeXt-Tiny	0.9924	0.9920	0.9917	0.9918
MobileNetV3-Large	0.9947	0.9945	0.9942	0.9943
ResNet50	0.9680	0.9672	0.9661	0.9664

Loss and validation curve

Figure 5 illustrates the training and validation loss and accuracy curves of the baseline ConvNeXt-Tiny model. Both curves demonstrate a rapid stabilization during the early training phase, with the loss decreasing sharply and reaching a steady level around the fifth epoch, while the accuracy curves gradually converged and remained closely aligned throughout the remaining epochs. This behavior indicates a stable training process with no evident overfitting or divergence between training and validation performance. Compared with prior studies employing ConvNeXt-based or other deep CNN architectures for image classification, which often report longer training schedules ranging from 30 to over 100 epochs depending on the dataset and imaging modality, the convergence observed within 20 epochs in this study suggests relatively fast learning dynamics. This behavior is primarily attributed to the use of transfer learning, which enables the ConvNeXt-Tiny model to exploit pretrained feature representations and reduces the need for extensive parameter updates. In addition, the relatively consistent structural characteristics of brain MRI images may facilitate efficient feature adaptation compared with natural image datasets.

Training and validation loss and accuracy curves for baseline model are presented in Figure 5.

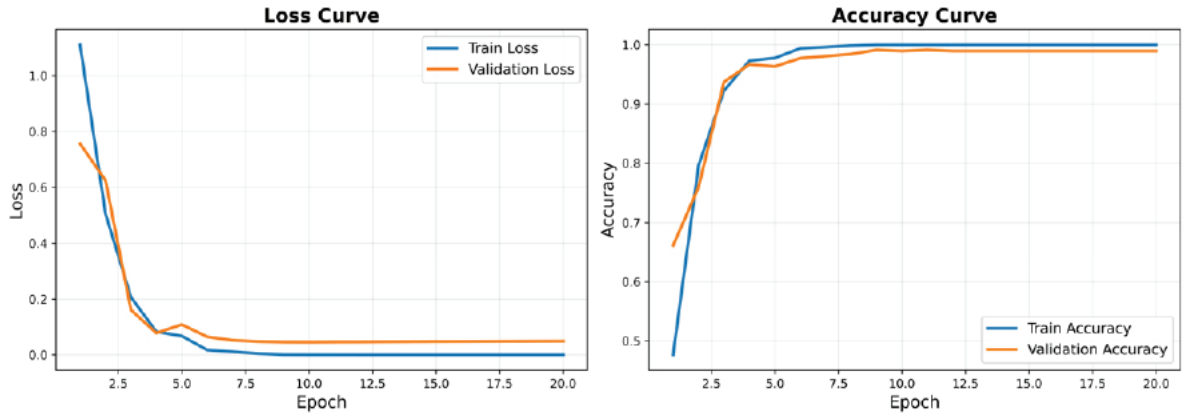


Figure 5. *Training and validation performance curves of the ConvNeXt-Tiny.*

Confusion matrix

To gain further insight into the classification performance of each model, a confusion matrix was constructed to illustrate class-wise predictions for the four brain tumor categories: glioma, meningioma, no tumor, and pituitary tumor.

Table 3 shows the confusion matrix of the baseline ConvNeXt-Tiny model on the test set. The model correctly classified most samples across all four categories, with only a few misclassifications observed between *glioma* and *meningioma*, and between *meningioma* and *pituitary*. The "no tumor" class was predicted with perfect consistency, indicating clear separability from tumor classes under the current dataset. These results suggest that the baseline model maintained balanced recognition among categories, with minor overlaps primarily occurring between histologically related tumor types.

Table 3. Confusion matrix of the ConvNeXt-Tiny model.

	Predicted glioma	Predicted meningioma	Predicted notumor	Predicted pituitary
Actual glioma	295	4	0	1
Actual meningioma	2	303	1	0
Actual no tumor	0	0	405	0
Actual pituitary	0	2	0	298

Learning curve

The learning curve (Figure 6) analysis indicates that both validation and test accuracy increase consistently with the number of training samples, demonstrating that the model benefits substantially from additional data. Early in training, with fewer than 200 samples, accuracy remains relatively low (~0.63-0.88), but a marked improvement is observed as the dataset grows to 500-1000 samples, reaching ~0.95. Beyond 1000 samples, gains in accuracy begin to plateau, indicating diminishing returns, with the model approaching its maximum performance at 5163 samples (~0.99). The close alignment between validation and test accuracy across all sample sizes suggests minimal overfitting and good generalization.

Stratified sampling was used to ensure each class contributed equally to the training subsets, preventing bias and supporting balanced learning across all classes.

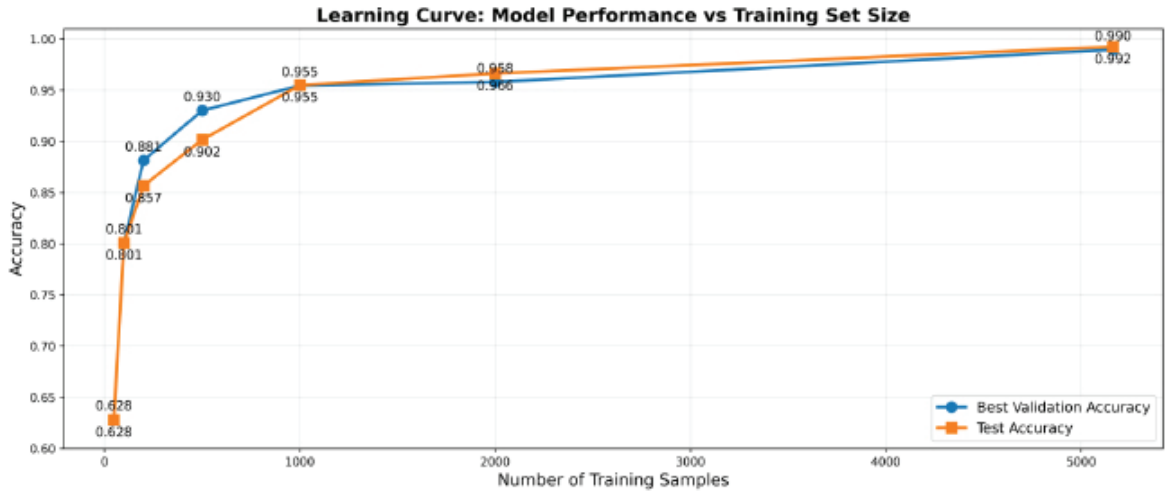


Figure 6. *Effect of training sample size on ConvNeXt-Tiny accuracy*

Statistical test

Null hypothesis (H_0)

There is no difference in the proportion of correctly classified samples between ConvNeXt-Tiny and the compared model.

Alternative hypothesis (H_1)

There is a significant difference in the proportion of correctly classified samples between ConvNeXt-Tiny and the compared model.

Table 4 summarized of the results of McNemar's tests comparing ConvNeXt-Tiny with other models indicate that for most comparisons, there is no statistically significant difference in performance. Specifically, comparisons with CNN+DenseNet169 (χ^2=0.0, p=1.0), Xception (χ^2=0.083, p=0.773), DenseNet169 (χ^2=3.125, p=0.077), and MobileNetV3 (χ^2=0.444, p=0.505) all yielded p>0.05, indicating that their performance differences with ConvNeXt-Tiny are not statistically significant. In contrast, the comparison with ResNet50 (χ^2=22.881, p<0.001) shows a highly significant difference, confirming that ConvNeXt-Tiny outperforms ResNet50. Overall, these results suggest that ConvNeXt-Tiny performs comparably to most modern architectures while significantly surpassing the older ResNet50 model.

Table 4. McNemar's test comparison of ConvNeXt-Tiny with other baseline models.

Baseline	Compared Model	Chi-square	*p* value	Significant (p<0.05)
ConvNeXt-Tiny	CNN+DenseNet169	0.0	1.0	No
ConvNeXt-Tiny	Xception	0.0833	0.7728	No
ConvNeXt-Tiny	DenseNet169	3.1250	0.0771	No
ConvNeXt-Tiny	MobileNetV3	0.4444	0.5050	No
ConvNeXt-Tiny	ResNet50	22.8810	1.7235×10^{-6}	Yes

The 95% confidence interval (CI)

According to Table 5, the six models show clear differences in both accuracy and Macro-F1. DenseNet169 achieved the highest performance (Accuracy 0.997±0.0053, Macro-F1 0.997±0.0054), indicating consistent and accurate classification across all classes. MobileNetV3 and Xception offer comparable results with low variability, suitable for lightweight or fast-inference applications. CNN+DenseNet169 and ConvNeXt-Tiny also performed well, slightly below DenseNet169. ResNet50 had the lowest accuracy and higher variability, suggesting a need for hyperparameter tuning or additional data augmentation. Overall, modern architectures provide superior accuracy and stability, with DenseNet169 representing the optimal choice for maximal classification performance.

Table 5. Model accuracy and Macro-F1 scores (95% CI).

Model	Accuracy (95% CI)	Macro-F1 (95% CI)
CNN + DenseNet169	0.992±0.0059	0.991±0.0055
Xception	0.994±0.0044	0.993±0.0051
DenseNet169	0.997±0.0053	0.997±0.0054
ConvNeXt-Tiny	0.992±0.0049	0.992±0.0051
MobileNetV3	0.995±0.0044	0.994±0.0045
ResNet50	0.968±0.0091	0.966±0.0090

Computational metrics

Table 6 summarizes the GPU memory usage and training time per epoch for all evaluated models. Among the models, MobileNetV3-Large required the least GPU memory (1,423.90 MB) and had the fastest training time per epoch (25.72 s), reflecting its lightweight architecture. ConvNeXt-Tiny, while achieving high overall performance, used moderate GPU memory (3,721.97 MB) and required 38.00 s per epoch for training. CNN+DenseNet169 consumed the most GPU memory (6,770.84 MB) with a training time of 40.10 s per epoch. Xception and DenseNet169 had moderate memory usage (4,419.10 MB and 5,037.70 MB, respectively) with training times of 30.19 s and 38.36 s per epoch. ResNet50 required 2,959.80 MB of GPU memory and 26.91 s per epoch. These results indicate that lightweight models like MobileNetV3-Large are highly efficient in terms of resource usage, while ConvNeXt-Tiny provides a good balance between classification performance and computational cost.

Table 6. Training time and memory usage of all models.

Model	GPU memory usage (MB)/epoch	Training time (s)/epoch
CNN + DenseNet169	6770.84	40.10
Xception	4419.10	30.19
DenseNet169	5037.70	38.36
ConvNeXt-Tiny	3721.97	38.00
MobileNetV3-Large	1423.90	25.72
ResNet50	2959.80	26.91

Subclass tumor

Table 7 and Table 8 demonstrate that ConvNeXt-Tiny achieved high and consistent metrics (accuracy 0.9923, precision 0.9940, recall 0.9923, F1-score 0.9931), indicating reliable identification of all subclasses. Specifically, the model correctly predicted glioma 298/300 (99.3%), meningioma 303/306 (99.0%), and pituitary 298/300 (99.3%), with only seven misclassifications among 906 test samples. The close alignment of precision, recall, and F1-score suggests balanced performance without bias toward any subclass, supporting its use as a baseline for further subclass classification studies.

Table 7. ConvNeXt-Tiny subclass prediction performance.

Model	Accuracy	Precision	Recall	F1-score
ConvNeXt-Tiny	0.9923	0.9940	0.9923	0.9931

Table 8. Confusion matrix (test set without no tumor).

	Predicted glioma	Predicted meningioma	Predicted pituitary
Actual glioma	298	2	0
Actual meningioma	2	303	1
Actual pituitary	0	2	298

No tumor vs tumor

Table 9 and Table 10 demonstrates that ConvNeXt-Tiny achieved very high and balanced performance (accuracy 0.9985, precision 0.9982, recall 0.9982, F1-score 0.9982) on the test set, which included 906 tumor samples and 405 no-tumor samples. The model correctly identified 905/906 tumor cases and 404/405 no-tumor cases, with minimal misclassifications. The close alignment of precision, recall, and F1-score indicates minimal bias toward either class, confirming its suitability as a reliable baseline model for binary classification with exceptional diagnostic performance.

Table 9. ConvNeXt-Tiny no tumor vs tumor prediction performance.

Model	Accuracy	Precision	Recall	F1-score
ConvNeXt-Tiny	0.9985	0.9982	0.9982	0.9982

Table 10. Confusion matrix (test set no tumor vs tumor).

	Predicted glioma	Predicted meningioma
Actual tumor	905	1
Actual no tumor	1	404

Grad-CAM

To further interpret model decisions, Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to the ConvNeXt-Tiny model.

The resulting heatmaps, illustrated in Figure 7, clearly highlight the tumor regions that contribute most strongly to the classification output. For *glioma* and *meningioma*, the model focuses accurately on irregular and dense tissue structures within the tumor boundaries, while for *pituitary tumors*, attention is centered around the sellar region. For *no-tumor* cases, activation is diffusely distributed, indicating the absence of focal lesions. These results verify that ConvNeXt-Tiny not only performs with high numerical accuracy but also exhibits interpretability consistent with clinical expectations. The localized activation patterns confirm that the model relies on medically relevant image regions rather than background artifacts, thereby enhancing its trustworthiness and potential for real-world diagnostic support.

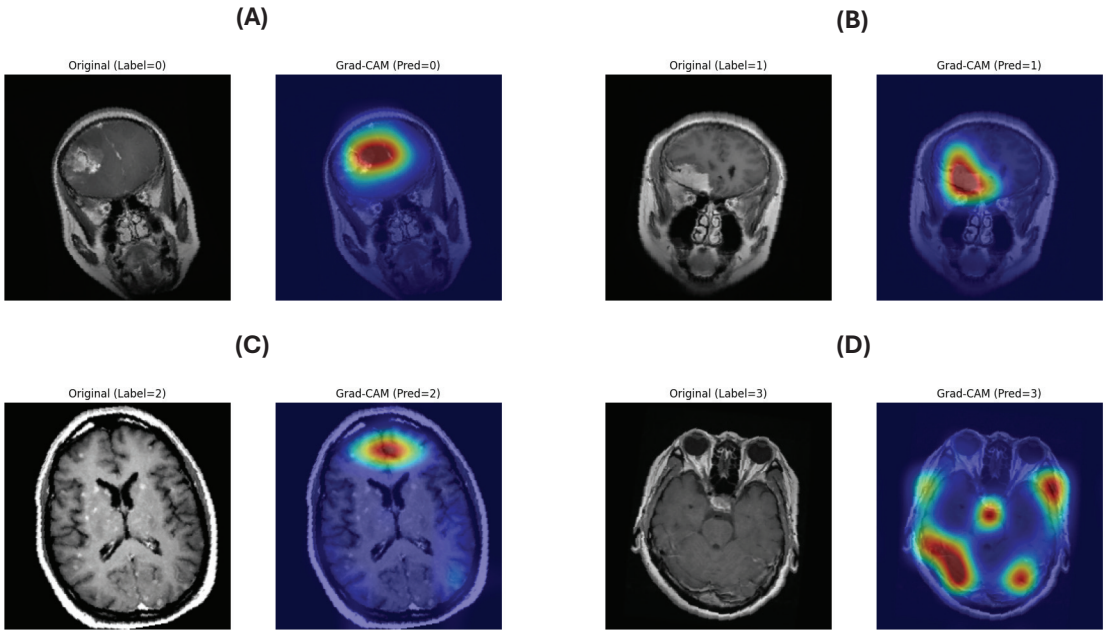


Figure 7. *Grad-CAM visualization of ConvNeXt-Tiny for each class. A: glioma, B: meningioma, C: no tumor, D: pituitary.*

Alternative dataset

To further evaluate the generalization ability of the baseline ConvNeXt-Tiny model, an alternative dataset was employed for testing. This dataset consists of images that were not included in the training process, providing an unbiased assessment of the model's performance on unseen data. The results summarized in Table 11.

Table 11. Performance of ConvNeXt-Tiny in alternative dataset.

Model	Accuracy	Precision	Recall	F1-score
ConvNeXt-Tiny	0.9984	0.9983	0.9982	0.9982

According to Table 11, the ConvNeXt-Tiny model demonstrates strong generalization capability when evaluated on the alternative dataset. The model achieves high performance across all evaluation metrics, with accuracy, precision, recall, and F1-score all exceeding 0.998. Although a slight reduction in accuracy is observed compared to the original training dataset, the overall performance remains consistently high, indicating the robustness and reliability of the proposed model under different data distributions.

Table 12 presents the confusion matrix of the ConvNeXt-Tiny model evaluated on the alternative brain tumor dataset. The results indicate that the model maintains strong generalization performance on unseen MRI scans, with most predictions concentrated along the main diagonal, reflecting a high level of classification accuracy across all four classes: glioma, meningioma, no tumor, and pituitary tumor. Only a small number of misclassifications are observed, mainly between glioma and meningioma, as well as between no-tumor and tumor classes in a few cases. These errors may be attributed to overlapping radiological features in certain MRI samples. Overall, the confusion matrix demonstrates the robustness of ConvNeXt-Tiny in accurately distinguishing no-tumor cases and reliably identifying pituitary tumors.

Table 12. Confusion matrix of the ConvNeXt-Tiny model.

	Predicted glioma	Predicted meningioma	Predicted notumor	Predicted pituitary
Actual glioma	804	1	0	0
Actual meningioma	0	544	0	1
Actual no tumor	0	1	481	1
Actual pituitary	0	0	0	610

Discussion

The comparative evaluation of six convolutional architectures demonstrates that ConvNeXt-Tiny provides a well-balanced trade-off among the evaluated architectures between classification accuracy, computational efficiency, and interpretability. This finding is consistent with prior studies employing ConvNeXt-based models in medical imaging. For example, recent work using pre-trained ConvNeXt on the BraTS 2019 dataset reported competitive or state-of-the-art performance when multi-sequence MRI inputs were available.[23] Although the present study focuses on single-sequence MRI images, ConvNeXt-Tiny achieves comparable performance, indicating its strong feature representation capability under constrained input settings.

While DenseNet169 achieved the highest raw accuracy, ConvNeXt-Tiny exhibited nearly equivalent performance with reduced computational cost and strong generalization on unseen data. These results align with previous MRI-based brain tumor classification studies using transfer learning, which typically report accuracies in the range of 96-98% depending on model complexity and dataset characteristics.[1,6,24] McNemar's test conducted on paired prediction outcomes indicated that performance differences between ConvNeXt-Tiny and other modern CNN architectures were not statistically significant ($p>0.05$), suggesting that ConvNeXt-Tiny should be regarded as competitive rather than strictly superior.

Grad-CAM visualizations showed that ConvNeXt-Tiny consistently focused on clinically relevant tumor regions, supporting findings from prior explainable AI studies in brain MRI analysis.[6] From a practical perspective, ConvNeXt-Tiny occupies a middle ground between heavy and lightweight models, highlighting its suitability for clinical scenarios where both performance and computational efficiency are critical.

Overall, the results position ConvNeXt-Tiny as an interpretable and generalizable baseline for multi-class brain tumor MRI classification. Consistent with prior literature, model selection should consider accuracy alongside efficiency and explainability. Future work should extend validation to multi-center clinical datasets and explore complementary interpretability techniques to further strengthen clinical reliability. To avoid potential metric inflation from the inclusion of the no-tumor class, subtype-specific performance was additionally evaluated, allowing a more realistic assessment of differential diagnosis performance.

Limitations

This study has several limitations. First, all experiments were conducted using publicly available MRI datasets, which may not fully reflect the variability of clinical imaging across different scanners or institutions. Second, the study focused solely on image-level classification without incorporating tumor segmentation, which limits direct clinical interpretability. Third, the interpretability analysis relied exclusively on Grad-CAM, providing only coarse localization rather than precise feature attribution. Finally, although the proposed model demonstrated promising accuracy, further investigation is needed to validate its feasibility for real-time deployment in practical clinical settings. Future work will aim to address these limitations.

Conclusion

This study evaluated the performance of multiple convolutional neural network architectures for multi-class brain tumor classification using MRI images. Among the evaluated models, ConvNeXt-Tiny was employed as the baseline model and demonstrated competitive performance, achieving high accuracy along with balanced precision, recall, and F1-scores, while maintaining a reasonable computational cost. In addition, the model provided interpretable Grad-CAM visualizations that corresponded well with relevant tumor regions, indicating its potential to support more transparent and explainable diagnostic modeling.

Although ConvNeXt-Tiny exhibited strong overall performance, the performance differences among modern architectures such as DenseNet169 and MobileNetV3-Large were relatively small. These observations were further supported by statistical analysis, which revealed no statistically significant differences among the models compared. This suggests that recent CNN-based architectures can achieve comparable classification performance when appropriately trained and optimized.

Overall, the findings highlight ConvNeXt-Tiny as a promising and efficient baseline for comparative research rather than a definitive solution for direct clinical deployment. Future work will focus on further validation in practical usage scenarios, incorporating segmentation-based analysis, and extending interpretability methods beyond Grad-CAM to enhance reliability and transparency in real-world medical imaging applications.

Ethical approval

Ethical approval was not required for this study because all data were obtained from publicly available.

Funding

There was no funding for this work.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Nitipon Pongphaw: conceptualization, software, validation, supervisor, writing: original draft, review and editing; **Prommin Buaphan**: data curation, data visualization, writing: review and editing.

Acknowledgements

The author would like to express sincere gratitude to Mr. Keerati Maneesai for valuable advice and assistance in preparing the manuscript.

References

[1] Deepa P, Narain P, Sreena V. Performance analysis of deep transfer learning approaches in detecting and classifying brain tumor from magnetic resonance images. Intelligent Data Analysis. 2023; 27(6):1759-80. doi:10.3233/ida-227321.

[2] Younis A, Li Q, Afzal Z, Adamu M, Kawuwa H, Hussain F, et al. Abnormal brain tumors classification using ResNet50 and its comprehensive evaluation. IEEE Access. 2024; 12: 78843-53. doi:10.1109/access.2024.3403902.

[3] Chen W, Sun P, Zhang Y. A brain tumor diagnosis approach based on deep separable convolutional networks and attention mechanisms. Applied and Computational Engineering. 2024; 67(1): 60-9. doi:10.54254/2755-2721/67/20240629.

[4] Oladimeji O, Ibitoye A. Brain tumor classification using ResNet50-convolutional block attention module. Applied Computing and Informatics. 2023. doi:10.1108/aci-09-2023-0022.

[5] Khan M, Khan A, Alhaisoni M, Alqahtani A, Alsubai S, Alharbi M, et al. Multimodal brain tumor detection and classification using deep saliency map and improved dragonfly optimization algorithm. Int J Imaging Syst Technol. 2023; 33(2): 572-87. doi:10.1002/ima.22831.

[6] Rastogi D, Johri P, Donelli M, Kumar L, Bindewari S, Raghav A, et al. Brain tumor detection and prediction in MRI images utilizing a fine-tuned transfer learning model integrated within deep learning frameworks. Life. 2025; 15(3): 327. doi:10.3390/life15030327.

[7] Li S. Classification of brain tumors based on magnetic resonance imaging using deep learning models. Applied and Computational Engineering. 2024; 40(1): 247-54. doi:10.54254/2755-2721/40/20230660.

[8] Türkoğlu M. Brain tumor detection using a combination of Bayesian optimization based SVM classifier and fine-tuned deep features. Eur J Sci Technol. 2021; 27: 251-258. doi:10.31590/ejosat.963609.

[9] Rehman A, Khan M, Saba T, Mehmood Z, Tariq U, Ayesha N. Microscopic brain tumor detection and classification using 3D CNN and feature selection

architecture. Microsc Res Tech. 2021; 84(1): 133-49. doi:10.1002/jemt.23597.

[10] Murugesan G, Nagendran P, Natarajan J. Advancing brain tumor diagnosis: deep Siamese convolutional neural network as a superior model for MRI classification. Brain-x. 2025; 3(2): e70028. doi:10.1002/brx2.70028.

[11] Kumar P, Bonthu K, Meghana B, Vani K, Chakrabarti P. Multi-class brain tumor classification and segmentation using hybrid deep learning network model. Scalable Comput Pract Exp. 2023; 24(1): 69-80. doi:10.12694/scpe.v24i1.2088.

[12] Ullah M, Khan M, Almujally N, Alhaisoni M, Akram T, Shabaz M. BrainNet: a fusion-assisted novel optimal framework of residual blocks and stacked autoencoders for multimodal brain tumor classification. Sci Rep. 2024; 14(1): 5895. doi:10.1038/s41598-024-56657-3.

[13] Joshi A, Aziz R. Deep learning approach for brain tumor classification using metaheuristic optimization with gene expression data. Int J Imaging Syst Technol. 2024; 34(2): e23007. doi:10.1002/ima.23007.

[14] Pei L, Vidyaratne L, Rahman M, Iftekharuddin K. Context-aware deep learning for brain tumor segmentation, subtype classification, and survival prediction using radiology images. Sci Rep. 2020;10(1): 19726. doi:10.1038/s41598-020-74419-9.

[15] Bacak A, Şenel M, Günay O. Convolutional neural network prediction on meningioma and glioma with TensorFlow. Int J Comput Exp Sci Eng. 2023; 9(2): 197-204. doi:10.22399/ijcesen.1306025.

[16] Ullah N, Hassan M, Khan J, Anwar M, Aurangzeb K. Enhancing explainability in brain tumor detection: a novel DeepEBTDNet model with LIME on MRI images. Int J Imaging Syst Technol. 2024; 34(1): e23012. doi:10.1002/ima.23012.

[17] Esmaeili M, Vettukattil R, Banitalebi H, Krogh N, Geitung J. Explainable artificial intelligence for human-machine interaction in brain tumor localization. J Pers Med. 2021; 11(11): 1213. doi:10.3390/jpm11111213.

[18] Goyal D, Sharma H. Brain tumor detection system using neural networks. Int J Commun Inf Technol. 2023; 4(1): 59-63. doi:10.33545/2707661x.2023.v4.i1a.61.

[19] Tonni S, Sheakh M, Tahosin M, Hasan M, Shuva T, Bhuiyan T, et al. A hybrid transfer learning framework for brain tumor diagnosis. Adv Intell Syst. 2025; 7(3): 2400495. doi:10.1002/aisy.202400495.

[20] Alanazi M, Ali M, Hussain S, Zafar A, Mohatram M, Irfan M, et al. Brain tumor/mass classification framework using MRI-based isolated and developed transfer deep-learning model. Sensors. 2022; 22(1): 372. doi:10.3390/s22010372.

[21] Nickparvar M. Brain tumor MRI dataset [Internet]. Kaggle; 2021 [cited 2025 Jan 3]. Available from: https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset

[22] Rostami A. Labeled MRI Brain Tumor Dataset Computer Vision Model [Internet]. Roboflow Universe; 2024 [cited 2025 Jan 3]. Available from: https://universe.roboflow.com/ali-rostami/labeled-mri-brain-tumor-dataset

[23] Mehmood Y, Bajwa UI. Brain tumor grade classification using the ConvNeXt architecture. Digit Health. 2024; 10: 20552076241284920. doi:10.1177/20552076241284920.

[24] Aamir M, Namoun A, Munir S, Aljohani N, Alanazi MH, Alsahafi Y, et al. Brain tumor detection and classification using an optimized convolutional neural network. Diagnostics. 2024; 14(16): 1714. doi:10.3390/diagnostics14161714.

Supplymentary

1. Dataset source

Dataset 1

The brain MRI dataset used in this study was compiled from multiple publicly available sources to form a four-class classification problem consisting of glioma, meningioma, pituitary tumor, and no-tumor categories. Images for the glioma, meningioma, and pituitary tumor classes were primarily obtained from the figshare dataset, which contains T1-weighted contrast-enhanced MRI scans acquired using clinical scanners from Nanfang Hospital (Guangzhou, China) and Tianjin Medical University General Hospital between 2005 and 2010. This dataset provides slice-level annotations for three tumor types and has been widely adopted in prior brain tumor classification studies. The no-tumor class was sourced from the Br35H dataset, which consists of brain MRI images from healthy subjects and is commonly used for binary and multi-class brain tumor classification tasks. Although the SARTAJ dataset was initially considered as an additional source for tumor images, the glioma subclass was excluded due to observed label inconsistencies, as indicated by prior studies and confirmed through our own experimental results. To ensure label reliability and reduce scanner-induced bias, glioma samples were therefore retained exclusively from the figshare dataset. The final curated dataset contains 7,023 MRI images with clearly defined class origins, enabling consistent multi-class training and evaluation.

Alternative Dataset

This project constructed a labeled brain MRI dataset for multi-class tumor analysis, covering four categories: pituitary tumor, meningioma tumor, glioma tumor, and no tumor. The dataset consists of a total of 2,443 MRI images, which were systematically divided into training, validation, and test subsets comprising 1,695, 502, and 246 images, respectively; however, in this study, all available images were additionally merged and used as an external test set to evaluate the generalization capability of the proposed models. All images are magnetic resonance imaging (MRI) scans, and each sample was annotated by medical experts following a standardized labeling protocol. The annotations include tumor presence and tumor type, with additional information on tumor location when applicable. This dataset is designed to support the development and evaluation of machine learning and deep learning models for automated brain tumor classification, with potential applications in assisting radiologists during clinical diagnosis and facilitating research toward improved diagnostic tools and treatment planning.

2. Loss curve

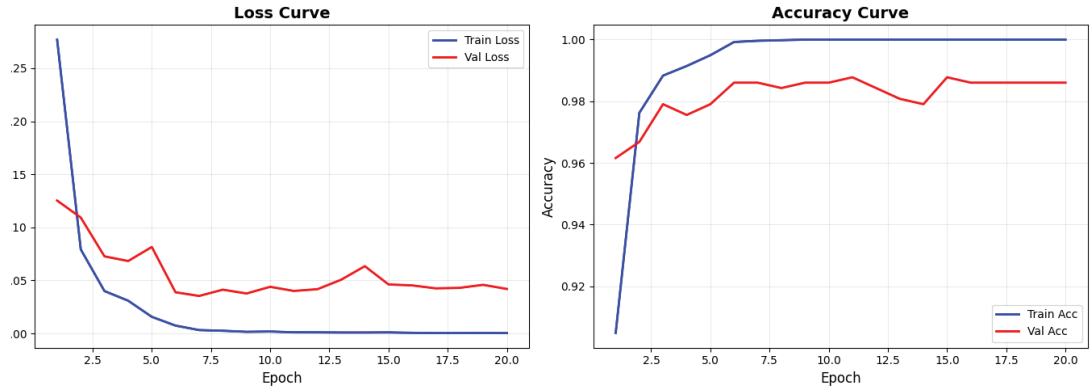


Figure S1. *Training and validation performance curves of the CNN+DenseNet169.*

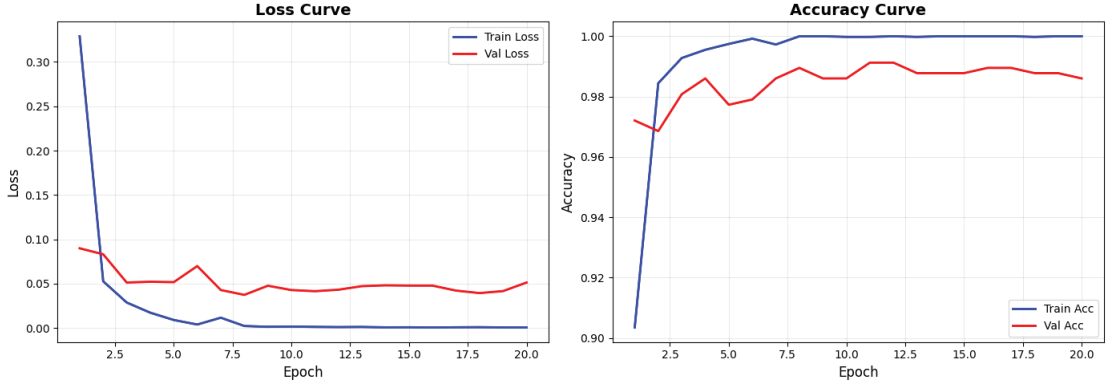


Figure S2. *Training and validation performance curves of the Xception.*

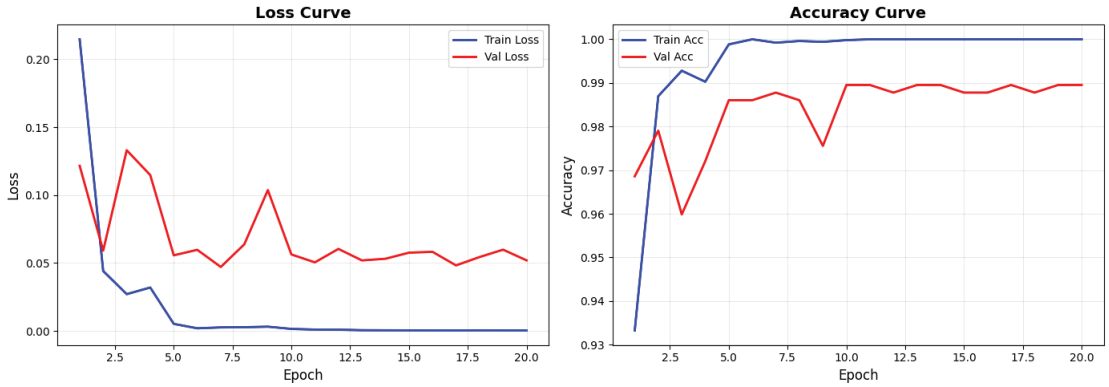


Figure S3. *Training and validation performance curves of the DenseNet169.*

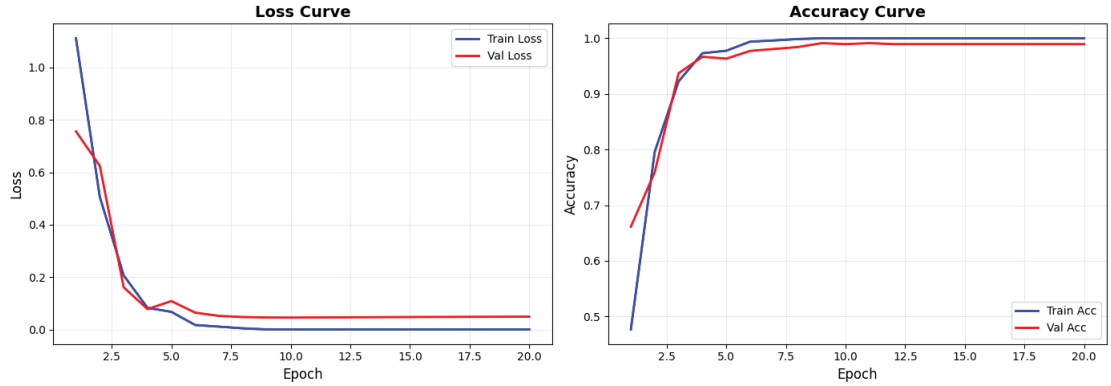


Figure S4. *Training and validation performance curves of the ConvNextTiny.*

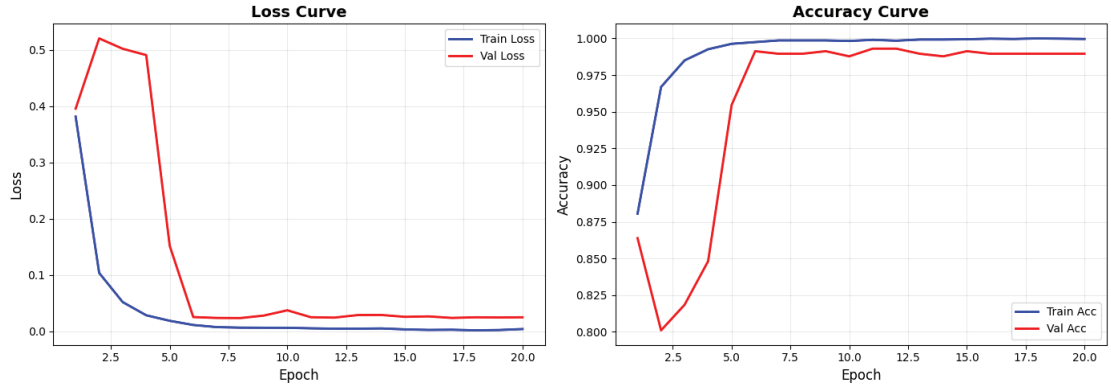


Figure S5. *Training and validation performance curves of the MobileNetV3.*

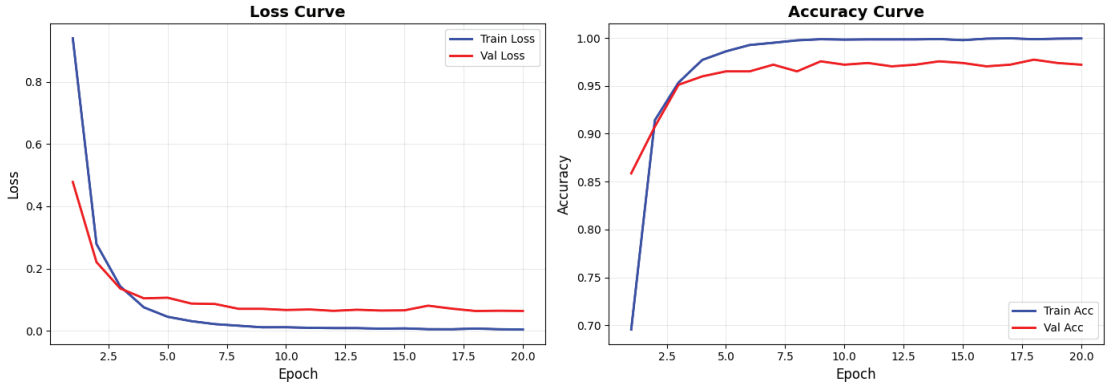


Figure S6. *Training and validation performance curves of the ResNet50.*

3. Confusion matrix

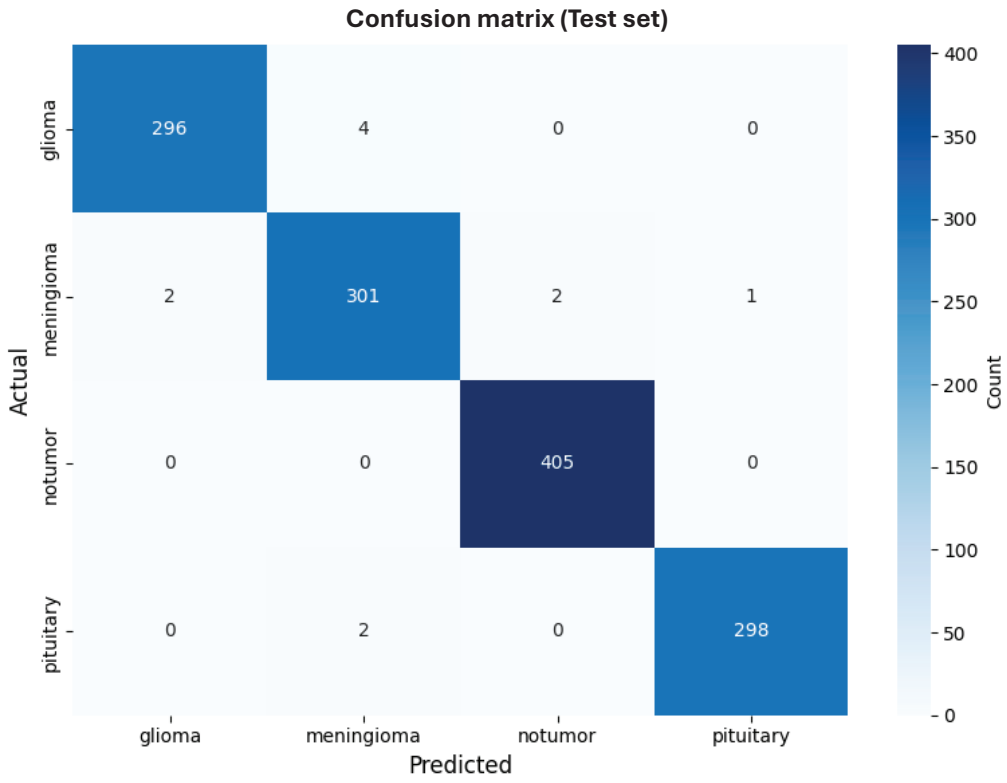


Figure S7. *Confusion matrix of the CNN+DenseNet169 model.*

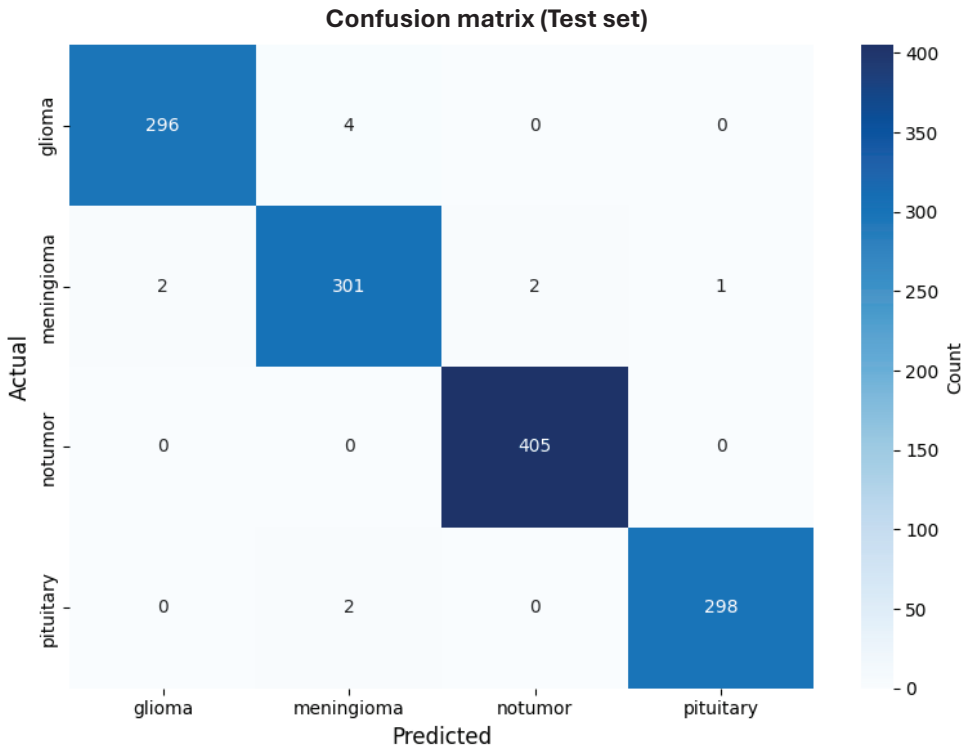


Figure S8. *Confusion matrix of the Xception model.*

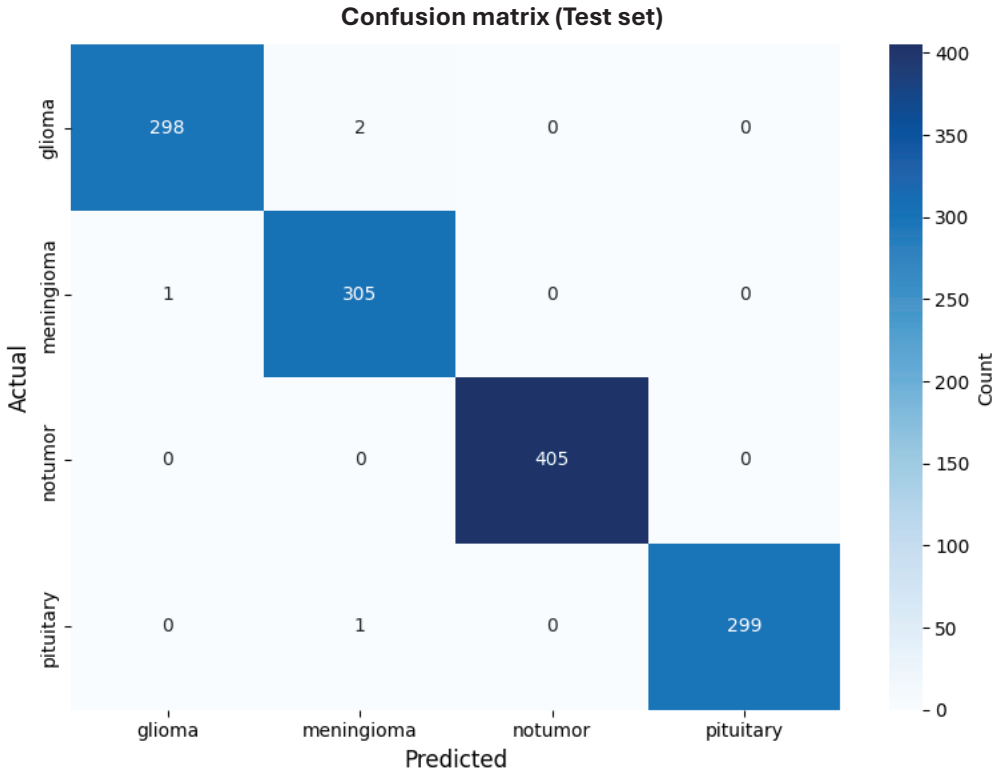


Figure S9. *Confusion matrix of the DenseNet169 model.*

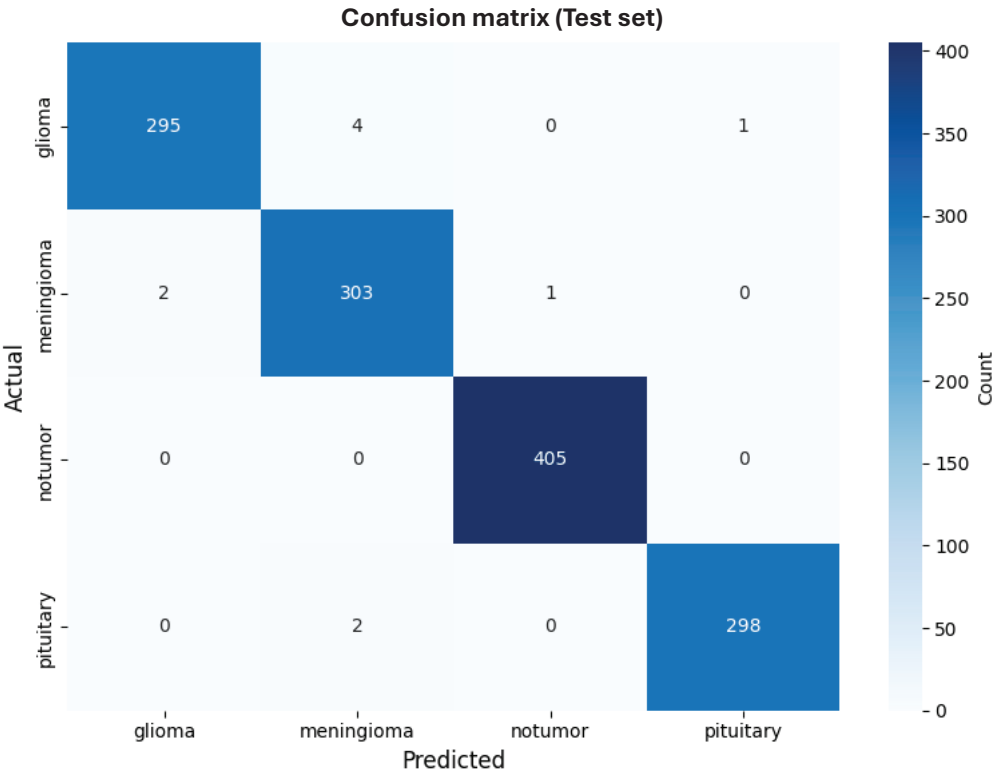


Figure S10. *Confusion matrix of the ConvNextTiny model.*

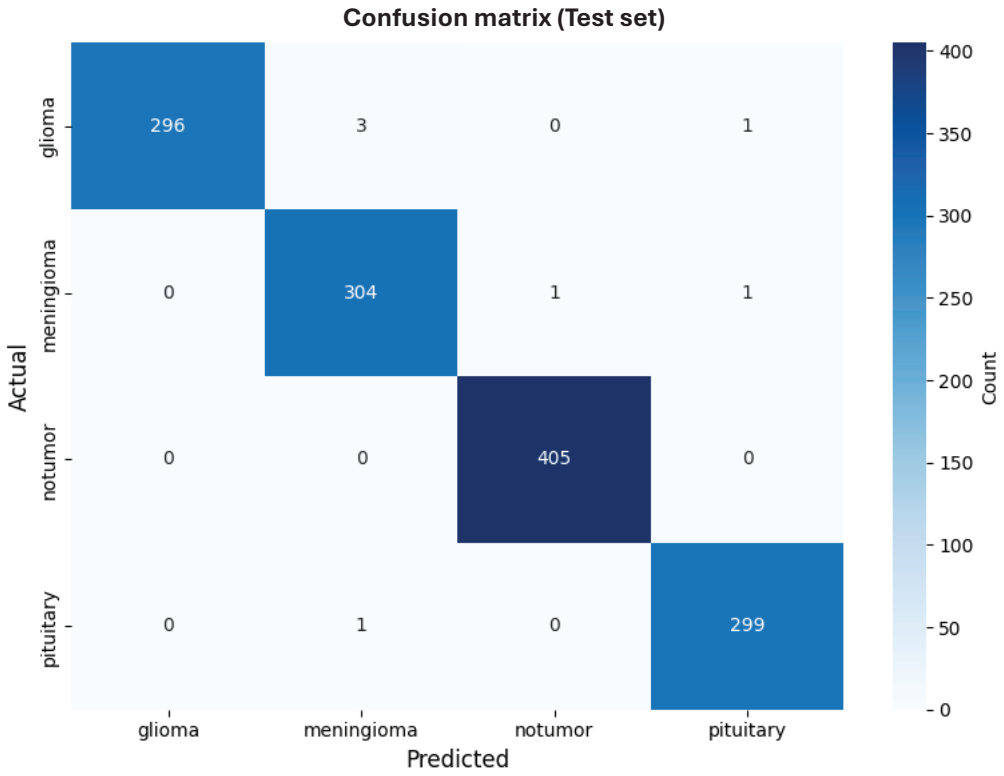


Figure S11. *Confusion matrix of the MobileNetV3 model.*

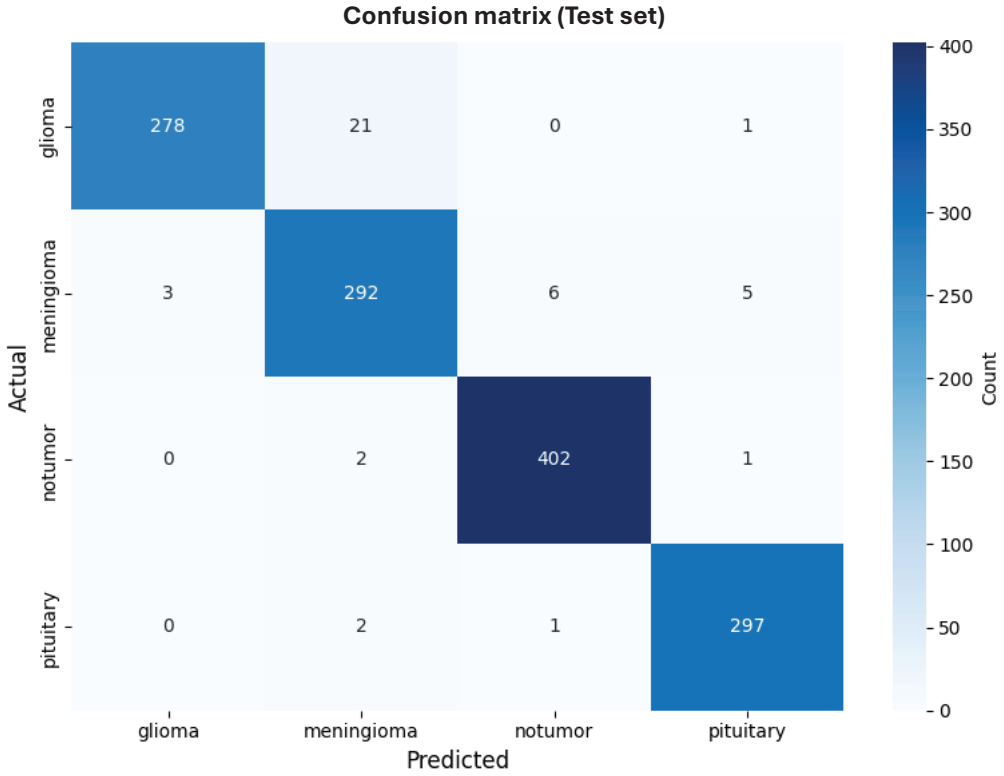


Figure S12. *Confusion matrix of the ResNet50 model.*

4. Subclass tumor

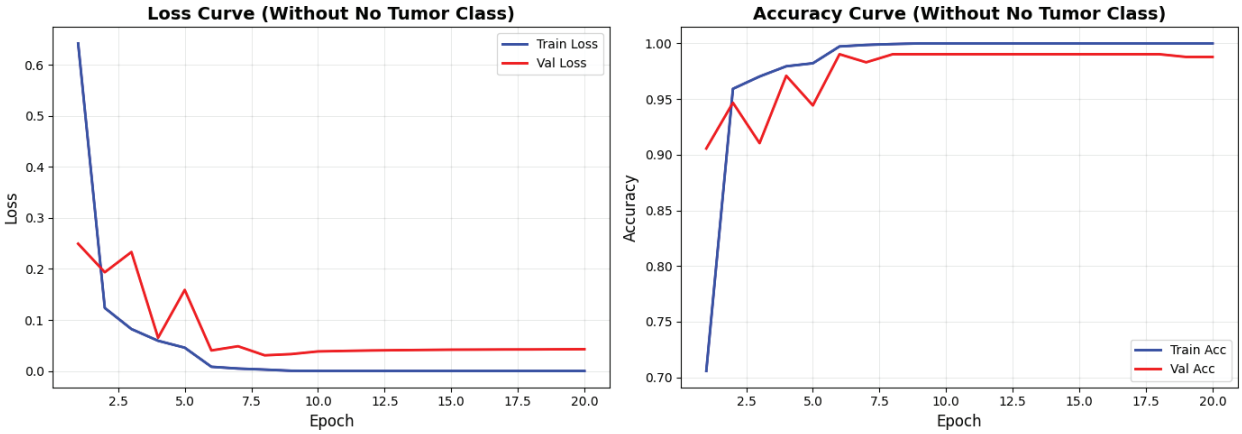


Figure S13. *Training and validation performance curves of the ConvNextTiny.*

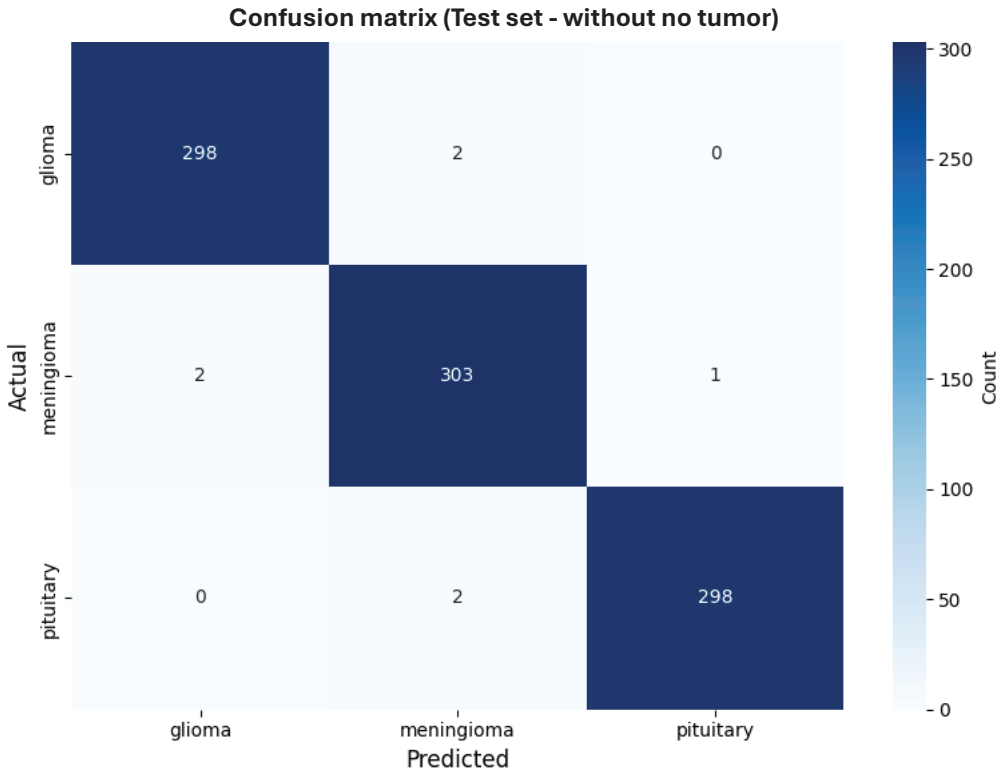


Figure S14. *Confusion matrix of the ConvNextTiny model.*

5. Tumor vs no tumor

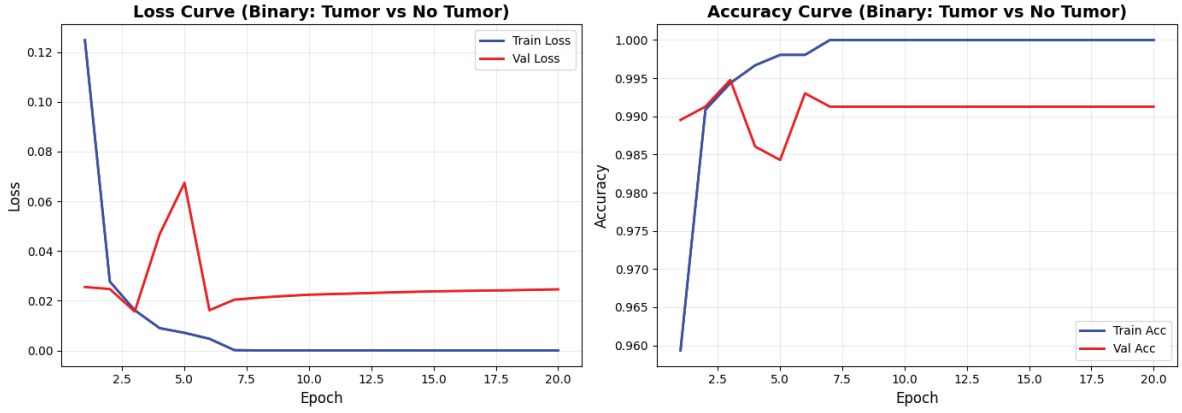


Figure S15. *Training and validation performance curves of the ConvNextTiny.*

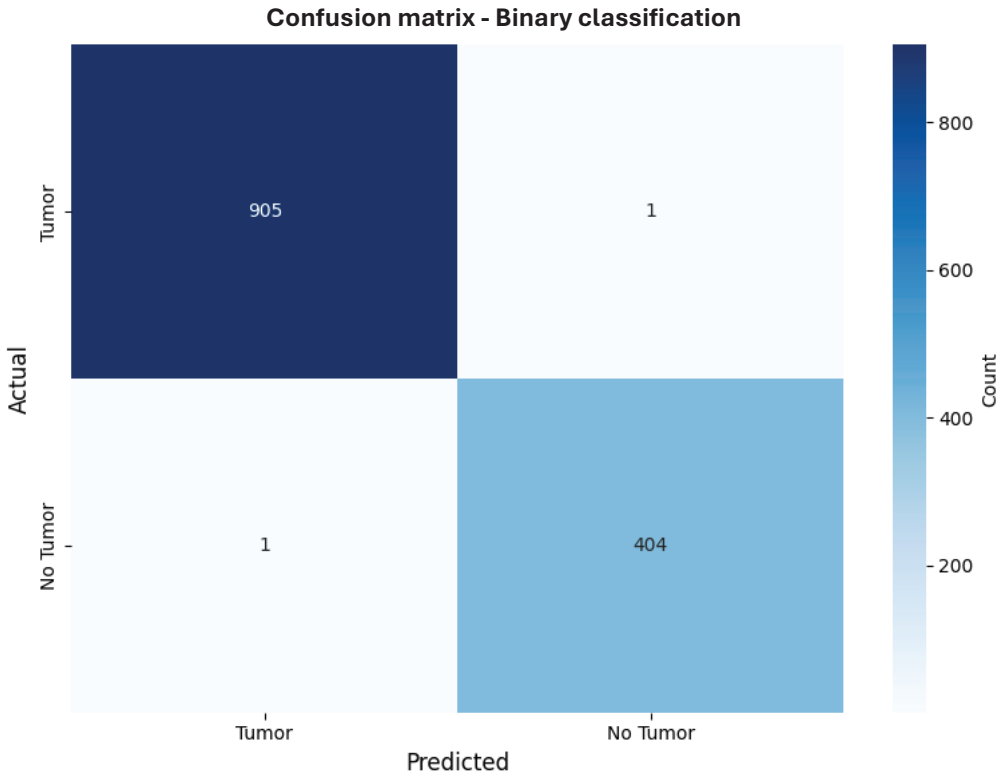


Figure S16. *Confusion matrix of the ConvNextTiny model.*

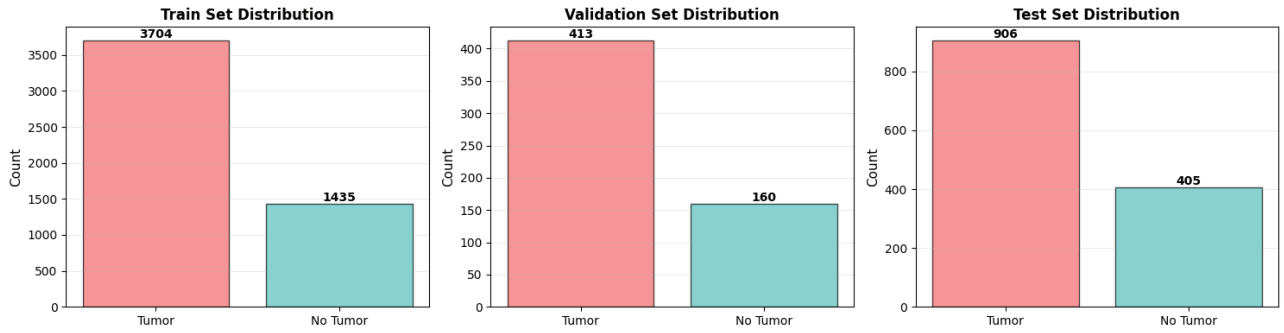


Figure S17. *Set distributions.*

Journal of Associated Medical Sciences 2026; 59(2): 63-71

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Effects of core strengthening exercises on inspiratory muscle strength and functional capacity among obese college students

Kaviya Muralidharan* and Santhosh Kumar Sankar

SRM College of Physiotherapy, Faculty of Medicine and Health Sciences (SRMIST), Kattankulathur, Tamil Nadu, India.

ARTICLE INFO

Article history:
Received 16 February 2025
Accepted as revised 14 January 2026
Available online 23 January 2026

Keywords:
Obesity, core strengthening exercise, inspiratory muscle training, maximal inspiratory pressure, functional capacity.

ABSTRACT

Background: Obesity weakens core muscles, impacting posture, balance, and respiratory function. Central obesity is particularly harmful due to its effect on visceral organs. Exercises targeting the core and respiratory muscles, such as inspiratory muscle training, are essential for managing obesity and enhancing functional capacity.

Objectives: To investigate the impact of core strengthening exercises on the strength of inspiratory muscles and overall functional capacity among obese college students.

Materials and methods: A 6-week experimental study was carried out involving 34 participants selected using convenience sampling methods. Individuals aged 18-25 years, regardless of gender, with a BMI exceeding 25 kg/m^2 were included in the study. Maximal inspiratory pressure was measured using a respiratory pressure meter, and the 6-minute walk test was used to measure functional capacity.

Results: The mean age of subjects was 20.85±1.63. The mean value of maximal inspiratory pressure (cmH_2O) pre-test was 95.09±9.58 and post-test was 95.85±9.39. The mean six-minute walk distance (metres) pre-test was 399.15±38.62 and post-test was 412.18±35.25. Both improvements were statistically significant (p=0.003 for MIP and p<0.001 for 6-MWTD)

Conclusion: A six-week core strengthening exercise programmme significantly increases the strength of inspiratory muscles and overall functional capacity among obese college students. These results support the clinical use of core exercises in obesity rehabilitation to enhance respiratory strength and physical function.

Introduction

Obesity is a critical global health issue characterized by the excessive accumulation of body fat in adipose tissue, leading to significant physiological and functional impairments. Risk factors for obesity include excessive caloric intake, alcohol consumption, physical inactivity, poor self-regulation, inadequate sleep, and chronic stress.[1] According to the World Health Organization (WHO), the prevalence of obesity has nearly tripled since 1975, affecting over 650 million adults globally. In India, 192 million people are currently affected. According to the WHO Asian classification, a BMI above 25 kg/m^2 qualifies as obese, with further subdivisions into Obese Class I (25-9.9) and Class II (>30).

Obesity is linked to a wide range of comorbidities, including hypertension, type 2 diabetes, cardiovascular diseases, obstructive sleep apnea, osteoarthritis,

* **Corresponding contributor.**
Author's Address: *SRM College of Physiotherapy, Faculty of Medicine and Health Sciences (SRMIST), Kattankulathur, Tamil Nadu, India.*
E-mail address: kaviyam3@srmist.edu.in
doi: 10.12982/JAMS.2026.037
E-ISSN: 2539-6056

anxiety, depression, and certain cancers. The pattern of fat distribution has a significant influence on disease risk.[2] Central or abdominal obesity, which involves the accumulation of visceral fat, poses a higher risk for metabolic and cardiovascular complications than peripheral obesity, which is more associated with subcutaneous fat. This makes central obesity particularly relevant when considering interventions that target functional and respiratory outcomes.[3]

Among the lesser-addressed consequences of obesity is its impact on the musculoskeletal and respiratory systems. The accumulation of abdominal fat can alter body mechanics and place excess strain on the core musculature. The core, comprising the abdominal, lumbar, pelvic, and deep spinal muscles, plays a pivotal role in maintaining posture, stability, and balance during movement. Weakness in these muscles may further compromise respiratory function, as they play a crucial role in supporting diaphragmatic movement and thoracic expansion. Strengthening the core is therefore not only essential for improving postural control but may also have implications for respiratory efficiency and endurance.[4]

Obesity also contributes to a restrictive pulmonary pattern, characterized by reduced lung volumes, limited diaphragmatic excursion, and increased airway resistance. These changes can impair inspiratory muscle strength (IMS) and reduce overall functional capacity. As excess abdominal mass elevates the diaphragm and reduces lung compliance, the respiratory muscles must work harder, often resulting in fatigue and decreased exercise tolerance. Maximal inspiratory pressure (MIP), an indicator of IMS, is commonly reduced in individuals with higher BMI, especially those classified as obese or morbidly obese.[5]

Recent evidence supports the use of respiratory muscle training (RMT) in improving breathing efficiency and exercise tolerance in individuals with various chronic conditions, including obesity. A systematic review in 2021 reported lower MIP and MEP values in obese populations, emphasizez the need for interventions targeting respiratory muscle performance.[6] Functional capacity is defined as the ability to perform daily activities and physical tasks and is often compromised in obese individuals due to limited mobility, joint stress, and reduced cardiovascular endurance. The 6-minute walk test (6MWT) and VO_2 max are commonly used to assess this capacity and are negatively influenced by excess weight and inactivity.[7]

While previous research has explored respiratory training independently, there is limited evidence on the combined effect of core strengthening exercises on inspiratory muscle strength and functional capacity in obese populations, particularly among young adults. The core musculature's anatomical and functional proximity to the diaphragm suggests a potential synergistic effect that has not been fully explored in current literature. Establishing this link is important, especially in populations such as college students, where sedentary behaviors and early-onset obesity are prevalent.

This study aims to investigate the effects of a structured core strengthening exercise programme on inspiratory muscle strength and functional capacity among obese college students. By addressing the gap, this preliminary study may inform rehabilitation and preventive strategies targeted at young adults. The findings may offer practical implications for designing exercise protocols that enhance respiratory performance and physical function in this demographic.

Materials and methods

This study was a quasi-experimental study conducted over six weeks among 34 participants selected through convenience sampling. The sample size was calculated using the formula $N=Z^2 \cdot p \cdot q/d^2$, where Z=1.96 (95% confidence level), p=0.5 (assumed prevalence), q = 1-p=0.5, and d=0.168 (margin of error). Substituting the values: $N=(1.96)^2 \times (0.5 \times 0.5)/(0.168)^2 = 3.8416 \times 0.25/0.0282 = 0.9604/0.0282 = 34.06$; hence, the sample size was 34 participants.

Ethical approval for this study was obtained from the Institutional Ethics Committee of SRM Medical College Hospital and Research Centre on 25/04/2024 (IEC Clearance No. SRMIEC-ST0224-1149). Before data collection, the trial was prospectively registered with the Clinical Trial Registry of India (CTRI) under the registration number CTRI/2024/07/070549. All procedures were conducted in accordance with the ethical principles outlined in the Declaration of Helsinki, and no changes were made to the study protocol after registration with CTRI.

The participant flow in this study was presented using the CONSORT 2010 Guidelines (Figure 1). Subjects aged 18-25 years, both male and female, with a BMI above 25 kg/m^2 based on the WHO Asian BMI classification[8] and an abdominal circumference greater than 102 cm for men and greater than 88 cm for women, were included in the study. Subjects with acute injuries (such as muscle strains, ligament sprains, or fractures), severe back pain, hypertension, diagnosed cardiovascular or pulmonary diseases, and recent abdominal or spinal surgeries were excluded from the study. Before the study, informed consent was obtained from the individuals after they were explained the procedure. Written informed consent was obtained from each participant following a comprehensive oral briefing about the study's aims, procedures, and ethical safeguards. The demographic data were obtained from the participants, including age, gender, height, weight, BMI, and waist circumference, which were documented prior to the commencement of the exercise protocol. The intervention sessions were supervised by trained physiotherapists. This was an open-label study with no blinding of participants, assessors, or intervention providers.

Prereading was taken by assessing the inspiratory muscle strength and functional capacity using a respiratory pressure meter and the 6MWT during the first week. Maximal inspiratory pressure was assessed with participants seated upright, feet flat on the floor, and wearing a nasal clip. Each participant performed three maximal inspiratory efforts starting from residual volume, with each effort lasting at least one second. A one-minute rest interval was provided between attempts, and the highest-pressure value was recorded.
Prediction equations for maximal respiratory pressure (PImax) [9,10]

Men: 142-(1.03×Age in years)

Women: -43+(0.71×Height in centimeter)

The six-minute walk test was carried out in a 30-metre hallway with cones marking each end, where participants were instructed to walk as far as possible within a six-minute period. To ensure consistency, standardized encouragement was provided, as it can influence performance, and the testing area was kept quiet with resuscitation equipment available for safety. At the end of the test, the total distance covered was recorded. Reference prediction equations proposed by Vaish *et al*. were used to interpret results, with the formula for females being 6MWD=856.55-16.08×age in years) and for males, 6MWD=681.97-(19.99×age in years)+(206×height in meters).[11] Following baseline testing, participants underwent a structured core strengthening program over six weeks, with three sessions per week, designed according to the FITT principle described by Kibler *et al*.[12] Each session lasted 15 to 30 minutes, performed at moderate intensity with progressive repetitions and sets. The exercise protocol focused on core strengthening movements, including knee planks and elbow planks (Figures 2. and 3.), pelvic bridges (Figure 4), crunches (Figure 5.), and reverse crunches (Figure 6). Before each session, a warm-up routine consisting of breathing exercises, spinal twists, cat-cow stretches, pelvic tilts, and arm circles was performed. The cool-down included cat-cow stretches, seated forward folds, and the child's pose to promote recovery. Details of the exercise protocol by week are presented in Table 1.

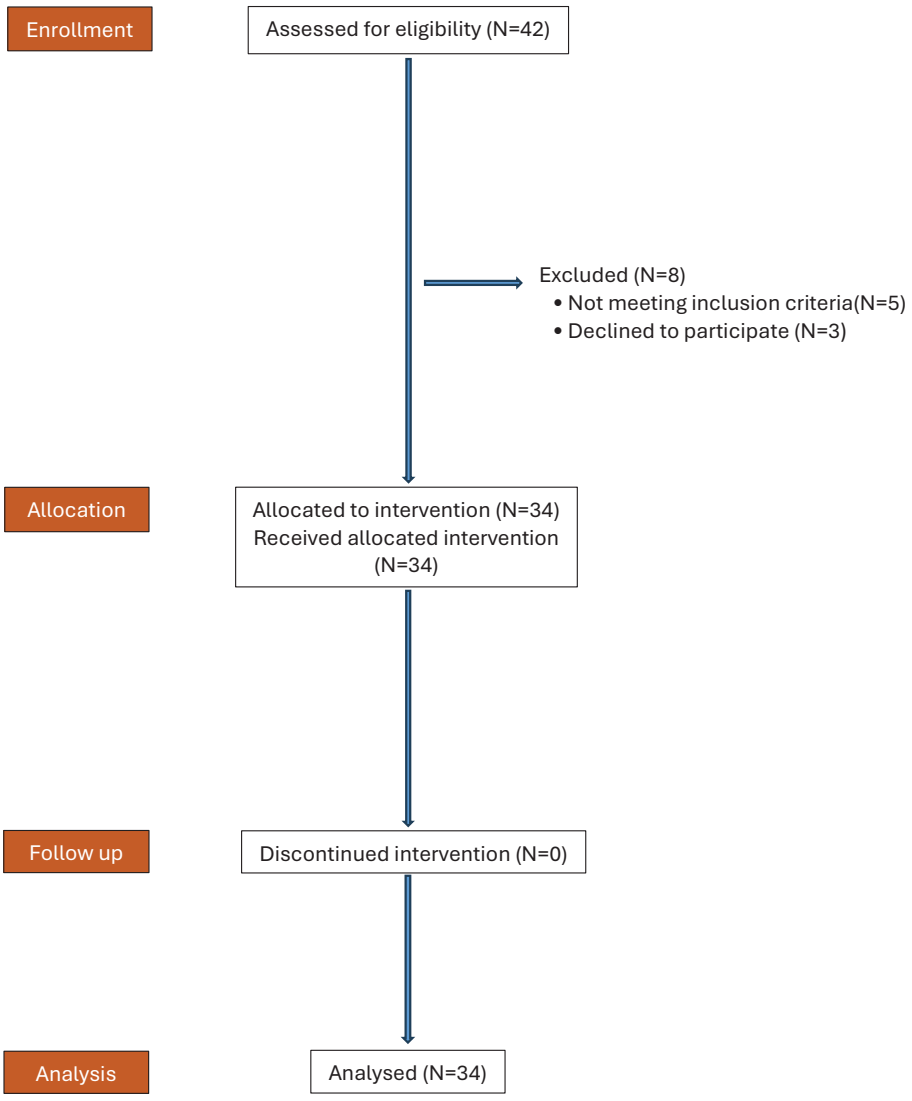


Figure 1. *The CONSORT diagram.*

Table 1. Core strengthening exercises protocol.

Week	Exercise	Rep	Sets	Rest time between sets (second)
1-2	Knee plank	10	2	60
	Bridges	12	2	60
	Crunches	15	2	60
	Reverse crunches	12	2	60
3-4	Elbow plank	12	3	60
	Bridges	15	3	60
	Crunches	20	3	60
	Reverse crunches	15	3	60
5-6	Elbow plank	15	3	60
	Pelvic bridges	18	3	60
	Crunches	20	3	60
	Reverse plank	12	3	60

Note: *Rep: repetitions, Sets: number of sets completed per exercise. All exercises were performed under physiotherapist supervision. Rest time between each set was standardized at 60 seconds. Exercise intensity and technique were monitored to ensure safety, and progression was introduced every two weeks as shown in the table.*

Statistical analysis

The collected data were analyzed using IBM SPSS v22.0 to assess all parameters, including descriptive statistics, mean, and standard deviation of maximal inspiratory pressure and the six-minute walk test. The paired t-test was used to compare pre- and post-intervention outcomes. The assumption for paired t-tests, including normality of distribution, were assessed using Shapiro Wilk tests and visual inspection of Q-Q plots. A $p<0.05$ was considered statistically significant. All relevant p-values are reported alongside outcome data in the results section.

Results

A total of 34 obese college students participated in the study, with a mean age of 20.85±1.63 years. The participants had a mean height of 173.50±5.34 cm, weight of 90.36±9.43 kg, and BMI of 30.03±2.94 kg/m^2. Most participants were classified as Class I obesity (61.76%), while 38.24% were classified as Class II obesity. Following the 6-week core strengthening program, a statistically significant improvement in maximal inspiratory pressure (MIP) was observed, increasing from 95.09±9.58 cmH_2O at baseline to 95.85±9.39 cmH_2O post-intervention (mean difference = -1.24; 95% CI: -3.25 to -0.28; p=0.003). Functional capacity also improved significantly, with the 6-minute walk distance increasing from 399.15±38.62 m to 412.18±35.25 m (mean difference=-16.03; 95% CI: -10.02 to -8.82; $p<0.001$) as illustrated in Table 3. Predicted values compared with actual pre- and post-test values are presented in Table 4 and demonstrate changes consistently with the effects of the intervention.



Figure 2. *Knee plank-starting position, engaging abdominal and core muscles. In a kneeling stance with hands shoulder width apart, the body is kept straight from knees to head while maintaining an isometric core contraction. This exercise primarily targets the rectus abdominis and transverse abdominis, and the hold is maintained for 20-60 seconds with gradual progression.*



Figure 3. *Elbow plank-advanced plank targeting deep core stabilizers. With forearms under the shoulders and legs extended, the body is held in a straight line while contracting the core, glutes, and shoulder stabilizers. This advanced plank targets deep core muscles and deltoids, ensuring hips do not pike or sag.*

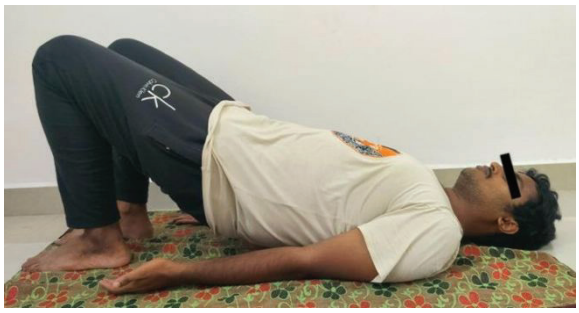


Figure 4. *Pelvic bridges-exercise strengthening gluteal and lower back muscles. Supine with knees bent and feet hip-width apart, lift hips to form a straight line from knees to shoulders, keeping the pelvis stable. Targets glutes, erector spinae, and hamstrings; exhale up, inhale down, 12-15 reps, 2-3 sets.*



Figure 5. *Crunches-Abdominal exercise to build core endurance. Supine position with knees bent and hands behind the head or on the chest, lift the shoulders by contracting the abdominals and return with control. This exercise targets the rectus abdominis and is performed for 15-20 repetitions.*



Figure 6. *Reverse crunches-movement focusing on lower abdominal activation. Lying supine with hips and knees at 90°, lift the pelvis and draw the knees toward the chest, then lower with control. This exercise targets the lower rectus abdominis and transverse abdominis, performed for 12-15 repetitions.*

Table 2. Demographic characteristics.

Variable	Value
Age (years)	20.85±1.63
Height (cm)	173.50±5.34
Weight (kg)	90.36±9.43
BMI (kg/m^2)	30.03±2.94
Waist circumference (cm)	106.56±5.14
Gender, N (%)	Male: 29 (85.3%), Female: 5 (14.7%)
Obesity Class, N (%)	Class I: 21 (61.76%), Class II: 13 (38.24%)

Note: *BMI: body mass index, mean±SD for continuous variables and frequency (%) for categorical variables.*

Table 3. Pre- and post- intervention comparison of the inspiratory muscle strength and functional capacity.

Outcome measure	Pre-test (Mean±SD)	Post-test (Mean±SD)	Mean difference	95% CI	t (df)	*p* value
MIP (cmH_2O)	95.09±9.58	95.85±9.39	-1.24	-3.25 to -0.28	33	0.003
6MWD (m)	399.15±38.62	412.18±35.25	-16.03	-10.02 to -8.82	33	<0.001

Note: *MIP: maximum inspiratory pressure, 6MWD: 6-minute walk test distance (meters), m: meter. Paired t-test, p<0.05 was considered statistically significant.*

Table 4. Predicted and actual values of MIP and functional capacity (6MWD).

Outcome measure	Predicted value	Actual pre-test	Actual post-test
MIP (cmH_2O)	114.16	95.09±9.58	95.85±9.39
6MWD (m)	611.95	399.15±38.62	412.18±35.25

Note: *MIP: maximum inspiratory pressure,6MWD: 6-minute walk distance, m: meter. Mean±SD.*

Discussion

The core strengthening exercises lead to improvement in the outcomes of Maximum inspiratory pressure and Functional capacity among obese individuals over 6 weeks**.** The observed improvements in MIP and functional performance may be attributed to the activation and enhanced coordination of the deep trunk muscles, facilitated by targeted core strengthening exercises, especially those muscles essential for supporting diaphragmatic movement and ensuring postural stability. With enhanced neuromuscular control of the core, participants presumably attained greater trunk stabilization efficiency, subsequently promoting improved breathing mechanics. This enhanced stability may have diminished dependence on compensatory and ineffective respiratory patterns frequently observed in obesity, facilitating a greater role for diaphragmatic breathing. As a result, the increased efficiency of inspiratory muscle activity most likely contributed to observable improvements in Inspiratory muscle strength and overall functional capacity, indicating greater integration of respiratory and postural functions during physical activity.

The overall respiratory compliance is diminished by as much as two-thirds of the standard value in individuals with obesity. This is partly caused by a reduction in lung compliance, which may be connected to the higher pulmonary blood volume observed in obese people. However, the primary cause is a decrease in chest wall compliance due to fat accumulation in and around the ribs, diaphragm, and abdomen. Compared to nonobese people, recumbency significantly lowers total respiratory compliance in obese people. This decrease is primarily attributable to reduced chest wall compliance, although an increase in respiratory resistance may also be a contributing factor.[13] Given that reduced chest wall compliance and increased respiratory resistance place a greater mechanical load on breathing, obese individuals often experience earlier fatigue and reduced movement efficiency, which directly influences their overall functional capacity.

Research has shown that obesity is associated with notable alterations in gait mechanics, including reduced walking speed, shortened stride length, and increased step breadth, which collectively reflect adaptations to the excess body mass and altered body biomechanics seen in this population. In addition to these changes, obese individuals often demonstrate impairments in postural stability, which can compromise balance control and contribute to a heightened risk of falls during both static and dynamic activities. These movement limitations are further compounded by the respiratory mechanical restrictions commonly observed in obesity, such as reduced chest wall

compliance and increased work of breathing. Together, these musculoskeletal and respiratory constraints may result in a decrease in overall functional performance and an increase in energy expenditure during daily activities[14], which can accelerate the onset of fatigue, limit participation in physical tasks, and negatively impact overall physical independence and quality of life.

The main findings of this study are that core strengthening exercises resulted in a statistically significant increase in MIP from 95.09±9.58 cmH_2O to 95.85±9.39 cmH_2O (p=0.003). Although the absolute change of 0.76 cmH_2O may appear modest, it reflects improved diaphragmatic activation and endurance, which are clinically relevant in populations at risk for respiratory complications. Even small increases in MIP can be meaningful in preventing progression to respiratory muscle fatigue, especially in obese individuals. The improvement in MIP aligns with findings reported in the literature, which have enhanced pulmonary capacity and respiratory muscle power following core strengthening exercises in adults with substance use disorder[15] and increased peak inspiratory and expiratory pressures in children with bronchiectasis following core stability exercises. These results support the physiological rationale that core musculature provides structural support to respiratory function by stabilizing the trunk and facilitating diaphragmatic movement.[16]

Beyond the respiratory mechanics, these physiological changes have significant effects on functional ability. Enhanced trunk stability and more effective diaphragmatic movement can decrease the perceived effort of breathing during exertion, enabling individuals to maintain activity for extended periods. As a result, these modifications may manifest as measurable enhancements in submaximal exercise performance, as evidenced by advancements in the 6-Minute Walk Test (6MWT), a broadly recognised indicator of functional capacity. Functional capacity, as assessed by the 6MWT, also showed significant improvement in this study. The mean distance increased from 399.15±38.62 meters to 35.25±412.18 meters (p<0.001), demonstrating enhanced endurance and physical performance. Obesity often limits mobility and reduces cardiopulmonary efficiency due to excess body weight and poor muscle function. Research findings reported that obese individuals, including children, exhibited reduced functional capacity compared to their non-obese peers, consistent with the baseline values seen in our participants.[17,18]

Our findings are supported by the literature, which demonstrated that a six-week core strengthening program markedly enhanced both inspiratory strength and functional capacity in patients with chronic kidney disease undergoing hemodialysis and individuals with asthma.[19,20] These studies emphasize the broader applicability of core-focused interventions in improving respiratory and functional outcomes across various clinical populations. Physiologically, core strengthening exercises enhance the coordination and strength of the diaphragm and intercostal muscles, reduce chest wall resistance, and improve postural control, all of which contribute to better respiratory function. Obesity imposes a restrictive load on the respiratory system, and improvements in MIP reflect a partial reversal of this dysfunction. Functional capacity also benefits from improved balance, mobility, and muscular endurance, which are outcomes associated with enhanced core stability.

Despite the positive findings, this study has several limitations. The sample size was relatively small and limited to a single age group (18-25 years) and educational setting, which reduced the generalizability of the results. The short intervention period (six weeks) may not capture long-term benefits or the sustainability of the improvements. The absence of a control group limits the ability to attribute all changes solely to the intervention. Future studies should consider larger, randomized controlled designs with longer follow-up periods, and include objective assessments of physical activity and adherence to exercise protocols.

In conclusion, the significant improvements in inspiratory muscle strength and functional capacity observed in this study highlight the value of core strengthening exercises as a non-pharmacological strategy in the management of obesity. These findings support the inclusion of core-focused physical activity in university wellness programmes, rehabilitation protocols, and public health initiatives aimed at reducing obesity-related complications and promoting functional independence in young adults.

Conclusion

The results of this study demonstrate that enhancements in core stability are associated with improved respiratory function and functional capacity in obese individuals. Specifically, improvements in maximal inspiratory pressure (MIP) and enhanced walking capacity on the 6MWT indicate significant physiological and functional advancements. According to these findings, core-focused therapies enhance respiratory strength and general functional performance, which have clinically significant advantages.

Ethical approval

Ethical approval was obtained from the Institutional Ethics Committee of SRM Medical College Hospital and Research Centre on 25.04.2024. The ethical clearance number is SRMIEC-ST0224-1149.

Clinical implication

The findings suggest that core strengthening exercises may serve as a practical and cost-effective intervention to improve inspiratory muscle performance and functional capacity in obese individuals. Incorporating such exercises into routine rehabilitation programs may help address common functional

limitations in this population and enhance overall movement efficiency and respiratory function.

Conflict of interest

The authors declare no conflict of interest with respect to the research, authorship and /or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

CRediT authorship contribution statement

Kaviya Muralidharan: conceptualisation, methodology, investigation, data curation, formal analysis, writing: original draft, review and edit, project administration, supervision; **Santhosh Kumar**: investigation, data curation, formal analysis, validation, writing: review and edit.

Acknowledgements

The authors would like to express their sincere gratitude to SRM College of Physiotherapy for providing the necessary facilities, institutional support, and a conducive academic environment to carry out this research. We also extend our heartfelt thanks to the students who willingly participated in this study for their time, cooperation, and valuable contribution, without which this research would not have been possible.

References

[1] World Health Organization. Obesity and overweight [Internet]. WHO; 2021. Available from: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight.

[2] GBD 2019 risk factors collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. Lancet. 2020; 396(10258): 1223-49. doi:10.1016/S0140-6736(20)30752-2.

[3] Després JP. Body fat distribution and risk of cardiovascular disease: an update. Circulation. 2012;126(10):1301-13. doi:10.1161/CIRCULATIONAHA.111.067264.

[4] McGill SM. Core training: evidence translating to better performance and injury prevention. Strength Cond J. 2010; 32(3): 33-46. doi:10.1519/SSC.0b013e3181df4521.

[5] Caicedo-Trujillo S, Torres-Castro R, Vasconcello-Castillo L, Solis-Navarro L, Sanchez-Ramirez D, Núñez-Cortés R, et al. Inspiratory muscle training in patients with obesity: a systematic review and meta-analysis. Front Med (Lausanne). 2023; 10: 1284689.doi:10.3389/fmed.2023.1284689.

[6] Torun E, Cakir E, Ozgüç F, Ozgen IT. The effect of obesity degree on childhood pulmonary function tests. Balkan Med J. 2014; 31(3): 235-8. doi: 10.5152/balkanmedj.2014.13101.

[7] Vincent HK, Vincent KR, Lamb KM. Obesity and mobility disability in the older adult. Obes Rev. 2010; 11(8): 568-79. doi: 10.1111/j.1467-789X.2009.00703.x.

[8] WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet. 2004; 363(9403): 157-63. doi:10.1016/S0140-6736(03)15268-3.

[9] Black LF, Hyatt RE. Maximal respiratory pressures: normal values and relationship to age and sex. Am Rev Respir Dis. 1969; 99(5): 696-702. doi:10.1164/arrd.1969.99.5.696.

[10] Evans JA, Whitelaw WA. The assessment of maximal respiratory mouth pressures in adults. Respir Care. 2009; 54(10): 1348-59.

[11] Vaish H, Gupta S, Sharma S. Six minute walk distance and six minute walk work in young adults aged 18-25 years. Int J Pharm Med Res. 2017;5(3):464-8. https://www.researchgate.net/publication/328274875.

[12] Kibler WB, Press J, Sciascia A. The role of core stability in athletic function. Sports Med. 2006; 36(3): 189-98. doi:10.2165/00007256-200636030-00001.

[13] Parameswaran K, Todd DC, Soth M. Altered respiratory physiology in obesity. Can Respir J. 2006; 13(4): 203-10. doi: 10.1155/2006/834786.

[14] Lai PP, Leung AK, Li AN, Zhang M. Three-dimensional gait analysis of obese adults. Clin Biomech (Bristol). 2008; 23(Suppl 1): S2-6. doi: 10.1016/j.clinbiomech.2008.02.004.

[15] Mustafaoglu R, Demir R, Demirci AC, Yigit Z. Effects of core stabilization exercises on pulmonary function, respiratory muscle strength, and functional capacity in adolescents with substance use disorder: randomized controlled trial. Pediatr Pulmonol. 2019; 54(7): 100211. doi:10.1002/ppul.24330.

[16] Vardar-Yagli N, Saglam M, Caik-Kutukcu E, Inal-Ince D, Arikan H, Kiper N. Effects of core stabilization training on balance, functional capacity, and respiratory muscle strength in children with bronchiectasis. Eur Respir J. 2014; 44(Suppl 58): P4279. doi:10.1183/13993003/erj.44.Suppl_58.P4279.

[17] Özgen İT, Çakır E, Torun E, Güleş A, Hepokur MN, Cesur Y. Relationship Between Functional Exercise Capacity and Lung Functions in Obese Chidren. J Clin Res Pediatr Endocrinol. 2015;7(3):217–21. doi:10.4274/jcrpe.1990.

[18] O'Donnell DE, Laveneziana P. The clinical importance of dynamic lung hyperinflation in COPD. COPD. 2006;3(4):219-32. doi:10.1080/15412550600977478.

[19] Waghmare PD, Das AK, Diwate AD. Effect of six-week core strengthening exercises on inspiratory

muscle strength and functional capacity in patients with chronic kidney disease undergoing haemodialysis. Int J Community Med Public Health. 2025; 11(3): 1311-3. doi.org/10.18203/2394-6040.ijcmph20240637.

[20] Develi E, Subasi F, Aslan GK, Bingol Z. The effects of core stabilization training on dynamic balance and pulmonary parameters in patients with asthma. J Back Musculoskelet Rehabil. 2021; 34(4): 639-48. doi:10.3233/BMR-191803.

Journal of Associated Medical Sciences 2026; 59(2): 72-83

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Comprehensive benchmarking of machine learning models for blood glucose classification and prediction: new approach for improved hyperglycemia and hypoglycemia detection and prediction

Houda El Bouhissi[1*], Tatiana Ermakova[2], Rabie A. Ramadan[3], Djamila Ouaret[1]

[1]*University of Bejaia, Faculty of Exact Sciences, LIMED Laboratory, Bejaia, Algeria.*
[2]*School of Computing, Communication and Business, Hochschule für Technik und Wirtschaft, University of Applied Sciences for Engineering and Economics, Germany.*
[3]*Department of Information Systems, College of Economics, Management, and Information Systems, Nizwa, Sultanate of Oman.*

ARTICLE INFO

Article history:
Received 22 August 2025
Accepted as revised 9 January 2026
Available online 27 January 2026

Keywords:
Diabetes, CRNN, CNN, LSTM, bioinspired algorithm.

ABSTRACT

Background: Diabetes mellitus affects 463 million people worldwide and necessitates continuous blood glucose monitoring. Current glucose prediction systems often lack efficiency, and real-time prediction is essential for timely clinical intervention.

Objectives: This study aims to develop and validate a novel Convolutional Recurrent Neural Network (CRNN) enhanced with bio-inspired algorithms to improve blood glucose prediction and enable real-time detection of hypoglycemia and hyperglycemia.

Materials and methods: The proposed framework employs a CRNN architecture that combines Convolutional Neural Networks (CNNs) for feature extraction with Long Short-Term Memory (LSTM) layers for temporal sequence learning. The model was trained and evaluated using the HUPA-UCM diabetes dataset. Additionally, the study benchmarks the proposed model against 19 traditional Machine Learning (ML) algorithms and compares it with state-of-the-art methods from the literature.

Results: The proposed approach demonstrates superior predictive capability, consistently delivering promising results across multiple evaluation frameworks. The model achieves clinically acceptable prediction intervals, confirming its effectiveness in enhancing the accuracy and reliability of blood glucose prediction for diabetes management.

Conclusion: The findings demonstrate that the proposed CRNN model, enhanced with bio-inspired algorithms, provides an effective and reliable solution for real-time blood glucose prediction. By outperforming conventional ML methods and achieving clinically acceptable accuracy levels, the model shows strong potential for integration into intelligent diabetes management systems to support timely clinical decisions and improve patient outcomes.

* **Corresponding contributor.**
Author's Address: *University of Bejaia, Faculty of Exact Sciences, LIMED Laboratory, Bejaia, Algeria.*
E-mail address: houda.elbouhissi@gmail.com
doi: 10.12982/JAMS.2026.038
E-ISSN: 2539-6056

Introduction

Medical background

Diabetes is a complex metabolic disorder that has become a major global health challenge. In 2014, it affected over 8.5% of adults aged 18 and older.[1] The disease is characterized by impaired blood glucose regulation and is associated with a high risk of complications, including cardiovascular disease, nephropathy, retinopathy, and neuropathy.[2,3] In 2019,

diabetes accounted for 1.5 million deaths worldwide, nearly half of which occurred before the age of 70; an additional 460,000 kidney disease-related deaths and approximately 20% of cardiovascular deaths were attributable to hyperglycemia.[4] Diabetes develops due to insufficient insulin secretion or reduced tissue sensitivity to insulin. Insulin is produced by pancreatic β-cells and facilitates glucose transport into cells, where it is used for energy or stored as glycogen.[5] Insulin deficiency or resistance leads to hyperglycemia, whereas excessive secretion can cause hypoglycemia.[6] In non-diabetic individuals, fasting glucose is <1.26 gm/L (7 mmol/L); values of 1.26-2 gm/L indicate prediabetes, and 2 gm/L confirm diabetes.[7-9] An HbA1c level <7% is widely recommended to minimize the risk of long-term complications.[10,11] Two main types of diabetes are recognized: type 1 (T1D) and type 2 (T2D). T1D results from autoimmune destruction of β-cells, leading to absolute insulin deficiency; T2D is characterized by insulin resistance and reduced secretion. Less common forms include gestational diabetes and other specific types.[12,13] This study focuses on T1D, the most severe form, often developing in childhood or adolescence and requiring lifelong insulin therapy.[14]

Motivation and research objectives

Continuous blood glucose monitoring is required for effective management.[15] Previous studies have combined CNNs and LSTM networks, for example, in CNN-LSTM hybrid architectures and two-headed LSTM models.[16] Dilated RNNs (DRNNs) have also been explored, but there has been no explicit application of Convolutional Recurrent Neural Networks (CRNNs) targeting both hypo- and hyperglycemia in real time.[17] Indeed, single models are primarily focused on either temporal dependencies or localized glucose patterns but rarely integrate both aspects in an effective manner. RNNs can capture long-term trends for modeling temporal dependencies.

This study integrates both and presents a CRNN-based predictive model enhanced with a bio-inspired optimizer, applied to real-time glucose monitoring. CRNNs are particularly well suited for sequential data such as time series and can learn both short- and long-term dependencies, making them ideal for detecting hyper-and hypoglycemia using time-series data. While CRNNs are not entirely new, having been initially used in speech recognition and video processing, their application to healthcare and glucose prediction remains relatively recent. They have shown promise in biomedical signal analysis, including ECG, EEG, and continuous glucose monitoring (CGM) data. In addition, researchers have widely adopted bioinspired algorithms for optimizing tasks such as CNNs. These algorithms efficiently navigate high dimensional hyperparameter spaces without resorting to exhaustive and computationally expensive exploration. Unlike many previous studies that focus solely on blood glucose prediction, this work also addresses classification of hypo and hyperglycemia in real time, which is an essential feature for practical deployment. This study develops and validates its model using the recent HUPA-UCM dataset[18] which provides rich, multivariate time series data including CGM readings and insulin doses, thus offering a more comprehensive input space for modeling. The research also includes extensive benchmarking, following prior works such as Tena *et al*.,[19] who compared two ensemble networks with 10 others; Ruan *et al*.,[20] who evaluated 18 Machine Learning (ML) algorithms including logistic regression, SGD, k-nearest neighbors, decision tree, Gaussian naive Bayes, Bernoulli NB, multinomial NB, support vector machine, quadratic discriminant analysis, random forest, extra trees, linear discriminant analysis, AdaBoost, bagging, gradient boosting, XGBoost, and multi-layer perceptron; and Shi *et al*., who compared six algorithms including logistic regression, random forest, gradient boosting machine, deep neural networks, XGBoost, and Rulefit.[21] Furthermore, the approach emphasized in this study is predicated on clinically relevant prediction horizons and actionable alerts, with a view to aligning model outputs with the practical management needs of individuals suffering from diabetes. This study makes a significant contribution by addressing methodological shortcomings and enhancing the applicability of predictive models in real-world settings, thereby advancing the goal of personalized and proactive diabetes care. Hypoglycemia classification was evaluated for two thresholds: mild (BG<4.0 mmol/L) and severe (BG<3.0 mmol/L), consistent with clinical definitions. The classification metrics (Precision, Recall, F1, Accuracy, AUROC) were computed for each threshold using the predicted labels vs. true labels. Regression performance for continuous glucose values was evaluated with MAE, RMSE, and R^2 on the predicted 30-min-ahead values.

The paper is organized as follows: Section 2 reviews related work. Section 3 outlines the proposed methodology, Section 4 presents and discusses the results, and Section 5 concludes the paper.

Related work

Tsichlaki *et al*. conducted a systematic literature for hypoglycemia detection in T1D patients.[22] In the following, we present the most important related works according to hyperglycemia and hypoglycemia into three main categories: DL, ML and reinforcement-based approaches.

Deep learning

Jaloli and Cescon propose a CNN-LSTM model to predict blood glucose (BG) level in people with T1D for 30-, 60, and 90-minute prediction horizons (PH), based on historical glucose readings, meal data, and insulin doses.[16] They evaluated the approach on two datasets: Replace-BG, reflecting free-living conditions, and DIAdvisor, representing an in- hospital setting. Aiello *et al*. introduced a two-headed LSTM-based Deep Glucose Forecasting approach for predicting interstitial

glucose levels in people with T1D, using CGM data, carbohydrate intake, and recommended insulin therapy.[23] The architecture was trained on data from 100 virtual adult patients generated by the UVA/Padova simulator and was evaluated on both virtual and real patient datasets, demonstrating strong generalization performance on previously unseen data. Zhu *et al*. presented a dilated recurrent neural network (DRNN) model to generate 30-minute forecasts of glucose levels in individuals with T1DM.[17]

Additionally, they apply transfer learning to leverage data from multiple subjects. On the OhioT1DM dataset, their model achieved a root mean square error (RMSE) of 18.9 mg/dL, outperforming autoregressive models (ARX, 20.1 mg/dL), support vector regression (SVR, 21.7 mg/dL), and conventional neural networks for predicting glucose (NNPG, 22.9 mg/dL). Tena *et al*.[19] introduced two ensemble neural network models for blood glucose prediction in people with T1DM at three prediction horizons (PHs)-30, 60, and 120 minutes and compares their performance against ten recently proposed neural networks. All these models are evaluated on the same OhioT1DM dataset. Karagoz *et al*. presented a comparative analysis of transformer models for multi-horizon BG prediction, evaluating forecasts up to 4 hours ahead using input histories of up to one week.[24] They used the publicly available DCLP3 dataset (N=112), split 80%-10%-10% for training, validation, and testing, and the OhioT1DM dataset (N=12) as an external test set. For short-term prediction, the Crossformer model, a patch-wise transformer architecture, achieved the best 30-minute forecast with an RMSE of 15.6 mg/dL on the OhioT1DM dataset. Munoz-Organero reviewed the current state of the art in blood glucose prediction for T1DM patients and proposed, implemented, and validated a novel hybrid model.[25] Specifically, the differential equations describing carbohydrate and insulin absorption are modeled using a RNN implemented with LSTM cells. The results show promising performance. Alvarado *et al*. presented a method for predicting and detecting hypoglycemic events over a 24-hour time PH.[26] The approach combines wavelet transform analysis of glucose time series with CNNs and was validated using real-world data from 20 individuals with T1D. The results demonstrate strong performance, with accuracy, sensitivity, specificity, and precision all exceeding 88. Mirshekarian *et al*. presented an RNN approach using LSTM units to learn a physiological model of blood glucose.[27] Sun *et al*. introduced a sequential model consisting of one LSTM layer, one bidirectional LSTM layer, and several fully connected layers to predict blood glucose levels across multiple PHs.[28] The model was trained and tested on 26 retrospectively analyzed datasets from 20 real patients. Idrissi *et al*. presented sequential neural network architecture with an LSTM layer followed by two fully connected layers for blood glucose prediction.[29] Another approach proposed by Iacono *et al*. highlights the importance of warning systems while examining the significant problem of preventing hypoglycemia and hyperglycemia in the management of T1D.[30] The proposed work has investigated the prediction of future blood glucose levels using DL models, such as personalized LSTM models. Using LSTM models, the individualized alarm system demonstrates promising predictive performance within 40-minutes prediction horizon, achieving an RMSE of 6.45. It attains an F-Score of 78/79% for hypoglycemia detection and 83.87% for hyperglycemia detection. In addition, the work proposed by Li et al. investigates the use of CRNN model for blood glucose level forecasting in the management of type 1 diabetes.[31] The proposal combined a CNN to extract robust features from tile-aligned inputs sauch as CGM readings, insulin boluses, and meals, with an RNN to capture temporal dynamics and generate predictions in 30 and 60 minutes ahead.

Machine learning and statistical models

Kavakiotis *et al*. explored the use of data mining and ML in diabetes research, highlighting the critical role these techniques play in turning massive volumes of data, including genetic and clinical data from Electronic Health Records (EHRs), into useful insights.[32] The authors discovered that ML algorithms, in particular supervised learning techniques (85% of the cases), were widely used in diabetes research through a systematic literature review (focusing on articles from 2011 to 2016). In addition, Rghioui *et al*. presented an innovative system for remotely monitoring diabetes patients, leveraging contemporary technology like artificial intelligence (AI) and smart devices to optimize monitoring procedures and minimize related costs.[33] The results show that the J48 algorithm has a remarkable 99.17% accuracy, 99.47% sensitivity, and 99.32% precision in classification. In addition, Kodama *et al*. looked at ML systems trained for hypoglycemia diagnosis or prediction.[34] The findings showed that while current ML algorithms (SVM, XG-Boost, RF, LR, and ANN) showed moderate potential in predicting impending hypoglycemia, with sensitivity and specificity averaging 0.80, they were limited in their capacity to detect continuing hypoglycemia. In addition, the research conducted by Shi *et al*. aimed to create an ML model that could forecast a probability that older persons with diabetes would experience severe hypoglycemia (SH) and require hospitalization.[21] To train the ML model, they selected 258 predictors related to medical history, drugs, laboratory tests, and demographics using EHR data from a sizable sample of older persons in Hong Kong. Six distinct ML algorithms were evaluated for performance, and the findings showed that the XG-Boost model performed the best (AUROC=0.978). This research confirms that older people who are at a high risk of experiencing severe hypoglycemia could benefit from preventive intervention by having this ML model integrated into electronic health record systems.

Reinforcement learning

The purpose of the study of Wang *et al*. is to enhance individualized insulin titration in the treatment of T2D

using a model-based reinforcement learning framework called RL-DITR.[35] This framework analyzes the benefits of glycemic status using model-patient interaction to determine the ideal insulin prescription. In the development stage, RL-DITR achieved a mean absolute error of 1.10±0.03 U, which was superior to other DL models and conventional clinical techniques for optimizing insulin titration. With a mean absolute error of 1.18±0.09 U, a step-by-step clinical validation showed that inpatient glycemic management was improved in comparison to junior and intermediate physicians.

To summarize, recent studies on predicting hyperglycemia and hypoglycemia have explored various ML and DL approaches. CGM forecasting is dominated by temporal deep models, such LSTM and CRNN. CRNNs, for instance, have demonstrated impressive performance by integrating convolutional layers for local pattern extraction with recurrent layers for modeling temporal dependencies, achieving a low root mean square error (RMSE) at 30-60-minute horizons. Overall, deep temporal architecture remains the gold standard for CGM- based short-term prediction, hybrid models offer incremental improvements, and interpretable models remain valuable in safety-critical clinical contexts.

Materials and methods

CRNNs have shown strong potential in blood glucose prediction by combining convolutional layers (for spatial feature extraction) with recurrent layers (for modeling temporal dependencies). This integration enables better modeling of both localized glucose fluctuations and long-term trends, outperforming models that address only one aspect. Hyperparameter tuning using metaheuristics such as Particle Swarm Optimization (PSO), Grey Wolf Optimizer (GWO), Cuckoo Search (CS) and Ant Colony Optimization (ACO) can further enhance performance.[36, 37]

The overall workflow of the proposed approach is presented in Figure 1 and involves four main steps: “Data description” where we introduce the used dataset by outlining its source, size, features, and target variables relevant to the study. “Data pre-processing” which is a crucial step for data cleaning, transforming, and normalizing to ensure consistency and suitability for model training. “Optimization” where GWO optimization algorithm is applied to fine-tune parameters or select features that enhance model performance. And finally, “Proposed CRNN architecture” The designed Convolutional Recurrent Neural Network integrates convolutional layers for feature extraction and recurrent layers for temporal pattern learning.

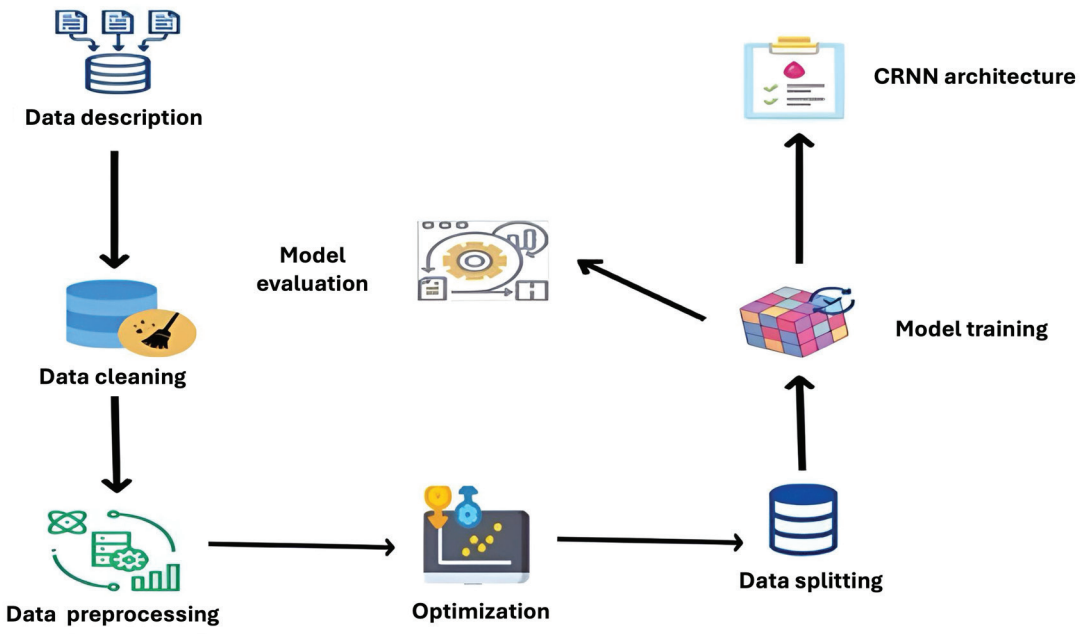


Figure 1. *Proposed approach architecture.*

Data description

Although our study relies on a single dataset HUPA-UCM diabetes dataset first made available in April 2024, comprises 309,392 rows and 8 columns (see Table 1).[17] This choice is justified by the fact that the dataset is recent, comprehensive, and high-quality, containing many samples and rich contextual information. Its size and diversity provide sufficient variability for robust model training and evaluation, reducing the risk of overfitting and enabling reliable generalization within the target population. Moreover, no other publicly available dataset offers the same level of completeness or alignment as the specific hyperglycemia-prediction task addressed in this work.

The dataset aggregates data from 25 type 1 diabetes (T1D) patients aged 18 to 65 years, collected between June 13, 2018, and May 18, 2022. Each record is time-stamped at 5-minute intervals, providing continuous time-series monitoring over several weeks per patient. The dataset contains glucose measurements obtained from a Continuous Glucose Monitor (CGM), physical activity indicators (steps, calories burned, and heart rate), insulin injections (basal rates and bolus volumes), and carbohydrate intake during the interval. Notably, the dataset shows initial glucose values above 300 mg/dL and later glucose values in the range of 110-128 mg/dL. The data enables forecasting tasks such as predicting blood glucose levels 30-60 minutes in advance and analyzing responses to diet and treatment. The temporal variable was converted to minutes to ensure consistency. No missing values are present, making the dataset well-suited for analytical and predictive tasks. The average glucose concentration is 141.43 mg/dL, with a standard deviation of 57.09, and values ranging from 40 to 444 mg/dL-indicating substantial glycemic variability. The average heart rate is 76.99 bpm (SD=85), and the mean number of steps per interval is 30.83. Mean energy expenditure is 8.81 calories per interval. In contrast, insulin delivery parameters (basal and bolus) and carbohydrate intake exhibit lower average values, consistent with typical clinical dosing and meal patterns.

Table 1. HUPA-UCM dataset description.

Attribute	Description
Time	Timestamp of each measurement in the format yyyy-MM-dd'T'HH:mm:ss
Glucose	Blood glucose concentration (mg/dL)
Calories	Calories burned during the time interval
Heart_rate	Heart rate (beats per minute)
Steps	Number of steps taken during the time interval
Basal_rate	Amount of basal insulin infused during the interval
Bolus_volume_delivered	Volume of bolus insulin injected during the interval
Carb_input	Carbohydrate intake during the interval (1 serving = 10 gm)

Data pre-processing

At this step, the dataset is prepared to ensure its quality and suitability for the modeling process. First, the missing values are identified and addressed using a statistical imputation method, the mean. Statistical measures such as the inter- quartile range (IQR) or z-score are used to detect outliers. These are then either removed or adjusted to reduce their influence. All variables are made comparable using either standardization or normalization techniques, which are used to perform feature scaling. This improves both model convergence and performance. Finally, the processed data is split into training and testing subsets, enabling the model to be evaluated impartially.

Optimization

Metaheuristic algorithms have been extensively adopted for optimizing CNNs due to their capacity to navigate high-dimensional hyperparameter spaces efficiently without resorting to exhaustive and computationally expensive searches. Recent studies have shown how well various metaheuristic algorithms can be used to optimize CNNs for healthcare tasks which improved accuracy and reduced error rates through optimized parameter selection. We have used GWO for our proposal because it is widely used for medical and prediction tasks. GWO optimization algorithm is a metaheuristic proposed by which mimics the leadership hierarchy and hunting mechanism of grey wolves in nature.[37] It is based on the hierarchy-based hunting characteristics of the Canidae family. The GWO principles focus on hunting behaviors highlight the dynamic nature of addressing convergence concerns. GWO algorithm uses the dynamic features of the Canidae population to solve this problem. This enhances the algorithm's capacity to attain superior global optimal convergence compared to other optimization techniques that often converge with local optimal solutions. Additionally, it fosters enhanced stability between the exploitation and exploration phases. When grey wolves are present, the hierarchy-based hunting phenomenon involves a leading grey wolf, a representative grey wolf, an obeying grey wolf and a scapegoat grey wolf. The leadership

hierarchy is simulated using four types of grey wolf: alpha, beta, delta and omega. In addition, three main steps of hunting are implemented.

These are searching for prey, encircling prey and attacking prey. These steps are used to perform optimization. The purpose of this step is to optimize the hyperparameters to achieve good accuracy in reduced time. In GWO, candidate solutions are modeled as wolves, and the best three solutions so far are considered α, β, and δ. Wolves encircle prey using position update equations that simulate hunting strategies. Two key coefficient vectors (A and C) control the movement toward prey and the balance between exploration and exploitation. The encircling is modeled by computing distances from each wolf to α, β, and δ, then estimating three possible positions. The new position is the average of these three positions, encouraging convergence toward optimal solutions. The parameter **a** decrease linearly from 2 to 0 over iterations, gradually shifting the search from exploration to exploitation. At each iteration, fitness is evaluated, and the top three wolves are updated.

Finally, the algorithm stops after a fixed number of iterations, returning the best-found solution (α). The hyperparameter search space for GWO was defined based on commonly reported ranges in the literature and guided by preliminary experiments to ensure efficient exploration and stable convergence. Specifically:

- Population size: 20-40 wolves, balancing search diversity and computational cost.
- Number of iterations: 30-50, allowing convergence without excessive runtime.
- Search bounds for model hyperparameters:
- Learning rate: 0.0001-0.01
- Number of neurons per layer: 16-256
- Batch size: 16-128

These ranges were chosen to cover realistic and meaningful values while avoiding unstable regions. We will clarify these details in the revised manuscript to ensure full transparency and reproducibility.

In this study, GWO is used exclusively for hyper parameter optimization rather than for training the model's weights. The CNN-LSTM hyperparameters—such as the number of filters, kernel size, LSTM units, dropout rate, and window length—are non-differentiable and form a highly non-convex search space, making gradient-based optimizers unsuitable. GWO provides a gradient-free global search strategy that can effectively explore this complex space and identify better-performing configurations, while the internal network weights are still trained using a standard gradient-based optimizer

Proposed CRNN architecture

CNN and RNN are combined into a powerful hybrid DL model. A main application of CNNs aims to learn to extract features from sequential data. CRNNs follow exactly this approach, as the connections of a CNN within the model learn relevant features in a timeseries input in a self-supervised manner. RNNs, mainly LSTMs: RNNs are used for modelling temporal dependencies that are part of the series. LSTMs solve the problem of vanishing gradients in RNNs so that long-term dependencies can be modeled. As shown in Figure 2, the CRNN model is built with a feedback mechanism for glucose prediction. The model receives the data via the input layer. This input passes through the CNN components. These components are designed to ex back mechanism refines the prediction process over time and allows the model to be used for real-time predictions.

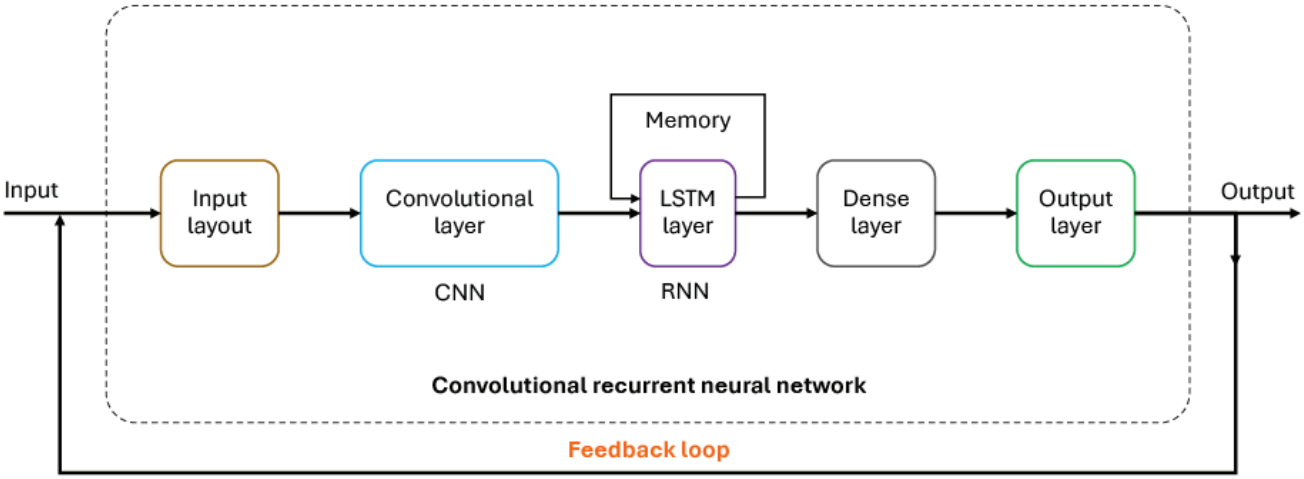


Figure 2. *The proposed CRNN architecture.*

Metrics

The following metrics were applied as the most used for evaluating classification models:

- Mean Absolute Error (MAE): This metric measures the average magnitude of errors without considering their direction.
- Mean Squared Error (MSE): This metric emphasizes larger errors by squaring the differences.
- Root Mean Squared Error (RMSE): These metric computes error in the same units as the target variable.
- Precision: The proportion of correctly predicted positive instances among all predicted positives.
- Accuracy measures the proportion of correctly predicted instances (both positive and negative) out of the total predictions made. It reflects the overall correctness of a classification model.
- Recall: measures the proportion of actual positive cases that are correctly identified by the model. It indicates how well the model detects true positives.
- F1-score: The F1-score is the harmonic mean of Precision and Recall, providing a balanced measure of a model's accuracy when both false positives and false negatives are important. It is especially useful for evaluating models on imbalanced datasets.
- Clarke Error Grid is a metric used to evaluate blood glucose predictions by classifying each predicted value based on its clinical impact on the patient. It distinguishes acceptable errors from potentially dangerous ones, assessing the safety of a model or device.

Results

In this section, the performance of the proposed model is evaluated in comparison to different baseline models in terms of hypoglycemia detection (classification) and prediction of continuous glucose values (regression). The training performance and model stability are analyzed. In addition, the current literature model is compared with clinical validations. The proposed model is trained and tested on the HUPA-UCM diabetes dataset. The model integrates layer-wise CNN layers for extracting short-term patterns and LSTM layers for capturing long-term dependencies in the time series data. The proposed architecture, as shown in Figure 3, is based on a 1D convolutional layer with 64 filters and a kernel size of 3. These layers process the input data, including the glucose measurements and the associated health metrics. This is followed by a max-pooling layer with a size of 2 to reduce dimensionality and explore salient features. In addition, the sequential structure contains two LSTM layers, with each layer consisting of 50 slices. The first 50 layers provide the sequences, and the second 50 layers provide the final output; they are designed to capture the temporal dependencies at different levels.

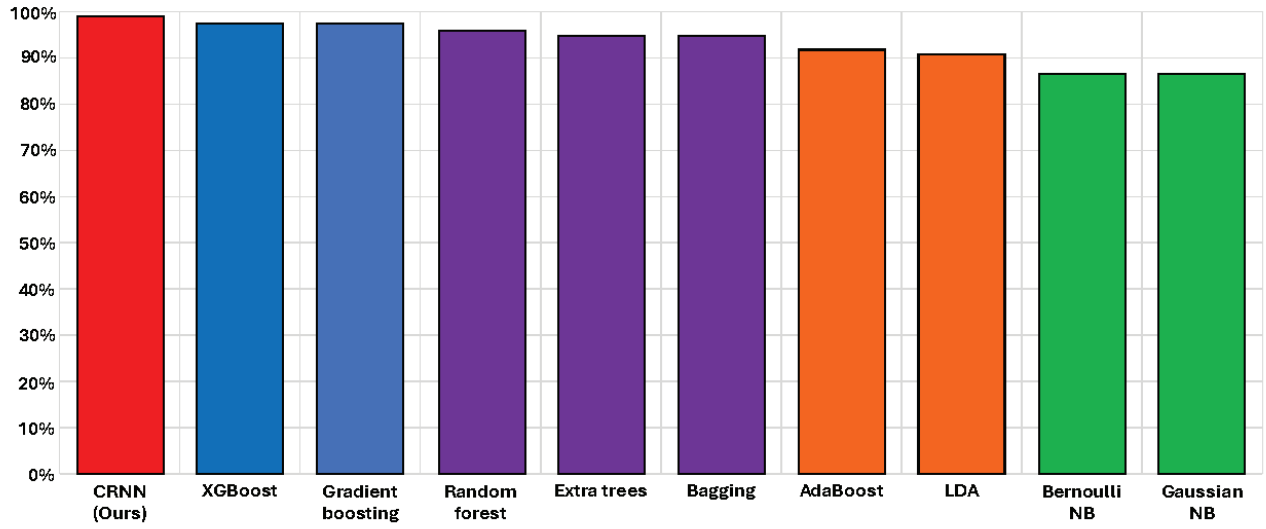


Figure 3. *Top 10 classification models.*

The proposed model (CRNN+GWO) is evaluated against 19 baseline ML algorithms and recent state-of-the-art methods demonstrating strong performance in both classification and regression tasks. As shown in Table 2, the proposal consistently outperforms all traditional ML approaches. As can be seen on Table 2, our model achieves improved precision (0.97), recall (0.98), and F1-score (0.98), which is a significant improvement over traditional ML approach. These improvements lead to an accuracy of 99%, outperforming all compared algorithms. This sets a new bench- mark for the classification of glucose events. In addition, the AUROC analysis showed that the hybrid model has a performance of 0.98, while the classification accuracy of CRNN improved by 1% compared to the best-performing ensemble methods (XGBoost and Gradient Boosting with 97%). Figure 3 shows the 10 most accurate models. The performance of ten models in the classification of hypoglycemia was compared with the proposed model demonstrating superiority in both accuracy and robustness. The area under the curve (AUROC) value for CRNN is 0.97, and the precision, recall, and F1 scores for both mild (BG<4) and severe (BG<3) hypoglycemia are all above 0.96. These results indicate that the proposed CRNN hybridized with GWO exhibits exceptional consistency and clinical reliability. It is noteworthy that an increase in the training epoch number from 50 to 100 results in a complete absence of change in the classification metrics. This serves to con- firm the stability and early convergence of the model.

Table 2. Classification performance comparison – hypoglycemia.

Model	AUROC	Precision (BG<4)	Recall (BG<4)	F1 score (BG<4)	Precision (BG<3)	Recall (BG<3)	F1 score (BG<3)	Classification accuracy
CRNN (50 Epochs) + GWO	0.98	0.97	0.98	0.98	0.98	0.98	0.98	0.99
CRNN (100 Epochs) + GWO	0.98	0.97	0.98	0.98	0.98	0.98	0.98	0.99
CRNN (50 Epochs)	0.97	0.96	0.97	0.96	0.96	0.97	0.96	0.98
CRNN (100 Epochs)	0.97	0.96	0.97	0.96	0.96	0.97	0.96	0.98
XGBoost	0.96	0.88	0.70	0.78	0.97	0.67	0.79	0.97
Gradient boosting	0.96	0.87	0.70	0.78	0.96	0.67	0.79	0.97
RF	0.94	0.86	0.67	0.75	0.96	0.66	0.78	0.95
Extra Trees	0.93	0.85	0.68	0.76	0.94	0.66	0.78	0.94
Bagging	0.93	0.84	0.70	0.76	0.93	0.67	0.78	0.94
AdaBoost	0.89	0.68	0.60	0.64	0.63	0.46	0.53	0.91
LDA	0.88	0.69	0.75	0.72	0.72	0.72	0.72	0.90
Bernoulli NB	0.82	0.60	0.60	0.60	0.47	0.67	0.55	0.86
Gaussian NB	0.81	0.47	0.68	0.56	0.33	0.81	0.47	0.85
Decision Tree	0.81	0.70	0.71	0.70	0.68	0.73	0.70	0.85
SVM	0.79	0.73	0.10	0.18	0.41	0.10	0.16	0.83
QDA	0.77	0.23	0.96	0.37	0.15	0.97	0.26	0.82
Passive aggressive	0.76	0.46	0.25	0.32	0.33	0.10	0.15	0.81
Multinomial NB	0.75	0.10	0.10	0.10	0.10	0.10	0.10	0.80
SGD	0.74	0.12	0.10	0.11	0.10	0.10	0.10	0.80
MLP	0.74	0.57	0.17	0.26	0.47	0.14	0.22	0.80
LR	0.73	0.48	0.10	0.17	0.39	0.10	0.16	0.79
KNN	0.62	0.40	0.18	0.25	0.30	0.15	0.20	0.70

In contrast, ensemble models such as XGBoost and Gradient Boosting demonstrate robust overall performance yet exhibit diminished recall values (0.67-0.70), particularly under BG3 conditions, signifying a diminished capacity to capture severe hypoglycemia events. Conventional methods, including RF, Extra Trees, and Bagging, demonstrate moderate performance, with Area Under the Receiver Operating Characteristic Curve (AUROC) values ranging from 0.93 to 0.94 and F1 scores approximating 0.75 to 0.78. This suggests a limited degree of sensitivity.

It has been demonstrated that simpler models, including SVM, LR, and NB variants, exhibit poor recall (<0.2) and unstable F1 scores. This renders them unreliable for real- world hypoglycemia detection. It is interesting to note that QDA achieves a high recall rate of approximately 0.97 but extremely low precision, which renders it susceptible to false positives. As a result, the

present study demonstrates that CRNN hybridized with GWO exhibits an optimal balance between detection capability and false alarm mitigation, rendering it highly conducive to the development of early hypoglycemia warning systems.

On the other hand, the regression analysis has shown that our model outperforms existing state-of-the-art approaches. Table 3 shows the comparison of regression performance with the current literature. As can be seen, our model achieved an improved RMSE of 10.0 mg/dL and an MAE) of 7.0 mg/dL with an R-squared value of 0.92, indicating good accuracy in predicting continuous glucose levels. In addition, clinical acceptability according to Clarke Error Grid Analysis is 92%, which is an excellent performance for continuous monitoring applications. The approach closest to our model is which gives an RMSE of 35.19 mg/dL for LSTM and 36.08 mg/dL for self-attention networks. The maximum clinical acceptability achieved by both algorithms was 91% using Clarke Error Grid analysis.

Table 3. Regression performance comparison – Blood glucose value prediction.

Model/Study	RMSE (mg/dL)	MAE (mg/dL)	R^2/Correlation	Dataset	Prediction horizon (minute)	Clinical acceptance	Real-time accuracy
Our CRNN (50 Epochs) + GWO	10	7	0.92	HUPA-UCM	Real-time (5-min)	92%	90%
Our CRNN (100 Epochs) + GWO	10	7	0.92	HUPA-UCM	Real-time (5-min)	92%	90%
[35]Bi-LSTM	19.49±5.42	14.93±4.20	0.43±0.2	Life-log data	Current time (0-min)	-	-
[33]LSTM	35.19	-	-	OhioT1DM	30-60	>91% (CEG)	-
[33]SAN	36.08	-	-	Multiple	30-60	>91% (CEG)	-
[33]CNN	~37-40	-	-	Multiple	30-60	>91% (CEG)	-
[33]TCN	~37-40	-	-	Multiple	30-60	>91% (CEG)	-

Table 4 shows the training performance analysis with 50 and 100 epochs. The analysis revealed optimal generalization properties where the improvement over RMSE during training was from 9.2 mg/dL (50 Epochs) to 8.8 mg/dL (100 Epochs). Also, the test set RMSE remained stable at 10.0 mg/dL for prediction tasks, which is an indicator of generalization without over-fitting. Comparing CRNN+GWO models trained for 50 and 100 epochs shows a high degree of model stability with only slight improvements in training performance. Training RMSE improved slightly from 9.2 mg/dL to 8.8 mg/dL (a 4.3% reduction), suggesting modest learning progress with additional epochs. However, the test RMSE and real-time RMSE remained unchanged at 10.0 and 15.0 mg/dL, respectively, strongly indicating that the model did not overfit de- spite the longer training period. Similarly, R-squared (0.92) stayed constant, confirming that the model's ability to explain variance was already optimized and consistent across epochs. The minor decrease in MAE (from 9.3 to 9.2 mg/dL) shows slightly better average prediction accuracy, but the effect is minimal. Classification metrics—F1 score, precision, recall, and accuracy—remained unchanged, all at strong values, con- firming that the model is highly stable in terms of event classification across both training configurations. The clinical acceptance score (0.92) and real-time accuracy (0.90) also showed no change, further reinforcing confidence in the model's performance when deployed in a clinical setting. Overall, the results suggest that the hybridized GWO based CRNN architecture achieves rapid convergence, with performance metrics stabilizing after 50 epochs. Extending training to 100 epochs does not degrade performance but offers diminishing returns, demonstrating that the model is efficient, stable, and robust under extended training.

Table 4. Training performance analysis.

Metric	CRNN (50 Epochs)	CRNN (100 Epochs)	Change
RMSE (Training set)	9.2 mg/dL	8.8 mg/dL	-0.4 (-4.3%)
RMSE (Test set – Prediction)	10.0 mg/dL	10.0 mg/dL	0.0 (0.0%)
RMSE (Real-time)	15.0 mg/dL	15.0 mg/dL	0.0 (0.0%)
R-squared (R^2)	0.92	0.92	0.00 (0.0%)
MAE (Prediction)	7.1 mg/dL	7.0 mg/dL	-0.1 (-1.4%)
MAE (Real-time)	10.0 mg/dL	10.0 mg/dL	0.0 (0.0%)
F1 Score	>0.95	>0.95	No significant change
Precision	0.97	0.97	No significant change
Recall	0.98	0.98	No significant change
Classification accuracy	0.99	0.99	No change
Clinical acceptance	0.92	0.92	No change
Real-time accuracy	0.90	0.90	No change

Table 5 compares our results with four recent studies using clinical validation metrics. The proposed CRNN model achieved 92% clinical acceptability and 99% accuracy for hypoglycemia detection, surpassing the best values reported in those studies. However, since our model was evaluated on the HUPA-UCM dataset whereas other works used different datasets, these comparisons are not one-to-one. We report them for context, but direct performance ranking should be interpreted with caution. Future evaluations on common datasets will be necessary for a fair comparison.

Table 5. Comparison of blood glucose prediction models.

Study & Model	Classification accuracy	Clinical acceptance	Real-time performance	Dataset	Key findings
Our CRNN (100 Epochs)+GWO	99% classification	92% clinical	90% real-time	HUPA-UCM	Superior across all evaluation modes
[33]LSTM	>91% (CEG Safe zone)	>91% (CEG)	Not reported	OhioT1DM	Best generalization across datasets
[33]SAN	>91% (CEG safe zone)	>91% (CEG)	Not reported	Multiple	Second-best performance
[35]LSTM	Physiologically sound	–	Not reported	OhioT1DM	SHAP-validated interpretability
[35]Bi-LSTM	–	Correlation: 0.43±0.2	Not reported	Life-log	Virtual CGM framework

Note: *CEG: Clarke error grid; CGM: ???, SAN: self-attention network, CRNN: convolutional recurrent neural network.*

An avenue for further improvement is the use of Transformer-based models, which have achieved state-of-the-art results in blood glucose prediction. These models leverage self-attention to capture long-term dependencies and have outperformed traditional CNN/LSTM in some scenarios. We did not evaluate Transformers in this study, as our focus was on CNN/LSTM hybrids; however, future work will compare the CRNN approach against Transformers to ensure the proposed model remains competitive with the latest deep learning techniques.

A limitation of this study is that the model was trained and tested on a single dataset (HUPA-UCM). No external or cross-population evaluation was performed, which may affect the generalizability of the results. In future work, we plan to validate the model on additional datasets (e.g., OhioT1DM) and across different patient populations, as well as exploring additional AI architectures and bio-inspired optimization algorithms to further enhance predictive performance.

Conclusion

In this paper, we present a CRNN architecture, enhanced with a GWO for hyperparameter tuning, applied to the task of predicting hyperglycemia and hypoglycemia. Among the tested configurations, the CRNN+GWO hybrid achieved the lowest RMSE and MAE, outperforming the baseline CRNN. While our model outperforms others in point estimates of accuracy, we acknowledge that we did not perform statistical significance testing. Thus,

differences, especially when within a few percentage points, should be interpreted with caution. Using a recent large-scale dataset, we benchmarked our results against recent state-of-the-art approaches. Our primary objective was to maximize accuracy in the early prediction of hypo and hyperglycemic events and to enable real-time detection. The proposed model achieved 99% accuracy on offline data with low mean squared and absolute errors, and nearly 90% accuracy in simulated real-time scenarios. It can trigger timely alerts for abnormal glucose levels, offering strong potential for improving patient health outcomes. The combination of CNN, RNN—particularly LSTM and GWO proved robust across DL tasks in this domain.

Funding

No funding was provided for this research.

Conflict of interest

There is no potential conflict of interest.

CRediT authorship contribution statement

Djamila Ouaret: conducted the experiments and analyzed the data; **Houda El Bouhissi**: supervised the project, drafted the manuscript, improved the contribution and served as the corresponding author; **Tatiana Emarkova**: drafted the manuscript, improved the contribution. **Rabie A. Ramadan**: drafted the manuscript, improved the contribution.

Acknowledgements

None

References

[1] ElSayed NA, McCoy RG, Aleppo G, Balapattabi K, Beverly EA, Briggs Early K, et al. Diagnosis and classification of diabetes: Standards of care in diabetes—2025. Diabetes Care. 2025; 48(Suppl 1): S27-S49. doi:10.2337/dc25-S002.

[2] Vaiyapuri T, Alharbi G, Muttipoll Dharmarajlu S, Bouteraa Y, Misra S, Venkata Naga Ramesh J, et al. IoT-Enabled early detection of diabetes diseases using deep learning and dimensionality reduction techniques. IEEE Access. 2024; 12: 143016-28.

[3] Albahli S. Predictive analytics for diabetic patient care: Leveraging AI to forecast readmission and hospital stays. Comput Model Eng Sci. 2025; 143(1): 1095-128. doi:10.32604/cmes.2025. 05882.

[4] Lin J, Song Z, Li Y, Chiang C, Hirakawa Y, Nakano Y, Hong YJ, Matsunaga M, Ota A, Tamakoshi K, Yatsuya H. Nonrestorative sleep and type 2 diabetes incidence: The Aichi workers' cohort study. J Epidemiol. 2024; 34(9): 428-33. doi:10.2188/jea.JE20230184.

[5] Bliss M. Who discovered insulin? In: The discovery of insulin. London: Palgrave Macmillan; 1982: pp 189-211.

[6] Umpierrez GE, Davis GM, ElSayed NA, Fadini GP, Galindo RJ, Hirsch IB, et al. Hyperglycemic crises in adults with diabetes: A consensus report. Diabetes Care. 2024; 47(8): 1257-75. doi:10.2337/dci24-0032.

[7] Yang T, Yu X, Tao R, Li H, Zhou J. Blood glucose prediction for type 2 diabetes using clustering-based domain adaptation. Biomed Signal Process Control. 2025; 105: 107629. doi:10.1016/j.bspc. 2025.107629.

[8] Mazgouti L, Laamiri N, Ben Ali J, El Amrani El Idrissi N, Di Costanzo V, Naeck R, et al. Optimization of blood glucose prediction with LSTM-XGBoost fusion and integration of statistical features for enhanced accuracy. Biomed Signal Process Control. 2025; 107: 107814. doi:10.1016/j.bspc.2025.107814.

[9] Gao J, Guo C, Liu Y, Li P, Zhang J, Liu M. Dynamic-static feature fusion with multi-scale attention for continuous blood glucose prediction. In: Proc IEEE Int Conf Acoust Speech Signal Process (ICASSP); 2025 Apr; Hyderabad; India: pp 1-5.

[10] Sharma SK, Zamani AT, Abdelsalam A, Muduli D, Alabrah AA, Parveen N, Alanazi SM. A diabetes monitoring system and healthmedical service composition model in cloud environment. IEEE Access. 2023; 11: 32804-19. doi:10.1109/ACCESS. 2023.3258549.

[11] Ghosh M, Bora V. Evolution in blood glucose monitoring: from invasive to noninvasive devices and sensors. Discover Med. 2025; 2: 74. doi:10.1007/ s44337025002731.

[12] El Bouhissi H, Al-Qutaish RE, Ziane A, Amroun K, Yaya N, Lachi M. Towards diabetes mellitus prediction based on machine learning. In: Proc Int Conf Smart Computing and Application (ICSCA); 2023: pp 1-6.

[13] El Bouhissi H, Ziane A, Rahmani L, Medbal M, Kostiuk M. RF-PSO: An optimized approach for diabetes prediction. In: Proc Int Conf Smart Technologies (ICST); 2023: pp 227-38.

[14] Tomic D, Harding JL, Jenkins AJ, Shaw JE, Magliano DJ. The epidemiology of type 1 diabetes mellitus in older adults. Nat Rev Endocrinol. 2025; 21(2): 92-104. doi:10.1038/s4157402401046z.

[15] Alfian G, Syafrudin M, Ijaz MF, Syaekhoni MA, Fitriyani NL, Rhee J. A personalized healthcare monitoring system for diabetic patients using BLEbased sensors. Sensors (Basel). 2018; 18(7): 2183. doi:10.3390/s18072183.

[16] Jaloli M, Cescon M. Longterm prediction of blood glucose levels in type 1 diabetes using a CNNLSTM based deep neural network. J Diabetes Sci Technol. 2023; 17(6): 1590-601. doi:10.1177/1932296822109 2785.

[17] Zhu T, Li K, Chen J, Herrero P, Georgiou P. Dilated recurrent neural networks for glucose forecasting in type 1 diabetes. J Healthc Inform Res. 2020; 4(3): 308-24. doi:10.1007/s41666020000682.

[18] Hidalgo JI, Alvarado J, Botella M, Aramendi A, Velasco JM, Garnica O. HUPAUCM diabetes dataset. Data Brief. 2024; 55: 110559. doi:10.1016/j.dib. 2024.110559.

[19] Tena F, Garnica O, Lanchares J, Fernándezde Vega

F, Hidalgo JI. Ensemble deep neural networks for blood glucose prediction. Sensors (Basel). 2021; 21(21): 7090. doi:10.3390/s21217090.

[20] Ruan Y, Bellot A, Moysova Z, Tan GD, Lumb A, Davies J, van der Schaar M, Rea R. Predicting the risk of inpatient hypoglycemia with machine learning using electronic health records. Diabetes Care. 2020; 43(7): 1504-11. doi:10.2337/dc191743.

[21] Shi M, Yang A, Lau ESH, Luk AOY, Ma RCW, Kong APS, Wong RSM, Chan JCM, Chan JCN, Chow E. A novel electronic health recordbased, machine-learning model to predict severe hypoglycemia leading to hospitalizations in older adults with diabetes: a territorywide cohort and modeling study. PLoS Med. 2024; 21(4): e1004369. doi:10. 1371/journal.pmed.1004369.

[22] Tsichlaki S, Koumakis L, Tsiknakis M. Type 1 Diabetes hypoglycemia prediction algorithms: Systematic review. JMIR Diabetes. 2022; 7(3): e34699. doi:10. 2196/34699.

[23] Aiello EM, Lisanti G, Magni L, Musci M, Toffanin C. Therapydriven deep glucose forecasting. Eng Appl Artif Intell. 2020; 87: 103255. doi:10.1016/j. engappai.2019.103255.

[24] Karagoz MA, Breton MD, Fathi AE. A comparative study of transformerbased models for multihorizon blood glucose prediction. arXiv. 2025: arXiv: 2505. 08821. doi:10.48550/arXiv.2505.08821.

[25] Munoz-Organero M. Deep physiological model for blood glucose prediction in T1DM patients. Sensors (Basel). 2020; 20(14): 3896. doi:10.3390/s2014 3896.

[26] Alvarado J, Velasco JM, Chavez F, Fernándezde Vega F, Hidalgo JI. Combining wavelet transform with convolutional neural networks for hypoglycemia prediction from CGM data. Chemometr Intell Lab Syst. 2023; 243: 105017. doi:10.1016/j.chemolab. 2023.105017.

[27] Mirshekarian S, Bunescu R, Marling C, Schwartz F. Using LSTMs to learn physiological models of blood glucose behavior. In: Proc IEEE EMBS Int Conf (EMBC); 2017; Jeju; Korea: pp 2887-91.

[28] Sun Q, Jankovic MV, Bally L, Mougiakakou SG. Predicting blood glucose with LSTM and Bi-LSTM deep neural networks. In: Proc Int Symp Neural Networks Appl (NEUREL); 2018; Belgrade: pp 1-5.

[29] El Idrissi T, Idri A, Abnane I, Bakkoury Z. Predicting blood glucose using an LSTM neural network. 2019 Federated Conference on Computer Science and Information Systems (FedCSIS). 2019: 35-41. doi:10.15439/2019F159.

[30] Iacono F, Magni L, Toffanin C. Personalized LSTM-based alarm systems for hypoglycemia and hyperglycemia prevention. Biomed Signal Process Control. 2023; 86: 105167. doi:10.1016/j.bspc.2023. 105167.

[31] Li K, Daniels J, Liu C, HerreroVinas P, Georgiou P. Convolutional recurrent neural networks for glucose prediction. IEEE J Biomed Health Inform. 2019; 24(2): 603-13. doi:10.1109/JBHI.2019.29084 88.

[32] Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine learning and data mining methods in diabetes research. Comput Struct Biotechnol J. 2017; 15: 104-16. doi:10.1016/j.csbj.2016.12.005.

[33] Rghioui A, Lloret J, Sendra S, Oumnad A. A smart architecture for diabetic patient monitoring using machine learning algorithms. Healthcare (Basel). 2020; 8(3): 348. doi:10.3390/healthcare8030348.

[34] Kodama S, Fujihara K, Shiozaki H, Horikawa C, Yamada MH, Sato T, et al. Ability of current machine learning algorithms to predict and detect hypoglycemia in patients with diabetes mellitus: a metaanalysis. JMIR Diabetes. 2021; 6(1): e22458. doi:10.2196/22458.

[35] Wang G, Liu X, Ying Z, Yang G, Chen Z, Liu Z, et al. Optimized glycemic control of type 2 diabetes with reinforcement learning: a proofofconcept trial. Nat Med. 2023; 29: 2633-42. doi:10.1038/ s41591023025529.

[36] LópezGómez JA, Romero FP, Angulo E. Bioinspired algorithms for the characterization of excellent performance in handball players: A datadriven methodology. Expert Syst Appl. 2025; 274: 126821. doi:10. 1016/ j.eswa.2025.126821.

[37] Mirjalili S, Saremi S, Mirjalili SM, dos S. Coelho L. Multiobjective grey wolf optimizer: a novel algorithm for multicriterion optimization. Expert Syst Appl. 2016; 47: 106-19. doi:10.1016/j.eswa. 2015.10.039.

Journal of Associated Medical Sciences 2026; 59(2): 84-90

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Radiation protection in dental radiology: Compliance, continuing education and equipment audit in imaging clinics

Emmanuel O. Esien-umo[1*], Valentine C. Ikamaise[2], Blessing S. Ibe[1], Akwa E. Erim[1], Christopher A. Ishiekwen[3], Kingsley C. Omeke[1], Ndubuisi O. Chiaghanam[1]

[1]*Department of Radiography and Radiological Science, University of Calabar, Nigeria.*
[2]*Department of Radiography and Radiation Science, University of Uyo, Nigeria.*
[3]*Department of Radiography and Radiation Science, Benue state University, Nigeria.*

ARTICLE INFO

Article history:
Received 28 September 2025
Accepted as revised 22 January 2026
Available online 28 January 2026

Keywords:
Radiation protection, compliance, continuing education, equipment availability.

ABSTRACT

Background: Continuing education and the availability of protective equipment are two key factors that significantly influence compliance with radiation protection measures in imaging. Their absence could lead to practices that expose patients, staff, and the public to unnecessary risks. This study assessed compliance with respect to the continuing education of dental professionals and the availability of equipment in selected dental imaging clinics.

Materials and methods: A cross-sectional study. Questionnaires were shared to 55 dental professionals working in eight dental imaging clinics, to assess their level of compliance, specifically regarding continuing education and equipment availability (self-audit), and physical audit of the protective equipment in these clinics by the authors.

Results: About 25.5% of the respondents demonstrated a good level of compliance and equipment availability in their centres, while 50.9% received further training. The multi-level aggregate analysis showed that two clinics had high equipment score, good compliance score and compliance level. Spearman's rank order correlation between compliance and equipment availability was statistically significant (p=0.001), while Cohen's Kappa agreement between self and physical audit ranged from 38%-100% (K: 0.091-1.00).

Conclusion: This study showed that though continuing training is essential for compliance, equipment availability significantly influenced compliance with radiation protection measures.

Introduction

Doses from dental imaging though low, could from multiple imaging procedures, potentially increase the risk of adverse health effects such as thyroid, salivary gland, head and neck tumors, underscoring the need for protection against ionizing radiation in dental imaging.[1-6] The International Commission for Radiation Protection (ICRP) recommends keeping radiation doses as low as reasonably achievable (ALARA principle), as there is no known safe dose of radiation. Compliance with this principle forms the basis for radiological protection against ionizing radiation.[7,8]

In radiation protection, continuous training of dental professionals and the availability of protective equipment play a significant role in compliance.[8,9]

* **Corresponding contributor.**
Author's Address: *Department of Radiography and Radiological Science, University of Calabar, Nigeria.*
E-mail address: emmyesien@unical.edu.ng
doi: 10.12982/JAMS.2026.039
E-ISSN: 2539-6056

The assessment of the continuing education of dental professionals and audit of equipment in dental clinics evaluates the human and material factors that influence compliance with radiation protection measures in dental imaging. The outdated practice of still using lead aprons and thyroid collars to protect patients[10] further underscores the need for continuing education. Studies have reported on the impact of continuing training of dental professionals in dental imaging, but continuing education alone does not translate into good compliance in the absence of protective equipment.[11-13] Also, the availability of protective equipment significantly improves the dental clinic's ability to maintain a safe working environment. However, studies on the impact of protective equipment on compliance in dental imaging are lacking in Nigeria as previous studies focused on the knowledge of radiation protection protocols by dental students and dental professionals.[10,14,15] This present study will assess compliance with respect to the continuing education of dental professionals and the availability of equipment in selected dental imaging centres in Calabar and Uyo cities of Cross River and Akwa Ibom States respectively, Nigeria.

Materials and methods

This study was a cross-sectional survey. A purposive convenience sampling method was used to obtain data from 55 respondents in four public and four private dental clinics in Calabar and Uyo cities. The clinics were selected based on high patient throughput (50 patients/week), access to respondents and the dental clinics. Pre-tested questionnaires were shared to dental professionals (Dentists, Dental assistants, and Radiographers) involved in carrying out dental imaging procedures (self-audit), while four authors (Radiographers trained on the checklist) audited the equipment in the dental clinics (physical audit). Questionnaires comprised two sections. Section A focused on the demographics while section B assessed compliance, equipment availability and continuing education. A three-point Likert scale was used to assess the respondents' level of compliance, self-and physical audit of equipment while a binary scale was used to assess continuing education. For compliance, checklist included adherence to ALARA principle, respondents standing behind a protective barrier, use of rectangular collimator, respondents standing at a distance of at least two metres, use of digital receptors, use of receptor holders, avoiding the use of lead aprons/thyroid collars and the use of a monitoring device. Adherence with 6-8 variables on the checklist was graded as good compliance (scored 3 points), 4-5 variables was graded as fair compliance (scored 2 points) and 0-3 variables was graded as poor compliance (scored 1point) on the Likert scale. For the self- and physical audit of equipment, checklist included availability of digital receptors, receptor holders, rectangular collimators, radiation signs/symbols, designated area for dental imaging, protective barrier, monitoring devices and quality control of dental X-ray units. A score of 6-8 was depicted as high availability (3 points), 4-5 as moderate availability (2 points), and 0-3 as low availability (1point). For continuing education, respondents who had received further training on radiation protection in the last three years were classified as trained and the others were classified as not trained. Multi-level aggregate analysis was used for the non-independence of data obtained from 55 respondents nested within 8 clinics. Spearman's rank-order correlation coefficient was used to assess the association between clinic-level variables and clinic compliance, while Cohen's Kappa was used to assess the agreement between the respondents' self-audit and authors' physical audits of protective equipment. Jamovi statistical software (version 2.6.4) was used to carry out all the analyses.

Results

Of the 62 questionnaires shared, only 55 respondents completed and returned, giving a response rate of 88.7%. Figure 1 showed that most of the respondents were within the 30-39 years age bracket (N=22, 40%), with more men (N=33, 60%) than women (N=22, 40%) involved in the study. Majority were dentists (N=40, 61.8%), while 1-10 years of work experience was the most common among the respondents (N=23, 41.8%). Periapical procedures (N=34, 61.8%) were the most prevalent imaging procedures carried out in these dental clinics.

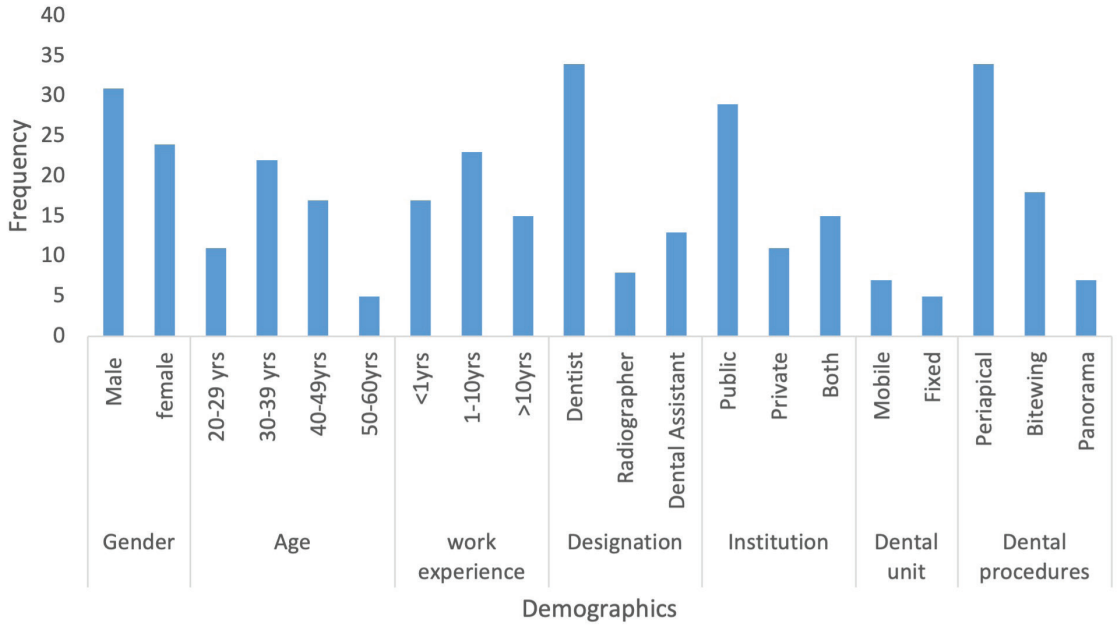


Figure 1. *Respondent's demographics*

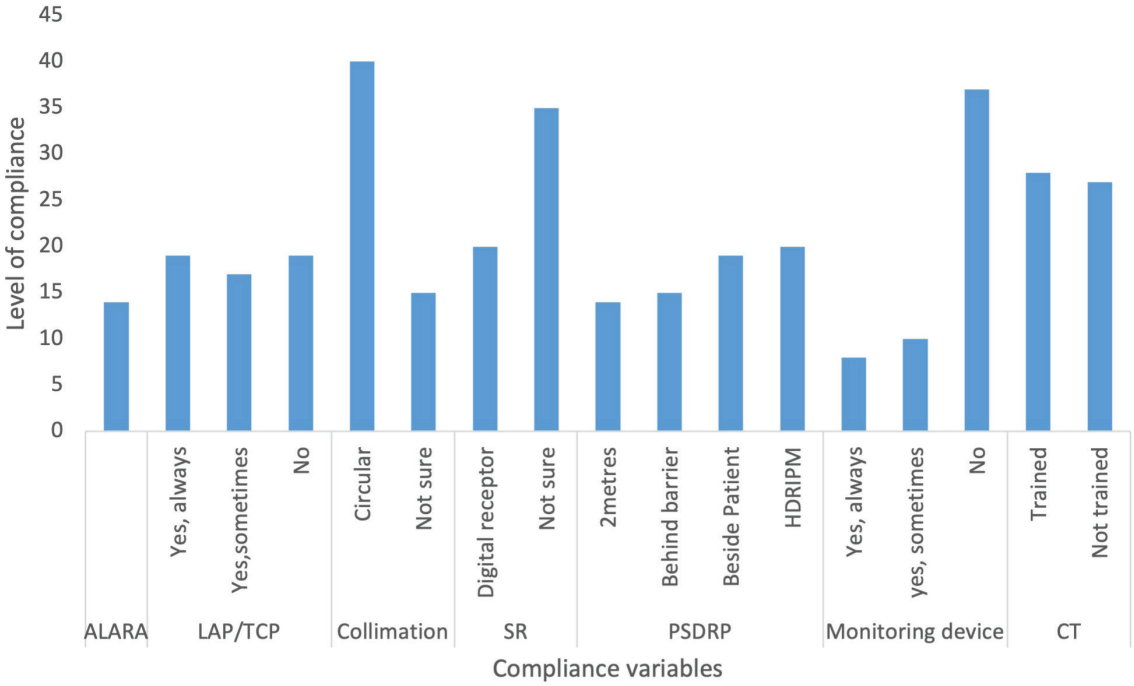


Figure 2. *Level of compliance with radiation protective measures. LAP/TCP: lead apron/thyroid collar protection, SR: speed of receptor, HDRIPM: holding of dental receptor in patient's mouth, PSDRP: position of staff during dental radiographic procedure, CE: continuing education.*

Figure 2 depicts the respondents' level of compliance with radiation protection measures. Only 14 of the 55 respondents adhered with ALARA principle giving a compliance rate of 25.5%. Majority of the respondents (N=36, 65%) used lead apron/thyroid collar for patients, while a quarter of the respondents (N=14, 25.5%), complied with the recommended two metres distance from the X-ray source. About 36.4% (N=20) of the respondents used digital receptors for dental imaging examinations, and 50.9% (N=28) received continuing education in the last three years. About 15.5% (N=8) always wore their monitoring device compared with the 18.2% (N=10) that wore theirs occasionally.

Table 1 displays the multi-level aggregate analysis of the respondents' responses in the eight imaging clinics. It showed that high continuing education score did not always result in a good compliance score and compliance level, rather clinics with high equipment score also had a good compliance score and compliance level. Clinics with high gender and work experience scores did not also affect the compliance

level. Dental clinic 2, had the highest number of respondents (12) and work experience score (8.88), while dental clinic 1 had the highest equipment and compliance scores. Five of the eight clinics showed poor compliance level, while two clinics showed good level of compliance with radiation protection measures.

Table 2 depicts the Spearman's rank order association between equipment availability and compliance was statistically significant (p=0.001), while it was not significant for continuing education, gender and work experience.

Table 3 showed that dental clinics 1 and 8 had the six of the eight items on the audit checklist and therefore scored 3-points (high equipment availability), while dental clinic 7 with four items on the checklist scored 2-points (fair equipment availability). Dental clinics 2, 3, 4, 5 and 6 had between 1-3 items on the checklist (1-point) and graded as having poor equipment availability. None of the clinics surveyed had rectangular collimator on their dental X-ray machines.

Table 4 depicts that the Cohen's Kappa analysis of agreement between the self-audit and the physical audits was perfect for warning signs/symbols and quality control procedures compared with substantial agreement shown by the other equipment. Digital receptor showed poor agreement.

Table 1. Multi-level aggregate analysis of the respondents' demographics, continuing education, equipment availability score and compliance score in the dental clinics.

Dental clinic	Number of respondents	Gender (female)	Work experience	Continuing education	Equipment score	Compliance score	Compliance level
1	6	33.3	5.17	66.7	79.17	6.33	Good
2	12	41.7	8.88	58.3	20.04	2.08	Poor
3	5	60.0	3.20	20.0	12.50	1.00	Poor
4	8	37.5	7.63	50.0	4.49	0.38	Poor
5	6	33.3	5.17	33.3	4.17	0.33	Poor
6	6	33.3	5.17	66.7	16.69	1.33	Poor
7	4	50.0	2.75	50.0	50.00	4.00	Fair
8	8	37.5	7.63	50.0	68.75	6.05	Good

Table 2. Spearman's rank order association between equipment availabilty, continuing education, gender, work experience and compliance.

Variables	Compliance	
	Spearman's rho	*p* value
Equipment availability	1.000	0.001
Continuing training	0.565	0.145
Gender	0.012	0.977
Work experience	0.012	0.977

Table 3. Physical audit of protective equipment and services in dental imaging clinics.

Dental clinic	Designed room	protective barrier	Rectangular collimator	Monitoring device	Warning signs	Digital receptor	Quality control for equipment	Digital receptor holder
1	A	A	NA	A	A	A	NA	A
2	NA	NA	NA	NA	NA	A	NA	A
3	NA	NA	NA	NA	NA	A	A	A
4	NA	NA	NA	NA	NA	A	NA	NA
5	A	NA	NA	NA	NA	A	NA	NA
6	A	NA	NA	NA	NA	A	NA	NA
7	A	A	NA	A	NA	NA	NA	A
8	A	A	NA	A	A	A	A	NA

Note: *A: availability, NA: not availability.*

Table 4. Summary of Cohen's Kappa results.

Equipment and services	Agreement (%)	Cohen's Kappa (K)	*p* value	Strength of agreement
Designated area for dental imaging	88	0.710	0.035	Substantial
Presence of protective barrier	88	0.750	0.028	Substantial
Presence of warning signs/symbols	100	1.000	0.005	Perfect
Digital receptor	38	0.091	0.537	Poor
Digital/film holders	88	0.750	0.028	Substantial
Monitoring device	88	0.750	0.028	Substantial
Quality control for dental units	100	1.000	0.005	Perfect

Discussion

Our study on the assessment of compliance with radiation protection measures in Calabar and Uyo cities revealed that many dental professionals could not identify some dental imaging equipment, such as collimators and dental receptors.

This study showed a low level of compliance with radiation protection measures among respondents (N=14, 25.5%), which could be attributed to the failure to adhere with the variables on the checklist for compliance. The low level of adherence with ALARA principle, from our survey was majorly due to the exposure settings of most of the dental X-ray units, which only allowed adjustable exposure time. The kilovoltage (kV) and milliamperes (mA) were fixed, restricting the flexibility to adjust to the required contrast and density. Thus, making the respondents to depend on the exposure time to compensate for varying thickness[16] which could result in motion blur, and implications, being increased radiation dose to the patient and dental professionals from repeat examination. Thyroid collars and lead aprons are now discouraged in imaging modality.[9] However, about 65.5% of the respondents still carry out this outdated practice. This practice is consistent with the study by Lawani *et al,*[10] indicating a gap in knowledge of the current standards in radiation protection, as lead aprons and thyroid collars are reported to introduce artefacts into dental images by obstructing the primary beam resulting in repeat examinations and do not protect against internal scatter radiation.[9,17] And the implication being more radiation dose to the patients and dental staff. While 36.4% of the respondents used digital receptors, the remaining 63.6% were not sure of the dental receptor, revealing a failure to identify the type of dental receptor used rather than a lack of equipment. This failure is confirmed by the physical audit report, that digital receptors were present in seven of the eight dental clinics but only 36.4% of the respondents reported their use, suggesting the need for continuing education for these respondents. Again, dental professionals are adviced to refrain from holding dental receptors during dental imaging procedures, as this could result in hand lesions.[18]

However, 36.4% of the respondents still hold dental receptors in the patient's mouth during the dental imaging procedure. This practice is consistent with studies by Samejo *et al.*,[5] and Lawani *et al.*,[10] with the implication being an increased likelihood of deterministic effects on the fingers.[5,18] The lack of dental receptor holders, fear of non-cooperation by children and claims of the holders making patients uncomfortable, were given as reasons for this practice by the respondents. Another 34.5% preferred to stand beside the patient during the dental examination, citing the need to reassure non-cooperative children and anxious patients as reasons, ignoring the scatter radiation and claiming that they are not in the path of the primary beam. The implication of this practice is increased exposure to the dental staff, because the closer the staff is to the patient, the more the radiation exposure.[7] Radiation monitoring device ensures that dose limits are not exceeded. About 67.3% of the respondents did not use monitoring devices during dental imaging procedures because the devices were not provided by their dental clinics, implying that dose limits are not maintained for the staff and facility.

Continuing education or training through workshops, conferences, and seminars has been recommended for dental professionals to stay updated on current standards in radiation protection.[19] Although 50.9% of the respondents had received further training on radiation protection, the level of compliance remained low, suggesting that continuing education may not be the reason for this level of adherence. Spearman's rank order correlation between continuing education and compliance was not significant, even with the knowledge gaps identified with respect to certain radiation protection measures, indicating that education did not affect compliance in this study. The reason continuing education did not affect compliance, as revealed in this study, may be due to the lack of protective equipment in most of the clinics. The availability of protective equipment allows for the practical application of knowledge gained from continuing education. When the equipment is absent, it becomes impossible to comply with the practice, meaning that the continuing education of the dental

professionals working in a poorly equipped clinic will not improve compliance.

This is evidenced in the correlation between equipment availability and compliance, which indicated a significant association. And further implies that dental professionals working in dental clinics with high equipment availability are more likely to demonstrate good compliance with radiation protection measures. This association is confirmed by the multi-level aggregate analysis in which clinics with the high equipment availability score also had good compliance score and compliance level. The Cohen's Kappa agreement between self-and physical audit showed that while self-audit was reliable in identifying structures and services, it was less reliable in identifying equipment like digital receptors. A perfect agreement depicted easy identification without misinterpretation, while substantial agreement indicated good identification but with minor uncertainties. Poor agreement indicated a high level of identification uncertainty due to the respondents' lack of awareness, showing a knowledge gap and the need for continuous education. The absence of rectangular collimators reported to reduce radiation dose to patient and dental staff by 60% [9] in all the dental clinics, the absence of protective barrier, warning signs/ symbols and monitoring device in many dental clinics surveyed affected compliance, suggesting that the respondents may be unnecessarily exposed to scatter radiation.[20]

Again, the lack of quality control measures in most dental clinics could result in undetected equipment malfunctions, leading to increased and unnecessary exposure from inconsistent radiation output.[8] Our survey revealed that the lack of enforcement of safety regulations and the high cost of radiation protection equipment and services were reasons given for the low equipment availability in the surveyed dental clinics. Gender and work experience of the respondents did not also affect compliance.

The overall implication of this poor level of compliance due to the lack of radiation protective equipment and the knowledge gaps identified, is the potential increase in ionizing radiation dose to the patients and dental professionals. This study also had its share of limitations, being the small sample size from eight dental facilities. However, these facilities have the highest number of dental imaging examinations in the two states. In addition, the survey did not carry out inter-rater reliability test as multiple authors did not audit the same dental centre. The discrimination of the respondents into professional category and dental clinics into public and private centres was not also explored.

Conclusion

This study showed that though, continuing education is essential, equipment availability significantly influenced compliance with radiation protection measures. Continuing education is recommended to address the knowledge gaps identified and the routine inspection of facilities by Regulatory Authority to enforce adherence with the minimum requirements for setting up dental imaging clinics in-order to ensure a safe working environment.

Ethics Approval

Ethical approval for this study was obtained from the Ethical Research Committee of the Department of Radiography and Radiological science, Faculty of Allied Medical Sciences of the University of Calabar with Reference number UC/ECRA/25/011.

Funding

This research did not receive any grant from funding agencies in the public, commercial or non-profit sectors.

Conflict of interest

The authors have no conflicts of interest to declared.

CRediT authorship contribution statement

Emmanuel O. Esien-umo: conceived the idea of 'Assessing the impact of continuing education and equipment availability on compliance with radiation protection measures in dental radiology in Calabar and Uyo metropolis', designed the study; generation, processing and analysis of the data, writing: final draft; **Valentine C. Ikamaise**: designed the study; generation, processing and analysis of the data, writing; final draft; **Blessing S. Ibe**: writing: first draft, designed the study, generation, processing and analysis of the data; **Akwa E. Erim**: writing: first draft and final draft; **Christopher A. Ishiekwen**: writing: first draft; **Kingsley C. Omeke**: conceived the idea of 'Assessing the impact of continuing education and equipment availability on compliance with radiation protection measures in dental radiology in Calabar and Uyo metropolis', designed the study; generation, processing and analysis of the data, writing: first draft; **Ndubuisi O. Chiaghanam**: designed the study; generation, processing and analysis of the data, writing: final draft.

Acknowledgements

Not applicable.

References

[1] Crane GD, Abbott PV. Radiation shielding in dentistry: an update. Aust Dent J. 2016; 61(3): 277-81. doi.10.1111/adj.12389

[2] Han MA, Kim JH. Diagnostic X-ray exposure and thyroid cancer risk: systematic review and meta-analysis. Thyroid. 2018; 28(2): 220-8. doi.org/ 10.1089/thy.2017.0159.

[3] Hwang SY, Choi ES, Kim YS, et al. Health effects from exposure to dental diagnostic X-ray. Environ Health Toxicol. 2018; 33(4): e2018017. doi.org/

10.5620/eht.e2018017.

[4] Memon A, Rogers I, Paudyal P, et al. Dental X-rays and the risk of thyroid cancer and meningioma: a systematic review and meta-analysis of current epidemiological evidence. Thyroid. 2019; 29(11): 1572-93. doi.org/10.1089/thy.2019.0105.

[5] Samejo I, Kumar B, Musharraf H, Ahmed J, Memon L, Bhatti R. Assessment of knowledge and perspective toward dental radiography among dental practitioners of Sindh Province of Pakistan. J Evolution Med Dent Sci. 2021; 10(29): 2186-92. doi.10.14260/jemds/2021/447.

[6] Elmorabit N, Obtel M, Azougagh M, Ennibi O Radiation protection knowledge and practices among Moroccan dentists: A cross-section study. Radiat med & Prot. 2024; 5(2): 131-8. doi.10.1016/j.radmp.2024.03.001.

[7] National Council for Radiation Protection & Measurements Radiation protection in dentistry 2003; Report No. 145 Bethesda, Md.

[8] International Atomic Energy Agency Radiation. Protection in Dental Radiology. 2022; Safety Report Series No. 108 Vienna. www.iaea.org/publications.

[9] Benavides E, Krecioch JR, Connolly RT, Allareddy T, et al. Optimizing radiation safety in dentistry. J Amer Dent Ass. 2024; 155(4): 280-93. doi.10.1016/j.adaj.2023.12.002

[10] Lawani U, Edetanlen E, Ogbebor I. A study on Awareness, knowledge and practice of exposure protection protocols by dentists in a Nigerian Teaching Hospital. J Global Oral Health 2023; 6(2): 101-6. doi.10.25259/JGOH_17_2023

[11] Sarman I, Che Hassan DH. Factors affecting radiographers' compliance with radiation protection on all areas of hospital settings worldwide-A meta-analysis. Int J Res Sc & Tech. 2016; 3(4): 433-38.

[12] Torresan TT, Rodrigues IC, Poletto MC, et al. Radioprotection in dentistry: knowledge and practices. Res Sov Dev. 2021; 10(14): e5831014 22429. doi.org/10.33448/rsd- v10i14.22429.

[13] Hussein SM, Abdelsalam N, Hashem N, Ibrahim BA. Knowledge, attitude and practice of radiation safety among dentists in Ismailia City, Egypt. J High Inst Pub Health. 2024; 54(1): 23-31. doi.10.21608/jhiph.2024.373484.

[14] Enabuelele J, Igbinedion B. An assessment of dental students' knowledge of radiation protection and practice. J Edu and Ethics in Dentistry. 2013; 3(2): 54-9. doi.10.4103/0974-7761.136044.

[15] Awosan KJ., Ibrahim MTO, Saidu SA, Ma'aji SM, et al. Knowledge of radiation hazards, radiation protection practices and clinical profile of health workers in a teaching hospital in Northern Nigeria. J Clin & Diag Res. 2016; 10(8): LC07-LC12. doi.10.7860/JCDR/2016/20398.8394.

[16] Praveen BN, Shubhasini AR, Bhanushree R, Sumsum PS, Sushma CN. Radiation in dental practice: Awareness, protection and recommendations. J Contem Dent Prac. 2013; 14(1):143-8. doi.10.5005/jp-journals-10024-1289

[17] Pauwels R. Cone beam CT for dental and maxillofacial imaging: dose matters. Radiat Prot Dosimetry. 2015; 165(1-4): 156-61. doi.org/10.1093/rpd/ncv057.

[18] Agrawal B, Dosi T, Hazari A, Maheshwari C, Rajput R, Yadav N. Evaluation of radiation awareness amongst general dental practitioners of Western Rajasthan in India. J Int Oral Health. 2015; 2(12): 51-5. https://www.ispcd.org/userfiles/rishabh/V7I12/V7I12A11.pdf.

[19] Alanazi AH, Alamoudi HI, Alqahtani SIS, Alshahrami FA, et al. Attitude and awareness of general dental practitioners towards radiation hazards and safety. J Pop Ther & Clin Pharm (JPTC) 2022; 29(4): 998-1002. doi.https://doi.org/10.53555/jptcp.v29i04.5639.

[20] Sudarshan R, RajVikram N, Rajakumar P, Ganesh R. Radiation safety practice in dentistry. Afr J Biomed Res. 2024; 27(3): 619-24. doi.https://doi.org/10.53555/AJBR.v27i3S.2072.

Journal of Associated Medical Sciences 2026; 59(2): 91-96

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Blood donation and blood components utilization in Suratthani Hospital

Weerasak Jornpakdee[1], Buchita Swatwan[2], Suwannee Jitueakul[2], Rungkarn Sangkaruk[1*]

[1]Department of Medical Technology, Faculty of Medicine, Western University, Pathum Thani Province, Thailand.
[2]Suratthani Hospital, Surat Thani Province, Thailand.

ARTICLE INFO

Article history:
Received 1 December 2025
Accepted as revised 26 January 2026
Available online 28 January 2026

Keywords:
Blood donation, blood utilization, crossmatch-to-transfusion ratio.

ABSTRACT

Background: Blood and its components are essential, irreplaceable biological resources required for a wide range of medical interventions, including surgery, trauma care, and chronic disease management. Despite ongoing improvements in blood services, many hospitals continue to experience challenges related to fluctuating donation rates, shortages of specialized components, and inefficiencies in blood utilization such as excessive crossmatching without transfusion.

Objectives: This study aimed to evaluate the efficiency of blood-donation management and the adequacy of blood-component supply relative to clinical demand at Suratthani Hospital. Additionally, it assessed the appropriateness of blood utilization using the crossmatch-to-transfusion (C/T) ratio as a key performance indicator.

Materials and methods: A retrospective descriptive study was conducted using data collected at Suratthani Hospital from January 1, 2020, to December 31, 2024. Information on blood donation types, blood-component preparation, utilization patterns, and C/T ratios was analyzed to assess supply-demand dynamics and rational use of blood products.

Results: On-site donation was the primary source of blood supply (49.9-65.5%), followed by off-site donation (32.4-47.7%) and single-donor platelet (SDP) donations (2.1-3.2%). Both blood donations and component utilization increased notably in 2024. Packed red cells (PRC) were the most frequently used component (37.6-44.7%) and constituted most transfusion requests (38.0-43.9%). Over five years, platelet concentrate (PC), fresh frozen plasma (FFP), and cryoprecipitate were consistently adequate, whereas leukocyte-depleted PRC (LDPRC), leukocyte-poor PRC (LPRC), and SDP remained insufficient, requiring additional units from the Thai Red Cross blood center. A total of 18,348 units were crossmatched, with annual C/T ratios ranging from 1.5 to 1.8 and showing no significant variation (p=0.121), indicating appropriate utilization..

Conclusion: Blood bank management at Suratthani Hospital demonstrated overall effectiveness, with adequate supply of PRC, FFP, and cryoprecipitate and appropriate utilization reflected by acceptable C/T ratios. However, persistent shortages of LDPRC, LPRC, and SDP highlight critical gaps in meeting increasing clinical needs. Strengthening donor recruitment strategies, expanding apheresis capacity, and enhancing regional collaboration are essential to ensuring a sustainable blood supply for southern Thailand.

*** Corresponding contributor.**
Author's Address: *Department of Medical Technology, Faculty of Medicine, Western University, Pathum Thani Province, Thailand.*
E-mail address: rungkarn.sa@western.ac.th
doi: 10.12982/JAMS.2026.040
E-ISSN: 2539-6056

Introduction

Blood donation is vital to public health because blood and its components are irreplaceable and cannot be artificially manufactured. They are essential for surgeries, trauma care, obstetrics, and chronic disease

management.[1] Therefore, maintaining an adequate and safe blood supply depends entirely on voluntary donations from healthy individuals. In Thailand, blood components like red blood cells, platelets, and plasma are produced from donated blood by the National Blood Centre (NBC) of the Thai Red Cross Society (TRCS), regional blood centers, and hospital blood bank service units. Efficient management of donated blood, blood component preparation, and utilization is essential for ensuring the availability of blood products for clinical use and for supporting effective hospital-based medical services, allowing for targeted transfusions based on patient needs.[2]

After collection, donated blood is processed into various components, including packed red cells, platelets, and plasma, to meet specific clinical indications and optimize resource utilization. Proper component preparation and inventory management can enhance the efficiency of blood usage, minimize wastage due to product expiration, and reduce the overall cost of blood bank operations. However, many hospitals continue to face challenges such as blood shortages during high-demand periods and excessive blood crossmatching without subsequent transfusion, both of which can adversely affect blood supply management and overall service efficiency.[3]

Effective blood bank management requires a strategic balance between maintaining sufficient inventory for clinical needs and minimizing wastage. This is complicated by the limited shelf life of products, the necessity of immunological compatibility, and the high rate of non-transfused requests.

In current clinical practice across many institutions, including Suratthani Hospital, multiple units of blood are frequently crossmatched as a precautionary reserve for surgical or high-risk patients. However, a significant number of these patients ultimately do not require the total volume of blood requested. This practice leads to excessive blood reservation, which escalates operational costs, specifically in terms of reagents, laboratory supplies, and staff workload. Furthermore, blood units reserved for a specific patient become unavailable for emergency use by others, creating an artificial shortage and increasing the complexity of inventory management. To assess the appropriateness of blood utilization, one of the most widely accepted quality indicators is the Crossmatch-to-Transfusion Ratio (C/T ratio), which represents the ratio of crossmatched blood units to the number of units actually transfused to patients.[4] Continuous monitoring and evaluation of this ratio provide valuable insights into clinical transfusion practices and support the development of strategies to promote the rational use of blood and improve blood inventory management in hospitals.

Considering these considerations, the present study aims to evaluate the efficiency of blood management from donated blood in relation to transfusion demand at Suratthani Hospital, and to assess the appropriateness and rationality of blood utilization using the Crossmatch-to-Transfusion Ratio (C/T ratio) as a key performance indicator.

Materials and methods

Study area

This study was done at Suratthani Hospital in Southern Thailand. The hospital has a bed capacity of 885, offers diagnostic and treatment for approximately 720,000 to 800,000 patients per year, and provides Surgical, Gynecology and Obstetrics, Neurosurgical, Orthopedics and Otolaryngological Surgery departments.

Study design

This retrospective study analyzed the number of blood donations, the types and quantities of blood products prepared, the types and quantities of blood products utilized, and labile blood products (LBPs) at Suratthani Hospital. All the blood is given by volunteer, nonremunerated donors. Donated whole blood is then made into transfusable components. The data was collected from blood bank departments from January 1, 2020, to December 31, 2024. The transfusable components that can be derived from donated blood including packed red blood cells (PRC), leukocyte-depleted PRC (LDPRC), leukocyte-poor PRC (LPRC), platelet concentrate (PC), leukocyte poor platelets (LPPC), single-donor platelets (SDP), fresh frozen plasma (FFP), and cryoprecipitate were calculated. The crossmatch-to-transfusion ratio (C/T ratio), which is used to evaluate the efficiency of blood utilization in transfusion services, was also calculated. It's calculated by dividing the number of blood units crossmatched which are prepared for transfusion by the number of units transfused. According to international standards, an optimal C/T ratio should be ≤2, reflecting efficient and rational blood usage, while a higher ratio may suggest unnecessary crossmatching. The utilization rates of blood components were not calculated.

Statistical analysis

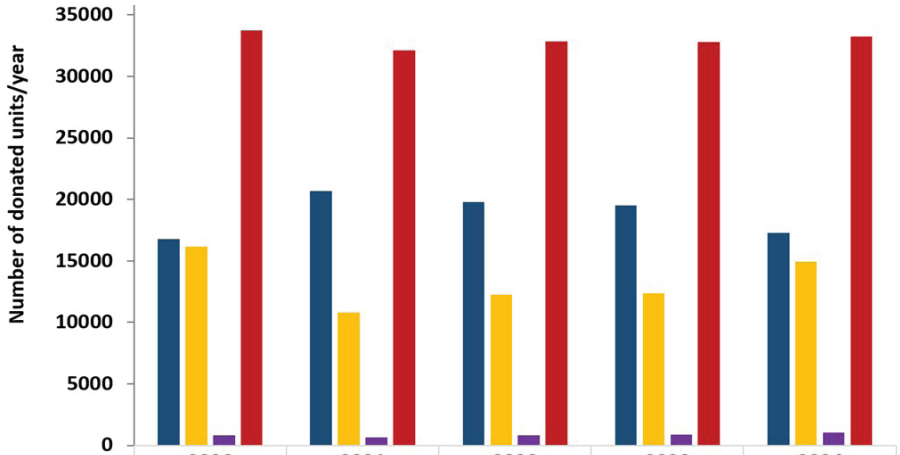
Data analysis was performed by using SPSS statistics Program version 27 from IBM, Chicago, IL, USA. All data were analyzed for normality using the Kolmogorov–Smirnov test. Due to the non-normal distribution of the data, the Kruskal-Wallis test was employed to evaluate the differences in C/T ratios among the study groups. A p-value less than 0.05 was considered statistically significant.

Results

Blood collection

The annual blood donation data from January 2020 to December 2024 revealed the number of donors as shown in Table 1. From 2020 to 2024, the numbers of blood donations per year were 33,729, 32,134, 32,867,

32,775, and 33,217 units, respectively (Figure 1). The major source of blood donation is typically on-site blood donation (49.9-65.5%), followed by off-site blood donation (32.4-47.7%) and SDP donations (2.1-3.2%), respectively.

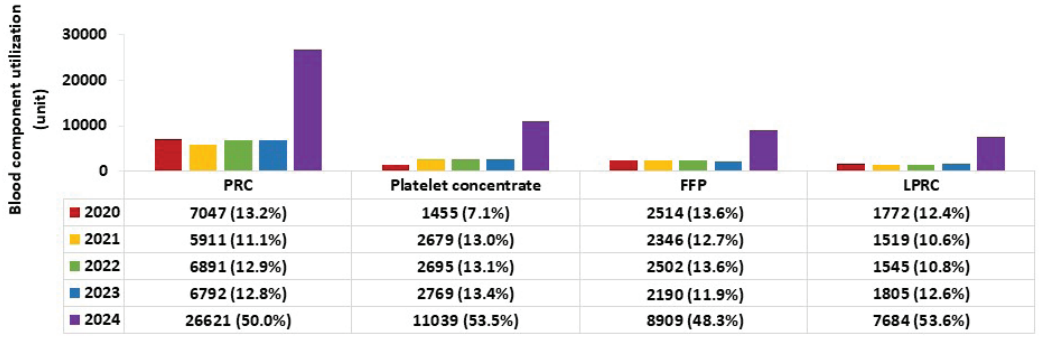


	2020	2021	2022	2023	2024
On-site blood donation	16759 (49.7%)	20682 (64.4%)	19779 (60.2%)	19501 (59.5%)	17270 (52.0%)
Off-site blood donation	16163 (47.9%)	10793 (33.6%)	12286 (37.4%)	12376 (37.8%)	14911 (44.9%)
Single donor platelets donation	807 (2.4%)	659 (2.0%)	802 (2.4%)	898 (2.7%)	1036 (3.1%)
Total number of donation/year	33729 (100%)	32134 (100%)	32867 (100%)	32775 (100%)	33217 (100%)

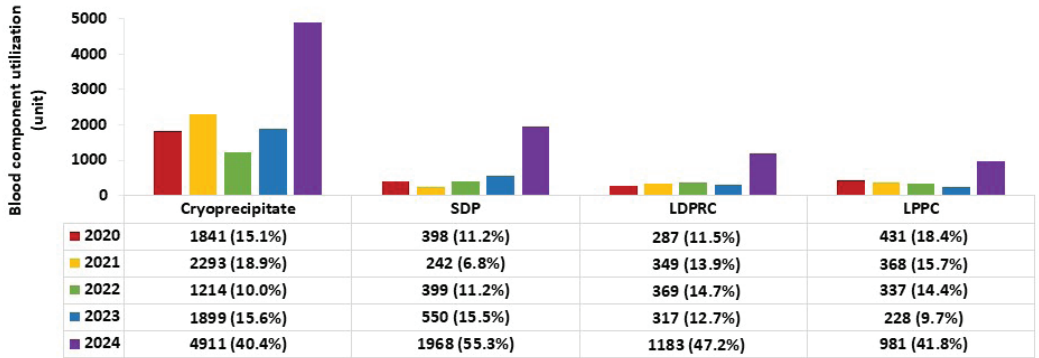
Figure 1. *Blood Collection in Suratthani Hospital 2020-2024.*

Blood utilization

Blood donation and blood component utilization obviously increased in 2024. The most utilized blood component was the PRC which accounts for 37.6-44.7%, similarly, most blood component requests were for PRC transfusions (38.0-43.9%) (Figure 2). Based on the five-year data, PC, FFP, and cryoprecipitate are near-complete fulfillment relative to requests. While LDPRC, LPRC, and SDP have been in insufficient supply for clinical use, with a consistently increasing demand observed annually. In response to the shortage, hospitals request blood components from the Thai Red Cross blood centers to treat patients (Figure 3).

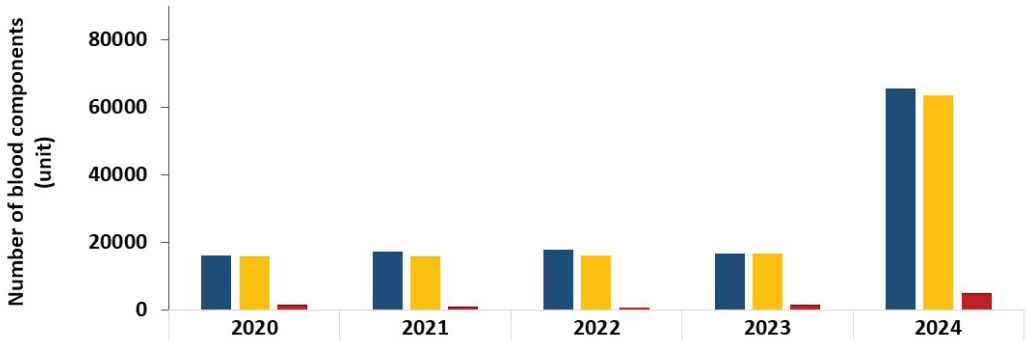


	PRC	Platelet concentrate	FFP	LPRC
2020	7047 (13.2%)	1455 (7.1%)	2514 (13.6%)	1772 (12.4%)
2021	5911 (11.1%)	2679 (13.0%)	2346 (12.7%)	1519 (10.6%)
2022	6891 (12.9%)	2695 (13.1%)	2502 (13.6%)	1545 (10.8%)
2023	6792 (12.8%)	2769 (13.4%)	2190 (11.9%)	1805 (12.6%)
2024	26621 (50.0%)	11039 (53.5%)	8909 (48.3%)	7684 (53.6%)



	Cryoprecipitate	SDP	LDPRC	LPPC
2020	1841 (15.1%)	398 (11.2%)	287 (11.5%)	431 (18.4%)
2021	2293 (18.9%)	242 (6.8%)	349 (13.9%)	368 (15.7%)
2022	1214 (10.0%)	399 (11.2%)	369 (14.7%)	337 (14.4%)
2023	1899 (15.6%)	550 (15.5%)	317 (12.7%)	228 (9.7%)
2024	4911 (40.4%)	1968 (55.3%)	1183 (47.2%)	981 (41.8%)

Figure 2. *Blood component utilization of Suratthani Hospital in 2020-2024. PRC: packed red cells, FFP: fresh frozen plasma, LPRC: leukocyte-poor PRC, SDP: single-donor platelet, LDPRC; leukocyte-depleted PRC, LPRC: leukocyte-poor PRC.*



	2020	2021	2022	2023	2024
Blood component preparation	16003 (48.3%)	17211 (51.0%)	17776 (51.6%)	16678 (48.1%)	65626 (49.0%)
Blood component utilization	15775 (47.6%)	15719 (46.6%)	15968 (46.3%)	16550 (47.7%)	63376 (47.3%)
Blood components received from Thai Red Cross blood center	1366 (4.1%)	786 (2.3%)	707 (2.1%)	1478 (4.3%)	4864 (3.6%)

Figure 3. *Number of blood component preparation and utilization of Suratthani Hospital and received blood components from Sirindhorn RC Health Station (TRC Health Station No.12 Thung Song), Nakhon Si Thammarat Province in 2020-2024.*

Appropriate utilization of blood transfusion

Over the past five years, a total of 18,348 units of RBC units were crossmatched for transfusion. From 2020 to 2024, the average of the C/T ratio were 1.6, 1.7, 1.7, 1.8, and 1.5, respectively, showing no significant with a p=0.121 (Table 1). The reported mean C/T ratio is specific to RBCs only and does not include other blood components.

Table 1. Appropriate utilization of blood transfusion in 2020-2024.

	2020	2021	2022	2023	2024
C/T ratio	1.6	1.7	1.7	1.8	1.5
Crossmatch units	3920	3686	3690	3724	3328
Transfused units	2513	2118	2158	2128	2161

Discussion

Over the past decade, Thailand has undergone continuous development of its national blood-transfusion system under the Thai Red Cross Society (TRCS), which is responsible for ensuring the safety, adequacy, and sustainability of the country's blood supply. Nevertheless, reports consistently indicate that the national volume of blood donations remains insufficient to fully meet clinical demand, leading to ongoing calls for increased donor recruitment and retention.[5] National reports show persistent challenges in balancing supply and demand due to demographic changes, seasonal fluctuations, and increasing clinical needs.[6,7] Data from Suratthani Hospital reflect similar patterns observed nationwide. On-site donation constituted the largest proportion of total blood collection (49.9-65.5%), followed by off-site donation (32.4-47.7%) and SDP donations (2.1-3.2%). This is comparable to nearby provincial hospitals such as Songklanagarind Hospital (74.5%)[8], where on-site collections also dominate. The low proportion of SDP donors, similar to other regional hospitals in southern Thailand, may present operational challenges as clinical demand for platelets and leukoreduced red-cell components continues to grow.[8,9] Specialized donations such as plateletpheresis require both donor suitability and hospital-level infrastructure, and the current shortage of these components indicates the need for improved donor engagement strategies and investment in apheresis capacity.

Blood-component utilization at Suratthani Hospital increased substantially in 2024, consistent with global and national trends showing increasing PRC utilization due to aging populations, increased surgical volumes, and higher survival rates in trauma and oncology care.[10,11] PRC remained the most commonly used component (37.6-44.7%), while PC, FFP, and cryoprecipitate supply were adequate relative to requests, demonstrating efficient stock management for these components. In contrast, LDPRC, LPRC, and SDP remained consistently insufficient, with annual demand increasing. This necessitated supplemental supply from the Sirindhorn RC Health Station (TRC Health Station No.12 Thung Song), Nakhon Si Thammarat Province. Similar shortages have been documented in other Thai provinces and emerging economies due to limited apheresis infrastructure and insufficient donor recruitment for component-specific donation.[5,12,13]

Assessment of transfusion appropriateness using the C/T ratio showed stable values ranging from 1.5 to 1.8 with no significant year-to-year variation (p=0.121). This falls within acceptable efficiency thresholds (C/T <2.0), indicating proper transfusion stewardship and

minimal unnecessary crossmatching.[4] Therefore, shortages of certain components in this setting are more attributable to insufficient supply than to inappropriate usage. The consistently acceptable C/T values further suggest that shortages of leukoreduced and platelet components are not due to inefficient usage but rather insufficient supply relative to true clinical need. Although C/T ratios are acceptable, continuous monitoring is needed to ensure minimizing unnecessary crossmatching and limiting wastage and to optimize component utilization.

This study possesses several notable strengths. Firstly, it draws upon a comprehensive five-year dataset, enabling the analysis of longitudinal patterns in both blood donation and blood-component utilization. Such an extended observation period enhances the reliability of trend interpretation and minimizes the influence of short-term fluctuations. Secondly, the study integrates both supply-side indicators, such as type and sources of donation and demand indicators, including component utilization profiles and C/T ratios. This dual-perspective approach provides a holistic assessment of blood-bank performance and facilitates a more accurate evaluation of supply and demand dynamics. Thirdly, the findings have substantial regional relevance, as Suratthani Hospital functions as a major referral center for multiple provinces in southern Thailand. Insights derived from this hospital are therefore valuable not only for internal quality improvement but also for informing broader regional policies and resource-allocation strategies. Despite these strengths, several limitations should be acknowledged. Comparative data from neighboring provincial hospitals were not available, constraining the ability to perform direct regional benchmarking. Furthermore, donor demographic information including age, sex, and repeat-donation frequency, was not collected. The absence of these data limits the ability to assess donor stability and to identify demographic groups that may require targeted recruitment. In future, detailed donor-profile analyses, and regional comparative studies, would provide stronger evidence for planning blood-supply strategies.

Overall, the findings indicate that while Suratthani Hospital maintains adequate supply for basic components such as PRC, FFP, and cryoprecipitate, its capacity to supply specialized components remains limited and increasingly mismatched with growing clinical demand, mirroring regional and national challenges.

Limitations

While this study provides insights into optimizing blood ordering and reducing wastage, it is limited by the unavailability of specific data required to calculate key performance indicators, transfusion probability (%T) and the transfusion index (Ti). The absence of these metrics may constrain the comprehensive evaluation of blood utilization efficiency and the precision of strategies aimed at minimizing blood expiration.

Conclusion

Suratthani Hospital has maintained generally effective blood-bank management over the past five years, with adequate availability of PRC, FFP, and cryoprecipitate, and reasonable indices of appropriate utilization. However, persistent shortages of specialized components, including LDPRC, LPRC, and SDP, highlight significant challenges in meeting increasing clinical demands. Strengthening donor recruitment, expanding apheresis capacity, and establishing regional collaborative networks are critical to ensuring long-term sustainability of blood supply in southern Thailand.

Ethical approval

This study has been reviewed and approved by Human Research Ethics Committee of Suratthani Hospital, Thailand (COA 013/2567).

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Weerasak Jornpakdee: formal analysis, writing: original draft; **Buchita Swatwan**: data curation, ethical approvals; **Suwannee Jitueakul**: resources and ethical approvals; **Rungkarn Sangkaruk**: conceptualization, methodology, writing: review and edit.

Acknowledgements

All facilities were supported by Western University and Suratthani Hospital.

References

[1] Gasparovic Babic S, Krsek A, Baticic L. Voluntary blood donation in modern healthcare: trends, challenges, and opportunities. Epidemiologia (Basel). 2024; 5(4): 770-84. doi: 10.3390/epidemiologia 5040052.

[2] Gammon RR, Dubey R, Gupta GK, Hinrichsen C, Jindal A, Lamba DS, et al. Patient blood management and its role in supporting blood supply. J Blood Med. 2023; 14: 595-611. doi: 10.2147/JBM.S387322.

[3] Chien JH, Yao CY, Chen HF, Ho TF. Trends in blood transfusion and causes of blood wastage: a retrospective analysis in a teaching hospital. BMC Health Serv Res. 2025; 25(1): 67-72. doi: 0.1186/ s12913-024-12170-x.

[4] Joon V, Robins RDL, Haran AH, I SK, James S. Evaluating the crossmatch-to-transfusion ratio as a tool for analyzing and optimizing blood bank resource utilization: a retrospective observational study. Cureus. 2024; 16(9): e69862. doi: 10.7759/ cureus.69862.

[5] Chaiwuttisak P, Honora S, Yue W, Chris P, Sakuldamrongpanich T, Pathomsiri S. Location of low-cost blood collection and distribution centres in Thailand. Operations Research for Health Care. 2015; 10: 14-18.
[6] Thai Red Cross Society, National Blood Centre. Annual Report 2023-2024: Blood Donation and Transfusion Services in Thailand. Bangkok: TRCS; 2024.
[7] World Health Organization. Global status report on blood safety and availability 2021. Geneva: WHO; 2022.
[8] Hasarin WYaS. Blood and blood reserve management before and after the COVID-19 pandemic: experiences from Songklanagarind Hospital. J Hematol Transfus Med. 2024; 34: 8-12.
[9] Narkpetch S, Neamkul N, Prapanwong P, Jiwbunchu T, Rattanapan Y. Recruitment and retention of voluntary blood donors during the COVID-19 pandemic. J Hematol Transfus Med. 2023; 33: 8-13.
[10] Carson JL, Guyatt G, Heddle NM, Grossman BJ, Cohn CS, Fung MK, et al. Clinical Practice Guidelines from the AABB: red blood cell transfusion thresholds and storage. JAMA. 2016; 316(19): 2025-35. doi: 10.1001/jama.2016.9185.
[11] Lertpaisankul S, Rodwihok T, Phomsila R, Tadsomboon S. Evaluation of donor recruitment integrated plan and blood component distribution effectiveness of the five lower-northern provinces of Thailand in 2018. J Hematol Transfus Med. 2020; 30: 6-11.
[12] Watanaboonyongcharoen P, Santipong S, Jaipian J, Siriaksorn J, Rojnuckarin P. Safety and efficacy of the Thai red cross society albumin replacement for therapeutic plasma exchange. J Med Assoc Thai. 2018; 101(11): 4-9.
[13] Nonsee N, Kidhen B, Chawalchitiporn S. Efficacies and cost evaluation of double dose plateletpheresis in additive solution by blood cell separator between continuous and intermittent flow system. Cancer Sci Res. 2020; 3(2): 6-10. doi: 10.33425/2639-8478.1048.

Journal of Associated Medical Sciences 2026; 59(2): 97-104

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Association of Intron 22 inversion and HLA-DRB1*15:01 allele with inhibitor development in hemophilia A: A genetic analysis of risk factors

Chitpon Seeljunya[1], Chollanot Kaset[1,2], Sirinart Chomean[1,2*]

[1]Graduate Program in Medical Technology, Faculty of Allied Health Sciences, Thammasat University, Pathum Thani, Thailand.
[2]Thammasat Research Unit in Medical Technology and Precision Medicine Innovation, Faculty of Allied Health Sciences, Thammasat University, Pathum Thani Province, Thailand.

ARTICLE INFO

Article history:
Received 2 December 2025
Accepted as revised 27 January 2026
Available online 29 January 2026

Keywords:
*Factor VIII inhibitor, genetic determinants, Intron 22 inversion, HLA-DRB1*15:01, personalized medicine.*

ABSTRACT

Background: Inhibitor development in hemophilia A is a major clinical challenge, leading to severe bleeding complications and reduced quality of life. This multifactorial process is influenced by genetic, immunological, and treatment-related factors.

Objectives: To determine the association between intron 22 inversion (Inv22) and the HLA-DRB1*15:01 allele with inhibitor development.

Materials and methods: This study investigated inhibitor development during initial and subsequent treatments, assessing the association between Inv22 and HLA-DRB1*15:01 with inhibitor formation in 40 hemophilia A patients, 40 non-hemophilia A patients, and 40 healthy controls.

Results: The study found no significant association between Inv22 or HLA-DRB1*15:01 allele and inhibitor development during initial (odds ratio (OR): 0.60, 95% CI: 0.11-3.43, p=0.697) or subsequent treatments (OR: 1.33, 95% CI: 0.30-5.93, p=1.000). Similarly, HLA-DRB1*15:01 allele showed no significant correlation (initial: OR: 0.25, 95% CI: 0.06-1.13, p=0.089; subsequent: OR: 0.82, 95% CI: 0.24-2.84, p=1.000). However, patients with both genetic factors had a significantly higher risk of developing inhibitors (OR: 7.90, 95% CI: 1.88-33.06, p=0.007).

Conclusion: This study found no significant association between Inv22 or the HLA-DRB1*15:01 allele and inhibitor development in hemophilia A patients during initial or subsequent treatments. However, their combined presence significantly increased the risk of inhibitor formation, suggesting a potential interactive effect requiring further investigation.

Introduction

Development of Factor VIII (FVIII) inhibitors represents the most formidable complication in Hemophilia A management today. While therapeutic advancements have significantly reduced mortality, the emergence of neutralizing antibodies—affecting approximately 25-30%[1,2] of patients globally and up to 17% of severe cases in Thailand[3]—renders standard replacement therapy ineffective. This failure of immune tolerance not only escalates the risk of severe bleeding and disability but also necessitates costly bypassing agents,[1] thereby imposing a profound clinical and economic burden. The etiology of inhibitor formation is

* **Corresponding contributor.**
Author's Address: *Graduate Program in Medical Technology, Faculty of Allied Health Sciences, Thammasat University, Pathum Thani, Thailand.*
E-mail address: sirinat.c@allied.tu.ac.th
doi: 10.12982/JAMS.2026.041
E-ISSN: 2539-6056

multifactorial, driven by a complex interplay between environmental exposures and genetic predisposition.[4] Central to this pathogenesis is the immune recognition of therapeutic FVIII as a neo-antigen, a process heavily influenced by the nature of *F8* gene mutations and the host's Major Histocompatibility Complex (MHC) profile.[4,5] Intron 22 inversion (Inv22), responsible for nearly half of severe Hemophilia A cases, causes a complete disruption of FVIII synthesis. In 2019, at Ramathibodi Hospital, the prevalence of Inv22 among patients with severe hemophilia A was 35.6%.[3] However, its independent role as a predictor for inhibitor development remains controversial, with prior studies yielding conflicting associations.[4] Concurrently, Class II Human Leukocyte Antigens (HLA), specifically the HLA-DRB1*15:01 allele, are implicated in efficient peptide presentation to T-cells. While HLA-DRB1*15:01 is a globally recognized risk factor,[5] local evidence indicates it is the only allele significantly associated with inhibitors in Thai patients (19.4% vs. 5.0% in controls, p=0.021),[6] exceeding the 8.09% background frequency observed in the general Thai population.[7] whereas alleles like DRB1*14 and DRB1*07 show no such significance.[6] These inconsistencies are frequently attributed to variations in sample sizes, study designs, and the diverse genetic backgrounds of different ethnic cohorts.[5] Yet, when analyzed in isolation, HLA-DRB1*15:01 often appears as a weak risk factor, suggesting that the mechanisms driving immunogenicity are likely synergistic rather than singular.[8] The synergy likely arises because Inv22 creates a 'null' state where FVIII is perceived as entirely foreign, which HLA-DRB1*15:01 then efficiently presents to the immune system.

This discrepancy in current literature highlights a critical gap in understanding how structural genetic defects interact with immune response genes. To advance the frontier of precision medicine, this study aims to elucidate the association and potential interaction between Inv22 and the HLA-DRB1*15:01 allele in the context of inhibitor development within the Thai population. For instance, identifying such high-risk genetic profiles can guide personalized prophylaxis or the early use of non-factor therapies like emicizumab to prevent inhibitor formation.[9]

Unraveling these genetic complexities is essential for constructing robust risk stratification models, ultimately enabling clinicians to personalize therapeutic strategies and optimize outcomes for high-risk patients.

Materials and methods

Participants and sample collection

This study included 120 participants, consisting of 40 hemophilia A patients (FVIII activity (FVIII:C) <50%), 40 non-hemophilia A patients (FVIII:C ≥50%), and 40 healthy individuals as controls. Among hemophilia A patients, the severity was classified into severe (N=27, 67.5%; FVIII:C<1%), moderate (N=10, 25.0%; FVIII:C 1-5%), and mild (N=3, 7.5%; FVIII:C >5%). Participant data included gender, age, FVIII activity (initial and post-treatment), age at first treatment, post-treatment inhibitor levels (Bethesda units; BU), coagulation test (aPTT), treatment type, and 2 mL collected EDTA blood samples. All samples were collected from Maharat Nakhon Ratchasima Hospital, Thailand, during treatment and follow-up. Inhibitor titer was defined by the Bethesda unit (BU),[10] representing the amount of inhibitor that neutralizes 50% of FVIII activity after 2-hour incubation at 37 °C with normal plasma; a titer 0.6 BU was considered positive. Inclusion criteria for "Non-hemophilia A" required a confirmed FVIII:C >50%, while "Healthy individuals" showed no coagulation test abnormalities. Exclusion criteria for all groups included individuals who had not received treatment for over one year. To reduce comparison bias, control groups were randomly selected without matching.

*Genotyping of the Inv22 and HLA-DRB1*15:01 allele*

Genomic DNA was extracted from EDTA-anticoagulated whole blood using the Wizard® Genomic DNA Purification Kit (Promega, USA) according to the manufacturer's instructions. Intron 22 inversion was genotyped using the inverse shifting (IS)-PCR method, modified from Rossetti *et al.*[11] Briefly, 20 µL of genomic DNA (100 ng/µL) was digested with 20 U of BclI restriction enzyme (Thermo Fisher Scientific, USA) at 55 °C for 4 hours in a 50 µL reaction volume. The reaction was heat-inactivated at 80 °C for 30 minutes. Circularization of the digested DNA was performed using 3 U of T4 DNA ligase (Thermo Fisher Scientific, USA) in a 400 µL reaction volume at 4 °C overnight. Ligated DNA was purified using the QIAquick® PCR & Gel Cleanup Kit (QIAGEN, Germany) and eluted in 20 µL of distilled water. Amplification was performed using the IS-PCR method in a 20 µL reaction containing 4 µL of circularized DNA, 2x Green PCR Master Mix (Biotechrabbit, Germany), and 0.5 µM of each primer (ID: 5'-ACATACGGTTTAGTCACAAGT-3', IU: 5'-CCTTTCAACTCCATCTCCAT-3', 2U: 5'-ACGTGT CTTTTGGAGAAGTC-3', 3U: 5'-CTCACATTGTGTTCTTGT AGTC-3'). PCR cycling conditions consisted of 35 cycles of denaturation at 95°C for 30 seconds, annealing at 55°C for 1 minute, and extension at 72 °C for 1 minute. Results were interpreted by band size: 487-bp for wild-type and 333-bp for Inv22.

For detection of the HLA-DRB1*15:01 allele, genomic DNA was amplified using the PCR-SSP method. as previously described by Delfan *et al.*[12] PCR was carried out in a final volume of 20 µL containing 100 ng of genomic DNA, 2x Green PCR Master Mix (Biotechrabbit, Germany), and 0.5 µM of each primer. The HLA-DRB1*15:01-specific primers used were forward 5'-TCCTGTGGCAGCCTAAGAG-3' and reverse 5'-CCGCGCCTGCTCCAGGAT-3'. The human growth hormone (HGH) gene served as an internal control, with forward primer 5'-TGCCTTCC CAACCATTCCCTTA-3' and reverse primer 5'-CCACT

CACGGATTTCTGTTGTGTTTC-3'. PCR amplification was performed for 35 cycles under the following conditions: denaturation at 95°C for 30 seconds, annealing at 60°C for 45 seconds, and extension at 72 °C for 30 seconds. All PCR products were separated on a 1.5% agarose gel and visualized under UV illumination. The 197-bp and 434-bp bands represented the HLA-DRB1*15:01 allele and the HGH internal control, respectively.

Association analysis of Inv22 and HLA-DRB1*15:01 allele influence on inhibitor development in hemophilia A

The Bethesda assay was used to measure inhibitors in hemophilia A patients, with levels quantified in BU. An inhibitor was defined as a level exceeding 0.6 BU.[10] The presence of the Inv22 and HLA-DRB1*15:01 allele was analyzed in relation to inhibitor formation at two time points: the first occurrence of an inhibitor (first inhibitor) and at any point during treatment. The relationship between Inv22, HLA-DRB1*15:01 allele, and inhibitor production in hemophilia A patients was tested using Fisher's exact test or Chi-squared test, with odds ratios (OR) and 95% confidence intervals (CI) calculated to quantify the strength of these associations. The Chi-squared test was used primarily, while Fisher's exact test was selected when any expected cell count was less than 5. This study also demonstrates the relationship between genetic determinants and inhibitor development following initial treatment and at any point during treatment. A $p<0.05$ indicated a statistically significant association between genetic mutations and inhibitor production. Statistical analyses were performed using IBM SPSS Statistics version 26 (SPSS Inc., Chicago, IL, USA).

Results

Participants data

A descriptive analysis was performed on gender, aPTT test, FVIII activity (initial and post-treatment), inhibitor level (BU), and treatment type (Table 1). The study enrolled 120 subjects: 40 Hemophilia A (HA) patients, 40 non-HA patients, and 40 healthy controls. The HA cohort was exclusively male. As anticipated, HA patients exhibited profound coagulopathy, characterized by markedly prolonged aPTT (>82 s) and severe FVIII deficiency (<5.2%) compared to controls. All HA patients received FVIII concentrate therapy. Notably, 30% (12/40) of HA patients presented with inhibitors initially, with the prevalence rising to 50% (20/40) cumulatively during treatment. Inhibitor titers were clinically substantial in positive cases (mean 16.12±21.57 BU), contrasting sharply with the negligible levels observed in non-inhibitor patients.

Table 1. Descriptive analysis of gender, age, FVIII Activity (%), inhibitor levels, and treatment among hemophilia A (inhibitor and non-inhibitor during initial or subsequent treatment), non-hemophilia A, and normal control groups.

Factors	**Hemophilia A patients** (First inhibitor) (N=40)		**Hemophilia A patients** (Inhibitors at any point during treatment) (N=40)		**Non-hemophilia A** (%) (N=40)	**Normal (%)** (N=40)
	Inhibitor (%)	Non-inhibitor (%)	Inhibitor (%)	Non-inhibitor (%)		
Number (%)	12 (30)	28 (70)	20 (50)	20 (50)	40 (100)	40 (100)
Sex						
- Male (%)	12 (100)	28 (100)	20 (100)	20 (100)	24 (60)	21 (52.5)
- Female (%)	0 (0)	0 (0)	0 (0)	0 (0)	16 (40)	19 (47.5)
Age (mean±SD)	23.54±10.51	14.26±12.33	18.68±10.67	15.21±14.08	33.60±25.05	62.5±17.70
Range	6-35	1-44	6-35	1-44	3-82	22-91
aPTT test (sec) (Mean±SD)	86.1±19.9	82.5±21.9	84.5±16.7	82.7±25.3	33.9±12.6	26.1±2.5
% FVIII activity (first time visiting; Mean±SD)	4.38±11.24	4.03±7.21	3.17±8.71	5.11±8.33	194.57±80.66	Not test
% FVIII activity (post-treatment; Mean±SD)	31.09±41.70	8.01±17.94	17.19±35.10	10.96±20.61	Not test	Not test
Inhibitor level (BU; Mean±SD)	16.12±21.57	0.41±0.02	9.84±13.21	0.41±0.02	Not test	Not test
Therapy by FVIII concentrate	12 (100)	28 (100)	20 (100)	20 (100)	Not treatment	Not treatment

Genotyping of the Inv22 and HLA-DRB1*15:01 allele

The detection of the Inv22 and HLA-DRB1*15:01 allele in blood samples was performed using the primers described above. Genotyping for Inv22 and HLA-DRB1*15:01 was successfully performed in all participants. Genotypes were determined based on the specific PCR product sizes as described in the Materials and Methods section. The results of this study demonstrated that the PCR product sizes corresponded to the expected genotypes (Figure 1), with Inv22 identified in 9 out of 40 hemophilia A patients (22.5%).

For the detection of the HLA-DRB1*15:01 allele, this study utilized the HGH gene as an internal control. The presence of the HLA-DRB1*15:01 allele was indicated by the specific PCR product shown in Figure 2. Based on the results, the HLA-DRB1*15:01 allele was detected in 19 of 40 hemophilia A patients (47.5%), 17 of 40 non-hemophilia A individuals (42.5%), and 12 of 40 healthy controls (30.0%).

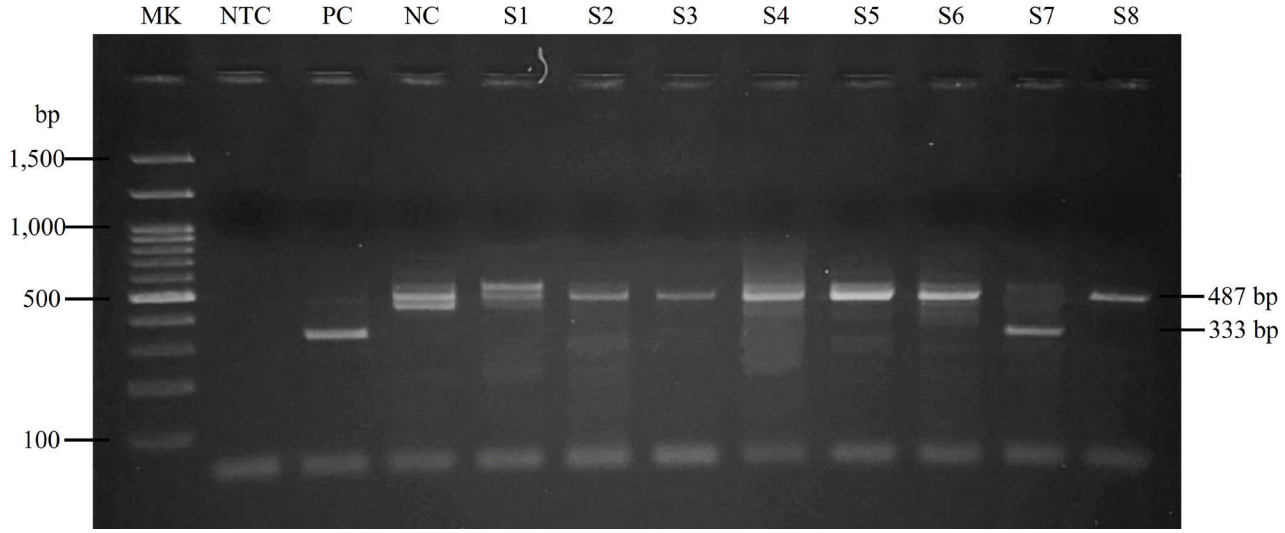


Figure 1. ***The genotyping of Inv22 gene in hemophilia A patients using the IS-PCR method.*** *MK: 100 bp DNA ladder, NTC: non-template control, PC: positive control containing the Inv22 gene, NC: DNA sample from a hemophilia A patient negative for the Inv22 gene, Lanes S1-S8: samples from hemophilia A patients 1 through 8.*

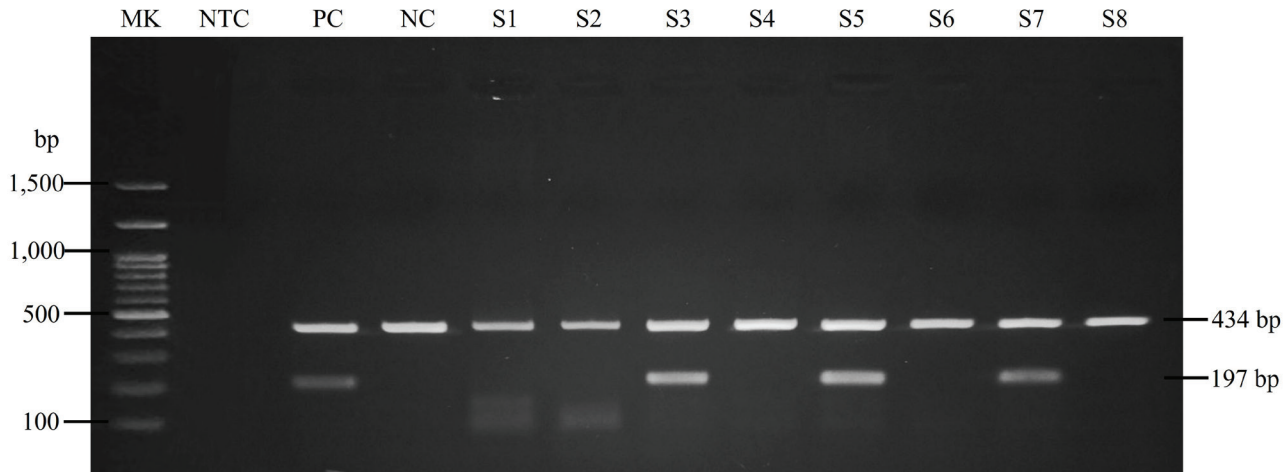


Figure 2. ***The genotyping of HLA-DRB1*15:01 allele using the PCR-SSP method in a cohort of 120 individuals.*** *MK: 100 bp DNA ladder, NTC: non-template control, PC: positive control containing the HLA-DRB1*15:01 allele, NC: DNA sample from an individual negative for the HLA-DRB1*15:01 allele, Lanes S1-S8 : samples from individuals 1 through 8.*

Association analysis of Inv22 and HLA-DRB1*15:01 allele influence on inhibitor development in hemophilia A

The genetic analysis of study groups in relation to first-inhibitor development in hemophilia A patients identified Inv22 in 7.5% (N=9) of participants, exclusively among those with hemophilia A (Table 2). It was present in 17% of inhibitor-positive and 25% of non-inhibitor patients, while absent in non-hemophilia A and normal individuals. The odds ratio for Inv22 in inhibitor-positive hemophilia A patients was 0.60 (95% CI: 0.11-3.43, p=0.697), indicating no significant association with inhibitor development.

For the HLA-DRB1*15:01 allele, 40% (N=48) of participants were positive (Table 2). It was found in 25% of inhibitor-positive and 57% of non-inhibitor hemophilia A patients, compared to 43% of non-hemophilia A and 30% of normal individuals. The odds ratio for HLA-DRB1*15:01 allele positivity in inhibitor-positive hemophilia A patients was 0.25 (95% CI: 0.06-1.13, p=0.089). No significant association was observed between this genetic marker and inhibitor formation (p>0.05). This result indicates that the presence of this genotype did not significantly influence the risk of inhibitor development in the studied population.

Table 2. Genetic analysis of the Inv22 and the HLA-DRB1*15:01 allele and their association with initial inhibitor development in hemophilia A patient's post-treatment.

Allele	**Hemophilia A (N=40)**		**Non-Hemophilia A** (N=40)	**Normal** (N=40)	**Total** (N=120)
	Inhibitor (N=12)	Non-inhibitor (N=28)			
Inv22 inversion positive (%)	2 (17)	7 (25)	0	0	9 (7.5)
Inv22 inversion negative (%)	10 (83)	21 (75)	40 (100)	40 (100)	111 (92.5)
Odd ratio (95% CI)	0.60 (0.11-3.43)		19.29 (0.86-432.99)	19.29 (0.86-432.99)	
p value	0.697		0.062	0.062	
HLA-DRB1*15:01 positive (%)	3 (25)	16 (57)	17 (43)	12 (30)	48 (40)
HLA DRB1*15:01 negative (%)	9 (75)	12 (43)	23 (57)	28 (70)	72 (60)
Odd ratio (95% CI)	0.25 (0.06-1.13)		0.45 (0.11-1.92)	0.78 (0.18-3.39)	
p value	0.089		0.330	1.000	

Genetic analysis based on inhibitor development at any treatment stage showed that Inv22 occurred in 7.5% (N=9) of participants, exclusively among hemophilia A patients (Table 3). Within this group, 25% of inhibitor-positive and 20% of non-inhibitor patients carried the inversion, whereas all non-hemophilia A and normal individuals were negative. The odds ratio for Inv22 in inhibitor-positive hemophilia A patients was 1.33 (95% CI: 0.30-5.93, p=1.00), indicating no significant association with inhibitor development. However, a significant difference was observed between hemophilia A and non-hemophilia A groups (p=0.026), suggesting a strong link between Inv22 and hemophilia A. For the HLA-DRB1*15:01 allele, 40% (N=48) of participants tested positive. Among hemophilia A patients, 45% of inhibitor-positive and 50% of non-inhibitor individuals carried this allele, compared to 43% of non-hemophilia A and 30% of normal individuals. The odds ratio for HLA-DRB1*15:01 allele positivity in inhibitor-positive hemophilia A patients was 0.82 (95% CI: 0.24-2.84, p=1.00), showing no significant association with inhibitor development. Similarly, no statistically significant differences were found between non-hemophilia A and normal groups.

Table 3. Genetic analysis of the Inv22 and the HLA-DRB1*15:01 allele and their association with the development of inhibitors in hemophilia A patient at any point during treatment.

Allele	Hemophilia A (N=40)		Non-Hemophilia A (N=40)	Normal (N=40)	Total (N=120)
	Inhibitor (N=20)	Non-inhibitor (N=20)			
Inv22 inversion positive (%)	5 (25)	4 (20)	0	0	9 (7.5)
Inv22 inversion negative (%)	15 (75)	16 (80)	40 (100)	40 (100)	111 (92.5)
Odd ratio (95% CI)	1.33 (0.30-5.93)		28.74 (1.50-551.18)	28.74 (1.50-551.18)	
p value	1.000		0.026*	0.026*	
HLA-DRB1*15:01 positive (%)	9 (45)	10 (50)	17 (43)	12 (30)	48 (40)
HLA DRB1*15:01 negative (%)	11 (55)	10 (50)	23 (57)	28 (70)	72 (60)
Odd ratio (95% CI)	0.82 (0.24-2.84)		1.11 (0.38-3.27)	1.91 (0.63-5.80)	
p value	1.000		1.000	0.268	

*Statistically significant

Genetic analysis of Inv22 and HLA-DRB1*15:01 allele in the study population showed an odds ratio (OR) of 3.34 (95% CI: 0.59-18.88) for the first inhibitor occurrence, though this association was not statistically significant (p=0.407) (Table 4). However, for inhibitors developed at any point during treatment, both intron 22 inversion and HLA-DRB1*15:01 allele were significantly associated with an increased risk, with an OR of 7.90 (95% CI: 1.88-33.06, p=0.007). These findings suggest that these genetic factors may influence inhibitor formation over the course of treatment rather than at the initial occurrence.

Table 4. Association of Inv22 and HLA-DRB1*15:01 allele with first-time and any-time inhibitor development in hemophilia A patients.

Inhibitor	Allele	OR (95%CI)	p value
First inhibitor	Inv22+HLA-DRB1*15:01 allele	3.34 (0.59-18.88)	0.407
Inhibitor at any-time inhibitor development	Inv22+HLA-DRB1*15:01 allele	7.90 (1.88-33.06)	0.007*

* Statistically significant

Discussion

Genetic factors play a significant role in inhibitor development among hemophilia A patients, particularly the type of mutation in the FVIII gene and variations in HLA class II alleles.[4] Severe, null mutations such as Inv22, large deletions, and nonsense mutations can lead to a complete absence of FVIII protein, heightening the likelihood that administered FVIII will be recognized as foreign and elicit an immune response. Conversely, milder mutations (e.g., missense mutations) are less immunogenic, resulting in a lower incidence of inhibitor formation.[4] In parallel, certain HLA class II alleles, including HLA-DRB1*15:01 allele, have been implicated in the antigen presentation process that underlies inhibitor development.[5] These alleles may predispose individuals to a more robust T-cell–mediated immune response against FVIII, thus increasing inhibitor risk.[13] This study analyzed whether the Inv22 and HLA-DRB1*15:01 allele influence inhibitor development in hemophilia A patients. As shown in Table 1, Factor VIII activity was higher in hemophilia A patients with inhibitors than in those without, regardless of whether inhibitors developed after initial treatment or at any point later. This reflects the treatment strategy of administering high-dose factor VIII concentrate to induce immune tolerance and suppress inhibitor formation.[13] This study found the frequency of the Inv22 to be 22.5% (9 of 40 hemophilia A patients), which is slightly lower than the 30-50% reported in previous studies.[14] Many other studies in Thailand found that the allele HLA-DRB1*15:01 is approximately 8%, within the general range previously reported to be 8-20%.[15,16]

This study found no statistically significant association between the Inv22 and HLA-DRB1*15:01 allele with inhibitor development in hemophilia A, either for the first inhibitor occurrence (Table 2) or at any point during treatment (Table 3), when compared to non-inhibitor hemophilia A patients, non-hemophilia A group, and normal controls. The results are consistent with previous studies, which reported Inv22 in 45.5% of hemophilia A patients with inhibitors (5/11), though this was not statistically significant when compared to

patients without inhibitors (p=0.532).[14] These findings are consistent with previous studies indicating that the Inv22 gene is not significantly associated with inhibitor development.[9,17] HLA-DRB1*15:01, DQB1*06:02, and DQA1*01:02 were frequently observed in hemophilia A patients with inhibitors. Notably, one study found a significant association for HLA-DQA1*01:02 (OR=2.7, 95% CI=1.2-4.9).[18] Nathalang O. and coworker reported an increased likelihood of HLA-DRB1*15 in hemophilia patients with inhibitors compared to healthy controls (OR=2.74, 95% CI=1.16-6.47, p=0.021), but no statistical significance when comparing inhibitor-positive and inhibitor-negative hemophilia patients (p=0.198).[6] Consequently, our study focused on HLA-DRB1*15:01 and Inv22 as primary risk factors in Thais; other alleles showed no significant association, likely due to ethnic genetic diversity. However, when examining the combined effect of the Inv22 and the HLA-DRB1*15:01 allele on inhibitor development, their co-occurrence was found to influence inhibitor formation at any point during treatment (Table 4). These findings regarding the influence of the Inv22 and HLA-DRB1*15:01 allele remain inconsistent, suggesting that while this allele may contribute to immunogenicity, it is not a sole determinant. Our results clarify that these inconsistencies likely stem from the complex synergy between genetic markers rather than isolated effects. By demonstrating that their co-presence significantly increases inhibitor risk throughout treatment (OR=7.90), this study highlights that a multi-genetic approach provides a more robust risk assessment than individual gene analysis alone. Taken together, these insights highlight the multifactorial nature of inhibitor formation, whereby both FVIII gene mutation severity and immune response genes intersect with additional clinical and environmental factors, such as treatment regimen and patient immune status. A better understanding of these genetic determinants will aid in identifying high-risk patients and guiding personalized therapeutic strategies, including prophylactic treatment and immune tolerance induction.[9]

Limitations

This study is limited by a relatively small sample size and the use of univariate analysis, which may not account for all clinical confounders. Additionally, we focused on a limited number of genetic markers (Inv22 and HLA-DRB1*15:01). Future studies with larger cohorts and broader genetic panels are necessary to validate these findings.

Conclusion

This study examined the association between intron 22 inversion (Inv22) and the HLA-DRB1*15:01 allele with inhibitor development in hemophilia A patients. The findings suggest that while Inv22 and HLA-DRB1*15:01 allele alone may not directly influence inhibitor formation, their combined presence may contribute to increased risk over time. Further studies with larger cohorts are warranted to validate these findings and better understand the genetic mechanisms underlying inhibitor development in hemophilia A.

Ethical approval

This cross-sectional study was approved by the Human Research Ethics Committee of Thammasat University (Science) according to COA (266/2565) and Maharat Nakhon Ratchasima Hospital (No.129/2022).

Funding

This study was supported by the Thammasat Research Unit in Medical Technology and Precision Medicine Innovation, Faculty of Allied Health Sciences, Thammasat University.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Chitpon Seeljunya: data collection, statistical analysis, writing: original draft, **Chollanot Kaset:** methodology, statistical analysis, data interpretation, **Sirinart Chomean**: conceptualization, methodology, investigation, data interpretation, writing: review and editing.

Acknowledgements

This research gratefully acknowledges the Department of Pediatrics, King Chulalongkorn Memorial Hospital, the Thai Red Cross Society, and the Laboratory for Pharmacogenomics, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, for providing control samples.

References

[1] Witmer C, Young G. Factor VIII inhibitors in hemophilia A: etiology and treatment. Ther Adv Hematol. 2013; 4(1): 59-72. doi: 10.1177/2040620712464379.

[2] El-Sheikh A, El-Ghamrawy M, Ibrahim M. Evaluation of factor VIII inhibitors among children with hemophilia A: a single-center study. Zagazig Univ Med J. 2024; 30(8.1): 4104-9. doi: 10.21608/zumj.2023.210924.2821.

[3] Chuansumrit A, Sosothikul D, Rungrungtham T, Sasanakul P, Tapaneeyakorn S, On-Takrai J, et al. The success of the national universal health coverage for hemophilia care in Thailand. Haemophilia. 2021; 27(1): 69-80. doi: 10.1111/hae.14144.

[4] Gouw SC, van den Berg HM, Oldenburg J, Astermark J, de Groot PG, Margaglione M, et al. F8 gene mutation type and inhibitor development in patients with severe hemophilia A: systematic review and meta-analysis. Blood. 2012; 119(12): 2922-34. doi: 10.1182/blood-2011-09-379453.

[5] Oldenburg J, Picard JK, Schwaab R, Brackmann HH, Tuddenham EG, Simpson E. HLA genotype of patients with severe Haemophilia A due to Intron 22 inversion with and without inhibitors of Factor VIII. Thromb Haemost. 1997; 77(2): 238-42. doi: 10.1055/s-0038-1655945

[6] Nathalang O, Sinsiriwong S, Pan-in T, O-charoen R, Chuansumrit A. Association between HLA-DRB1 alleles and FVIII inhibitors in Thai patients with severe hemophilia A. Turk J Hematol. 2012; 29(1): 34-9. doi: 10.5505/tjh.2012.87126.

[7] Satapornpong P, Jantararoungtong T, Konyoung P, Khunkaew S, Sukasem C. Epidemiology of HLA-B*15:02 and other HLA alleles as pharmacogenetic markers in Thais. Front Pharmacol. 2020; 11: 78. doi: 10.3389/fphar.2020.00078.

[8] Astermark J, Oldenburg J, Pavlova A, Berntorp E, Lefvert AK. Polymorphisms in the IL10 and TNF-alpha gene promoters and inhibitor development in patients with hemophilia A. Blood. 2006; 107(8): 3167-72. doi: 10.1182/blood-2005-09-3811.

[9] Mancuso ME, Santagostino E. The role of emicizumab in hemophilia A patients with inhibitors. J Thromb Haemost. 2012; 10(5): 781-90. doi: 10.1111/ j.1538-7836.2012.04683.x.

[10] Verbruggen B. Diagnosis and quantification of factor VIII inhibitors. Haemophilia. 2010; 16(102): 20-4. doi: 10.1111/j.1365-2516.2010.02261.x.

[11] Rossetti LC, Radic CP, Larripa IB, De Brasi CD. Genotyping of hemophilia A rearrangements using long-distance PCR. J Thromb Haemost. 2008; 6(5): 830-6. doi: 10.1111/j.1538-7836.2008.02927.x.

[12] Delfan N, Galehdari H, Kazeminejad SR, Alikamali M, Ahmadi M. Analysis of HLA-DRB1*15:01 in multiple sclerosis patients: a case-control study. Zahedan J Res Med Sci. 2017; 19(3): e5647. doi: 10.5812/zjrms.5647.

[13] Pavlova A, Delev D, Lacroix-Desmazes S, Schwaab R, Mende M, Fimmers R, et al. Impact of polymorphisms of the major histocompatibility complex class II on inhibitor development in severe hemophilia A. J Thromb Haemost. 2009; 7(12): 2006-15. doi: 10.1111/j.1538-7836.2009.03636.x.

[14] Abu Arra CM, Arra IA, El-Gharbawy FA, Mansour S. Factor VIII intron 22 inversion in severe hemophilia A patients in Palestine. Scientifica (Cairo). 2020; 2020: 3428648. doi: 10.1155/2020/3428648.

[15] Chuansumrit A, Sasanakul P, Seksarn P, Sosothikul D, Rungrungtham T, Tapaneeyakorn S, et al. Genotype and phenotype of haemophilia A in Thai patients. Haemophilia. 2003; 9(2): 179-86. doi: 10.1046/j.1365-2516.2003.00727.x.

[16] Romphruk AV, Romphruk A, Incheun R, Butthep P, Buaklee N, Jatuasitha S, et al. HLA class I and II alleles and haplotypes in ethnic Northeast Thais. Tissue Antigens. 2010; 75(6): 701-11. doi: 10.1111/j.1399-0039.2010.01449.x.

[17] Abdulqader AMR, Mohammed AI, Rachid S, Ghoraishizadeh P, Mahmood SN. Identification of the intron 22 and intron 1 inversions of the factor VIII gene in Iraqi Kurdish patients with hemophilia A. Clin Appl Thromb Hemost. 2020; 26: 1076029619888293. doi: 10.1177/1076029619888293.

[18] Hay CR, Ollier W, Pepper L, Cumming A, Keeney S, Goodeve AC, et al. HLA class II profile: a weak determinant of factor VIII inhibitor development in severe haemophilia A. Thromb Haemost. 1997; 77(2): 234-7. doi: 10.1055/s-0038-1655944.

Journal of Associated Medical Sciences 2026; 59(2): 105-112

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Fall risk, fear of falling and their association with depressive symptoms among institutionalized older adults: A cross-sectional study

Bowas S, Loganathan S*, Raghuram P, Jaivignesh S

Faculty of Occupational Therapy, Sri Ramachandra Institute of Higher Education and Research (DU), Porur, Chennai, India.

ARTICLE INFO

Article history:
Received 23 September 2025
Accepted as revised 22 January 2026
Available online 4 February 2026

Keywords:
Fall risk, fear of falling, depression, institutionalized elderly, cross-sectional study.

ABSTRACT

Background: Falls, fear of falling, and depression are major geriatric syndromes that threaten independence and overall well-being. Older people living in institutions may be particularly vulnerable due to physical, environmental, and psychosocial factors. This study aimed to examine the association among fall risk, fear of falling and depressive symptoms among Institutionalized Older Adults.

Materials and methods: A cross-sectional study was carried out among 90 residents from two selected old age homes in Chennai. Participants aged more than 60 years were recruited using purposive sampling. Fall risk, fear of falling and depressive symptoms were assessed using the Berg Balance Scale (BBS), Falls Efficacy Scale-International (FES-I) and Geriatric Depression Scale (GDS-15), respectively. The data was analyzed with IBM SPSS Statistics version 27, employing descriptive statistics, ANOVA, Chi-Square test and Pearson correlation.

Results: High fear of falling was reported by 62% of participants while 67% were presented with depressive symptoms. The results obtained were statistically significant associations between fall risk and fear of falls (p=0.032), fall risk and depressive symptoms (p<0.001), and fear of falls and depressive symptoms (p=0.042). Age showed a significant association with fall risk and fear of falling, but not with depression.

Conclusion: The study shows that fall risk and fear of falling are highly prevalent among the institutionalized older adults and are closely linked to depressive symptoms. These findings emphasize the need for integrated intervention addressing both physical balance and psychological well-being to reduce fall risk, alleviate fear of falling, and improve mental health in residential care settings.

Introduction

Ageing results from the progressive accumulation of cellular and molecular damage throughout the lifespan, leading to a gradual decline in physical and cognitive functions, increased vulnerability to disease, and ultimately, death.[1] These biological changes are not strictly determined by chronological age and vary among individuals. The global demographic shift toward an ageing population is particularly evident in high-income countries, where the proportion and absolute number of older adults are expected to rise sharply in the coming decades.[2] As ageing is closely associated with the need for long-term care, the number

* **Corresponding contributor.**
Author's Address: *Faculty of Occupational Therapy, Sri Ramachandra Institute of Higher Education and Research (DU), Porur, Chennai, India.*
E-mail address: loganathanot@sriramachandra.edu.in
doi: 10.12982/JAMS.2026.042
E-ISSN: 2539-6056

of older individuals requiring institutionalization is projected to increase substantially.[3] Nevertheless, most older adults prefer to remain in their own homes to preserve familiar surroundings and maintain social connections for as long as possible.[4] Depression is a common but often underrecognized mental health condition among older adults, with approximately 15% of community-dwelling older persons reporting significant depressive symptoms[5]. Despite its high prevalence, depression is not an inevitable part of the ageing process. It is frequently underdiagnosed and undertreated, especially among institutionalized or residential care populations, where symptoms may remain unnoticed or misattributed to ageing itself.[6] Falls represent another major public health concern in geriatric populations, being one of the leading causes of injury, hospitalization, and mortality among older adults.[7] Beyond the physical consequences, falls can result in profound psychological impacts. Fear of falling (FoF) is a common psychological outcome that may occur even in individuals without a history of falls. This fear can lead to activity restriction, social withdrawal, and decreased mobility, which in turn contribute to physical deconditioning, increased fall risk, and functional dependence.[8] Fall-related problems and psychological distress are particularly prevalent among institutionalized older adults, potentially compromising both physical independence and mental well-being. Importantly, fall risk reflects not only physical frailty but may also be closely linked to emotional factors such as anxiety and depression. Although fall risk, FoF, and depressive symptoms are clinically interconnected, their relationships have not been adequately explored among institutionalized populations. Understanding these associations is crucial for designing effective preventive and rehabilitative interventions that promote both physical stability and psychological resilience in older adults. Therefore, the present study aimed to examine the associations among fall risk, fear of falling, and depressive symptoms among institutionalized older adults in Chennai, India. By identifying the interrelationships among these variables, the study seeks to contribute to the development of evidence-based strategies that address both the physical and mental health needs of older adults in institutional care settings.

Materials and methods

Study design and participants

This cross-sectional study was conducted over a period of four weeks among elderly residents of two old age homes in Chennai-Anandam, a free home for senior citizens, and Akshaya Trust, Velappanchavadi. A total of 90 participants were enrolled, comprising 48 residents from Anandam and 42 from Akshaya Trust. Both institutions were selected for their comparable residential environments and caregiving structures to minimize potential sampling bias. All participants were provided with detailed information about the purpose and procedures of the study in their preferred language (Tamil or English). Written informed consent was obtained prior to data collection. For participants with limited literacy, the consent form was read aloud in the presence of a witness, and a thumb impression was obtained to indicate voluntary participation. Only individuals with adequate cognitive ability (Six-Item Cognitive Impairment Test score < 8) were included to ensure capacity to provide informed consent. Confidentiality of personal data was maintained throughout the study, and participants were assured that they could withdraw at any point without any impact on their care or living arrangements.

Sample size and sampling method

Purposive sampling was used to recruit 90 older adults. The sample size was determined using the prevalence of fear of falling (90,5%) reported by Pereira *et al.*,[9] assuming a 7% relative precision, a 95% confidence level, and a 10% non-response rate, giving a total of 90 participants.

Eligibility criteria

Inclusion criteria

1. Age more than 60 years
2. Minimum 3 months residence in the institution
3. Independent mobility with or without aids
4. Six-Item Cognitive Impairment Test (6CIT) score less than 8, and
5. Ability to understand Tamil or English.

Exclusion criteria

1. History of neurological or psychiatric disorders
2. Recent trauma or surgical procedures.

Data collection tools

Screening was first performed using the Six-Item Cognitive Impairment Test (6CIT) to ensure cognitive eligibility. Participants who met the inclusion criteria were then assessed using the following standardized instruments:

1. Berg Balance Scale (BBS): The BBS is a performance-based tool used to assess functional balance through 14 common tasks such as standing, reaching, and turning. Each item is scored on a 5-point scale (0-4), with a total possible score of 56. Scores below 40 indicate a higher risk of falls. The BBS has demonstrated high inter-rater reliability (ICC=0.98) and internal consistency (Cronbach's α=0.96).

2. Falls Efficacy Scale-International (FES-I): The FES-I is a self-reported questionnaire designed to evaluate an individual's concern about falling during daily activities. It consists of 16 items scored from 1 (not at all concerned) to 4 (very concerned), with total scores ranging from 16 to 64. Higher scores indicate greater fear of falling. The FES-I has shown excellent internal consistency (Cronbach's α=0.96) and good

test-retest reliability.

3. Geriatric Depression Scale-Short Form (GDS-15): The GDS-15 is a validated screening instrument for depressive symptoms in older adults. It comprises 15 yes/no questions reflecting mood, motivation, and life satisfaction. Scores range from 0-15, with 0-4 indicating normal mood, 5-8 mild depression, 9-11 moderate depression, and 12-15 severe depression. The tool demonstrates high internal consistency (Cronbach's α=0.80-0.90) and is widely used for geriatric populations.

Preparatory steps

To ensure accuracy and reliability of data collection, assessors underwent structured training in the standardized administration and scoring of all instruments. The tools were reviewed for cultural and linguistic relevance, and a pilot test was conducted among a small group of older adults to confirm clarity and feasibility. All assessments were carried out under supervision, with measures taken to ensure participant safety and comfort throughout the procedure.

Each assessment was administered individually in a quiet, private setting within the institution to ensure comfort and accuracy of responses.

Statistical analysis

All data were entered into Microsoft Excel 2021 and subsequently analyzed using IBM SPSS Statistics version 27. Descriptive statistics, including means, standard deviations, frequencies, and percentages, were used to summarize participants' demographic characteristics and baseline study variables. Inferential analyses were conducted to examine the relationships among fall risk (Berg Balance Scale), fear of falling (Falls Efficacy Scale-International), and depressive symptoms (Geriatric Depression Scale). Chi-square tests were used for categorical variables, one-way analysis of variance (ANOVA) for continuous variables across groups, and Pearson's correlation coefficient to assess bivariate associations. $p<0.05$ was considered statistically significant.

Results

A total of 90 older adults aged between 61 and 90 years were enrolled in the study based on the selection criteria. The mean age of the participants was 74.24 years (SD=7.85). Distribution by gender showed a slight predominance of females, with 54.4% (N=49) being female and 45.6% (N=41) male. Majority were married (64.4%, N=58), while 35.6% (N=32) were unmarried. The mean duration of stay in the institution was 3.72 years (SD=4.70), ranging from 0.33 years (4 months) to 17 years. Educational attainment varied considerably. The largest group (44.4%, N=40) had completed SSLC/PUC, followed by 25.6% (N=23) with incomplete schooling. A smaller proportion had completed higher secondary or diploma (10%, N=9), college-level education (11.1%, N=10), while 8.9% (N=8) reported having no formal education. Overall, the findings suggest that most participants had at least secondary-level education. Regarding fall history, 26.7% (N=24) experienced at least one fall previously, while 73.3% (N=66) had no history of falls. Most of the participants (83.3%, N=75) did not use any device, 15.6% (N=14) used a walking cane or stick, and only one participant (1.1%) reported the use of a walker. For past orthopaedic history, 43.3% (N=39) had been previously diagnosed with or treated for an orthopaedic condition, while 56.7% (N=51) reported no such history. Details of the socio-demographic characteristics of participants are presented in Table 1.

Table 1. Socio-demographic characteristics of participants (N=90)

Characteristic	Category	Frequency (N)	Percentage (%)
Age (years)	61–70	30	33.3
	71–80	41	45.6
	81–90	19	21.1
	Mean±SD	74.24±7.84	
Gender	Male	41	45.6
	Female	49	54.4
Marital status	Married	58	64.4
	Unmarried	32	35.6
Duration of stay (years)	Less than 1	33	36.7
	1 to 5	34	37.8
	More than 5	23	25.5
	Mean±SD	3.71±4.69	

Table 1. Socio-demographic characteristics of participants (N=90) (continued)

Characteristic	Category	Frequency (N)	Percentage (%)
Educational history	No schooling	8	8.9
	Incomplete schooling	23	25.6
	SSLC/PUC	40	44.4
	HSC/Diploma	9	10.0
	College	10	11.1
History of falls	Yes	24	26.7
	No	66	73.3
Assistive/adaptive devices	Walking cane or stick	14	15.6
	Walker	1	1.1
	None	75	83.3
Past orthopaedic history	Yes	39	43.3
	No	51	56.7

Note: *SD: Standard Deviation*

On the Berg Balance Scale (BBS), 56.7% (N=51) of participants were classified as low fall risk, 36.7% (N=33) as medium fall risk, and 6.6% (N=6) as high fall risk, with a mean score of 41.29±11.62. Falls Efficacy Scale-International (FES-I) indicated that 62.2% (N=56) of participants reported high fear of falling, 27.8% (N=25) reported moderate fear, and 10.0% (N=9) reported low fear. The mean FES-I score was 30.94±8.96. Depressive Symptoms were distributed as follows, 32.2% (N=29) of participants had normal scores, 32.2% (N=29) had mild depression, 22.2% (N=20) had moderate depression, and 13.3% (N=12) had severe depression, with a mean score of 6.92±4.00. The distribution of participants across levels of fall risk, fear of falling, and depressive symptoms is summarized in Table 2.

Table 2. Levels of fall risk, fear of falling, and depressive symptoms among participants.

Variable	Level	Frequency	Percentage	Mean±SD
Fall risk (BBS score)	Low fall risk	51	56.7	41.29±11.62
	Medium fall risk	33	36.7	
	High fall risk	6	6.6	
Fear of falling (FES-I score)	Low fear of falling	9	10.0	30.94±8.96
	Moderate fear of falling	25	27.8	
	High fear of falling	56	62.2	
Depressive symptoms (GDS score)	Normal	29	32.2	6.92±4.00
	Mild depression	29	32.2	
	Moderate depression	20	22.2	
	Severe depression	12	13.3	

Note: *BBS: Berg Balance Scale, FES-I: Falls Efficacy Scale-International, GDS: Geriatric Depression Scale.*

A statistically significant difference was observed using analysis of variance (ANOVA) in mean age among the Falls Efficacy Scale-International (FES-I) categories ($F=3.746$, $p=0.028$). The mean age was lowest in the low fear group (68.44±7.038 years), followed by the moderate fear group (73.24±7.172 years), and highest in the high fear group (75.63±7.882 years). ANOVA also indicated a significant difference in mean age across the Berg Balance Scale (BBS) fall risk categories ($F=8.681$, $p<0.001$). Participants classified as low fall risk had a mean age of 71.67±6.716 years, those with medium fall risk had a higher mean age of 78.39±7.701 years, while those with high fall risk had a mean age of 73.33±9.136 years. These comparisons of mean age by fear of falling and fall risk categories are shown in Table 3.

Table 3. Comparison of mean age by fear of falling (FES-I) and fall risk (BBS) categories.

Categories		N	Mean age±SD	F	*p* value
Fear of falling (FES-I)	Low fear of falling	9	68.44±7.038	3.746	0.028*
	Moderate fear of falling	25	73.24±7.172		
	High fear of falling	56	75.63±7.882		
Fall risk (BBS)	Low fall risk	51	71.67±6.716	8.681	<0.001*
	Medium fall risk	33	78.39±7.701		
	High fall risk	6	73.33±9.136		

Note: *FES-I: Falls Efficacy Scale-International, BBS: Berg Balance Scale.*

A Chi-square test showed a significant association between fall risk (BBS) and fear of falling (FES-I) (χ^2=10.525, p=0.032). Among low fall risk participants, 49.0% reported high fear, 39.2% moderate, and 11.8% low fear, while 75.8% of the medium fall risk group and all the high fall risk group reported high fear. Fall risk was also significantly associated with depressive symptoms (GDS) (χ^2=32.724, p<0.001). In the low fall risk group, 39.2% were normal and 9.8% severely depressed, whereas 83.3% of the high fall risk group had severe depression. A further significant association was observed between fear of falling and depression (χ^2=13.085, p=0.042). Among those with low fear, 66.7% were normal and none severely depressed, while in the high fear group 37.5% had mild, 25.0% moderate, and 17.9% severe depression. The associations between these variables are presented in Table 4.

Table 4. Associations between fall risk, fear of falling, and depressive symptoms.

Association	Categories	Distribution (N, %)	χ^2	*p* value
Fall risk and fear of falling (BBS × FES-I)	Low fall risk	Low fear: 6 (11.8%) Moderate fear: 20 (39.2%) High fear: 25 (49.0%)	10.525	0.032*
	Medium fall risk	Low fear: 3 (9.1%) Moderate fear: 5 (15.2%) High fear: 25 (75.8%)		
	High fall risk	Low fear: 0 (0.0%) Moderate fear: 0 (0.0%) High fear: 6 (100.0%)		
Fall risk and depressive symptoms (BBS × GDS)	Low fall risk	Normal: 20 (39.2%) Mild: 13 (25.5%) Moderate: 13 (25.5%) Severe: 5 (9.8%)	32.724	<0.001*
	Medium fall risk	Normal: 9 (27.3%) Mild: 16 (48.5%) Moderate: 6 (18.2%) Severe: 2 (6.1%)		
	High fall risk	Normal: 0 (0.0%) Mild: 0 (0.0%) Moderate: 1 (16.7%) Severe: 5 (83.3%)		
Fear of falling and depressive symptoms (FES-I × GDS)	Low fear of falling	Normal: 6 (66.7%) Mild: 1 (11.1%) Moderate: 2 (22.2%) Severe: 0 (0.0%)	13.085	0.042*
	Moderate fear of falling	Normal: 12 (48.0%) Mild: 7 (28.0%) Moderate: 4 (16.0%) Severe: 2 (8.0%)		
	High fear of falling	Normal: 11 (19.6%) Mild: 21 (37.5%) Moderate: 14 (25.0%) Severe: 10 (17.9%)		

Note: *BBS: Berg Balance Scale, FES-I: Falls Efficacy Scale-International, GDS: Geriatric Depression Scale, χ^2: Chi-square value, *Significant at p<0.05.*

Pearson's correlation analysis was carried out to determine the relationships among age, fall risk (BBS), fear of falls (FES-I), and depressive symptoms (GDS). Age was significantly correlated with BBS (r=0.290, *p*=0.006) and FES-I (r=0.275, *p*=0.009), but not with GDS (r=0.034, *p*=0.752). BBS was negatively correlated with FES-I (r=-0.281, *p*=0.007) and positively correlated with GDS (r=0.271, *p*=0.010). FES-I demonstrated a positive correlation with GDS (r=0.310, *p*=0.003). These findings suggest that fear of falling is strongly linked with both fall risk and depression, while age influences balance and fear of falling but not depressive symptoms, underscoring the need for integrated interventions addressing both mobility and psychological health in older adults. The correlation matrix illustrating relationships among age, fall risk, fear of falling, and depressive symptoms is displayed in Table 5.

Table 5. Pearson correlation matrix: Relationships between age, fall risk, fear of falling, and depressive symptoms.

Variables		Age	BBS score	FES-I score	GDS score
Age	Pearson correlation	1.000	0.290**	0.275**	0.034
	p value	-	0.006	0.009	0.752
BBS score	Pearson correlation	0.290**	1.000	-0.281**	0.271**
	p value	0.006	-	0.007	0.010
FES-I score	Pearson correlation	0.275**	0.281**	1.000	0.310**
	p value	0.009	0.007	-	0.003
GDS score	Pearson correlation	0.34	0.271**	0.310**	1.000
	p value	0.752	0.010	0.003	-

Note: *BBS: Berg Balance Scale, FES-I: Falls Efficacy Scale-International, GDS: Geriatric Depression Scale.*

Discussion

The present study demonstrated a statistically significant association between the risk of falling and fear of falling among institutionalized older adults (*p*=0,032). All participants identified as having a high risk of falls also exhibited high levels of fear of falling, underscoring the strong interconnection between physical balance impairment and psychological apprehension. This finding corroborates the work of Ellmers *et al*. who reported that fear of falling independently predicts future fall incidence and is associated with reduced balance performance and activity restriction.[10]

Furthermore, the results align with Martínez-Arnau *et al*. who observed that frailty markers such as slowness and exhaustion are positively related to fear of falling, particularly among older adults with reduced physical activity.[11] The overlap between these findings suggests that fear of falling is not merely a psychological response to instability, but also a precipitating factor that can exacerbate physical deconditioning, diminish self-efficacy, and ultimately lead to a loss of independence. This bidirectional interaction highlights the necessity of addressing both physical and psychological domains in fall prevention strategies.[11]

A highly significant relationship was also observed between fall risk and depressive symptoms (*p*<0,001). Participants with a higher fall risk demonstrated markedly elevated depression levels, with 83,3% classified as severely depressed. This observation echoes Biderman *et al*. who emphasized that depression and falls share overlapping risk factors such as reduced physical strength, limited engagement in daily activities, and slower gait speed.[12] The anticipation or experience of falls can lead to functional limitations, fostering dependence and emotional distress.[12] The "DS-falls pathway" proposed by Hoffman *et al*. further explains this mechanism, linking depressive-related fatigue, inattention, and impaired coordination with an increased likelihood of falls.[13] Accordingly, comprehensive fall risk assessments should incorporate screening for depressive symptoms, acknowledging the intertwined nature of emotional and physical decline.[13]

The study also revealed a significant association between fear of falling and depressive symptoms (*p*=0,042), supporting the bidirectional model proposed by Pereira *et al*.[9] Fear of falling may restrict mobility, increase social withdrawal, and reduce autonomy, while depression can heighten perceived vulnerability and reduce confidence in physical abilities, forming a "vicious cycle". Integrative interventions-combining psychological support, cognitive-behavioral strategies, and environmental modifications-are therefore essential to break this cycle and enhance both physical safety and psychological well-being among older adults.[9] Lee and Kim also noted that fear-based avoidance behaviors often result in social disengagement, which contributes to depressive symptoms and diminished well-being, further reinforcing the interplay between psychological and behavioral factors.[14]

Lastly, age was significantly correlated with fall risk (*p*=0.006) and fear of falling (*p*=0.009), but not

with depressive symptoms (p=0,752). This pattern suggests that aging exerts a stronger influence on physical vulnerability than on emotional well-being among institutionalized elderly populations. This aligns with Gazibara *et al*. and Gale *et al*. who reported that age-related declines in muscle strength, gait stability, and confidence contribute significantly to fall risk.[15,16] In contrast, the lack of association between age and depression corresponds with Trevisan *et al*. who highlighted that psychosocial and gender-related variables exert a greater influence than chronological age in predicting depressive symptoms among older adults.[17]

In summary, the interrelationships among fall risk, fear of falling, and depression observed in this study highlight the multifactorial nature of geriatric health. These findings underscore the need for multidisciplinary interventions integrating physical rehabilitation, psychological support, and environmental safety measures to promote holistic well-being and preserve autonomy in institutionalized elderly individuals.

Limitations

This study has several limitations that should be acknowledged. Firstly, the use of convenience sampling may limit the representativeness of the sample, as participants were recruited from a single institutional setting. Consequently, the findings may not be generalizable to older adults living independently in the community. Secondly, the reliance on self-reported measures, including the Geriatric Depression Scale (GDS) and the Falls Efficacy Scale-International (FES-I), introduces potential response bias. Participants' mood, cognitive capacity, and willingness to disclose personal information could have influenced their responses, which should be considered when interpreting the results. Thirdly, the cross-sectional design precludes causal inference between fall risk, fear of falling, and depressive symptoms. Future studies should adopt longitudinal designs to clarify temporal relationships among these variables. In addition, employing random sampling and recruiting a larger, more diverse population from multiple institutions and community settings would enhance the external validity of future research.

Moreover, institutional living conditions may shape perceptions of autonomy, safety, and emotional well-being, which could confound the observed associations. Comparative studies between institutionalized and community-dwelling older adults are therefore recommended.

Finally, future research should explore the potential mediating or moderating roles of psychosocial factors-such as social support, loneliness, and cognitive decline-to provide a more comprehensive understanding of the interplay between physical and psychological health in aging populations.

Conclusion

This study demonstrates a significant interrelationship among fall risk, fear of falling, and depressive symptoms in institutionalized older adults. A substantial proportion of participants reported high levels of fear of falling, and varying degrees of depressive symptoms were observed. Notably, individuals at greater risk of falls exhibited more severe depressive symptoms, and elevated fear of falling was associated with both increased fall risk and heightened depressive symptoms.

These findings underscore the intertwined nature of physical and psychological health concerns among elderly individuals residing in old age homes. The results highlight the importance of comprehensive interventions that address both balance and mental well-being to reduce fall risk, alleviate fear of falling, and mitigate depressive symptoms. Future research and clinical practice should consider integrative strategies that simultaneously target physical stability, emotional support, and environmental safety to promote holistic health and independence in institutionalized older adults.

Ethical approval

Approved by the Institutional Ethics Committee of Sri Ramachandra Institute of Higher Education and Research (DU), Chennai, India.

Funding

No external funding was received.

Conflict of interest

The authors declare no conflicts of interest.

CRediT authorship contribution statement

Bowas S: conceptualization, methodoly, investigation, writing original draft, review and edit; **Loganathan S, Raghuram P**, and **Jaivignesh S**: supervision, writing: review and edit.

Acknowledgements

The authors thank the management of the participating old age homes and all participants for their cooperation.

References

[1] World Health Organization. Ageing and health [Internet]. Geneva: World Health Organization; 2024 Oct 1 [cited 2025 Sep 16]. Available from: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

[2] Kinsella KG, He W. An aging world: 2008 [Internet] International Population Reports P95/09-1. Washington (DC): U.S. Census Bureau; 2009 [cited 2025 Sep 21]. Available from: https://www.census.gov/library/publications/2009/demo/p95-09-1.html

[3] Channon A, Matthews Z, Van Lerberghe W. Will there be enough people to care? Notes on workforce implications of demographic change 2005-2050. Geneva: World Health Organization; 2006.
[4] Del Duca GF, Silva SG, Thumé E, Santos IS, Hallal PC. Predictive factors for institutionalization of the elderly: a case-control study. Rev Saude Publica. 2012; 46(1): 147-53. English, Portuguese. doi: 10.1590/s0034-89102012000100018.
[5] Beekman AT, Deeg DJ, van Tilburg T, Smit JH, Hooijer C, van Tilburg W. Major and minor depression in later life: a study of prevalence and risk factors. J Affect Disord. 1995; 36(1-2): 65-75. doi: 10.1016/0165-0327(95)00061-5.
[6] Carreira L, Botelho MR, Matos PCB, Torres MM, Salci MA. Prevalence of depression in institutionalized elderly. Rev Enferm UERJ. 2011; 19(2): 268-73.
[7] Milos V, Bondesson Å, Magnusson M, Jakobsson U, Westerlund T, Midlöv P. Fall risk-increasing drugs and falls: a cross-sectional study among elderly patients in primary care. BMC Geriatr. 2014; 14: 40. doi: 10.1186/1471-2318-14-40.
[8] Brouwer B, Musselman K, Culham E. Physical function and health status among seniors with and without a fear of falling. Gerontology. 2004; 50(3): 135-41. doi:10.1159/000076771.
[9] Pereira MEA, Santos GS, Almeida CR, Nunes KCS, Silva MCM, José H, et al. Association between falls, fear of falling and depressive symptoms in community-dwelling older adults. Healthcare (Basel). 2024; 12(16): 1638. doi:10.3390/healthcare 12161638.
[10] Ellmers TJ, Ventre JP, Freiberger E, Hauer K, Hogan DB, Lim ML, et al. Does concern about falling predict future falls in older adults? A systematic review and meta-analysis. Age Ageing. 2025; 54(4): afaf089. doi:10.1093/ageing/afaf089.
[11] Martínez-Arnau FM, Prieto-Contreras L, Pérez-Ros P. Factors associated with fear of falling among frail older adults. Geriatr Nurs. 2021; 42(5): 1035-41. doi:10.1016/j.gerinurse.2021.06.007.
[12] Biderman A, Cwikel J, Fried AV, Galinsky D. Depression and falls among community-dwelling elderly people: a search for common risk factors. J Epidemiol Community Health. 2002; 56(8): 631-6. doi:10.1136/jech.56.8.631.
[13] Hoffman GJ, Hays RD, Wallace SP, Shapiro M, Ettner SL. Depressive symptomatology and fall risk among community-dwelling older adults. Soc Sci Med. 2017; 178: 206-13. doi:10.1016/j.socscimed.2017.02.020.
[14] Lee TH, Kim W. Is fear of falling and the associated restrictions in daily activity related to depressive symptoms in older adults? Psychogeriatrics. 2021; 21(3): 304-10. doi:10.1111/psyg.12664.
[15] Gazibara T, Kurtagić Pećanin I, Kišić-Tepavčević D, Nurković S, Kovačević N, et al. Falls, risk factors and fear of falling among persons older than 65 years of age. Psychogeriatrics. 2017; 17(4): 215-23. doi:10.1111/psyg.12217.
[16] Gale CR, Cooper C, Sayer AA. Prevalence and risk factors for falls in older men and women: the English Longitudinal Study of Ageing. Age Ageing. 2016; 45(6): 789-94. doi:10.1093/ageing/afw129.
[17] Trevisan FK, Silva RH, Reis SFA, Giehl MWC. Prevalence of depressive symptoms and associated factors in Brazilian older adults: 2019 Brazilian National Health Survey. Cad Saude Publica. 2024;40(12):e00006124.

Journal of Associated Medical Sciences 2026; 59(2): 113-124

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Factors increasing telemedicine adoption across adult non-healthcare professionals: A systematic review of observational studies published between 2021 and 2025

Trisakti Halimah Delimasari*, Elsa Tursina, Annisa' Arifatul Hikmah, Riska Fajar Fatony

Department of Health Information Management, Faculty of Health, State Polytechnic of Jember, East Java, Indonesia.

ARTICLE INFO

Article history:
Received 30 July 2025
Accepted as revised 9 January 2026
Available online 6 February 2026

Keywords:
Adult non-healthcare professionals, determinants, factors, digital medical services, telehealth, telemedicine.

ABSTRACT

Background: Telemedicine has emerged as a key digital healthcare platform, offering convenient access to medical services for the public. Despite its potential, widespread adoption remains hindered by lingering consumer skepticism.

Objectives: The aim of this systematic review was to investigate the factors increasing telemedicine adoption across adult non-healthcare professionals.

Materials and methods: This was a systematic review conducted of primary observational studies published between 2021 and 2025. Adhering to PRISMA guidelines, the authors systematically searched three online databases (PubMed, ProQuest, and Science Direct) and screened articles using the PRISMA Flow Chart to identify eligible studies based on predefined inclusion criteria. The Joanna Briggs Institute's (JBI) critical assessment techniques were used to evaluate the quality of the research findings.

Results: From the reviewed secondary data derived from primary studies, 22 original articles were selected for analysis. Four major themes emerged as drivers of telemedicine utilization, including: 1) socioeconomic-demographic, such as age, education, income, employment, marital status, and social influence, 2) user's perceived benefits factors like perceived usefulness, benefits, ease of use, and trust, 3) health-related problems, such as frequent healthcare demands, caregiving responsibilities, having chronic disease, and mental health conditions, and 4) internal factors or technological factors of telemedicine systems: price, service, and system quality.

Conclusion: In conclusion, telemedicine adoption is influenced by a multifaceted interplay of socioeconomic and demographic factors, a user's perceived benefits, individual health needs, and systemic features.

Introduction

The digital revolution has fundamentally transformed various sectors, including healthcare. One of the most prominent advancements in recent years is the rise of telemedicine, which refers to the provision of medical services remotely using information and communication technologies.[1] Telemedicine encompasses a wide range of applications from real time video consultations and remote patient monitoring to mobile health apps and virtual health education platforms. This innovation has redefined how healthcare is accessed and delivered, particularly for populations with geographical, physical, or systemic barriers to in person care.[2]

* **Corresponding contributor.**
Author's Address: *Department of Health Information Management, Faculty of Health, State Polytechnic of Jember, East Java, Indonesia.*
E-mail address: trisakti.delimasari@polije.ic.id
doi: 10.12982/JAMS.2026.043
E-ISSN: 2539-6056

The urgency of adopting telemedicine became especially apparent during the COVID-19 pandemic, where lockdowns, physical distancing measures, and overwhelmed health facilities necessitated a shift towards remote healthcare solutions. In this context, telemedicine not only functioned as an emergency response but also demonstrated long term potential in enhancing the resilience, efficiency, and inclusivity of health systems. Consequently, governments, health organizations, and private sectors have increasingly invested in telemedicine infrastructure, policy frameworks, and public awareness campaigns.[3]

However, despite the growing infrastructure and policy support, the adoption of telemedicine remains inconsistent and fragmented across different populations, regions, and health contexts. Numerous studies have shown that while some groups embrace telemedicine rapidly, others exhibit hesitation, low usage rates, or complete avoidance.[4] This disparity raises crucial questions: What drives or hinders telemedicine adoption? Are there specific sociodemographic, psychological, or contextual factors that consistently predict usage? How do perceptions of technology usability, healthcare needs, or trust influence decision-making? Understanding these factors is essential because the success of telemedicine relies not only on technological availability but also on its acceptance and meaningful utilization by intended users.[5]

Given this background, there is a pressing need to synthesize findings from existing quantitative studies to identify robust, evidence-based determinants of telemedicine use across different demographics. This study aimed to identify the nature and scope of available research literature on factors increasing telemedicine adoption across adult non-healthcare population.

Materials and methods

This study was a systematic review. The articles were obtained from several online databases such as PubMed, ScienceDirect, and ProQuest. The basic query developed for the literature search was: ("adult non-healthcare professionals") AND ("determinants" OR "factors") AND ("telehealth" OR "telemedicine") AND ("observational study"). The Boolean operators in this query were adjusted to fit the specifications of each database.

We followed five steps in conducting a systematic review, which included 1) framing questions for a review, 2) identifying relevant work, 3) assessing the quality of studies, 4) summarizing the evidence, and 5) interpreting the findings.

Studies were included if they met the following criteria: 1) full-text articles examining factors influencing the adoption, use, or preference for telemedicine services among individuals or patients in the general population (≥18 years old), regardless of their specific health needs; 2) studies involving non–healthcare provider populations (i.e., general community members or patients); 3) articles published in English; 4) publications from the last five years (2021-2025); 5) studies that applied quantitative analyses, including at least univariate or bivariate analysis.

Studies were excluded if they met any of the following conditions: 1) articles with themes unrelated to telemedicine utilization as defined in this study, 2) studies with anonymous authorship or unclear authorship, 3) articles that did not employ quantitative analyses, such as those lacking univariate or bivariate results, 4) studies where the dependent variable focused on broader digital health concepts (e.g., e-health, digital health, mobile health, virtual clinics) instead of telemedicine specifically, 5) review articles, case reports, or purely descriptive studies, and 6) studies in which respondents were healthcare providers (e.g., physicians, nurses, or other medical professionals) rather than telemedicine users or the general population.

The authors followed the PRISMA flow chart guidelines for the search process and selection of eligible articles based on the inclusion criteria. The conclusions of this study were derived from secondary data obtained from the primary studies reviewed.

The study selection process was carried out by three independent reviewers, consisting of three authors of this review (except the third author). Each reviewer was assigned one online database to screen. All reviewers extracted the initial search results into separate Excel spreadsheets to ensure systematic organization and transparency.

After individual screening, the reviewer authors compared their extracted lists using Mendeley reference management software to identify duplicates and consolidate the records. Any discrepancies or conflicts regarding study relevance, duplication status, or eligibility were discussed collectively among the reviewers. In cases where disagreement persisted, the team referred to the predefined inclusion and exclusion criteria as well as the study appraisal guidelines to reach a consensus. This multi-reviewer approach strengthened the accuracy and rigor of the study selection and data extraction process.

A comprehensive search was conducted across five online databases, yielding a total of 1,515 records. Specifically, the search identified 654 from PubMed, 273 from ProQuest, and 296 from ScienceDirect. During the identification phase, all retrieved records were screened for duplication. A total of 1,086 duplicate records were removed. After this removal, 429 unique articles remained and were taken forward for title and abstract screening. The 429 articles underwent screening based on their titles and abstracts. From this process, 385 articles were excluded because they did not meet the eligibility criteria. The remaining 44 articles appeared to fit the review objectives and were moved to full-text assessment. Full-text screening was conducted on 44 articles. After a thorough review, 22 articles were excluded for various reasons. After

applying all eligibility criteria, 22 studies met the inclusion requirements. These studies were ultimately included in the qualitative synthesis for the systematic review (Figure 1).

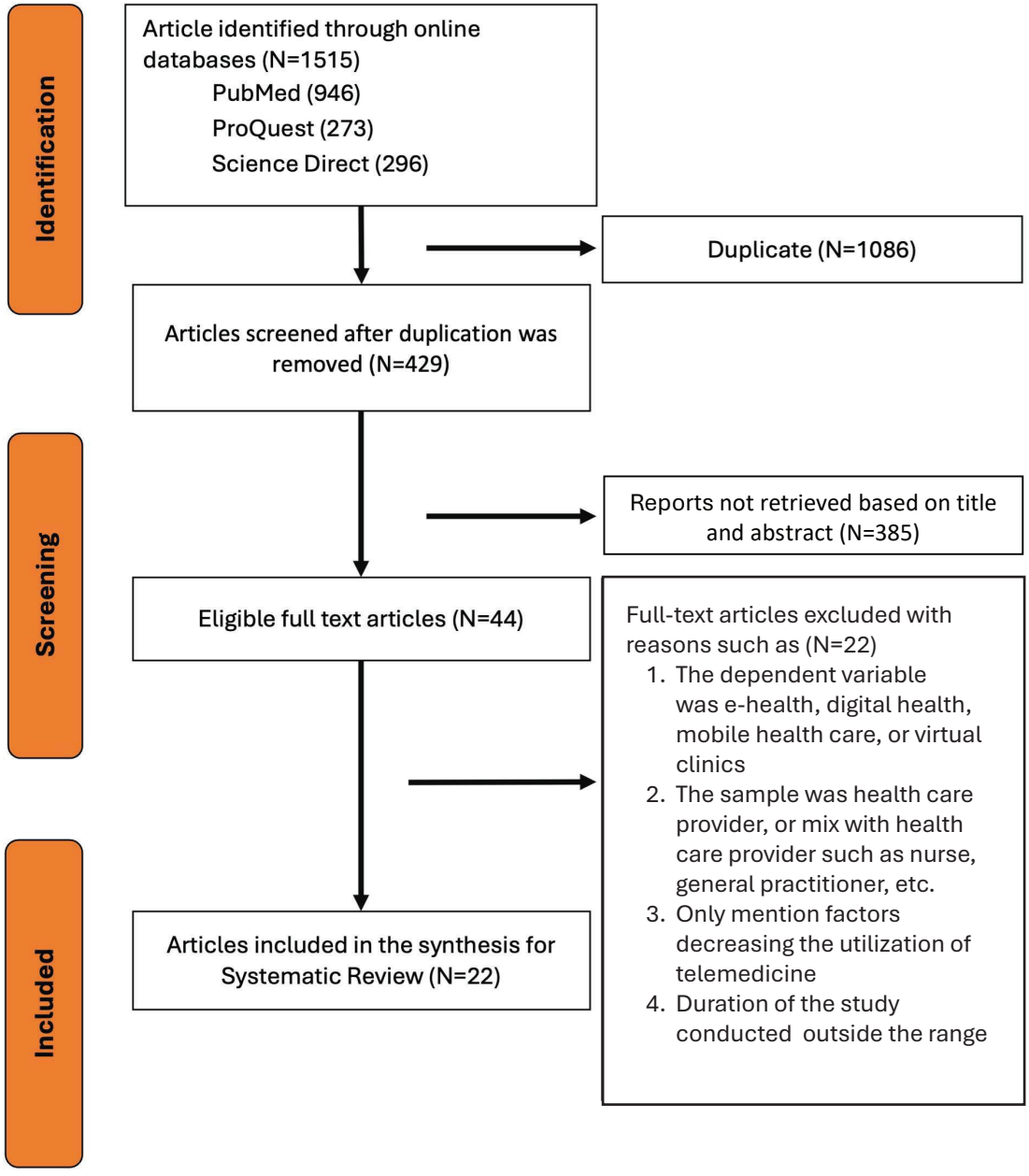


Figure 1. *PRISMA diagram of factors associated with the adoption of telemedicine.*

To assess the quality of the studies, we used the JBI critical appraisal tool for observational studies (Table 1) and for cohort studies (Table 2). The appraisal of the cross-sectional studies showed that 19 studies had a low risk of bias, while one study demonstrated a moderate risk of bias. Both cohort studies were assessed as having a low risk of bias.

Table 1. Risk of bias for cross-sectional studies using the JBI cross-sectionals studies checklist.

	Were the criteria for inclusion in the sample clearly defined?	Were the study subjects and the setting described in detail?	Was the exposure measured in a valid and reliable way?	Were objective, standard criteria used for measurement of the condition?	Were confounding factors identified?	Were strategies to deal with confounding factors stated?	Were the outcomes measured in a valid and reliable way?	Was appropriate statistical analysis used?	Overall appraisal
An *et al.*[17]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Anil *et al.*[12]	Y	Y	Y	Y	NA	NA	Y	Y	7/9 Moderate Risk
Athavale *et al.*[22]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Bhuyan *et al.*[5]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Cobb *et al.*[11]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Jeon *et al.*[13]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Kabir *et al.*[1]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Li *et al.*[18]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Lin *et al.*[19]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Mackwood *et al.*[8]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Meraya *et al.*	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Muehlensiepen *et al.*[6]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Odebunmi *et al.*[7]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Praha *et al.*[10]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Shao *et al.*[9]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Tan *et al.*[4]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Tsatsos *et al.*[15]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Wang *et al.*[16]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Yan *et al.*[20]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk
Yao *et al.*[3]	Y	Y	Y	Y	Y	Y	Y	Y	9/9 Low Risk

Note: Y: yes, N: no, NA: not applicable

Table 2. Risk of bias for cohort studies using the JBI cohort studies checklist.

	Mao *et al.*[21]	**Williams *et al.*[2]**
1. Were the two groups similar and recruited from the same population?	Y	Y
2. Were the exposures measured similarly to assign people to both exposed and unexposed groups?	Y	Y
3. Was the exposure measured in a valid and reliable way?	Y	Y
4. Were confounding factors identified?	Y	Y
5. Were strategies to deal with confounding factors stated?	Y	Y
6. Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)?	Y	Y
7. Were the outcomes measured in a valid and reliable way?	Y	Y
8. Was the follow up time reported and sufficient to be long enough for outcomes to occur?	Y	Y
9. Was follow up complete, and if not, were the reasons to loss to follow up described and explored?	Y	Y
10. Were strategies to address incomplete follow up utilized?	N	N
11. Was appropriate statistical analysis used?	Y	Y
Overall appraisal	**8/9 Low Risk**	**8/9 Low Risk**

Note: Y: yes, N: no, NA: not applicable

Results

A total of 24 studies were included, covering diverse geographical settings. Most were conducted in China (5 studies) and the United States (7 studies), followed by South Korea (2 studies) and Thailand (2 studies), while Bangladesh, Greece, Ireland, Malaysia, Germany, and Saudi Arabia each contributed one study. The combined sample encompassed 2,885,667 individuals (Table 3).

Table 3. Primary article Included in qualitative analysis.

No	Author (year)	Study Design	Study Location	Sample Size	Population	Intervention	Outcome (Factors Increasing Telemedicine Utilize)
1	An *et al.*[17]	Cross-sectional	South Korea	471	Individuals aged ≥30 years	Acceptance of telemedicine	Perceived usefulness, Perceived ease
2	Anil *et al.*[12]	Cross-sectional	USA	215	Individuals >18 years of age	Utilize telemedicine	Economic stability
3	Athavale *et al.*[22]	Cross-sectional	USA	256	Patients aged 18 years and older with a clinical diagnosis of cirrhosis who had at least 1 hepatology clinic visit within the prior 6 months	Utilize telehepatology	Participants with greater pandemic stress
4	Bhuyan *et al.*[5]	Cross-sectional	USA	4,502	Individuals diagnosed with depression	Utilize telemedicine	Individuals diagnosed with depression, women, being married or living with a partner, frequent healthcare visits, and health insurance coverage
5	Cobb *et al.*[11]	Cross-sectional	USA	252	Adults aged 18 years or older	Utilize telemedicine	Having regular doctor and having chronic disease
6	Jeon *et al.*[13]	Cross-sectional	South Korea	552	Participants aged 19-69 years	Utilize telemedicine and teleconsultation	Noncapital residents
7	Kabir *et al.*[1]	Cross-sectional	Bangladesh	1,038	Bangladeshi population	Utilize telemedicine	Perceived benefit, perceived concern, predisposition, age < 20 years, age 20-29 years, unmarried, graduate, higher secondary school certificate, student, urban
8	Li *et al.*[18]	Cross-sectional	China	400	Urban seniors aged 60 and above	Utilize telemedicine	Perceived usefulness
9	Lin *et al.*[19]	Cross-sectional	China	1,006	Not specified population	Utilize telemedicine	Perceived usefulness and perceived ease of use
10	Mackwood *et al.*[8]	Cross-sectional	USA	19,280	Patiens with Cancer	Utilize telemedicine	Patient age (increased use in age < 45 and 85 years and older) and urban residence
11	Mao *et al.*[21]	Longitudinal	Ireland	2,607	Community-dwelling adults aged 50 years and over	Utilize telemedicine	People with more chronic conditions, poorer mental health, and having private health insurance
12	Meraya *et al.*[14]	Cross-sectional	Saudi Arabia	466	Adults aged 60 or above	Utilize telemedicine	Perceived usefulness

Table 3. Primary article Included in qualitative analysis (continued).

No	Author (year)	Study Design	Study Location	Sample Size	Population	Intervention	Outcome (Factors Increasing Telemedicine Utilize)
13	Muehlensiepen *et al*.[6]	Cross-sectional	Germany	438	Patients with RMD	Utilize telemedicine	The possibility that telemedicine services were offered by a rheumatologist, younger age, internet access at home and having RMD diseases.
14	Odebunmi *et al*.[7]	Cross-sectional	USA	1,056	Patients aged between 45 and 75 years who were eligible for colorectal cancer screening	Utilize telemedicine and telepharmacy	Adults younger than 55 years, those with a regular provider health care and long travel times, people who rated pharmacy service quality higher
15	Praha *et al*.[10]	Cross-sectional	Thailand	159	Patients Undergoing Continuous Ambulatory Peritoneal Dialysis	Utilize telemedicine	Older age, higher perceived usefulness
16	Shao *et al*.[9]	Cross-sectional	China	514	Patients with diabetes who are aged ≥18 years	Utilize telemedicine	Social influence, age, education, and family income, patients with a higher family income
17	Tan *et al*.[4]	Cross-sectional	Malaysia	119	Adults aged ≥60 years	Utilize telemedicine	Perceived usefulness
18	Tsatsos *et al*.[15]	Cross-sectional	Greece	90	Adult participants with eye diseases in a single practice office	Utilize telemedicine Care	Perceived usefulness
19	Wang *et al*.[16]	Cross-sectional	China	150	Participants with rehabilitation	Utilize telerehabilitation	Health condition, performance expectancy, and perceived trust
20	Williams *et al*.[2]	Retrospective Observational	USA	2,850,831	Adults aged 19 years and older	Utilize telemedicine	Patients who were employed full-time, patients with a bachelor's degree or higher, patients in the 18-44-year age group
21	Yan *et al*.[20]	Cross-sectional	Thailand	385	Thai individuals aged 18-65 years	Utilize telemedicine	Trust, ease of use, system quality, benefits of use, price, and service quality
22	Yao *et al*.[3]	Cross-sectional	China	880	Middle-aged and older patients with chronic conditions	Utilize telemedicine	Patients with chronic respiratory diseases, education level, monthly income, patients who required caregiving

Most of the studies employed cross-sectional designs (20 studies), with additional longitudinal and retrospective observational designs. Several studies examined telemedicine within specific service areas such as telehepatology, telepharmacy, and telerehabilitation.

Thematic synthesis identified four major domains influencing telemedicine use: 1) socioeconomic and demographic factors; 2) the user's perceived benefits; 3) health-related needs and clinical conditions; and 4) internal or technological system factors.

Socioeconomic and demographic factors

Socioeconomic and demographic characteristics such as age, education, income, employment, marital status, and social influence were examined in 13 studies, [1–13] making this the most frequently assessed category. Several factors consistently supported greater telemedicine uptake. For example, higher educational attainment was positively associated with telemedicine use. [1-4,9]

Evidence regarding geographical location showed that both urban and non-capital (rural) residents demonstrated higher use of telemedicine, as reported by Jeon *et al.*[13] and Kabir *et al.*[1], respectively. Additional enabling factors included having health insurance,[5,21] home internet access,[6] higher or stable household income,[4,5,9,12] full-time employment,[2] and longer travel times to health facilities.[7]

Findings on age were mixed. While Mackwood *et al.*[8] reported no significant age effect, Praha *et al.*[10] found that older adults were more likely to use telemedicine. In contrast, several studies documented higher uptake among middle-aged adults or younger age groups.[1,2,6,7,9]

Marital status also yielded inconsistent findings. Bhuyan *et al.*[5] identified marriage as a facilitator of telemedicine use, whereas Kabir *et al.* reported lower uptake among married individuals.[1] Social influence emerged as an important determinant.[9] Participants who received recommendations from their physician[6] or who had a regular doctor [7,11] were more likely to use telemedicine services.

User's perceived benefits

User perceptions were highlighted in 10 studies[1,4,10,14–20] and represented one of the strongest thematic drivers of telemedicine adoption. Key perceptual determinants included perceived usefulness[4,10,14,15,17-19] perceived benefits,[1,20] ease of use,[17,19,20] and trust in the system or provider.[16,20] All these factors were associated with increased likelihood of telemedicine use.

Health-related needs and clinical conditions

Health-related needs were identified in six studies[3,5-6,11,21,22] and demonstrated a strong positive association with telemedicine uptake. Individuals with chronic conditions including diabetes, respiratory disorders, and rheumatic or musculoskeletal diseases reported higher use of telemedicine.[1,3,6,21] Similarly, those experiencing mental health conditions such as stress or depression[5,21,22] were more likely to use telemedicine services. High healthcare utilization, such as frequent appointments or ongoing medical management, also increased engagement with telemedicine.[3,5]

Internal technological system factors

Technology-related determinants were reported in one study,[20] which identified system quality, service quality, and cost or price considerations as influential factors shaping telemedicine adoption.

Discussion

This systematic review synthesized evidence from 22 studies conducted across multiple countries to identify factors influencing telemedicine adoption. The findings demonstrate that telemedicine utilization is shaped not only by access to technology but also by a multidimensional interplay of socioeconomic-demographic, perceptual, clinical, and systemic determinants. Taken together, these factors reveal that telemedicine adoption is a behavioral, structural, and contextual phenomenon not simply a technological one.

Socioeconomic position consistently emerged as a strong predictor across studies. Individuals with higher education,[1-4] greater income stability[3,4,12] and insurance coverage[5,21] were more likely to adopt telemedicine, reflecting longstanding patterns described in the digital divide framework, where economic and educational advantages translate into greater digital participation. These patterns indicate that telemedicine may inadvertently reinforce existing inequities if socioeconomic barriers are not addressed.

Geographical differences offer an important contrasting picture. Although urban residence often facilitated telemedicine use due to better digital infrastructure,[23] this pattern did not hold across all settings. For example, higher uptake among non-capital residents in South Korea[13] contrasted sharply with findings from Bangladesh where urban populations dominated telemedicine use.[1] This divergence suggests that geographical effects probably are deeply intertwined with system-level variables such as national investment in rural telehealth, local digital literacy, and the maturity of healthcare networks rather than geography alone. Thus, geographical findings cannot be interpreted uniformly without considering broader policy, and infrastructural contexts.

Age, often be assumed to a straightforward determinant, also yielded mixed results. Instead of age itself driving adoption, the evidence suggests that it functions as a proxy for more salient constructs such as digital literacy,[24] prior technology experience,[25] and perceived usefulness key components of technology acceptance model (TAM) and unified theory of acceptance and use of technology (UTAUT).[26] Younger adults often used telemedicine more due to higher

familiarity with digital tools[1,2,6,7,9] yet older adults in certain clinical contexts (e.g., peritoneal dialysis) showed higher adoption when telemedicine addressed specific health needs.[10] These nuances imply that telemedicine adoption is less about chronological age and more about the alignment between the technology and the user's perceived functional value and self-efficacy.

Marital status also showed contradictory associations. In populations dealing with depression, partnered individuals demonstrated higher telemedicine use, likely due to emotional and logistical support that reduces barriers to using digital platforms.[5] In contrast, studies from general community samples found higher uptake among unmarried individuals, possibly due to autonomy in decision-making or lack of alternative health support.[1] These mixed patterns demonstrate that the role of marital status is context-dependent and shaped by health needs, cultural norms, and household responsibilities, rather than acting as a universal facilitator or barrier.

Differences in methodological rigor further help explain these inconsistencies. The vast majority of included studies were low risk, providing relatively stable evidence on socioeconomic and perceptual determinants. The only moderate-risk study, conducted by Anil *et al.* examined economic stability. Its isolated focus and higher risk level may reflect limitations in sampling and confounder control, but importantly, it did not intersect with domains where the strongest contradictions appeared (e.g., age, geography, marital status).[12] This suggests that the inconsistencies observed are more likely driven by genuine contextual variation than by differences in methodological quality.

Beyond demographics, recommendations from trusted healthcare providers consistently facilitated telemedicine use. This aligns with UTAUT's construct of social influence, wherein endorsement from credible sources significantly shapes behavioral intention.[27] Telemedicine systems integrated into primary care workflows might be particularly effective, reinforcing the importance of clinician buy-in for sustained adoption.

Perceived benefit,[7,25] perceived usefulness, convenience,[20,21,25] and, trust,[25,27] emerged as one of the most powerful predictors of telemedicine uptake. These findings are central to both TAM and UTAUT, which emphasize that individuals adopt health technologies when they perceive them as efficient, easy to use, and reliable.

Clinical needs also had a strong influence. Individuals with chronic conditions, frequent healthcare visits, or mobility limitations were more likely to adopt telemedicine because the technology reduces logistical burdens, and, supports ongoing monitoring.[3,5,6,11,21] Mental health populations similarly demonstrated higher uptake, suggesting that telemedicine helps address barriers such as stigma, transportation, or appointment availability.[5,21,22] These patterns illustrate that telemedicine adoption is partly driven by how well the technology aligns with the user's clinical needs a core principle of technology task fit theory.

Finally, system-level features such as cost, service, and system quality played a decisive role. Positive experiences smooth scheduling, responsive consultations, and stable platforms reinforced adoption. In contrast, poor system performance or additional financial burdens quickly eroded trust, and led users to abandon telemedicine.[20] These findings emphasize that telemedicine implementation must prioritize affordability, reliability, and seamless user experience to ensure sustainable adoption.

Conclusion

Telemedicine adoption is influenced by multifaceted socioeconomic and demographic factors, user's perceived benefits, health-related needs and clinical conditions and internal or technological system factors. Furthermore, persistent challenges (such as financial constraints, inadequate digital literacy, privacy concerns, and weak system integration) are continuing to limit equitable implementation. Advancing telemedicine necessitates equity-focused measures, including affordable broadband access, structured digital literacy programs, enhanced data-protection mechanisms, workflow alignment, and culturally responsive design. Policy directions should emphasize digital equity and provider support, while future research need to assess long-term usability, and contextual variations across diverse populations to ensure telemedicine develops into an inclusive, reliable, and sustainable element of modern healthcare systems.

Ethical approval

Not applicable. The study intends to perform a systematic review of available literature and requires formal ethics committee approval. Participants do not need to give their informed consent. This technique reflects the ideals established in the Declaration of Helsinki, supporting ethical standards in literature reviews.

Funding

None

Conflict of interest

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

CRediT authorship contribution statement

Trisakti Halimah Delimasari: conceptualization, study design, methodology, protocol development, literature search strategy, data curation, formal analysis, data synthesis and interpretation, writing: original draft, visualization, project administration, supervision; **Elsa Tursina**: literature screening and selection, data

extraction, quality assessment of included studies, writing: review and edit; **Annisa' Arifatul Hikmah**: data extraction, risk of bias assessment, evidence synthesis support, writing: review and editing; **Riska Fajar Fatony**: literature screening, data validation, reference management, writing: review and edit.

Acknowledgements

We respectfully appreciate the authors of the included papers.

References

[1] Kabir H, Hasan MK, Akter N, et al. Factors associated with the intention of telehealth service utilization among Bangladeshi people: a cross-sectional study. F1000Res. 2022; 11: 996. doi: 10.12688/f1000research.124410.2

[2] Williams C, Shang D. Telehealth Usage Among Low-Income Racial and Ethnic Minority Populations During the COVID-19 Pandemic: Retrospective Observational Study. J Med Internet Res. 2023; 25: e43604. doi: 10.2196/43604

[3] Yao L, Li Q, Li Q, et al. Factors influencing the adoption of telemedicine services among middle-aged and older patients with chronic conditions in rural China: a multicentre cross-sectional study. BMC Health Serv Res. 2025; 25: 775. doi: 10.1186/s12913-025-12931-2

[4] Tan S-H, Yap Y-Y, Tan S-K, et al. Determinants of Telehealth Adoption Among Older Adults: Cross-Sectional Survey Study. JMIR Aging. 2025; 8: e60936.doi: 10.2196/60936

[5] Bhuyan SS, Kalra S, Mahmood A, et al. Motivation and Use of Telehealth Among People with Depression in the United States. J Prim Care Community Health. 2024; 15: 21501319241266515. doi: 10.1177/21501319241266515

[6] Muehlensiepen F, Petit P, Knitza J, et al. Prediction of the acceptance of telemedicine among rheumatic patients: a machine learning-powered secondary analysis of German survey data. Rheumatol Int 2024; 44: 523-34. doi: 10.1007/s00296-023-05518-9

[7] Odebunmi OO, Hughes TD, Waters AR, et al. Findings from a national survey of older US adults on patient willingness to use telehealth services: Cross-sectional survey. J Med Internet Res 2024; 26: e50205. doi: 10.2196/50205

[8] Mackwood MB, Tosteson TD, Alford-Teaster JA, et al. Factors Influencing telemedicine use at a Northern New England Cancer Center during the COVID-19 pandemic. JCO Oncol Pract 2022; 18: e1141-e1153. doi: 10.1200/OP.21.00750

[9] Shao H, Liu C, Tang L, et al. Factors Influencing the behavioral intentions and use behaviors of telemedicine in patients with diabetes: Web-based survey study. JMIR Hum Factors 2023; 10: e46624. doi: 10.2196/46624

[10] Praha N, Sriyuktasuth A, Puwarawuttipanit W, et al. Factors influencing telehealth service use and health outcomes in patients undergoing continuous ambulatory peritoneal dialysis: Cross-sectional study. J Med Internet Res 2023; 25: e48623. doi: 10.2196/48623

[11] Cobb S, Dillard A, Yaghmaei E, et al. Telehealth perceived benefits and self-efficacy do not mediate the effects of demographic, health, and social determinants on telehealth use of low-income African American and Latino residents of public housing in Los Angeles. Healthcare 2025; 13: 286. doi: 10.3390/healthcare13030286

[12] Anil Kumar Vaidhyam S, Huang K-T. Social determinants of health and patients' technology acceptance of telehealth during the COVID-19 pandemic: Pilot survey. JMIR Hum Factors 2023; 10: e47982. doi: 10.2196/47982

[13] Jeon H, Lee J, Jang J, et al. Willingness to use and pay for telemedicine and teleconsultation across five clinical domains in South Korea: Cross-sectional survey. J Med Internet Res 2025; 27: e65304. doi: 10.2196/65304

[14] Meraya AM, Khardali A, Ahmad S, et al. Telehealth perceptions and associated factors among older adults with chronic conditions in Saudi Arabia: a comparative study of users and non-users. Front Public Health 2025; 13: 1542974. doi: 10.3389/fpubh.2025.1542974

[15] Tsatsos M, Rodafinos A, Athanasiadis IK, et al. Patients' intentions to use telemedicine for ophthalmic medical care during the COVID-19 pandemic. Cureus 2024; 16: e54709. doi: 10.7759/cureus.54709

[16] Wang M-Y, Chen H, Gong C, et al. Understanding the use intention and influencing factors of telerehabilitation in people with rehabilitation needs: a cross-sectional survey. Front Public Health 2023; 11: 1274080. doi: 10.3389/fpubh.2023.1274080

[17] An MH, You SC, Park RW, et al. Using an extended technology acceptance model to understand the factors influencing telehealth utilization after flattening the COVID-19 curve in South Korea: Cross-sectional survey study. JMIR Med Inform 2021; 9: e25435. doi: 10.2196/25435

[18] Li W, Guo J, Liu W, et al. Effect of older adults' willingness on telemedicine usage: an integrated approach based on technology acceptance and decomposed theory of planned behavior model. BMC Geriatr 2024; 24: 765. doi: 10.1186/s12877-024-05361-y

[19] Lin Y, Xu X, Liu Y, et al. Perception and acceptance of telemedicine use in health care among the general public in China: Web-based cross-sectional survey. J Med Internet Res 2024; 26. Doi: 10.2196/53497

[20] Yan C, Samphanwattanachai B, Ratarat C, et al. Factors influencing user decision of telemedicine applications in Thailand. PLoS One 2025; 20:

e0325512. doi: 10.1371/journal.pone.0325512

[21] Mao L, Mohan G, Normand C. Use of information communication technologies by older people and telemedicine adoption during COVID-19: a longitudinal study. J Am Med Inform Assoc 2023; 30: 2012–2020. doi: 10.1093/jamia/ocad165

[22] Athavale P, Wong RJ, Satre DD, et al. Telehepatology Use and Satisfaction Among Vulnerable Cirrhosis Patients Across Three Healthcare Systems in the Coronavirus Disease Pandemic Era. Gastro hep advances 2024; 3: 201-9. doi: 10.1016/j.gastha.2023.11.006

[23] Curtis ME, Clingan SE, Guo H, et al. Disparities in digital access among American rural and urban households and implications for telemedicine-based services. The Journal of Rural Health 2022; 38: 512-8. doi: 10.1111/jrh.12614

[24] Kim S, Chow BC, Park S, et al. The Usage of Digital Health Technology Among Older Adults in Hong Kong and the Role of Technology Readiness and eHealth Literacy: Path Analysis. J Med Internet Res 2023; 25: e41915. doi: 10.2196/41915

[25] Han J, Ko D. Mobile Financial Service Adoption Among Elderly Consumers: The Roles of Technology Anxiety, Familiarity, and Age. FinTech 2025; 4: 36. doi: 10.3390/fintech4030036

[26] Jokisch MR, Schmidt LI, Doh M. Acceptance of digital health services among older adults: Findings on perceived usefulness, self-efficacy, privacy concerns, ICT knowledge, and support seeking. Front Public Health; 10. Epub ahead of print 13 December 2022. doi: 10.3389/fpubh.2022.1073756.

[27] Ronzani P, Panizza F, Martini C, et al. Countering vaccine hesitancy through medical expert endorsement. Vaccine 2022; 40: 4635-43. doi: 10.1016/j.vaccine.2022.06.031

Appendix A. Search strategy

A1. Databases searched

The literature search was conducted across the following electronic databases:

- PubMed
- ScienceDirect
- ProQuest

The search strategy was adapted for each database to accommodate specific indexing systems and search interface requirements.

A2. Search terms and boolean operators

The core search string used was:

("adult non-healthcare population") AND ("determinants" OR "factors") AND ("telehealth" OR "telemedicine") AND ("observational study")

For each database, Boolean operators, truncation symbols, and field tags were modified as needed. The search included both keywords and subject headings (MeSH) when available.

A3. Filters applied

Across databases, the following filters were applied consistently:

- Publication year: 2021–2025
- Language: English
- Population: Adults (≥18 years)
- Article type: Quantitative or observational studies (cross-sectional, cohort, case-control)

A4. Date of last search

- Last search conducted: 30 June 2025

A5. Number of records retrieved

Below is the number of articles retrieved from each database before deduplication:

Database	Search Results Retrieved
PubMed	946
ScienceDirect	273
ProQuest	296
Total Retrieved	1,515

A6. Deduplication

All records were imported into a reference manager (Mendeley) and duplicate entries were removed prior to screening with total after deduplication: 1,086

Journal of Associated Medical Sciences 2026; 59(2): 125-133

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



The effects of motor-cognitive training on balance and working memory in young adult women with low physical activity: A randomized controlled trial

Supanee Chounchay*, Pinpapha Retsuk, Parintorn Mantam, Kannika Seetha, Warangkana Arsadee, Aparat Chanchuphon, Nattida Pantub, Supitcha Wongphasuksathaporn

Faculty of Physical Therapy, Huachiew Chalermprakiet University, Samut Prakan Province, Thailand.

ARTICLE INFO

Article history:
Received 25 March 2025
Accepted as revised 2 February 2026
Available online 9 Feburary 2026

Keywords:
Balance, working memory, dual-task training, physical activity.

ABSTRACT

Background: Young adult women with low physical activity levels often experience a decline in muscle mass and balance, which can negatively impact both physical and mental health. Cognitive functions, particularly working memory, play a crucial role in maintaining balance during various activities. A reduction in working memory capacity may impair the ability to manage dual-task activities, increasing the risk of instability and falls.

Objectives: To investigate the effects of motor-cognitive training combining balance exercises with working memory tasks on balance performance and working memory skills in women aged 18-25 years with low physical activity levels.

Materials and methods: A randomized controlled trial was conducted with 46 female participants aged 18-25 years who had low physical activity levels. Participants were randomly assigned to either an experimental group (N=23) or a control group (N=23). The experimental group underwent motor-cognitive training, which integrated balance training with working memory tasks, for three consecutive days. The control group received general information on improving balance and working memory. Pre- and post-training assessments were conducted using the Biodex Balance System to measure static balance (sway index values from the Balance Error Scoring System), dynamic balance (limit of stability), and working memory were assessed using the digit span test. Data were analyzed statistically, with significance set at $p<0.05$.

Results: Following the training, the experimental group exhibited significantly greater improvements in working memory skills compared to the control group. Additionally, the experimental group demonstrated enhanced dynamic balance performance, particularly in the forward/right and backward/left directions, compared to the control group.

Conclusion: Motor-cognitive training enhances balance and working memory in young adult women with low physical activity over the short period. This dual-task training approach presents a promising strategy for promoting balance and cognitive function in this population.

Introduction

The prevalence of low physical activity among young adults has been steadily increasing,[1] particularly among women.[2] Recent research suggests that physical activity significantly influences both physical and mental health.[3] A reduction in physical activity is associated with an increased risk of metabolic syndrome, decreased cardiorespiratory fitness, as well

*** Corresponding contributor.**
Author's Address: *Faculty of Physical Therapy, Huachiew Chalermprakiet University, Samut Prakan Province, Thailand.*
E-mail address: supanee.cho@hcu.ac.th
doi: 10.12982/JAMS.2026.044
E-ISSN: 2539-6056

as musculoskeletal[4] and cognitive decline.[5] Even a short period of physical inactivity such as 14 days can lead to reduced skeletal muscle protein synthesis, resulting in muscle mass loss, decreased muscle strength, and ultimately, skeletal muscle atrophy.[4] Moreover, low physical activity or sedentary behavior has been shown to have detrimental effects on postural balance,[6,7] which refers to the ability to maintain the body's center of mass within its base of support. Balance consists of both static and dynamic components,[8] regulated through various sensory inputs[9] and cognitive demands.[10] Several cognitive domains contribute to balance maintenance, particularly executive functions, which play a crucial role in an individual's work-life success.[11] Among these, inhibitory control and working memory are especially important.[12] Working memory is responsible for temporarily storing and processing information necessary to perform tasks.[13] It is known to develop rapidly between the ages of 18 and 30,[14] playing a vital role in postural control.[15] Notably, previous studies have found an association between balance dysfunction and mild cognitive impairment in young adults.[16] Given that young adult women have not received as much attention in physical activity promotion,[17] developing an intervention that enhances both balance and working memory in this population is essential.

Engaging in physical activity or exercise has been shown to have positive effects on postural balance and cognitive function.[5,6] A study investigating balance performance in young adults aged 17–25 years found that balance improved following two weeks of wobble board training, performed three days per week. Researchers suggested that such training enhances balance by improving sensorimotor processing for maintaining stability in unexpected situations.[18] Furthermore, balance exercises have been identified as a form of physical activity that can enhance both balance performance and cognitive function, particularly in memory and spatial cognition, although executive function remained unchanged.[19] However, a recent meta-analysis revealed that combining exercise with working memory training provides greater benefits for improving executive function, particularly working memory and inhibitory control.[20] A study examining two-day dual-task training found that integrating dynamic balance exercises on a perturbation platform with a cognitive task, such as serial three subtractions, effectively enhanced both balance and cognitive performance compared to single-task training.[21] According to the American College of Sports Medicine (ACSM) guidelines, balance training should be performed for 20–30 minutes per day, at least a few days per week, to achieve optimal benefits.[22]

Therefore, this study aimed to investigate the effects of motor-cognitive training, which combined balance exercises and working memory tasks over three consecutive days, on balance performance and working memory in young adult women aged 18–25 years with low physical activity levels. We hypothesize that three consecutive days of motor-cognitive dual-task training will lead to statistically significant improvements in both balance performance and working memory in young adult women.

Materials and methods

This study was conducted as a randomized controlled trial with an assessor-blinded design. Based on a previous study,[18] the sample size was calculated using G*Power 3.1.9.7. The calculation was based on the difference between two independent means (two groups) at an α level of 0.05, power of 0.95, and an effect size of 1.21, which determined that 20 participants per group were required. To account for an estimated 20% dropout rate, the total number of participants was increased to 48 (24 per group). All participants provided written informed consent before participation. Ethical approval was obtained from the Human Research Ethics Committee of Huachiew Chalermprakiet University, Thailand (No. 1307/2566), and the study was registered under TCTR20250104001.

The inclusion criteria were as follows: 1) females aged 18-25 years,[23] 2) BMI of 18.50-22.99, and 3) low physical activity (<600 MET-minute/week), assessed using the Thai version of International Physical Activity Questionnaire short form (Thai IPAQ-SF).[24] The Thai IPAQ-SF is a last 7 days self-administered questionnaire, consisted of 7 items involved in a time spent on physical activity such as leisure, recreation activity, walking, and sitting. The total time spent (minute/week) was computed to a total metabolic equivalent score (MET-minute/week). Participants with a history of neurological or musculoskeletal diseases or those with balance impairments due to illness were excluded. The simple random sampling was used, and participants were randomly assigned using a closed-envelope method. Two participants dropped out during the training as represented in Figure 1 (CONSORT flow diagram), leaving a total of 46 participants who completed the study. The study assessed static balance, dynamic balance, and working memory skills. Static balance was measured using the sway index from the Balance Error Scoring System (BESS), assessed via the Biodex Balance System (Biodex Balance System™ SD Version 4.x, Biodex Medical System Inc., USA). Participants performed six balance conditions for 30 seconds each, barefoot, on firm and foam surfaces in the following stances: Double-leg stance (firm/foam), Single-leg stance (firm/foam), and Tandem stance (firm/foam). Dynamic balance was evaluated using the Limit of Stability (LOS) test, expressed as a percentage of standard performance, which automatically computed by the Biodex software with age-matched normative data. Participants stood barefoot and shifted their weight as far as possible without losing balance in eight directions: forward, forward/right, right, backward/

right, backward, backward/left, left, and forward/left. Working memory skills were assessed using the Digit Span Test, which included forward and backward digit recall.[25] The participants were asked to listen to the set of 2-9 digits and repeat it in the same sequences (forward digit span), and in the reverse sequences (backward digit span). The total score of forward and backward digit spans was counted from correctly repeated sequences and transformed to percentile equivalents from norms, based on Martin Turner and Jacky Ridsdale. All test data were averaged for analysis. The motor-cognitive training consisted of three progressive steps, each lasting 90 seconds per session for seven sessions per step: 1) balancing on a wobble board (7 sessions), 2) standing on a firm surface while performing backward digit recall (7 sessions), and 3) balancing on a wobble board while performing backward digit recall (7 sessions). A 15-second rest period was provided between sessions. The wobble board used in training was a 15-degree tilted circular platform. A total of 21 sessions (approximately 30 minutes/day) were conducted over three consecutive days. The experimental group (N=23) received the motor-cognitive training. The control group (N=23) were not received the training; however, they were got some knowledge about dual-task training to promote a balance and working memory after the study.

Demographic and baseline data were collected before training and compared between group using Mann-Whitney U Test. Post-training assessments were conducted within 24 hours after the last training session. Since the data were not normally distributed, which verified using the Shapiro-Wilk test, non-parametric statistical tests were used. Wilcoxon Signed-Rank Test was used to compare pre- and post-training effects within each group. Mann-Whitney U Test was used to compare the change scores (post-training minus pre-training) between the two groups. All significant differences were tested at a level of 0.05.

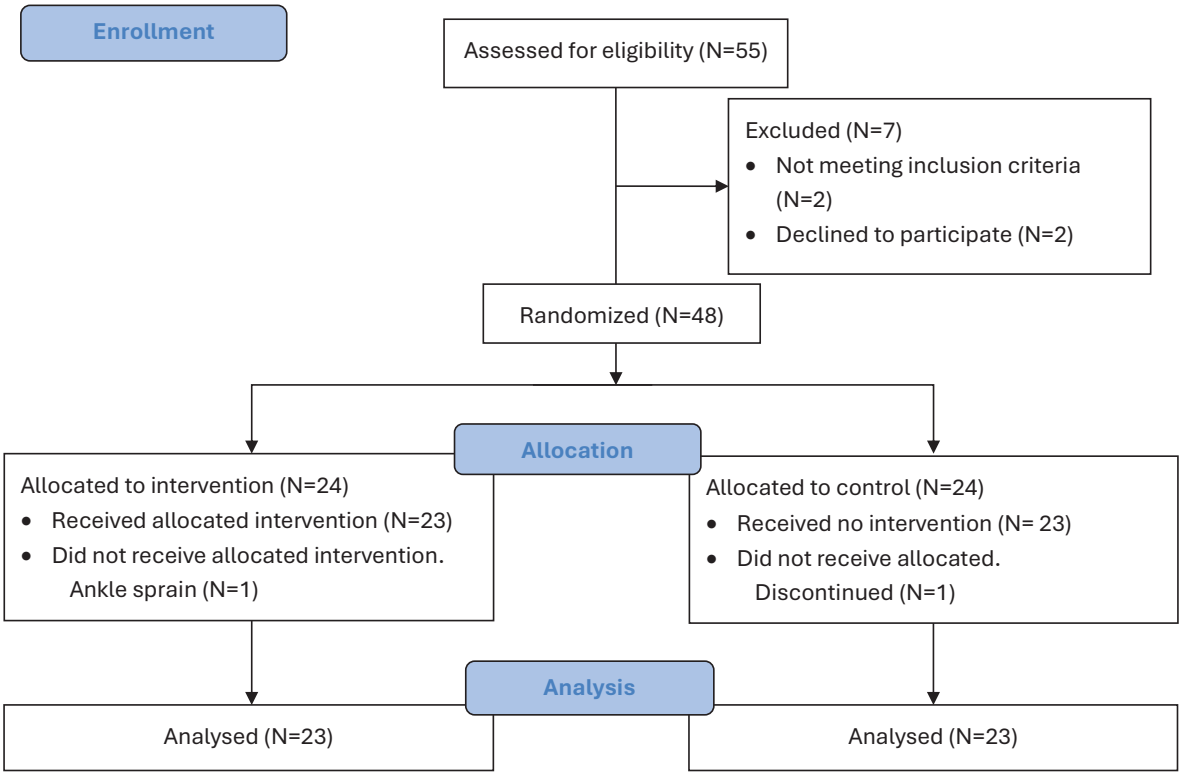


Figure 1. *The CONSORT flow diagram of the participants*

Results

There were no significant differences in demographic characteristics, including age, height, weight, body mass index (BMI), metabolic equivalents (METs), and dominant limbs, as represented in Table 1.

Balance performance results for both groups are presented in Tables 2 and 3. Static balance, assessed using the Balance Error Scoring System (BESS) and represented by the sway index (Table 2), showed no significant difference between the experimental and control groups. Furthermore, within the control group, a significant decrease in the sway index was observed in the tandem stance on a firm surface post-training (*p* 0.02).

Dynamic balance, represented by the percentage of the standard limit of stability (LOS) (Table 3), showed significant difference between the experimental and control groups in the forward direction, both pre-training and post-training when using Mann Whitney test. However, the change score demonstrated there were no significant differences between group all directions. Furthermore, within the experimental group, a significant increase in the % standard of LOS was observed in the forward/right direction (p=0.03), backward/left direction (p=0.02), and the average score (p=0.04) post-training. Similarly, within the control group, a significant increase in the % standard of LOS was noted in the right (p=0.03), backward/right (p=0.00), and backward (p=0.02) directions, as well as in the average score (p=0.01) post-training.

Table 4 presented the percentile equivalent of working memory skills between the experimental and control groups. The experimental group demonstrated a significantly higher percentile equivalent compared to the control group after three consecutive days of training (p<0.001). In addition, the total forward and backward scores, as well as the standard score of working memory skills, were significantly higher in the experimental group than in the control group. However, within-group analysis revealed that all parameters significantly increased after three consecutive days of training in both groups.

Table 1. Baseline characteristics of all participants (N=46).

Characteristics	Experimental group (N=23)	Control group (N=23)	*p* value
Age (year)	20.65±1.22	20.52±1.16	0.71
Height (centimeter)	158.43±4.74	159.74±5.22	0.45
Weight (kilogram)	52.28±5.16	50.98±4.78	0.37
BMI (kilogram/meter^2)	20.80±1.56	19.96±1.38	0.08
METs	142.34±196.93	106.94±182.87	1.00
Dominant hand			0.32
right, N (%)	21 (91.30)	21(91.30)	
left, N (%)	2 (8.70)	2 (8.70)	
Dominant leg			0.60
right, N (%)	22 (95.65)	23 (100.00)	
left, N (%)	1 (4.35)	0	

Note: *Significant differences tested by Mann Whitney Test, abbreviation; METs: metabolic equivalents.*

Table 2. Static balance performance in the experimental and control groups.

Static balance performance (BESS; sway index)	**Groups**	**Pre-training** (mean±SD)	**Pre-training** (mean±SD)	***p* value**[a]	**Change score** (mean±SD)
Firm surface					
Double leg stance	Experimental	0.33±0.36	0.23±0.06	0.96	-0.10±0.34
	Control	0.28±0.16	0.27±0.08	0.42	-0.01±0.15
	p value[b]	0.28	0.14		0.62
Single leg stance	Experimental	0.81±0.66	0.65±0.16	0.11	-0.16±0.62
	Control	0.67±0.15	0.67±0.13	0.62	0.01±0.15
	p value[b]	0.59	0.69		0.15

Table 2. Static balance performance in the experimental and control groups (continued).

Static balance performance (BESS; sway index)	**Groups**	**Pre-training** (mean±SD)	**Pre-training** (mean±SD)	***p* value**[a]	**Change score** (mean±SD)
Tandem stance	Experimental	0.51±0.17	0.47±0.10	0.62	-0.03±0.17
	Control	0.54±0.12	0.49±0.09	0.02*	-0.05±0.10
	p value[b]	0.14	0.54		0.37
Foam surface					
Double leg stance	Experimental	0.52±0.22	0.46±0.11	0.10	-0.06±0.18
	Control	0.54±0.15	0.50±0.10	0.20	-0.04±0.16
	p value[b]	0.26	0.19		0.79
Single leg stance	Experimental	1.03±0.68	0.84±0.24	0.44	-0.19±0.66
	Control	0.94±0.23	0.86±0.21	0.13	-0.07±0.24
	p value[b]	0.25	0.57		0.60
Tandem stance	Experimental	0.77±0.36	0.67±0.15	0.39	-0.09±0.39
	Control	0.70±0.17	0.72±0.18	0.62	0.01±0.18
	p value[b]	0.92	0.51		0.34
Average score	Experimental	0.66±0.37	0.56±0.10	0.25	-0.10±0.36
	Control	0.64±0.16	0.59±0.10	0.08	-0.05±0.13
	p value[b]	0.28	0.25		0.61

Note: **Significance tested at 0.05, [a]significant difference tested by Wilcoxon signed-rank Test, [b]significant difference tested by Mann-Whitney Test.*

Table 3: Dynamic balance performance in the experimental and control groups.

Dynamic balance performance (LOS; % of Standard)	**Groups**	**Pre-training (mean±SD)**	**Post-training (mean±SD)**	***p* value** [a]	**Change score (mean±SD)**
Forward	Experimental	62.57±9.78	62.70±14.56	0.41	0.13±8.97
	Control	57.57±6.99	59.39±5.96	0.15	1.83±5.34
	p value[b]	0.04*	0.04*		0.88
Forward/Right	Experimental	83.91±9.69	87.17±11.95	0.03*	3.26±5.99
	Control	82.17±9.30	83.13±10.02	0.41	0.96±8.91
	p value[b]	0.39	0.19		0.47
Right	Experimental	92.70±11.19	98.70±11.02	0.30	6.00±8.32
	Control	94.57±12.46	99.65±12.54	0.03*	5.09±11.16
	p value[b]	0.50	0.70		0.50
Backward/Right	Experimental	128.43±17.19	126.78±19.00	0.34	-1.65±18.91
	Control	123.52±22.99	133.17±14.59	0.00**	9.65±19.04
	p value[b]	0.33	0.42		0.25
Backward	Experimental	113.04±20.08	118.96±14.26	0.12	5.91±15.66
	Control	109.91±15.64	117.17±13.19	0.02*	7.26±14.04
	p value[b]	0.22	0.65		

Table 3: Dynamic balance performance in the experimental and control groups (continued).

Dynamic balance performance (LOS; % of Standard)	Groups	Pre-training (mean±SD)	Post-training (mean±SD)	*p* value [a]	Change score (mean±SD)
Backward/Left	Experimental	127.43±20.16	132.96±16.70	0.02*	5.52±18.95
	Control	123.70±18.14	130.87±15.23	0.06	7.17±20.08
	p value[b]	0.28	0.65		0.87
Left	Experimental	92.13±16.01	96.96±12.61	0.06	4.83±14.98
	Control	95.74±14.11	100.39±9.25	0.19	4.65±13.72
	p value[b]	0.50	0.44		0.28
Forward/Left	Experimental	81.35±15.00	84.70±16.95	0.16	3.35±16.08
	Control	80.96±10.87	80.00±7.50	0.64	-0.96±11.43
	p value[b]	0.88	0.06		0.15
Average score	Experimental	97.74±11.39	101.09±11.56	0.04*	3.35±8.91
	Control	95.96±9.09	100.48±8.00	0.01*	4.52±7.56
	p value[b]	0.37	0.65		0.83

Note: **Significance tested at 0.05, **significance tested at 0.001, [a]significant difference tested by Wilcoxon signed-rank Test, [b]significant difference tested by Mann-Whitney Test.*

Table 4: Working memory skills in the experimental and control groups.

Working memory skill (Digit span test)	Groups	Pre-training (mean±SD)	Post-training (mean±SD)	*p* value[a]	Change score (mean±SD)
Total forward and backward	Experimental	17.26±3.74	23.04±2.83	<0.001**	5.78±3.19
	Control	16.91±3.01	18.09±3.36	0.005*	1.17±1.75
	p value[b]	0.66	<0.001**		<0.001**
Percentile equivalent	Experimental	46.39±27.30	83.89±14.81	<0.001**	37.50±24.08
	Control	39.65±22.58	49.13±24.74	0.005*	9.48±15.11
	p value[b]	0.43	<0.001**		<0.001**

Note: **Significance tested at 0.05, **significance tested at 0.001, [a]significant difference tested by Wilcoxon signed-rank Test, [b]significant difference tested by Mann- Whitney Test.*

Discussion

Our findings indicated that motor-cognitive training can enhance both dynamic balance performance and working memory skills in young adult women within three days. The control group also exhibited a potential increase in balance and working memory over the following three days after baseline data collection. This may be due to the continuous development of working memory in young adults,[14] which is related to balance performance.[15] Consequently, the control group demonstrated significant differences after the short-term intervention.

Previous research reported that the 30-minute motor-cognitive training was not more effective than the single task training. They suggested that both single-task and dual-task intervention can improve attention, shifting and working memory ability in young men.[26] Similarly, our results showed that working memory after the short period of motor-cognitive training was improved in both the experimental and control group. Few studies have identified the effect of short-term dual-task training on executive function. However, dual-task training has been shown to influence both motor and cognitive performance.[27] Repeated dual-task practice can enhance working memory organization.[28] One study on young adults aged 18-25 years found that participants who walked and maintained balance while performing a visuospatial working memory task demonstrated improved working memory related to dual-task gait interference.[29] The digit span test, which includes both forward and backward digit recall, has been used in cognitive training. Notably, backward digit span training has been found to be more effective in enhancing working memory than forward digit span training or no intervention. This improvement was associated with increased event-related potential (ERP) activity in specific brain areas, providing a more direct measure of neural activity.[30] The activation of

the prefrontal cortex and cerebellum plays a crucial role in both voluntary movement and cognitive tasks.[31] The cerebellum is primarily involved in sensory-motor function,[32] while the prefrontal cortex is responsible for higher-order cognitive processing.[33] So, another aspect of executive function and cognition is necessary for exploring.

Previous study reported that combining wobble board balance training with cognitive training has been proved to improve both static and dynamic balance performance. This improvement is likely due to enhanced proprioceptive sensation at the ankle joint and increased duration of center of pressure maintenance after training.[34] Although the training tended to decrease the sway index after three days, the significant result indicated greater stability in tandem stance within the control group. The practice effect may occur at baseline test over the short period of study, which is a result of rapid neural adaptation[35] and reduces cognitive load.[36]

A complex balance training on a moveable platform has been shown to influence brain structure and function. Gray matter in the prefrontal cortex increased continuously.[37] Furthermore, unstable surfaces have been found to positively impact dynamic balance following balance exercises.[38] In contrast, the current study did not show improvements in dynamic balance over the short period of motor-cognitive training. This may be due to the cognitive interference imposed by performing a balance task concurrently with a cognitive task. The negative effect of motor-cognitive training on static and dynamic balance performance in youth, as cognitive load increases during postural control, making it difficult to allocate attentional resources to both tasks.[39] Moreover, the dynamic balance training relies on movement of body's position and rapid movement in any direction.[40] The current study used a wobble board, which may not have enough external perturbations to challenge the center of pressure move outside the base of support.

Although motor-cognitive training is an effective intervention for promoting both balance and working memory, certain study limitations should be considered. Factors such as activities of daily living during data collection may have influenced the results. Additionally, using the same set of digit recall tasks for the digit span test in both pre-training and post-training assessments for the control group may have introduced learning and retention effects in short-term intervention. Future research should investigate the long-term effects of motor-cognitive task training and consider controlling for confounding factors that may influence dual-task training outcomes.

Conclusion

The findings indicate that motor-cognitive training was superior to the control condition, as it led to greater improvements in balance and working memory among young adult women with low physical activity levels. Thus, the training may have potential as a strategy for promoting balance and working memory in this population.

Ethical approval

The study was approved by the Human Research Ethics Committee of Huachiew Chalermprakiet University, Thailand (No. 1307/2566).

Funding

No funding support

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Supanee Chounchay: conceptualization, formal analysis, methodology, writing-original draft, review and editing, **Pinpapha Retsuk:** methodology, data collection, **Parintorn Mantam:** methodology, data collection, **Kannika Seetha:** methodology, data collection, **Warangkana Arsadee:** methodology, data collection, **Aparat Chanchuphon:** methodology, data collection, **Nattida Pantub:** methodology, data collection, **Supitcha Wongphasuksathaporn:** methodology, data collection.

Acknowledgements

We would like to thankfulness for our participants, who kind and willing to join the study.

References:

[1] Corder K, Winpenny E, Love R, Brown HE, White M, van Sluijs E. Change in physical activity from adolescence to early adulthood: a systematic review and meta-analysis of longitudinal cohort studies. Br J Sports Med. 2019; 53(8): 496-503. doi.org/10.1136/bjsports-2016-097330.

[2] Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. Lancet Glob Health. 2018; 6(10): e1077-86. doi.org/10.1016/S2214-109X(18)30357-7.

[3] Liu R, Menhas R, Saqib ZA. Does physical activity influence health behavior, mental health, and psychological resilience under the moderating role of quality of life? Front Psychol. 2024; 15: 1349880. doi.org/10.3389/fpsyg.2024.1349880.

[4] Bowden Davies KA, Pickles S, Sprung VS, Kemp GJ, Alam U, Moore DR, et al. Reduced physical activity in young and older adults: metabolic and musculoskeletal implications. Ther Adv Endocrinol Metab. 2019; 10: 1-15. doi.org/10.1177/2042018819888824.

[5] Cox EP, O'Dwyer N, Cook R, Vetter M, Cheng HL, Rooney K, et al. Relationship between physical

activity and cognitive function in apparently healthy young to middle-aged adults: A systematic review. J Sci Med Sport. 2016; 19(8): 616-28.doi.org/10.1016/j.jsams.2015.09.003.

[6] Onofrei RR, Amaricai E. Postural balance in relation with vision and physical activity in healthy young adults. Int J Environ Res Public Health. 2022; 19(9): 5021. doi.org/10.3390/ijerph19095021.

[7] Petroman R, Rata AL. Balance performance in sedentary and active healthy young individuals-a cross-sectional study. Phys Educ Stud. 2020; 24(2): 115-9. doi.org/10.15561/20755279.2020.0207.

[8] Susan B O'sullivan, Thomas J Schmitz, George Fulk. Physical rehabilitation. 7th Ed. Philadelphia: F.A. Davis; 2019.

[9] Woollacott MH, Shumway-Cook A. Changes in posture control across the life span-a systems approach. Phys Ther. 1990; 70(12): 799-807. doi.org/10.1093/ptj/70.12.799.

[10] Lajoie Y, Teasdale N, Bard C, Fleury M. Attentional demands for static and dynamic equilibrium. Exp Brain Res. 1993; 97(1): 139-44. doi.org/10.1007/bf00228824.

[11] Welsh M, Schmitt-Wilson S. Executive function, identity, and career decision-making in college students. SAGE Open. 2013; 3(4): 2158244013505755. doi.org/10.1177/2158244013505755.

[12] Divandari N, Bird M-L, Vakili M, Jaberzadeh S. The association between cognitive domains and postural balance among healthy older adults: A systematic review of literature and meta-analysis. Curr Neurol Neurosci Rep. 2023; 23(11): 681-93. doi.org/10.1007/s11910-023-01305-y.

[13] Baddeley A. Working memory. Science. 1992; 255 (5044): 556-9. doi.org/10.1126/science.1736359.

[14] Andre J, Picchioni M, Zhang R, Toulopoulou T. Working memory circuit as a function of increasing age in healthy adolescence: A systematic review and meta-analyses. NeuroImage Clin. 2016; 12: 940-8. doi.org/10.1016/j.nicl.2015.12.002.

[15] Doumas M, Rapp MA, Krampe RT. Working memory and postural control: adult age differences in potential for improvement, task priority, and dual tasking. J Gerontol B Psychol Sci Soc Sci. 2009; 64(2): 193-201. doi.org/10.1093/geronb/gbp009.

[16] Calderón-Garcidueñas L, Torres-Solorio AK, Kulesza RJ, Torres-Jardón R, González-González LO, García-Arreola B, et al. Gait and balance disturbances are common in young urbanites and associated with cognitive impairment. Air pollution and the historical development of Alzheimer's disease in the young. Environ Res. 2020; 191(110087): 110087. doi.org/10.1016/j.envres.2020.110087.

[17] Peng B, Ng JYY, Ha AS. Barriers and facilitators to physical activity for young adult women: a systematic review and thematic synthesis of qualitative literature. Int J Behav Nutr Phys Act. 2023; 20(1): 23. doi.org/10.1186/s12966-023-01411-7.

[18] Yalnaz U, Sarwar U, Quresh S, Saeed HS, Manzoor I, Rasheed N. Effectiveness of wobble board exercises on dynamic balance among Undergraduate students. J Univ Med Dent Coll. 2022; 13(2): 391-5. doi.org/10.37723/jumdc.v13i2.737.

[19] Rogge A-K, Röder B, Zech A, Nagel V, Hollander K, Braumann K-M, et al. Balance training improves memory and spatial cognition in healthy adults. Sci Rep. 2017; 7(1): 5661. doi.org/10.1038/s41598-017-06071-9.

[20] Ludyga S, Held S, Rappelt L, Donath L, Klatt S. A network meta-analysis comparing the effects of exercise and cognitive training on executive function in young and middle-aged adults. Eur J Sport Sci. 2023; 23(7): 1415-25. doi.org/10.1080/17461391.2022.2099765.

[21] Kiss R, Brueckner D, Muehlbauer T. Effects of single compared to dual task practice on learning a dynamic balance task in young adults. Front Psychol. 2018; 9: 311. doi.org/10.3389/fpsyg.2018.00311.

[22] Pescatello LS, Arena R, Riebe D, Thompson PD. ACSM's guidelines for exercise testing and prescription. 9th ed. Philadelphia (PA): American College of Sports Medicine; 2014.

[23] Arnett JJ. Emerging adulthood: A theory of development from the late teens through the twenties. Am Psychol. 2000; 55(5): 469-80. doi.org/10.1037/0003-066x.55.5.469.

[24] Rattanawiwatpong P, Khunphasee A, Pongurgsorn C, Intarakamhang P. Validity and reliability of the Thai version of short format international physical activity questionnaire (IPAQ). J Thai Rehabil. 2006; (16): 147-60.

[25] Turner M, Ridsdale J. The digit memory test. Dyslexia Int. 2004.

[26] Zhang W, Liu H, Zhang T. Immediate and short-term effects of single-task and motor-cognitive dual-task on executive function. PLoS One. 2023; 18(8): e0290171. doi.org/10.1371/journal.pone.0290171.

[27] Beurskens R, Brueckner D, Muehlbauer T. Effects of motor versus cognitive task prioritization during dual-task practice on dual-task performance in young adults. Front Psychol. 2020; 11: 581225. doi.org/10.3389/fpsyg.2020.581225.

[28] Schubert T, Kübler S, Strobach T. A mechanism underlying improved dual-task performance after practice: Reviewing evidence for the memory hypothesis. Psychon Bull Rev. 2024; 31(5): 2005-21. doi.org/10.3758/s13423-024-02498-0.

[29] Hocking DR, Fritsche S, Farhat H, Atkinson A, Bendak H, Menant J. Working memory is a core executive function supporting dual-task locomotor

performance across childhood and adolescence. J Exp Child Psychol. 2020; 197: 104869. doi.org/10.1016/j.jecp.2020.104869.

[30] Li Y, Fu W, Zhang Q, Chen X, Li X, Du B, et al. Effects of forward and backward span trainings on working memory: Evidence from a randomized controlled trial. Psychophysiology. 2023; 60(1): e14154. Available from: doi.org/10.1111/psyp.14154.

[31] Koziol LF, Budding D, Andreasen N, D'Arrigo S, Bulgheroni S, Imamizu H, et al. Consensus paper: the cerebellum's role in movement and cognition. Cerebellum. 2014; 13(1): 151-77. doi.org/10.1007/s12311-013-0511-x.

[32] Ito M. Control of mental activities by internal models in the cerebellum. Nat Rev Neurosci. 2008; 9(4): 304-13. doi.org/10.1038/nrn2332.

[33] Diamond A. Executive functions. Annu Rev Psychol. 2013; 64(1): 135-68. doi.org/10.1146/annurev-psych-113011-143750.

[34] Taghavi Asl A, Shojaedin SS, Hadadnezhad M. Comparison of effect of wobble board training with and without cognitive intervention on balance, ankle proprioception and jump landing kinetic parameters of men with chronic ankle instability: a randomized control trial. BMC Musculoskelet Disord. 2022; 23(1): 888. doi.org/10.1186/s12891-022-05706-x.

[35] Bakker, L.B.M., Lamoth, C.J.C., Vetrovsky, T. et al. Neural correlates of balance skill learning in young and older individuals: A systematic review and meta-analysis. Sports Med Open. 2024; 10(1): 3. doi.org/10.1186/s40798-023-00668-3.

[36] Haith AM, Krakauer JW. The multiple effects of practice: skill, habit and reduced cognitive load. Curr Opin Behav Sci. 2018; 20: 196-201. doi.org/10.1016/j.cobeha.2018.01.015.

[37] Taubert M, Draganski B, Anwander A, Müller K, Horstmann A, Villringer A, et al. Dynamic properties of human brain structure: learning-related changes in cortical areas and associated fiber connections. J Neurosci. 2010; 30(35): 11670-7. doi.org/10.1523/JNEUROSCI.2567-10.2010.

[38] Patel JK, Patel VG, Desai NR, Patel NR, Mukkhi MS, Ribadiya PK. Effect of 4 weeks balance training program in healthy young adults: A randomized clinical trial study. Indian J Physiother Occup Ther. 2020; 14(2): 74-9. doi.org/10.37506/ijpot.v14i2.2611.

[39] Lüder B, Kiss R, Granacher U. Single- and dual-task balance training are equally effective in youth. Front Psychol. 2018; 9: 912. doi.org/10.3389/fpsyg.2018.00912.

[40] Petró B, Papachatzopoulou A, Kiss RM. Devices and tasks involved in the objective assessment of standing dynamic balancing - A systematic literature review. PLoS ONE. 2017; 12(9): E018 5188. doi.org/10.1371/journal.pone.0185188.

Journal of Associated Medical Sciences 2026; 59(2): 134-141

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



The study of maximum electromyogram of masseter muscle before and during modified electroconvulsive therapy in psychiatric patients

Namsukh Kongkalai*

Dental Department, Suanprung Psychiatric Hospital, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 17 November 2025
Accepted as revised 5 February 2026
Available online 9 February 2026

Keywords:
Electroconvulsive therapy, masseter muscle, electromyogram, bite guard, oral complications.

ABSTRACT

Background: Electroconvulsive therapy (ECT) is generally used to treat various severe and treatment resistant psychiatric disorders. Nowadays modified ECT is used, injury to teeth and other oral structure remain a risk in ECT despite the use of muscle relaxants due to the electrodes are placed on the temporal, quite close to the masseter muscle. The use of a bite guard is essential to prevent dental and soft tissue injuries; however, limited evidence exists regarding electromyogram (EMG) of the masseter muscle before and during modified ECT.

Objectives: To compare the maximum EMG of the masseter muscle under three conditions: 1) maximal clenching without a bite guard, 2) maximal clenching with a bite guard, and 3) during modified ECT with anesthesia, muscle relaxants and a bite guard and to evaluate post procedural oral complications.

Materials and methods: This cross-sectional study included 40 psychiatric inpatients who undergoing modified ECT at Suanprung Psychiatric Hospital. Surface EMG signals of the masseter muscle were recorded using the Thymatron® System IV (Somatics LLC, USA) under the three conditions: before anesthesia without a bite guard (NB), before anesthesia with a bite guard (B), and during ECT with anesthesia, muscle relaxants and a bite guard (ECT). Differences among conditions were analyzed using repeated measures ANOVA, and independent samples t-tests were used to compare EMG values between groups (<5,451.1 μV vs ≥5,451.1 μV). Significance was set at $p<0.05$. Post procedural oral examinations were performed to identify any trauma or soft-tissue injuries.

Results: All 40 participants completed the study (mean age 35.63±13.17 years). Schizophrenia was the most prevalent diagnosis (45%), and 67.5% of patients were classified as ASA Class II. The maximum EMG amplitude during ECT was significantly higher than in the NB and B conditions ($p<0.001$). No significant differences in EMG amplitude were found between the NB and B conditions ($p>0.05$). Across all participants, EMG values differed significantly among all three conditions ($p<0.001$). Notably, no oral complications were observed with the use of the newly designed silicone bite guard.

Conclusion: Modified ECT induces significantly higher masseter muscle EMG activity approximately 15-fold greater than pre-ECT clenching. The use of newly designed silicone bite guard effectively prevents oral injury during ECT, supporting its routine clinical application to enhance patient safety and minimize oral complications.

* **Corresponding contributor.**
Author's Address: *Dental Department, Suanprung Psychiatric Hospital, Chiang Mai Province, Thailand.*
E-mail address: Dentkawila@hotmail.com
doi: 10.12982/JAMS.2026.045
E-ISSN: 2539-6056

Introduction

Treatment for psychiatric patients primarily involves three aspects: psychological therapy, such

as counseling, encouragement, and understanding the patient; milieu therapy, which includes creating a comfortable and relaxing environment for the patient; and pharmacological therapy. Electroconvulsive therapy (ECT) is a procedure used to treat psychiatric patients with severe symptoms, those who may pose a danger to themselves or others, or those for whom other treatments have been ineffective, or who experience intolerable side effects from medication, or cannot wait for the medication to take effect. An example is patients with severe suicidal ideation who require rapid stabilization of their mental state.

Electroconvulsive therapy is a method that uses electrical current to stimulate the brain, inducing a seizure. During the seizure, increased cerebral blood flow, elevated metabolism, and higher oxygen and glucose consumption are observed, leading to changes in neurotransmitters[1,2] and a balance in neurotransmission.[3] It is generally used in psychiatric patients with severe symptoms, especially those unresponsive to medication, those with medication administration difficulties, those for whom other methods have been ineffective, and those with various complications.[3,4] It is an effective and rapid treatment.[5] There are two types of ECT: Unmodified ECT, which is performed without anesthesia and muscle relaxants. In this type, the patient remains conscious during the procedure, and when the electrical current passes through the patient's temporal, they experience tonic-clonic seizures followed by relaxation. The other type is Modified ECT, which involves the use of anesthetics and muscle relaxants. In this type, the patient is unconscious during the procedure, and the convulsive movements are less severe.[3]

During ECT, patients experience muscle and joint contractions throughout the body, including the temporomandibular joint (TMJ) and jaw muscles. This can result in higher-than-normal biting forces,[6] which may lead to oral problems such as trismus (jaw locking), temporomandibular joint pain, fractured or broken teeth, tooth avulsion, or oral mucosa lacerations.[7] Dentists play an important role in assessing patients' oral health before ECT and by using a bite guard during the procedure, which helps to reduce the incidence and severity of oral and jaw complications.[8]

A study by Watt and colleagues found that the most common complications of ECT were tooth and tongue injuries, often attributed to incorrect bite guard placement or the absence of a bite guard during ECT. A total of 73,440 ECT treatments were administered, resulting in 12 recorded injuries associated with the dental and oral regions; 5 cases involving an injury to a tooth, while 7 cases were related to lacerations of the tongue or lip. In each case, either incorrect bite guard placement or failure to use a bite guard was cited as the reason for the injury. One patient had a minor laceration to his arm because of incomplete paralysis and striking his arm on a bed rail.[9]

Ogami and colleagues found that the average occlusal force before anesthesia without bite guard was 844±412 N. The average occlusal force during ECT was 989±433 N. The average occlusal force before anesthesia with bite guard was 313±150 N. The average occlusal force during ECT was 365±166 N. The main finding of the study was that the use of a custom-made bite guard resulted in an average occlusal force reduction of 58%±22% during modified ECT. This study concluded that occlusal force significantly decreased both before anesthesia and during ECT when a bite guard was used. The bite guard was exceptionally effective at preventing procedure related complications. Without the bite guard, a high incidence of adverse effects was observed: 18 out of 22 patients (82%) reported post-procedural pain and 5 out of 22 patients experienced bleeding in the mouth. In contrast, when the bite guard was used, no complications neither pain nor bleeding.[6]

A positive linear relationship between electromyogram (EMG) and bite force has been reported.[10-14] A study by Gonzalez and colleagues found that the relationship between EMG and bite force of the temporalis and masseter muscles during molar and incisor area clenching was reliable. The research found that the reliability of the slopes describing the relationship between EMG activity and bite force was generally acceptable for the masseter and temporalis muscles, especially for molar biting tasks compared to incisor biting tasks. Test-retest reliabilities were generally above acceptable levels for masseter and temporalis muscles, with molar biting showing more reliable results than incisor biting. However, suprahyoid muscles showed more variable results for both molar and incisor biting tasks and often did not meet reliability criteria. The reliability of masseter and temporalis muscles for molar biting was excellent and ranged from acceptable to excellent for incisor biting, with the potential for further improvements using electrode position templates. The study demonstrated reproducibility of EMG versus bite force data and emphasized the importance of consistent biting conditions for reliability.[15]

A study of oral complication incidence in ECT patients at a psychiatric hospital (Suanprung Psychiatric Hospital, Chiang Mai) shows that a complication cause was the inappropriate bite guard size for the patient's mouth or a bite guard shape that hindered the procedure. Suanprung Psychiatric Hospital has therefore developed and manufactured silicone bite guards in various size; small (S), medium (M), and large (L), with clearly color-coded, to suit the mouth size of each patient and to facilitate ease of use for the physician. A specific type is also available for the patients who require an endotracheal tube, also in S, M, and L sizes. Descriptively, it was found that the incidence of oral complications in ECT patients is decreased from 5.45% to 0.12%.[16] Suanprung Psychiatric Hospital has undertaken a continuous quality improvement initiative for patients

receiving Modified ECT since 2022. Analysis of procedure related oral injuries revealed that in 2022, among 4,106 ECT sessions, 10 cases of oral injury occurred, including one tongue laceration. In 2023, there were 4,751 procedures with 4 oral injuries, again including one tongue laceration. By July 2024, 3,827 procedures had been performed, with 2 oral injuries reported; one case involved a severe tongue laceration requiring suturing.

These findings indicate that tongue lacerations remain a significant clinical concern. In response, a new bite guard was specifically designed to address tongue biting injuries (Figure 1). This design was introduced in August 2024, and no oral injuries have been reported since its implementation. The development of this device was based on clinical experience at Suanprung Psychiatric Hospital, without reference to biomechanical or force related evidence from the scientific literature.

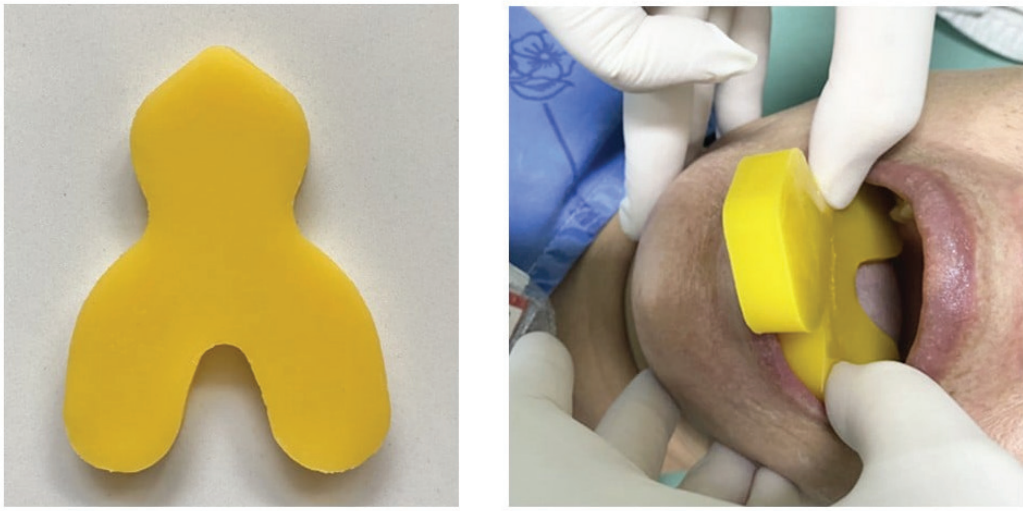


Figure 1. *A new bite guard design with a shape of silicone provides a space for the tongue to prevent biting the tongue.*

Currently, Suanprung Psychiatric Hospital exclusively uses Modified ECT with the Thymatron® System IV (Somatics, LLC, USA), which includes monitoring for electroencephalogram (EEG), electrocardiogram (ECG), and electromyogram (EMG). Surface electrodes are used for EMG signal recording. Although patients undergoing Modified ECT receive anesthetic agents and muscle relaxants before the electrical current is applied and are unconscious with reduced convulsive movements during the procedure. To our best knowledge, there are no study examining the relationship of masseter muscle EMG values in patients during Modified ECT compared to when they clench their teeth before receiving anesthetics and muscle relaxants.

Therefore, The purpose of this study was to compare the maximum EMG of the masseter muscle under three conditions: 1) prior to administration of anesthetic agents and muscle relaxants, during maximal clenching without a bite guard, 2) prior to administration of anesthetic agents and muscle relaxants, during maximal clenching with a bite guard, and 3) during ECT, after administration of anesthetic agents and muscle relaxants, while the patient was wearing a bite guard and to evaluated the complications during ECT.

Materials and methods

This study was conducted in the inpatient department of Suanprung Psychiatric Hospital. Ethical approval was obtained from the Human Ethics Committee for Psychiatric Clinical Research at Srithanya Hospital (Approval No. STY.COA007/2568).

Inclusion criteria

1. Psychiatric inpatients scheduled for modified electroconvulsive therapy (ECT) at Suanprung Psychiatric Hospital
2. Both male and female patients, irrespective of psychiatric diagnosis
3. Age ≥18 years
4. Possessing ≥20 natural teeth
5. Able to communicate and follow instructions

Exclusion criteria

1. Patients unable to cooperate due to psychiatric symptoms or poor comprehension
2. Fewer than 20 teeth present
3. Presence of Grade 3 tooth mobility (horizontal mobility >2 mm and/or vertical mobility)

Data collection

All data were collected by a single officially trained researcher. In accordance with standard ECT patient care protocols, all participants underwent an oral health examination before treatment. Eligible patients received a verbal explanation of the study, and written informed consent was obtained prior to participation. Participation was entirely voluntary.

The following data were collected:

1. Dental examination: A complete clinical examination was performed, and findings were recorded in a standardized dental chart.
2. Personal data: Demographic and clinical information were obtained from the patients' medical records.
3. EMG monitoring: EMG of the masseter muscle was recorded using the Thymatron® System IV (Somatics LLC, USA), which also provides EEG and ECG monitoring. Surface EMG electrodes (Somatics Cat. #EEDS) were placed randomized unilaterally over the masseter muscles following skin preparation (cleansing with alcohol swabs and drying). Electrodes were connected to the EMG input channel. The device's maximum measurable EMG amplitude was 5,451.1 μV.
4. Measurement protocol: Masseter EMG activity was recorded under three conditions (Figure 2):
   - (NB) Prior to administration of anesthetic agents and muscle relaxants, without a bite guard: patients were instructed to clench maximally for 3 seconds.
   - (B) Prior to administration of anesthetic agents and muscle relaxants, with a bite guard inserted: patients performed maximal clenching for 3 seconds.
   - (ECT) During the ECT procedure: after receiving propofol (0.75-1.0 mg/kg) and succinylcholine (0.3-1.0 mg/kg), a bite guard was inserted, electrical stimulus was applied, and the maximum EMG value during the convulsion was recorded.
5. Post-procedural oral examination: After ECT, the bite guard was removed and the oral cavity examined for trauma or soft tissue injury. Findings were recorded.

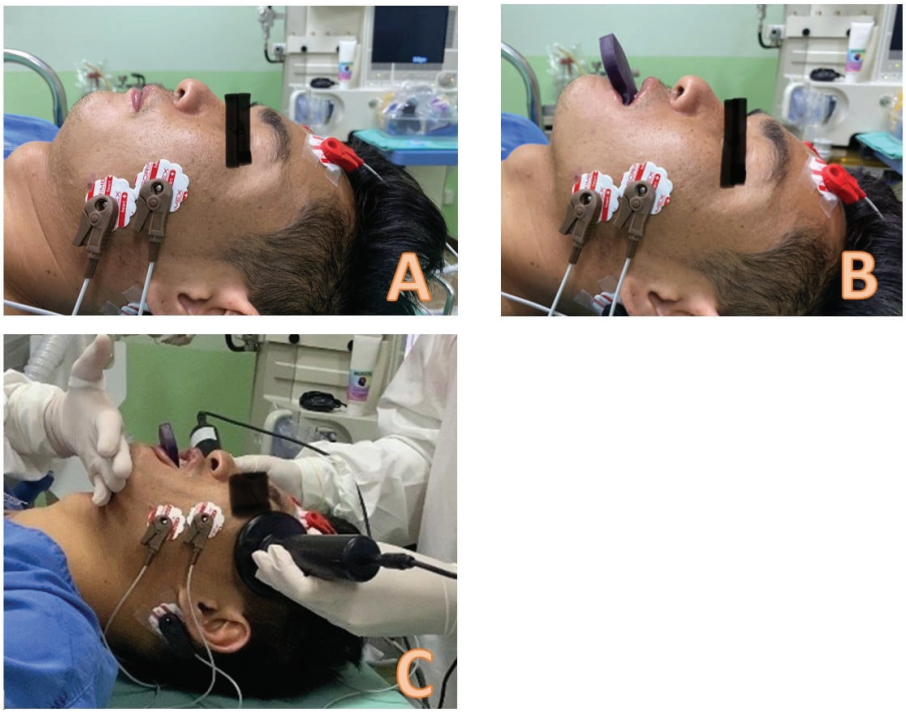


Figure 2. *Measurement protocol. A: prior to administration of anesthetic agents and muscle relaxants, maximal clenching without a bite guard, B: prior to administration of anesthetic agents and muscle relaxants, maximal clenching with a bite guard, and C: during ECT, after administration of anesthetic agents and muscle relaxants, while the patient was wearing a bite guard*

Data analysis

Data normality was evaluated using the Shapiro-Wilk test. Differences in EMG amplitudes across conditions (NB, B, ECT) were analyzed using repeated measures ANOVA with Pairwise comparison post hoc. Independent samples t-tests were used to compare EMG values between groups (<5,451.1 μV vs ≥5,451.1 μV). All analyses were performed using SPSS version 29.0 (IBM Corp., USA), and statistical significance was set at $p<0.05$.

Results

A total of 40 patients (20 males and 20 females) participated in the study. Participant ages ranged from 20 to 65 years (mean=35.63±13.17 years). The most common age group was 20-29 years (45%). Body weight ranged from 39 to 80 kg (mean=57.18±10.44 kg).

The most prevalent diagnosis was schizophrenia (F20.0) in 45% of patients, followed by bipolar disorder (F31.2) in 10%. According to the American Society of Anesthesiologists (ASA) physical status classification, 67.5% of participants were classified as ASA Class II.

The mean number of teeth was 27.7±2.33. Demographic data are summarized in Table 1.

Participants were divided into two groups based on maximum masseter EMG amplitude during ECT:

- Group 1: <5,451.1 μV
- Group 2: ≥5,451.1 μV

Independent samples t-tests compared the maximum masseter EMG values across three conditions: before anesthesia without bite guard (NB), before anesthesia with bite guard (B), and during ECT with anesthesia, muscle relaxants, and bite guard (ECT).

Table 1. Demographic data of study participants.

Characteristic	Category	Number (%)	Mean±SD	Min	Max
Sex	Male	20 (50.0)			
	Female	20 (50.0)			
Age (years)	20-29	18 (45.0)	35.63±13.17		
	30-39	8 (20.0)			
	40-49	9 (22.5)			
	50-59	2 (5.0)			
	60-69	3 (7.5)			
Weight (kg)			57.181±0.44	39	80
Diagnosis	Multiple drug use	3 (7.5)			
	Schizophrenia	18 (45.0)			
	Bipolar disorder	4 (10.0)			
	Depression	3 (7.5)			
	Other	12 (30)			
ASA classification	I	13 (32.5)			
	II	27 (67.5)			
Number of teeth			27.70±2.33	23	32

Under the NB and B conditions, no significant differences in EMG amplitude were observed between the two groups (p>0.05). However, in the ECT condition, group 2 exhibited a significantly higher mean EMG amplitude compared to group 1 (p=0.001), with a mean difference of approximately 532.4 μV (Figure 3).

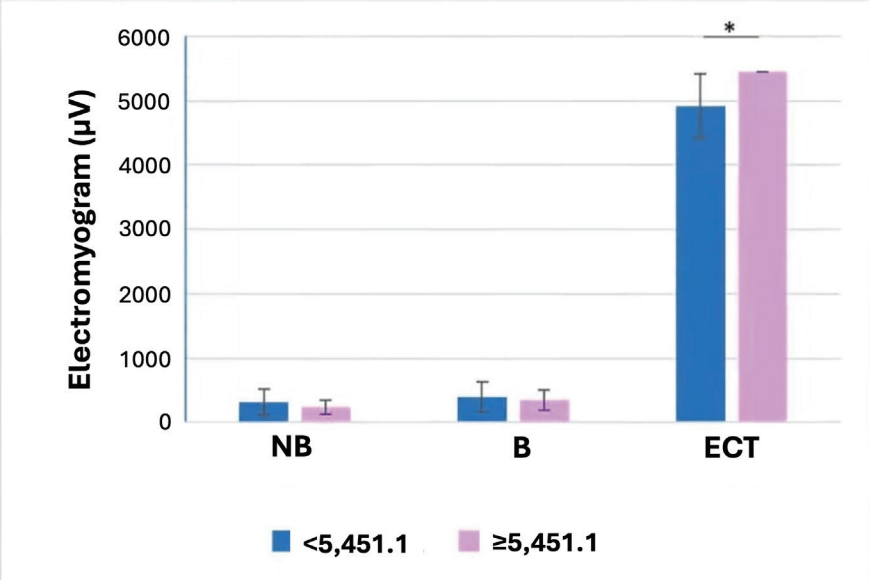


Figure 3. *The bar graph of EMG values before anesthesia and muscle relaxants without a bite guard (NB), before anesthesia and muscle relaxants with a bite guard (B), and during ECT with anesthesia, muscle relaxants and a bite guard (ECT) between both groups (*p=0.001).*

Further analysis revealed that maximum EMG values across all three conditions (NB, B, and ECT) differed significantly ($p<0.001$) in all participants (Figure 4).

Note: with the use of the newly designed bite guard, no oral complications were observed following any ECT session.

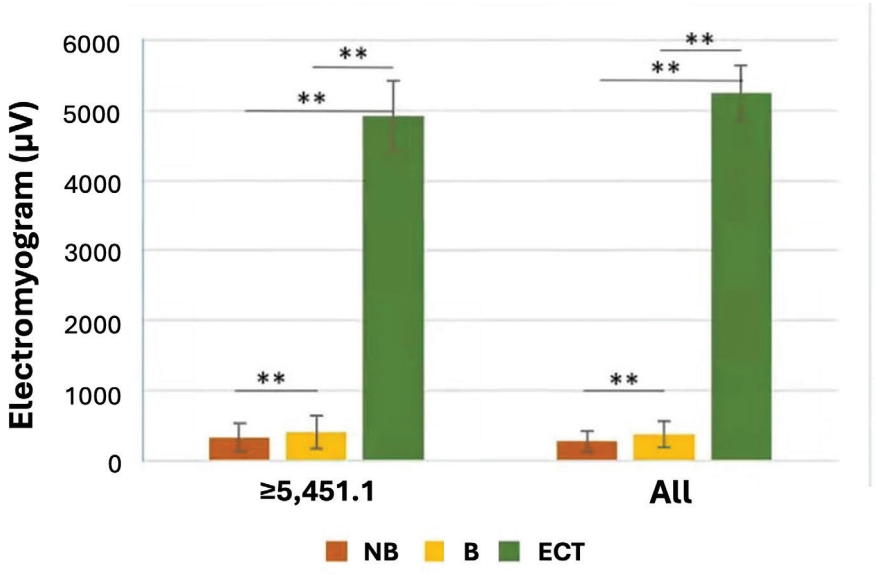


Figure 4. *The bar graph of electromyogram values. NB: before anesthesia and muscle relaxants without a bite guard, B: before anesthesia and muscle relaxants with a bite guard, ECT: during modified ECT with anesthesia and muscle relaxants with bite guard (ECT) in patients with values less than 5,451.1 μV and all patients (**p<0.001).*

Discussion

Electroconvulsive therapy (ECT) is generally used to treat various severe and treatment resistant psychiatric disorders. Nowadays modified ECT is used, this procedure involves administering an anesthetic and a muscle relaxant before passing a controlled amount of electrical current from an ECT machine into a specific brain region for a limited time to induce a generalized tonic–clonic (grand mal) seizure. This seizure causes widespread muscle contractions, including those of the masticatory muscles particularly the temporalis and masseter which are located close to the electrode placement site and may be directly stimulated by the electrical current, leading to a severe, sudden jaw closure.

Even with the use of muscle relaxants, the normal neuromuscular feedback mechanism that prevents excessive biting force does not take place in anesthetized patients. Consequently, injuries to oral and jaw structures may occur. Such injuries include tooth avulsion (which carries a risk of aspiration), tooth fractures, lacerations to the cheek, lips, or tongue, bleeding, and temporomandibular joint pain.[6,17,18] Watt and colleagues reported that the most common complications of ECT were injuries to the mouth, including dental and tongue trauma.[9]

Pre-ECT oral assessment by dentists and the use of appropriate mouth protection have been recommended to prevent injury to the teeth and other oral structures.[19,20] To reduce the risk of oral and dental injury during ECT, the American Psychiatric Association[21] recommends the use of oral protection devices made of flexible materials that cover the molar region.[22,23] However, various types and sizes of bite guards should be available to accommodate the individual characteristics of each patient's oral cavity.[24]

An ideal bite guard should prevent dental injuries without obstructing the airway,[20] be easy to insert and remove, not interfere with the placement or seal of the oxygen mask, be inexpensive, and be easy to clean and sterilize.[17] The severity of the bite reaction in ECT patients, including bite force and EMG activity, remains unclear when compared with maximum voluntary bite force.

In the present study, most patients were diagnosed with schizophrenia and classified as ASA Class II, like a study conducted at Siriraj Hospital, which utilized a rolled gauze bite guard. That study analyzed 217 ECT sessions and reported a significant oral injury incidence of 10.1%. The most common type of injury was mucosal or lip abrasion, accounting for 72.8% of all incidents. A key finding was a direct correlation between the dosage of the muscle relaxant succinylcholine and the likelihood of injury. Patients who received a lower dose (0.9 mg/kg versus 1.0 mg/kg) were significantly more likely to experience oral injury ($p=0.009$). The authors concluded that the incidence of oral injury was not negligible and could be reduced through a multidisciplinary approach involving appropriate anesthetic dosing, the use of properly sized and softer bite guards, and comprehensive pre-procedural dental assessment. [25] The bite guard used in our study was a customized bite guard fabricated from food-grade silicone (Silicone Food Grade USA SF820). No oral injuries were observed

among patients in this study. Similarly, Watt and colleagues demonstrated that custom-made bite guards fabricated individually from 3-mm ethyl-vinyl-acetate (EVA) sheets using impressions and pressure forming to ensure a precise fit were highly effective in preventing ECT-related oral complications. In their study, a high incidence of adverse events occurred without bite guards, while no complications, pain, or bleeding were reported when bite guards were used in 22 patients.[6]

It has been thoroughly searched for the study about the comparison the effectiveness of different bite guard types, there are no any study comparing the effectiveness of bite guard's types, the further study is required, however, the ethical concerns are taken into the considerations.

The bite guards are still recommended to use in edentulous patients, as the bite force is still very high and can potentially cause injury to the temporomandibular joint, lips, and tongue.

A positive linear relationship between EMG activity and bite force has been reported. [10-14] González and colleagues found that EMG activity of the temporalis and masseter muscles during clenching in the molar and incisor regions was strongly correlated with bite force.[15] Bogdanov and colleagues reported that the EMG and bite force relationship varied between the temporalis and masseter muscles; the masseter exhibited a more consistent linear relationship bilaterally, suggesting that masseter EMG can serve as a reliable indirect indicator of bite force in situations where direct measurement is unsafe.[26]

Accordingly, the results of our study, which assessed unilateral masseter muscle activity, demonstrated a significant increase in EMG values during the ECT procedure compared with normal biting force, consistent with previous findings. These results support the use of optimized, customized bite guards to prevent oral injuries. The findings also confirm that maximum EMG activity of the masseter muscle is significantly reduced when a bite guard is used and further altered during ECT under anesthesia. Therefore, the routine use of bite guards in ECT is strongly recommended to minimize oral injury. Moreover, the study emphasizes the need for properly designed, size appropriate, and material optimized bite guards for psychiatric patients undergoing ECT. Further improvements in the design of bite guards for Modified ECT require evidence-based guidance on biomechanical factors. Specifically, research is needed to determine the optimal material, shape, size, and thickness. While minimizing thickness is desirable for patient comfort, the device must still provide adequate force absorption to prevent temporomandibular joint complications and ensure overall safety. Such evidence will be essential for the future development and planning of Modified ECT services.

Limitations

Bite force measurement can be performed using devices equipped with electrical sensors. However, this method is unsuitable during ECT because the device's electrical activity could interfere with the ECT process,[6] representing the first limitation of our study. Another limitation involves the Thymatron® System IV (Somatics, LLC, USA), which monitors EMG activity from the extensor digitorum brevis muscle on the dorsal surface of the foot a site distant from the electrode placement and records a maximum measurable value of 5,451.1 μV. In contrast, our study monitored EMG activity from the masseter muscle, located closer to the electrode site and directly stimulated by the current. This proximity can cause severe, sudden jaw closure, occasionally exceeding the measurable range of the device.

Further research is needed to optimize bite guard design to improve ease of use and ensure secure intraoral placement during ECT. Studies with larger sample sizes are also required to establish average bite force values in ECT patients, particularly if safe and compatible bite force measuring devices can be developed for use during ECT.

Conclusion

Modified ECT induces masseter muscle EMG activity approximately 15 times greater than that observed during pre-ECT maximal clenching. Despite the use of anesthesia and muscle relaxants, substantial masticatory muscle activation persists during ECT.

In summary, EMG monitoring provides valuable objective evidence of masticatory muscle activity during ECT and supports the routine integration of customized silicone bite guards as a standard preventive measure in clinical practice.

Funding

This research did not receive any specific grant or funding from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

Authors declared no conflict of interest.

Acknowledgements

The authors would like to thank study participants, Suanprung Psychiatric Hospital staffs involved and especially Dr. Pised Methapatara, neurologist and Dr. Thanapat Sastraruji, Dental Research Center, Faculty of Dentistry, Chiang Mai University, for supporting statistical data analysis.

References

[1] Singh A, Kar SK. How electroconvulsive therapy works?: understanding the neurobiological mechanisms. ClinPsychopharmacol Neurosci. 2017; 15(3): 210-21. doi: 10.9758/cpn.2017.15.3.210.

[2] Saipanit R. Electroconvulsive therapy (ECT). Articles from Ramamental. 2011, accessed 8 December.
[3] Clinical Dental Practice Guideline for Psychiatric Patients 2014 ISBN 978-974-296-825-0.
[4] Jaffe R. The practice of electroconvulsive therapy: recommendations for treatment, training, and privileging: a task force report of the American Psychiatric Association. American Journal of Psychiatry. 2002; 159(2): 331. doi.org/10.1176/appi.ajp.159.2.331.
[5] What is electroconvulsive therapy (ECT)? [Internet]. 2023, accessed September 2024.
[6] Ogami S, Yamada M, Kanazawa M, Takeda K, Kimura N, Mizutani H, et al. The effectiveness of a mouth guard to protect against strong occlusion caused by modified electroconvulsive therapy. Dent Traumatol. 2014; 30(5): 368-73. doi: 10.1111/edt.12097.
[7] Muzyka BC, Glass M, Glass OM. Oral health in electroconvulsive therapy: a neglected topic. J ECT. 2017; 33(1): 12-5. doi: 10.1097/YCT.0000000000000351.
[8] Varotto BLR, Neves RRN, Nasri-Heir C, Nápole RCDO, Antequera R, Siqueira JTTd. Mouth guards and intra-oral evaluation for patients. 2021. Rev Bras Odontol. 2021; 78: e1952. doi.org/10.18363/rbo.v78.2021.e1952.
[9] Watts BV, Groft A, Bagian JP, Mills PD. An examination of mortality and other adverse events related to electroconvulsive therapy using a national adverse event report system. J ECT. 2011; 27(2): 105-8. doi: 10.1097/YCT.0b013e3181f6d17f.
[10] Bakke M, MICHLER L, Han K, Möller E. Clinical significance of isometric bite force versus electrical activity in temporal and masseter muscles. Scand J Dent Res. 1989; 97(6): 539-51. doi: 10.1111/j.1600-0722.1989.tb00929.x.
[11] Blanksma N, Van Eijden T, Weijs W. Electromyographic heterogeneity in the human masseter muscle. J Dent Res. 1992; 71(1): 47-52. doi: 10.1177/00220345920710010801.
[12] Van Eijden T, Brugman P, Weijs W, Oosting J. Coactivation of jaw muscles: recruitment order and level as a function of bite force direction and magnitude. J Biomech. 1990; 23(5): 475-85. doi: 10.1016/0021-9290(90)90303-k.
[13] Koolstra J, Van Eijden T. Combined finite-element and rigid-body analysis of human jaw joint dynamics. J Biomech. 2005; 38(12): 2431-9. doi: 10.1016/j.jbiomech.2004.10.014.
[14] Uchida S, Iwasaki L, Marx D, Yotsui Y, Inoue H, Nickel J. Variations in activities of human jaw muscles depend on tooth-tipping moments. Arch Oral Biol. 2008; 53(2): 199-205. doi: 10.1016/j.archoralbio.2007.10.002.
[15] Gonzalez Y, Iwasaki L, McCall Jr W, Ohrbach R, Lozier E, Nickel J. Reliability of electromyographic activity vs. biteforce from human masticatory muscles. ur J Oral Sci. 2011; 119(3): 219-24. doi: 10.1111/j.1600-0722.2011.00823.x
[16] Kongkalai N. Suanprung bite guard for use in electroconvulsive therapy (ECT) in psychiatric patients. Dental Council News. 2020; 3: 26-8.
[17] Kiran S, Bala R, Singh T. Dental protection during modified electroconvulsive therapy using roll-gauze mouth gag. J ECT. 2009; 25(1): 74-5. doi: 10.1097/YCT.0b013e3181770224.
[18] Woo S-W, Do S-H. Tongue laceration during electroconvulsive therapy. Korean J Anesthesiol. 2012; 62(1): 101-2. doi: 10.4097/kjae.2012.62.1.101.
[19] Beli N, Bentham P. Nature and extent of dental pathology and complications arising in patients receiving ECT. Psychiatric Bulletin. 1998; 22(9): 562-5. doi.org/10.1192/pb.22.9.562.
[20] Morris AJ, Roche SA, Bentham P, Wright J. A dental risk management protocol for electroconvulsive therapy. J ECT. 2002;18(2): 84-9. doi: 10.1097/00124509-200206000-00004.
[21] The Practice of ECT: Recommendations for treatment, training and privileging. Convuls Ther. 1990; 6(2): 85-120. PMID: 11659302
[22] Minneman SA. A history of oral protection for the ECT patient: past, present, and future. J ECT. 1995; 11(2): 94-103. PMID: 7552060
[23] Paparone P, Ee P-L, Kellner CH. Oral protection in electroconvulsive therapy: Modified technique using 2 bite blocks. J ECT. 2019; 35(4): 224. DOI:10.1097/YCT.0000000000000585
[24] Martin D SB. Dental Issues Related to ECT. Cambridge: Cambridge University Press; 2019.
[25] Jirakulsawat A, Siriussawakul A, Triyasunant N, Wangdee A, Jiraroch Y, Chiripu M. Incidence of oral injury and risk factors associated with oral injury in psychiatric patients undergoing electroconvulsive therapy in Siriraj Hospital. Siriraj Med J. 2012; 64(5): 145-8.
[26] Bogdanov V. Type of correlation between bite force and EMG activity of the temporalis and masseter muscles during maximal and submaximal clenching. Folia Med. 2023; 65(6): 975-85. doi: 10.3897/folmed.65.e107180.

Journal of Associated Medical Sciences 2026; 59(2): 142-152

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Fabrication and evaluation of 3D-printed customizable boluses for optimized radiotherapy applications

Chitchaya Suwanraksa[1], Athitan Khiewded[2], Gadesuda Jankaew[1], Thirawut Rojchanaumpawan[1], Surapong Chatpun[3], Pornchai Phukpattaranont[2*]

[1]*Department of Radiology, Faculty of Medicine, Prince of Songkla University, Songkhla Province, Thailand.*
[2]*Department of Electrical and Biomedical Engineering, Faculty of Engineering, Prince of Songkla University, Songkhla Province, Thailand.*
[3]*Department of Biomedical Sciences and Biomedical Engineering, Faculty of Medicine, Prince of Songkla University, Songkhla Province, Thailand.*

ARTICLE INFO

Article history:
Received 17 November 2025
Accepted as revised 5 February 2026
Available online 9 February 2026

Keywords:
3D printing, radiotherapy, bolus fabrication, polylactic acid.

ABSTRACT

Background: Commercial radiotherapy boluses often fail to conform accurately to patient anatomy, resulting in air gaps and surface dose inaccuracy. Three-dimensional (3D) printing using fused deposition modeling with polylactic acid (PLA) offers a customizable alternative, but clinical implementation requires a reproducible fabrication workflow that ensures both radiological consistency and geometric conformity.

Objectives: To evaluate the effect of infill density on radiological properties and assess the feasibility of a consistent fabrication approach for both standard and patient-specific boluses.

Materials and methods: Standard PLA boluses (10×10×1 cm) were fabricated at 75-95% infill density using a line infill pattern, while customized boluses were created from a head-and-neck phantom and printed at 90-95% infill density. All boluses underwent CT imaging and were analyzed in the Eclipse treatment planning system (TPS) to determine Hounsfield unit (HU), mass density, and relative electron density (RED). Air gap measurements and dose distributions were evaluated in the TPS with a 6 MV photon beam.

Results: As infill density increased from 75%-95%, HU values rose from -184.6 to +48.5, with 90–94% infill yielding HU values within the acceptable ±50 HU range. Standard and customized boluses demonstrated comparable radiological behavior, with mass density and RED near unity. Customized boluses reduced air gaps (0.21±0.02 cm vs 0.54±0.16 cm) and improved surface-dose build-up and isodose coverage in complex regions.

Conclusion: Optimized infill density settings support a reproducible 3D printing workflow. Both standard and customized PLA boluses can achieve radiological accuracy and surface conformity suitable for clinical implementation.

***Corresponding contributor.**
Author's Address: *Department of Electrical and Biomedical Engineering, Faculty of Engineering, Prince of Songkla University, Songkhla Province, Thailand.*
E-mail address: pornchai.p@psu.ac.th
doi: 10.12982/JAMS.2026.046
E-ISSN: 2539-6056

Introduction

Radiotherapy is a fundamental treatment for malignant and non-malignant tumors that utilizes high-energy ionizing radiation to eliminate cancer cells and inhibit tumor growth.[1] Recently, modern radiotherapy techniques have advanced to maximize tumor control while minimizing damage to surrounding healthy tissues.[2] Linear accelerators generate megavoltage photon beams that have the characteristic of delivering peak radiation doses deposited below the skin at

the depth of the maximum dose (Dmax), called a skin-sparing effect.[3] While this characteristic is advantageous for treating deep-seated tumors, it poses challenges for delivering precise surface doses for superficial tumors, including skin cancers, post-mastectomy chest wall recurrences, and superficial lymph node involvement. Achieving adequate dose coverage for these superficial lesions remains a significant challenge.[4,5] To overcome this limitation, a bolus, which is a tissue-equivalent material, is applied to the patient's skin to increase the surface dose by shifting the Dmax toward the skin surface.[6-8] Commercial boluses are mass-produced and lack the flexibility to provide patient-specific design options. They often fail to conform to complex body contours, leading to air gaps between the bolus and the skin surface. These air gaps result in dose inhomogeneity and compromise the accuracy of surface dose delivery.[9,10] In some cases, commercial boluses are manually cut or modified to fit specific patient anatomies, which can lead to material waste and still fail to achieve the desired level of precision.[11] In addition, manual fabrication increases preparation time and operator workload, adding to overall treatment costs and reducing clinical throughput in busy radiotherapy departments. These limitations highlight the need for a fabrication method that is both geometrically accurate and operationally efficient.

In treatment planning, the radiological properties of bolus materials expressed by Hounsfield unit (HU), mass density, and relative electron density (RED) directly influence dose calculation accuracy. Therefore, maintaining consistent HU/RED characteristics between the planned and delivered setup is essential for reliable surface-dose estimation and quality assurance.[12]

While commercial boluses are suitable for routine treatments, advances in additive manufacturing have introduced opportunities to improve workflow efficiency, reproducibility, and customization through patient-specific design. Three-dimensional (3D) printing enables the fabrication of boluses that closely match individual anatomy and maintain stable radiological properties.[13]

Among various additive manufacturing techniques, fused deposition modeling (FDM) has emerged as a particularly promising method for producing patient-specific boluses, constructing physical objects directly from digital models by layering extruded thermoplastic materials.[14] Polylactic acid (PLA) is a thermoplastic widely used in medical applications, including radiotherapy bolus fabrication. Its excellent printability enables the precise and stable production of complex shapes. PLA's mechanical strength preserves structural integrity during repeated treatments, and its radiological properties can be adjusted through infill density to achieve HU values within a clinically acceptable range and with low variability, supporting dose consistency in treatment planning.[15] The adaptability of PLA facilitates the creation of 3D-printed boluses that support accurate surface dose delivery during radiotherapy as planned in treatment simulations.[16–18] The combination of FDM printing technology and PLA material provides a cost-effective, customizable, and scalable solution for producing high-precision, patient-specific boluses. The use of a commercially available 3D printer demonstrates strong potential for translating this methodology into routine clinical workflows, offering a practical solution to both technical and economic challenges in bolus production.[19] Despite significant advancements in 3D-printed boluses for radiotherapy, critical gaps remain in their clinical adoption. Although recent studies have examined HU accuracy, dose distribution, and air-gap reduction,[20-22] a streamlined and systematic evaluation of design, fabrication, and integration processes is still lacking. Key challenges include ensuring material integrity, dose uniformity, geometric conformity, and consistency during the transition from standard to patient-specific bolus shapes.

To address these gaps, this study systematically evaluated how infill density in FDM-printed boluses affects key radiological parameters (HU, RED, and mass density) and physical characteristics (surface conformity and geometric consistency), to develop a reproducible fabrication workflow for clinical use. By benchmarking against virtual and commercial boluses, this research aims to provide a robust foundation for the clinical implementation of 3D-printed bolus technology.

Materials and methods

Figure 1 illustrates the methodology employed in this study, comprising two stages: A) fabrication of standard and customized boluses using 3D printing, and B) performance evaluation of the fabricated boluses.

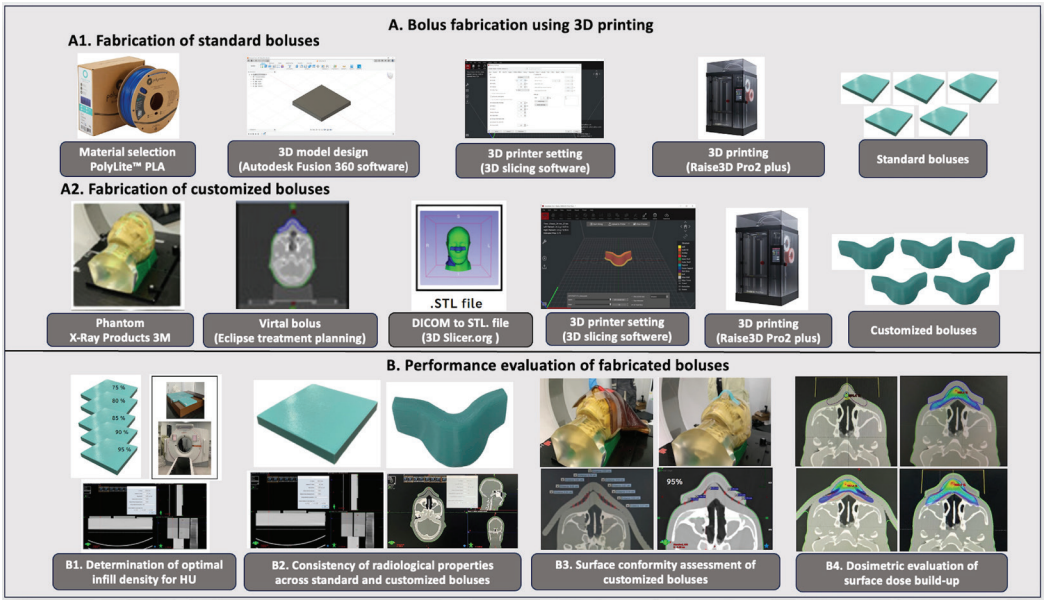


Figure 1. *Schematic representation of the fabrication and evaluation process for 3D-printed customizable boluses, developed to optimize radiotherapy applications.*

3D printing bolus fabrication

Two types of boluses were fabricated for this study: standard and customized. We used the standard bolus as a reference to evaluate the impact of varying infill densities on radiological properties and dosimetric performance. In contrast, the customized bolus was designed to conform to patient-specific anatomical contours, aiming to enhance surface conformity and reduce air gaps. The fabrication procedures for both bolus types are described in detail below.

Fabrication of standard boluses

Standard boluses were digitally designed as rectangular plates measuring 10×10×1 cm using 3D design software (Fusion 360; Autodesk Inc., San Rafael, CA, USA). The 3D model of the rectangular plate was exported as an STL file, which was used as an input for 3D printing. Boluses were fabricated via FDM using a Raise3D Pro2 Plus 3D printer (Raise3D Technologies, Irvine, CA, USA). The PLA filament (PolyLite™ PLA; Polymaker, Shanghai, China) was selected for its high stiffness and tensile strength, radiological equivalence, and ease of handling during printing.[23] The filament, with a diameter of 1.75 mm and density of 1.19 gm/cm^3, enabled consistent extrusion and dimensional stability, which are critical for ensuring the quality and accuracy of the printed boluses. The slicing parameters were configured using a 3D slicing software (IdeaMaker, Raise3D Technologies, Irvine, CA, USA), with settings adjusted to balance print quality and fabrication efficiency while minimizing structural defects. Key parameters, including layer height, print speed, extruder temperature, and infill pattern, were carefully selected to ensure precision and reproducibility. The printing parameters are summarized in Table 1.

Table 1. Printing parameters configured using 3D slicing software.

Printing parameters	Values
Layer height	0.2 mm
Number of shells	2
Nozzle diameter	0.4 mm
Nozzle temperature	230 °C
Bed temperature	60 °C
Printing speed	120 mm/s
Infill pattern	Line

A total of five boluses were fabricated for each infill density (75%, 80%, 85%, 90%, 95%) to evaluate measurement reproducibility. All boluses were printed and stored under controlled environmental conditions to minimize the moisture sensitivity of PLA before CT imaging. Each HU and RED measurement was repeated three times on separate CT acquisitions to confirm consistency.

To systematically investigate the effect of infill density on radiological properties, standard boluses were printed with infill densities incrementally adjusted in 5% steps: 75%, 80%, 85%, 90%, and 95%. Additional boluses were fabricated for precise tuning with density adjustments in 1% steps within the range closely approximating water equivalence (0 HU). Each fabricated bolus was visually inspected for structural integrity, dimensional accuracy, and surface uniformity before being subjected to further testing. This procedure ensured that only boluses meeting quality standards were included in subsequent evaluations.

Fabrication of customized boluses

Customized boluses were designed to replicate patient-specific anatomical contours to improve surface conformity and minimize air gaps. An anthropomorphic head-and-neck phantom manufactured by the X-Ray Products Division of 3M (USA) was scanned using a Siemens SOMATOM GoPro CT Scanner (Siemens Healthineers, Erlangen, Germany) at 120 kVp with a slice thickness of 1 mm for accurate imaging. CT images were imported into the Eclipse treatment planning system (TPS) (Version 16.0, Varian Medical Systems, Palo Alto, CA, USA) to generate a virtual bolus with HU of 0, mass density of 1.0081 gm/cm^3, RED of 1.0061, and uniform thickness of 1 cm. The DICOM file of the virtual bolus was converted to an STL file using 3D Slicer (3D Slicer.org, Boston, Massachusetts, USA) to ensure compatibility with the 3D printing workflow.

The customized boluses were fabricated using the same 3D printing process as that for the standard boluses. The printing parameters listed in Table 1 were applied to ensure consistency in fabrication quality. Infill densities were varied between 90% and 95% in 1% steps to evaluate performance. Following fabrication, the customized boluses were inspected for dimensional accuracy, structural integrity, and surface conformity. Alignment with the phantom's anatomical contours was verified to ensure their suitability for radiotherapy applications. Figure 2 shows examples of the 3D-printed standard and customized boluses.

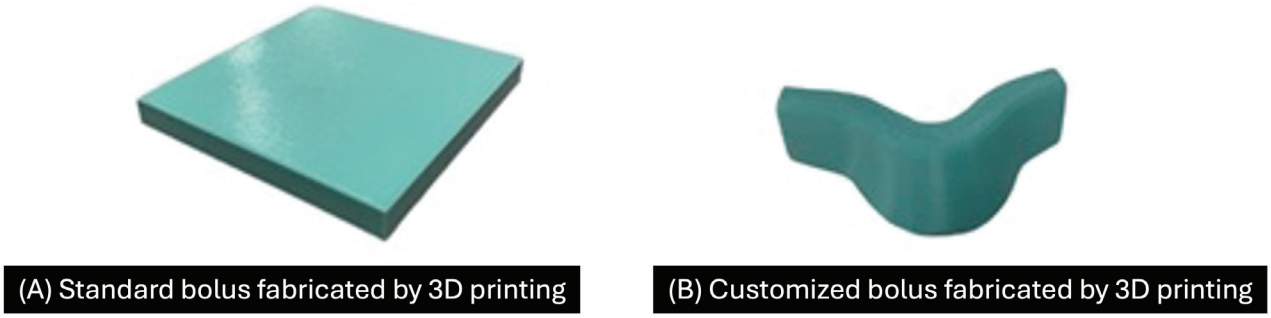


Figure 2. Examples of 3D-printed bolus fabrication in this study. (A) standard bolus, and (B) customized bolus.

Performance evaluation of fabricated boluses

The performance of the fabricated boluses was systematically analyzed through four evaluations, focusing on radiological properties, surface conformity, and dosimetric accuracy, to ensure that the fabricated boluses meet the requirements for radiotherapy applications.

Determination of optimal infill density for HU

This evaluation aimed to determine the infill density that yielded a HU value within a clinically acceptable range close to 0, representing radiological behavior consistent with that assumed in treatment planning. Standard boluses with infill densities of 75%, 80%, 85%, 90%, and 95% were each fabricated in triplicate and CT-scanned using a Siemens SOMATOM GoPro scanner at 120 kVp, 200 mA, 1-mm slice thickness, and a 250-mm field of view. The boluses were stored at room temperature before CT imaging. HU values were measured from a centrally placed region of interest on the mid-plane of each bolus and averaged across three consecutive slices to minimize local variation. The mean±SD of HU, mass density, and RED were determined for each infill density to assess measurement reproducibility.

Consistency of radiological properties across standard and customized boluses

After identifying the infill densities that yielded HU values within a clinically acceptable range close to 0, both standard and customized boluses were fabricated using identical infill densities to ensure comparable material characteristics irrespective of geometry. Each bolus was CT-scanned to determine HU, mass density, and RED values. The measured parameters were compared to assess the consistency of radiological properties between the two bolus designs.

Surface conformity assessment of customized boluses

This evaluation aimed to assess surface conformity and air-gap reduction of customized boluses applied to a head-and-neck phantom. Customized boluses with optimized infill densities were fitted onto the phantom and CT-scanned. For comparison, a commercially available bolus commonly used in clinical practice was positioned at the same location on the phantom and scanned under identical imaging conditions. Air gaps between each bolus and the phantom surface were quantified using the Eclipse TPS. The air-gap distance was defined as the perpendicular separation between the outer contour of the phantom and the inner surface of the bolus, identified in axial CT images where discontinuities were observed. The maximum and mean gap distances were measured across the contact region and compared between the customized and commercial boluses to evaluate surface conformity and setup reproducibility.

Dosimetric evaluation of surface dose build-up

The dose distributions of customized boluses were compared with those of three alternatives no bolus, virtual bolus, and a commercially available clinical bolus to evaluate their impact on surface-dose build-up. Treatment planning was performed in the Eclipse treatment planning system (Version 16.0, Varian Medical Systems) using the Anisotropic Analytical Algorithm (AAA) with a 6 MV photon beam generated by a TrueBeam linear accelerator. A single anterior photon field was applied, with the field size adjusted to cover the entire bolus area at a source-to-surface distance (SSD) of 100 cm, simulating a typical setup for superficial irradiation. Dose distributions were analyzed qualitatively using the color-wash display at the 100% isodose level to compare surface coverage under different bolus conditions. The analysis focused on the positional shift of the 100% isodose line toward the phantom surface and the uniformity of coverage across regions with complex geometry. Observations from the Eclipse TPS were used to compare dose build-up and conformity among the no-bolus, virtual-bolus, commercial-bolus, and 3D-printed-bolus configurations.

Results

Determination of optimal infill density for HU

Standard boluses with infill densities ranging from 75% to 95% were evaluated for HU, mass density, and RED values using the TPS. The detailed results presented in Table 2 indicated that boluses with infill densities between 90% and 95% exhibited HU values approaching 0, falling within the clinically acceptable water-like range relative to the virtual bolus used in treatment planning. Based on these findings, additional boluses were fabricated with 1% increments within the 90-95% range to refine the high-infill configuration and verify the consistency of radiological properties near water equivalence. As infill density increased, HU values progressively shifted from negative to positive, reflecting increased material density and X-ray attenuation. HU values close to those of water (≈0 HU) were observed at higher infill densities, accompanied by increases in mass density and RED. The boluses produced at higher infill levels exhibited mean mass density and RED values close to unity, supporting their approximation to water-equivalent behavior. These findings indicate that high-infill configurations fall within the clinically acceptable range for water-equivalent bolus fabrication, suggesting their suitability for accurate dose calculation in radiotherapy applications. In contrast, lower infill densities (<90%) produced substantially lower HU values, indicating under-density and reduced reliability in dose modeling. Therefore, high-infill 3D-printed boluses are recommended to achieve clinically acceptable HU conformity and radiological consistency.

Table 2. HU, mass density (MD), and RED values for 3D-printed standard boluses with varying infill densities (IDs) in percent.

ID (%)	HU	MD (gm/cm^3)	RED
75	–184.63±8.23	0.82±0.008	0.81±0.008
80	–156.83±6.57	0.87±0.012	0.85±0.008
85	–107.50±9.59	0.92±0.011	0.91±0.009
90	–43.83±6.09	0.98±0.005	0.97±0.004
91	–31.33±2.56	0.99±0.006	0.97±0.007
92	–19.67±3.68	0.99±0.001	0.98±0.002
93	–1.17±4.81	1.00±0.003	0.99±0.004
94	13.00±3.00	1.01±0.004	1.01±0.003
95	48.50±8.92	1.06±0.010	1.04±0.007

Consistency of radiological properties across standard and customized boluses

To ensure consistent material properties despite geometric variations, customized boluses were fabricated with 90–95% infill in 1% increments. A head-to-head comparison with the standard design showed a mean HU increase of approximately 21 HU, along with average rises of 0.02 gm/cm^3 in mass density and 0.02 in RED (Table 3). The clinically acceptable HU range was defined as ±50 HU around water (0 HU). Within this high-infill band, HU values entered the acceptable range at the lower end and increased further with higher infill, occasionally exceeding the range. Despite minor differences, the overall HU, density, and RED trends between the standard and customized boluses were consistent, indicating stable radiological behavior across geometries. Corresponding increases in mass density and RED approached unity, supporting water-equivalent behavior at higher infill rates, whereas sub-90% infill produced clearly under-dense, sub-water HU values.

Table 3. HU, mass density (MD), and RED values for 3D-printed customized boluses with varying infill densities (IDs) in percent.

ID (%)	HU	MD (gm/cm^3)	RED
90	–38.00±13.32	0.98±0.009	0.97±0.009
91	–23.70±6.38	0.99±0.003	0.98±0.004
92	17.80±11.39	1.02±0.011	1.01±0.010
93	33.55±10.00	1.04±0.013	1.03±0.009
94	44.00±11.55	1.05±0.014	1.04±0.010
95	60.00±8.78	1.07±0.006	1.05±0.005

Surface conformity assessment of customized boluses

Air gaps between the bolus and phantom were analyzed in the TPS to evaluate surface conformity and the effectiveness of reducing air gaps. Figure 3 shows representative measurements comparing commercial and 3D-printed customized boluses. The 3D-printed boluses demonstrated improved conformity to the phantom surface compared with commercial boluses. The mean air-gap distance for 3D-printed customized boluses was 0.21±0.02 cm (range: 0.19-0.25 cm), which was smaller than that of commercial boluses, 0.54±0.16 cm (range: 0.27-0.76 cm). These results confirm the enhanced surface conformity and effective air-gap minimization achieved with 3D-printed customized boluses.

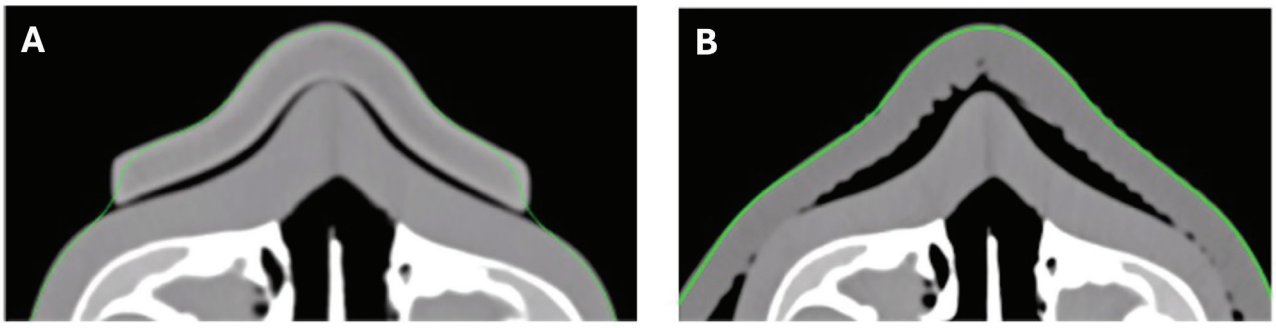


Figure 3. *Comparison of air-gap measurements between the (A) 3D-printed customized bolus and (B) commercial bolus.*

Dosimetric evaluation of surface dose build-up

This analysis evaluated surface dose build-up under different bolus configurations using a 6 MV photon beam in an AP field, as shown in Figure 4. The 100% isodose distribution (yellow region) was assessed under four conditions: no bolus, virtual bolus, commercial bolus, and 3D-printed bolus. Without a bolus, the 100 % isodose line extended deeper into the tissue, demonstrating insufficient surface dose due to the skin-sparing effect of megavoltage photons. The virtual bolus shifted the isodose line toward the surface, simulating an idealized treatment-planning scenario. The commercial bolus improved surface dose uniformity, however, air gaps limited full 100 % isodose coverage, particularly in regions with complex geometry, resulting in potential underdosage. In contrast, the 3D-printed bolus effectively enhanced surface-dose build-up and minimized air gaps. Its customized design provided superior conformity to irregular contours, optimizing dose distribution across the treatment area. Isodose analysis confirmed that the 3D-printed bolus achieved the most complete surface-dose coverage, particularly in complex geometries. These findings demonstrate the feasibility of 3D-printed boluses as customizable alternatives to commercial products, offering enhanced dose conformity and treatment precision.

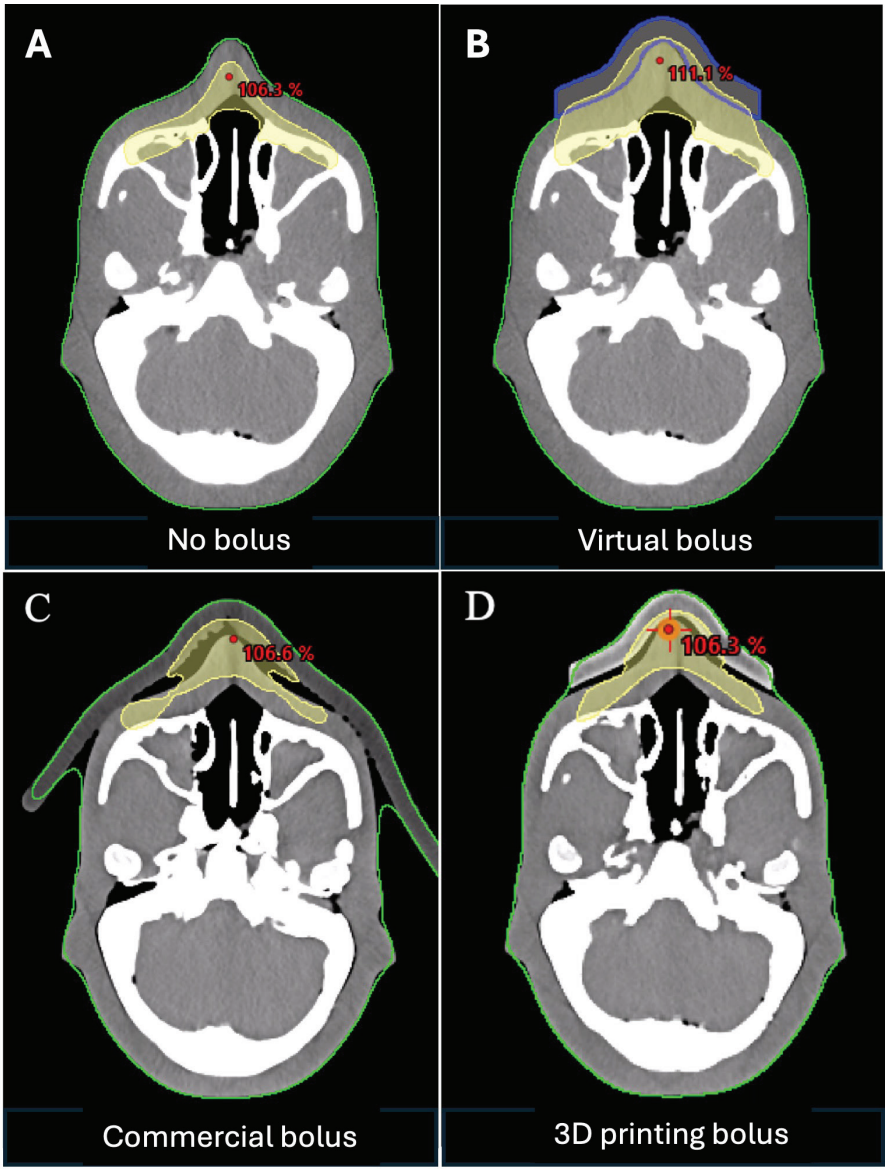


Figure 4. *Comparison of 100% isodose distributions for different bolus configurations using a 6 MV photon beam in an anterior–posterior field. A: without bolus, B: virtual bolus, C: commercial bolus, D: 3D-printed bolus.*

Discussion

This study demonstrated that 3D-printed boluses, when fabricated with optimized infill density, achieved radiological properties within the clinically acceptable water-equivalent range and improved surface conformity compared with commercial boluses. The fabrication process also showed high reproducibility and geometric precision, consistent with previous studies using FDM for radiotherapy applications.[14,15] Fabricating standardized and patient-specific boluses offers enhanced flexibility in treatment planning and reduces inter-operator variability commonly encountered in conventional manual bolus shaping.[24] Our findings confirmed a strong positive correlation between infill density and CT HU, with HU values shifting progressively from negative to positive as infill density increased. In radiotherapy, water-equivalent materials, characterized by HU values near 0, RED≈1.00, and by physical density ≈1.00 g/cm^3, are considered the standard reference for dose calculation in TPS.[8] This standard also reflects common clinical practice, where virtual boluses are typically assigned a HU of 0 during plan optimization. Aligning the HU of physical boluses with this value ensures dosimetric consistency between the planned and delivered dose distributions.[25]

In this study, both standard and customized bolus designs demonstrated similar HU-infill dependence, achieving water-like radiological behavior when printed with dense infill settings. Across the 90-94% infill range, the HU distribution remained within a practical water-equivalent interval (approximately -50 to +50 HU), consistent with prior reports on radiologically water-like 3D-printed bolus materials used in radiotherapy applications.[26,27] In contrast, infill levels below 90% yielded markedly lower HU values, indicating reduced material density and attenuation. Minor HU differences between customized and standard boluses likely reflect geometry-related differences in material deposition rather than true radiological deviations.[18,28,29] These findings suggest that using a dense-infill configuration can ensure radiological consistency across bolus designs, although optimal settings may vary with printer model, filament composition, and infill pattern.

Many studies have employed various approaches to assess the radiological equivalence of 3D-printed boluses. Some targeted HU values near 0 to emulate water-like density,[30] whereas others referenced conventional bolus materials such as Superflab or paraffin wax (approximately -30 to -100 HU) that still achieve RED≈1.00.[31] In addition to HU-based evaluation, several studies have validated dosimetric performance experimentally under photon and electron beams.[18,32] In contrast, the present study relied solely on TPS-based photon dose calculations, with interpretation restricted to the build-up region. No direct physical surface-dose measurements were performed. In addition, uncertainties associated with TPS modeling in this region may therefore influence absolute surface-dose estimation. This reliance on TPS calculations represents a key limitation of the current work. Future work will address this limitation through direct surface-dose measurements using radiochromic film, complemented by parallel-plate ionization chambers for absolute dose verification in the build-up region. In addition, electron-beam dosimetric evaluation will be incorporated to provide more comprehensive clinical validation across beam modalities. The correlation observed between HU and radiological density in this study aligns with previous findings in FDM-printed materials.[17,33] However, acceptable HU tolerances vary across TPS algorithms and clinical settings,[34,35] highlighting the value of density-based QA criteria. For example, Kairn *et al.*[12] proposed ±5% deviation from water-equivalent density as clinically acceptable and ±10 % as an action threshold, providing a practical framework for in-house QA of patient-specific boluses. Customized boluses significantly reduced air gaps compared with commercial boluses. Although bolus placement beneath thermoplastic masks can improve conformity in head-and-neck treatments, this method remains operator-dependent and potentially variable between fractions.

In this study, CT-derived air gap measurements provided a practical quantitative assessment of bolus–skin conformity. Other quantitative air gap measurement methods, such as CT-based HU profile analysis and volumetric air gap quantification using image segmentation, have been reported to provide more detailed geometric characterization of the bolus–skin interface, particularly in regions with complex surface geometry.[36,37] Fixed-geometry 3D-printed boluses may therefore improve reproducibility, setup efficiency, and patient comfort—especially in regions with complex topology, aligning with reports of improved surface-dose build-up with customized bolus fabrication.[38] Isodose analysis further demonstrated improved surface coverage with 3D-printed boluses, supporting their integration into photon-based workflows.[39,40] While electron beams remain standard for very superficial lesions, contemporary photon techniques such as VMAT increasingly rely on bolus for chest-wall and nodal treatments; therefore, photon-based assessment is clinically relevant. Additionally, our results support previous reports that 3D printing offers a potentially cost-efficient alternative to commercial boluses, primarily through reduced material consumption and in-house fabrication.[31] Although the degree of savings depends on local workflows and resources, the ability to individualize bolus design without substantial added expense enhances the scalability of this approach in clinical practice. Material usage (103.5-120.5 gm per standard bolus) and printing time (~4 hours) from this study demonstrate predictable manufacturing demands consistent with prior workflows.[41,42] This supports the practical feasibility of routine clinical deployment.

Limitations

This study used PLA as the primary material for 3D-printed bolus fabrication, which was selected based on a balance of radiological equivalence, mechanical performance, and manufacturing efficiency. While PLA offers advantages such as low ultrafine particle emissions compared with ABS and avoids the release of potentially hazardous thermal decomposition byproducts, alternative materials may provide distinct benefits depending on clinical priorities.[43] For example, silicone-based polymers processed via direct ink writing offer improved flexibility and tissue conformity, but their slow build speeds and limited spatial resolution constrain routine clinical use.[16] Additionally, previous comparative studies have shown that although PLA demonstrates favorable radiation transmission characteristics, materials such as PETG may provide superior mechanical strength and flexibility.[15] Accordingly, future work will systematically evaluate a broader range of printable materials to better balance radiological equivalence, mechanical performance, and manufacturing efficiency.

Another limitation of this study is the use of only one FDM printer model and a specific PLA filament. Variations in printer hardware, such as nozzle calibration, extrusion control, and thermal stability, as well as differences in filament composition and additives across manufacturers, may influence effective infill density and measured HU values. As a result, while the optimal infill density provides a useful benchmark, the reported HU values are specific to the printer–material combination used. Future studies will aim to improve generalizability through validation across multiple FDM systems and filament brands, including multi-center investigations using identical STL geometries. Reporting measured mass and volume to derive physical density, rather than relying solely on slicer-defined infill percentages, may also provide a more universal and reproducible metric.

Finally, this study evaluated only a line infill pattern. However, infill geometry is known to influence internal density distribution, mechanical properties, and manufacturing time. Prior work comparing triangle, honeycomb, honeycomb-gyroid, and gyroid infill patterns in PLA has demonstrated differences in both ultimate tensile strength and HU values.[15] Future investigations will explore diverse infill geometries to identify optimal pattern–density combinations that balance internal density distribution, mechanical properties, and manufacturing time, particularly for boluses with complex anatomical shapes.

Conclusion

This study demonstrated that PLA 3D-printed boluses fabricated with an infill density above 90% achieved radiological properties within the clinically acceptable range for water-like materials and provided superior surface conformity relative to commercial boluses. The fabrication process was highly reproducible, cost-efficient, and suitable for routine in-house clinical implementation. These findings indicate that high-infill 3D-printed boluses can serve as reliable, customizable alternatives to conventional boluses, supporting improved workflow consistency, surface-dose accuracy, and patient comfort in modern radiotherapy practice.

Ethical approval

Ethical approval: REC 68-230-7-2

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare no conflict of interest for this publication.

CRediT authorship contribution statement

Chitchaya Suwanraksa: conceptualization, methodology, investigation, data collection and analysis, writing: original draft, supervision; **Athitan Khiewded**: methodology, investigation, data collection, and analysis. **Gadesuda Jankaew**: investigation, data collection, and analysis; **Thirawut Rojchanaumpawan**: methodology, data collection, and analysis; **Surapong Chatpun**: methodology, data collection, and analysis. **Pornchai Phukpattaranont**: conceptualization, methodology, investigation, data collection and analysis, writing: original draft, supervision. All authors contributed to reviewing and editing the manuscript and approved the final version for submission.

Acknowledgements

The authors would like to acknowledge the Radiation Oncology Unit, Faculty of Medicine, Prince of Songkla University, for their support and technical assistance throughout this study.

References

[1] Bernier J, Hall EJ, Giaccia A. Radiation oncology: a century of achievements. Nat Rev Cancer. 2004; 4: 737-47. doi:10.1038/nrc1451.

[2] Garibaldi C, Jereczek-Fossa BA, Marvaso G, Dicuonzo S, Rojas DP, Cattani F, et al. Recent advances in radiation oncology. Ecancermedicalscience. 2017; 11: 785. doi:10.3332/ecancer.2017.785.

[3] Khan FM, Gibbons JP Jr. Khan's. The physics of radiation therapy. Philadelphia: Lippincott Williams & Wilkins; 2014: pp 584.

[4] Dogan N, Glasgow GP. Surface and build-up region dosimetry for obliquely incident intensity modulated radiotherapy 6 MV X-rays. Med Phys. 2003; 30: 3091-6. doi:10.1118/1.1625116.

[5] Turner JY, Zeniou A, Williams A, Jyothirmayi R. Technique and outcome of post-mastectomy adjuvant chest wall radiotherapy—the role of tissue-equivalent bolus in reducing risk of local recurrence.

Br J Radiol. 2016; 89(1064): 20160060. doi:10.1259/bjr.20160060.

[6] Hsu SH, Roberson PL, Chen Y, Marsh RB, Pierce LJ, Moran JM. Assessment of skin dose for breast chest wall radiotherapy as a function of bolus material. Phys Med Biol. 2008; 53: 2593-606. doi:10.1088/0031-9155/53/10/010.

[7] Healy E, Anderson S, Cui J, Beckett L, Chen AM, Perks J, et al. Skin dose effects of postmastectomy chest wall radiation therapy using brass mesh as an alternative to tissue-equivalent bolus. Pract Radiat Oncol. 2013; 3: e45-e53. doi:10.1016/j.prro.2012.05.009.

[8] Vyas V, Palmer L, Mudge R, Jiang R, Fleck A, Schaly B, et al. On bolus for megavoltage photon and electron radiation therapy. Med Dosim. 2013; 38: 268-73. doi:10.1016/j.meddos.2013.02.007.

[9] McCallum S, Maresse S, Fearns P. Evaluating 3D-printed bolus compared to conventional bolus types used in external beam radiation therapy. Curr Med Imaging Rev. 2021; 17: 820-31. doi:10.2174/1573405617666210202114336.

[10] Pollmann S, Toussaint A, Flentje M, Wegener S, Lewitzki V. Dosimetric evaluation of commercially available flat vs. self-produced 3D-conformal silicone boluses for the head and neck region. Front Oncol. 2022; 12: 881439. doi:10.3389/fonc.2022.881439.

[11] Davis TM, Luca K, Sudmeier LJ, Buchwald ZS, Khan MK, Yang X, et al. Total scalp irradiation: A study comparing multiple types of bolus and VMAT optimization techniques. J Appl Clin Med Phys. 2024; 25: e14260. doi:10.1002/acm2.14260.

[12] Kairn T, Talkhani S, Charles PH, Chua B, Lin CY, Livingstone AG, et al. Determining tolerance levels for quality assurance of 3D printed bolus for modulated arc radiotherapy of the nose. Phys Eng Sci Med. 2021; 44: 1187-99. doi:10.1007/s13246-021-01054-7.

[13] Fujimoto K, Shiinoki T, Yuasa Y, Hanazawa H, Shibuya K. Efficacy of patient-specific bolus created using three-dimensional printing technique in photon radiotherapy. Phys Med. 2017; 38: 1-9. doi:10.1016/j.ejmp.2017.04.023.

[14] Pugh R, Lloyd K, Collins M, Duxbury A. The use of 3D printing within radiation therapy to improve bolus conformity: a literature review. J Radiother Pract. 2017; 16: 319-25. doi:10.1017/S146039 691 7000115.

[15] Phantawong K, Singthuan N, Pinichsai K, Taertulakarn S. The feasibility of fabricated 3D-printed immobilization masks for radiotherapy: mechanical and dosimetric analysis. J Assoc Med Sci. 2025; 58(3): 361-75.

[16] Lu Y, Song J, Yao X, An M, Shi Q, Huang X. 3D printing polymer-based bolus used for radiotherapy. Int J Bioprint. 2021; 7: 414. doi:10.18063/ijb.v7i4.414.

[17] Van Der Walt M, Crabtree T, Albantow C. PLA as a suitable 3D printing thermoplastic for use in external beam radiotherapy. Australas Phys Eng Sci Med. 2019; 42: 1165-76. doi:10.1007/s13246-019-00818-6.

[18] Ricotti R, Ciardo D, Pansini F, Bazani A, Comi S, Spoto R, et al. Dosimetric characterization of 3D-printed bolus at different infill percentage for external photon beam radiotherapy. Phys Med. 2017; 39: 25-32. doi:10.1016/j.ejmp.2017.06.004.

[19] Ashenafi M, Jeong S, Wancura JN, Gou L, Webster MJ, Zheng D. A quick guide on implementing

[20] and quality assuring 3D printing in radiation oncology. J Appl Clin Med Phys. 2023; 24: e14102. doi:10.1002/acm2.14102.

[21] Şahin T, Çavdar Karaçam S, Tunçman D, Şahin Ş, Ergen ŞA, Çolpan Öksüz D, et al. Thermal, dosimetric and thermo-mechanical characterization of ABS and PLA polymer materials via plastic injection molding for innovative bolus materials in radiotherapy applications. Radiat Phys Chem. 2024; 223: 111901. doi:10.1007/s42247-024-00706-0.

[22] Jreije A, Keshelava L, Ilickas M, Laurikaitiene J, Urbonavicius BG, Adliene D. Development of patient-specific conformal 3D-printed devices for dose verification in radiotherapy. Appl Sci. 2021; 11: 8657. doi:10.3390/app11188657.

[23] Ciobanu AC, Petcu LC, Járai-Szabó F, Bálint Z. Exploring the impact of filament density on the responsiveness of 3D-printed bolus materials for high-energy photon radiotherapy. Phys Med. 2024; 127: 104849. doi:10.1016/j.ejmp.2024.104849.

[24] 3DTechSupplies. PolyLite™ PLA. Available from: http://www.3dtechsupplies.com.au/Products/PolyMaker/PolyLite-PLA. Accessed 6 Jan 2026.

[25] Canters RA, Lips IM, Wendling M, Kusters M, Van Zeeland M, Gerritsen RM, et al. Clinical implementation of 3D printing in the construction of patient-specific bolus for electron beam radiotherapy for non-melanoma skin cancer. Radiother Oncol. 2016; 121: 148-53. doi:10.1016/j.radonc.2016.07.011.

[26] Olch AJ, Gerig L, Li H, Mihaylov I, Morgan A. Dosimetric effects caused by couch tops and immobilization devices: report of AAPM Task Group 176. Med Phys. 2014; 41(6 Part 1): 061501. doi:10.1118/1.4876299.

[27] Mutic S, Palta JR, Butker EK, Das IJ, Huq MS, Loo LND, et al. Quality assurance for computed-tomography simulators and the computed-tomography-simulation process: Report of the AAPM Radiation Therapy Committee Task Group No. 66. Medical Physics. 2003; 30(10): 2762-92.

[28] Burleson S, Baker J, Hsia AT, Xu Z. Use of 3D printers to create a patient-specific 3D bolus for external beam therapy. J Appl Clin Med Phys. 2015; 16(3): 5247. doi: 10.1120/jacmp.v16i3.5247.

[29] Savi M, Potiens MPA, Cechinel CM, Silveira LC, Soares FAP. Relationship between infill patterns in 3D printing and Hounsfield Unit. In: Proceedings of the International Joint Conference RADIO 2017: Goiania 30 Years Later; 2017 Oct; Goiania, Brazil.

[30] Biltekin F, Yazici G, Ozyigit G. Characterization of 3D-printed bolus produced at different printing parameters. Med Dosim. 2021; 46: 157-63. doi:10.1016/j.meddos.2020.10.005.

[31] Jeong S, Yoon M, Chung WK, Kim DW. Preliminary study of the dosimetric characteristics of 3D-printed materials with megavoltage photons. J Korean Phys Soc. 2015; 67(1): 189-94. doi:10.3938/jkps.67.189.

[32] Albantow C, Hargrave C, Brown A, Halsall C. Comparison of 3D-printed nose bolus to traditional wax bolus for cost-effectiveness, volumetric accuracy and dosimetric effect. J Med Radiat Sci. 2020; 67: 54-63. doi:10.1002/jmrs.378.

[33] Kim SW, Shin HJ, Kay CS, Son SH. A customized bolus produced using a 3-dimensional printer for radiotherapy. PLoS One. 2014; 9: e110746. doi:10.1371/journal.pone.0110746.

[34] Dancewicz OL, Sylvander SR, Markwell TS, Crowe SB, Trapp JV. Radiological properties of 3D-printed materials in kilovoltage and megavoltage photon beams. Phys Med. 2017; 38: 111-8. doi:10.1016/j.ejmp.2017.05.051.

[35] Burleson S, Baker J, Hsia AT, Xu Z. Use of 3D printers to create a patient-specific 3D bolus for external beam therapy. J Appl Clin Med Phys. 2015; 16: 166-78. doi:10.1120/jacmp.v16i3.5247.

[36] Park SY, Choi CH, Park JM, Chun M, Han JH, Kim JI. A patient-specific polylactic acid bolus made by a 3D printer for breast cancer radiation therapy. PLoS One. 2016; 11: e0168063. doi:10.1371/journal.pone.0168063.

[37] Dipasquale G, Poirier A, Sprunger Y, Uiterwijk JWE, Miralbell R. Improving 3D-printing of megavoltage X-rays radiotherapy bolus with surface-scanner. Radiat Oncol. 2018; 13(1): 203. doi: 10.1186/s13014-018-1148-1.

[38] Gugliandolo SG, Pillai SP, Rajendran S, Vincini MG, Pepa M, Pansini F, et al. 3D-printed boluses for radiotherapy: influence of geometrical and printing parameters on dosimetric characterization and air gap evaluation. Radiol Phys Technol. 2024; 17(2): 347-59. doi: 10.1007/s12194-024-00782-1.

[39] Łukowiak M, Boehlke M, Lewocki M, Jezierska K, Piatek-Hnat M, Lewocki M, et al. Use of a 3D printer to create a bolus for patients undergoing tele-radiotherapy. Int J Radiat Res. 2016; 14: 287-95. doi:10.18869/acadpub.ijrr.14.4.287.

[40] Łukowiak M, Jezierska K, Boehlke M, Więcko M, Łukowiak A, Podraza W, et al. Utilization of a 3D printer to fabricate boluses used for electron therapy of skin lesions of the eye canthi. J Appl Clin Med Phys. 2017; 18: 76-81. doi:10.1002/acm2.12013.

[41] Sasaki DK, McGeachy P, Alpuche Aviles JE, McCurdy B, Koul R, Dubey A. A modern mold room: Meshing 3D surface scanning, digital design, and 3D printing with bolus fabrication. J Appl Clin Med Phys. 2019; 20: 78-85. doi:10.1002/acm2.12703.

[41] Basaula D, Hay B, Wright M, Hall L, Easdon A, McWiggan P, Yeo A, Ungureanu E, Kron T. Additive manufacturing of patient-specific bolus for radiotherapy: large-scale production and quality assurance. Phys Eng Sci Med. 2024;47:551–61. doi:10.1007/s13246-024-01385-1.

[42] Ehler E, Sterling D, Dusenbery K, Lawrence J. Workload implications for clinic workflow with implementation of three-dimensional printed customized bolus for radiation therapy: a pilot study. PLoS One. 2018;13(10):e0204944. doi:10.1371/journal.pone.0204944.

[43] Stephens B, Azimi P, El Orch Z, Ramos T. Ultrafine particle emissions from desktop 3D printers. Atmos Environ. 2013; 79: 334–339. doi: 10.1016/j.atmosenv.2013.06.050.

Journal of Associated Medical Sciences 2026; 59(2): 153-162

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Effectiveness of guided imagery to improve sleep quality among young adults with post-traumatic stress disorder (PTSD)

Vikraman G.* **and Benitta Paulin Giftlin R.**

Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India.

ARTICLE INFO

Article history:
Received 24 September 2025
Accepted as revised 4 February 2026
Available online 10 February 2026

Keywords:
Guided Imagery, post-traumatic stress disorder (PTSD), sleep quality, young adults.

ABSTRACT

Background: Post-traumatic stress disorder often leads to ongoing sleep difficulties such as insomnia and nightmares, which can interfere with mood, concentration, and daily activities. Guided imagery, a relaxation method that uses calming, multisensory visualizations, has shown potential to improve sleep and reduce distress in people experiencing high stress or trauma. This study evaluates the effectiveness of guided imagery in improving sleep quality among young adults with post-traumatic stress disorder (PTSD) compared to conventional occupational therapy.

Objectives: To evaluate and compare the effectiveness of guided imagery and conventional occupational therapy in improving sleep quality among young adults (18-25 years) with post-traumatic stress disorder.

Materials and methods: A quasi experimental design was conducted with 30 young adults aged 18-25 years, recruited through convenience sampling and divided into two groups: experimental (guided imagery intervention) and control (conventional occupational therapy). The potential of the post-traumatic stress disorder was accomplished by the diagnosis of PTSD using Clinician-Administered PTSD Scale for DSM-5, Past-Month version. Both interventions were delivered over three months, with 36 sessions held three times per week, each lasting 45 minutes. The state of sleep prior to intervention and right after the intervention was critically explained with the aid of the Pittsburgh Sleep Quality Index (PSQI). Statistical analysis was performed to compare pre- and post-test scores within and between groups.

Results: The experimental group (N=15) had a mean pre-test PSQI score of 14.40 (SD=2.41) and a post-test score of 8.73 (SD=2.05), showing a highly significant improvement (t=15.707, p=0.000). The control group (N=15) had a mean pre-test PSQI score of 14.07 (SD=1.53) and a post-test score of 13.47 (SD=1.60), also showing significant improvement (t=2.358, p=0.033). Post-test scores between the groups revealed a significantly greater improvement in the experimental group compared to the control group (t=7.050, p=0.000).

Conclusion: Guided imagery was found to greatly improve sleep quality in young adults living with post-traumatic stress disorder, offering better results than conventional occupational therapy. These results highlight the value of using gentle, non-drug approaches like guided imagery to ease sleep difficulties and support overall well-being in this group.

* **Corresponding contributor.**
Author's Address: *Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India.*
E-mail address: vikramang.scot@saveetha.com
doi: 10.12982/JAMS.2026.047
E-ISSN: 2539-6056

Introduction

Occupational therapy is a client-centred profession aimed at enabling individuals to participate in meaningful daily activities by enhancing skills or modifying environments.[1-6] In mental health, it focuses on restoring routines, improving life skills, and

fostering emotional and social well-being, which is vital for individuals with psychiatric conditions. Sleep disturbances are common in various mental health disorders, especially in post-traumatic stress disorder, where up to 87% of affected individuals report issues like insomnia and nightmares that significantly impact mood, cognition, and daily life.[7,8]

Post-traumatic stress disorder develops after exposure to traumatic events and is characterised by intrusive memories, avoidance behaviours, negative mood changes, and hyperarousal. High-risk groups include young adults, who are more likely to encounter trauma, making them vulnerable to long-term functional impairments.[9-13] Sleep problems in this population often persist even after core PTSD symptoms are addressed, highlighting the need for targeted interventions.[14-16]

Guided imagery is a mind–body technique that uses positive mental visualization to promote relaxation and well-being. It engages multiple senses, not just vision, making it effective across different learning styles. Sessions typically begin with relaxation and breathing exercises, followed by calming imagery that counters distressing thoughts. Regular practice helps manage anxiety, arousal, and intrusive symptoms in PTSD by fostering grounding and self-regulation. Its non-invasive nature makes it a valuable alternative for individuals hesitant to pursue conventional therapy.[17-21]

Materials and methods

Research design

A quasi-experimental study design was employed to investigate the effectiveness of guided imagery on improving sleep quality in young adults with post-traumatic stress disorder.

Sampling technique

The study utilized a convenience sampling technique to select participants.

Sample size

A total of 30 subjects participated in this study. The participants were divided into two groups equally,

- **Experimental group:** 15 subjects underwent guided imagery interventions.
- **Control group:** 15 subjects received conventional occupational therapy.

Sample size justification

The sample size was determined based on feasibility within the available study duration and clinical setting. As this was a time-bound interventional study, the selected sample size was considered adequate to examine preliminary effects of guided imagery on sleep quality and to generate clinically relevant findings for future large-scale research.

Sample characteristics

The study sample comprised 30 young adults aged between 18 and 25 years diagnosed with post-traumatic stress disorder based on DSM-5 criteria and confirmed using the Clinician-Administered PTSD Scale for DSM-5 (CAPS-5). Both male and female participants were included. All participants reported sleep disturbances related to traumatic experiences and were clinically stable at the time of recruitment. Participants were able to understand and follow verbal instructions and actively engage in structured therapy sessions. Individuals with severe cognitive impairments, neurological conditions, or comorbid psychiatric disorders that could interfere with participation were excluded to ensure sample consistency.

Sample setting

The study was conducted at Grace Rehab Clinic, Purasawalkam.

Selection criteria

Inclusion criteria

1. Young adults aged 18-25 years.
2. Clinical diagnosis of post-traumatic stress disorder (PTSD) (based on DSM-5 criteria).
3. Reports of sleep disturbances.
4. History of psychological trauma.
5. Using sleep medication, no more than twice per week, with stable dosage.
6. Both male and female participants were included.

Exclusion criteria

1. Receiving other PTSD or sleep disorder treatments, except occasional sleep medication as allowed.
2. Cognitive impairments or neurological conditions that may interfere with participation.
3. Comorbid psychiatric conditions that impair study participation.

Tools used

1. Clinician-Administered PTSD Scale for DSM-5 (CAPS-5) – Past Month Version

The CAPS-5 is a structured clinical interview considered the gold standard for assessing post-traumatic stress disorder. It evaluates PTSD symptoms across four core clusters: intrusion, avoidance, negative alterations in cognition and mood, and alterations in arousal and reactivity.

Scoring method

Total symptom severity score is calculated by summing the severity scores for all PTSD symptoms.[22]

2. Pittsburgh Sleep Quality Index (PSQI)

The PSQI is a self-report questionnaire assessing sleep quality and disturbances over the past month. It evaluates seven components: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleep medication, and daytime dysfunction.

Scoring method

Each component is scored from 0 to 3, with higher

scores indicating worse sleep quality. The global PSQI score is calculated by summing the seven component scores, with scores above 5 indicating poor sleep quality.[23]

Intervention procedures

The study was conducted over 12 weeks, with participants attending three sessions per week, each lasting 45 minutes. The intervention was structured as follows:

- Introduction and Settling: 10 minutes (establishing a comfortable environment, brief check-in).
- Guided Imagery Session: 20 minutes (structured imagery script focusing on relaxation, grounding, and positive visualization).
- Closure and Reflection: 15 minutes (gentle return to present awareness, participant feedback).

Conventional occupational therapy for control group

The therapy sessions focused on providing standard sleep hygiene education and behavioural strategies to promote better sleep patterns among young adults with PTSD. A structured approach was followed to address common factors affecting sleep quality.

- **Sleep hygiene education:** Included guidance on establishing a consistent bedtime and wake-up routine, creating a restful sleep environment, and avoiding caffeine, nicotine, and heavy meals before bedtime.
- **Relaxation techniques:** Participants were introduced to basic relaxation practices such as slow breathing and gentle stretching to be performed before sleep.
- **Lifestyle modification advice:** Provided recommendations on limiting screen time before bed, engaging in regular physical activity during the day, and managing evening light exposure.
- **Stimulus control strategies:** Encouraged participants to use the bed only for sleep, get out of bed if unable to sleep within 20 minutes, and return only when sleepy.
- **Sleep scheduling:** Advised on maintaining regular sleep and wake times, even on weekends, to stabilize the circadian rhythm.
- **Psychoeducation on sleep and PTSD:** Shared information on the relationship between trauma, hyperarousal, and disrupted sleep, aiming to increase self-awareness and adherence to sleep-promoting behaviours.[24]

Session duration

Forty-five minutes, three times per week. Pre- and post-intervention assessments were conducted using the Pittsburgh Sleep Quality Index (PSQI) to measure changes in sleep quality

Guided imagery therapy for experimental group

The guided imagery sessions were designed to help participants relax, regulate emotions, and improve sleep quality. The intervention followed a structured format aimed at reducing arousal and promoting a calm mental state conducive to sleep.[25]

Session process

Each session began with simple grounding and slow breathing exercises to help participants settle and focus. This was followed by guided imagery involving calm, safe, and soothing scenes, encouraging participants to engage their senses and experience a sense of comfort and relaxation. Imagery was gradually directed toward releasing emotional tension and supporting the transition into sleep.

Session duration

Sessions were conducted for approximately 45 minutes, three times per week.

Outcome measurement

Changes in sleep quality were assessed before and after the intervention using the Pittsburgh Sleep Quality Index (PSQI). A detailed session-wise description of the guided imagery intervention is presented in Appendix A.

Results

The effectiveness of guided imagery on sleep quality in young adults with post-traumatic stress disorder was evaluated at Grace Rehab Clinic, Purasawalkam. A total of 30 participants were included in the study, divided into two groups: experimental (N=15) and control (N=15). The control group received conventional occupational therapy, whereas the experimental group received guided imagery interventions.

The experimental group consisted of 15 participants aged 18-25 years. The mean pre-test PSQI score in this group was 14.40 (SD=2.41), while the post-test mean score was 8.73 (SD=2.05). The control group also included 15 participants within the same age range. The mean pre-test PSQI score in the control group was 14.07 (SD=1.53), while the post-test mean score was 13.47 (SD=1.60). Both groups were comparable in terms of age and baseline PSQI scores, ensuring the validity of the comparison. The pre-test and post-test scores provide a basis for evaluating the effectiveness of the intervention in both groups.

As shown in Table 1 and Figure 1, a statistically significant difference was observed between the pre-test and post-test PSQI scores in the control group. The mean pre-test score was 14.07 (SD=1.53), while the post-test mean score was 13.47 (SD=1.60). The t-value for the comparison was 2.358, and the *p* value was 0.033, which is less than the significance level of 0.05. This indicates a statistically significant improvement in sleep quality, as measured by PSQI scores, among participants in the control group following conventional occupational therapy.

Table 1. Difference between pre-test and post-test scores in control group.

No.	Test	N	Mean	SD	T Value	Df	*p* value
1	Pre-test	15	14.07	1.534	2.358	14	0.033*
2	Post-test	15	13.47	1.598			

Note: **significant at 1% level, *significant at 5% level.

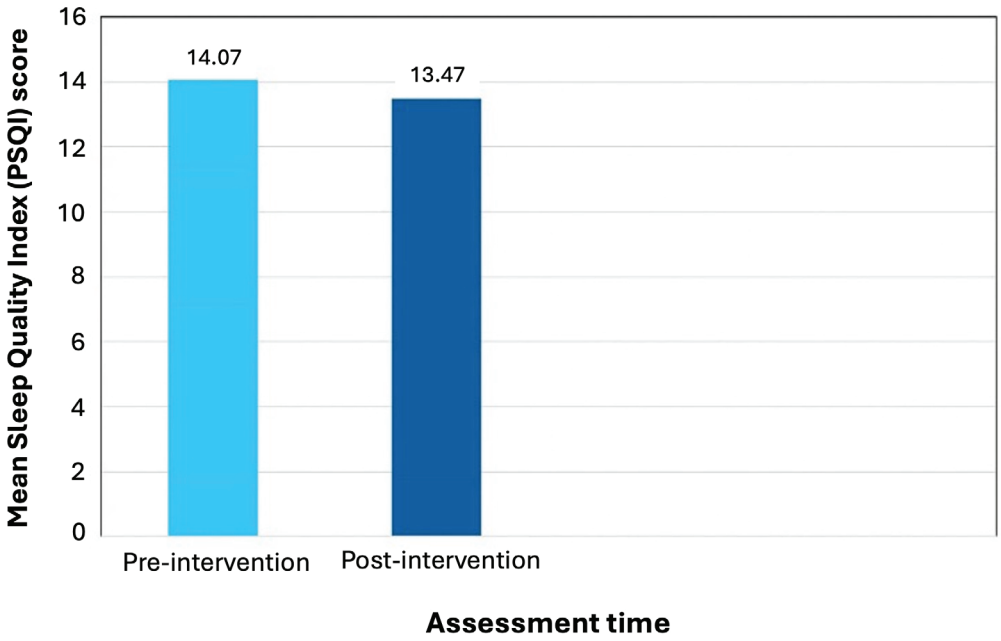


Figure 1. *Comparison of mean sleep quality index score between pre- and post-test in the control group.*

In the experimental group, a highly significant difference was observed between the pre-test and post-test scores. The mean pre-test score was 14.40 (SD=2.41), while the post-test mean score was 8.73 (SD=2.05). The t-value was 15.707, and the *p* value was 0.000, which is less than the 0.01 level of significance. This demonstrates that guided imagery produced a highly significant improvement in sleep quality for the experimental group.

In comparison between the post-test scores of the control and experimental groups, a statistically highly significant difference was found. The mean post-test score of the control group was 13.47 (SD=1.60), while that of the experimental group was 8.73 (SD=2.05). The t-value for the comparison was 7.050, with a p-value of 0.000, which is less than the 0.01 level of significance. This indicates that guided imagery was more effective than conventional occupational therapy in improving sleep quality among young adults with post-traumatic stress disorder (PTSD).

In the Control group, since the *p* value 0.033 is less than 0.05 and is significant at 5% level , an alternate hypothesis is accepted. Hence, there is statistically significant difference in the Control group between pre-test and post-test of PSQI test scores. This suggests that the intervention received by the Control group had improvement due to Conventional Occupational Therapy.

As shown in Table 2 and Figure 2, the experimental group exhibited a clear and statistically meaningful reduction in PSQI scores from pre-test to post-test ($p<0.01$). This improvement reflects the positive impact of guided imagery in enhancing sleep quality among young adults with post-traumatic stress disorder.

Table 2. Difference between pre-test and post-test scores in experimental group.

No.	Test	N	Mean	SD	T Value	Df	*p* value
1	Pre test	15	14.40	2.414	15.707	14	0.000**
2	Post test	15	8.73	2.052			

Note: **significant at 1% level, *significant at 5% level.

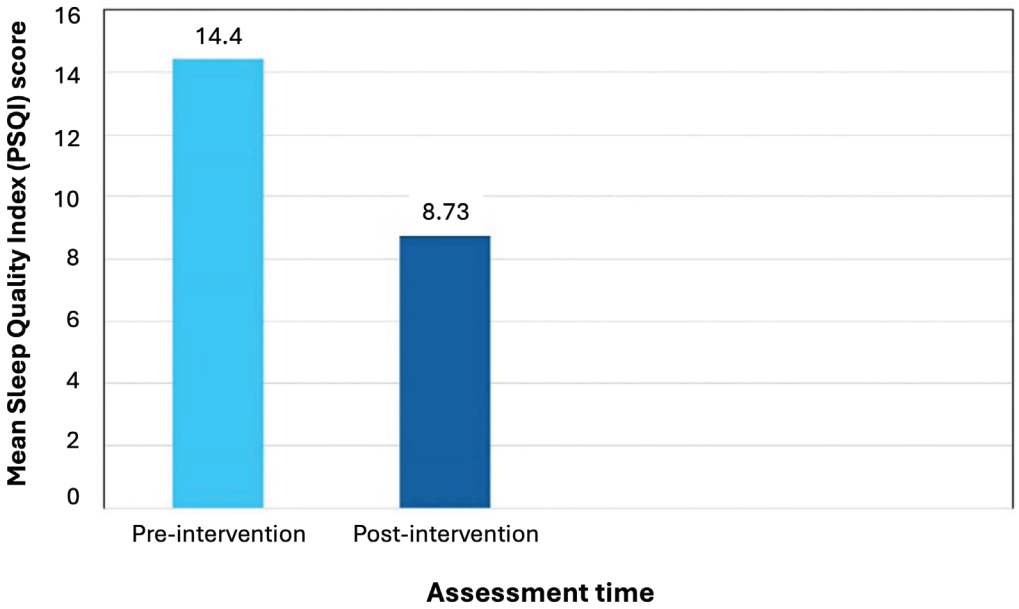


Figure 2. *Comparison of mean sleep quality index score between pre- and post-test in the experimental group.*

As presented in Table 3 and Figure 3, a statistically significant difference was observed between the post-test PSQI scores of the experimental and control groups (p<0.01). Participants who received guided imagery demonstrated greater improvement in sleep quality compared to those who received conventional occupational therapy.

Table 3. Difference between post-test scores of control group and experimental group.

No.	Test	N	Mean	SD	T Value	Df	*p* value
1	Control	15	13.47	1.598	7.050	28	0.000**
2	Experimental	15	8.73	2.052			

Note: **significant at 1% level, *significant at 5% level.

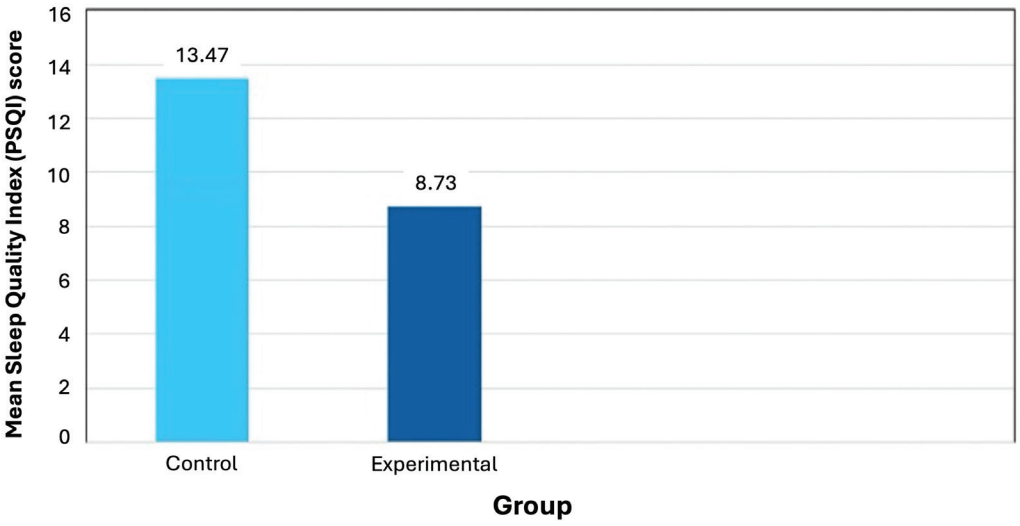


Figure 3. *Comparison of mean sleep quality index score between control and experimental group.*

Discussion

Guided imagery served as an effective intervention for addressing sleep disturbances associated with post-traumatic stress disorder. The structured sessions promoted relaxation, grounding, and positive mental imagery, helping participants manage hyperarousal and intrusive thoughts that interfere with sleep. This aligns with previous research by Schaffer *et al.*,[25] and Afshar *et al.*,[26] which reported improvements in sleep quality and reductions in psychological distress following imagery-based interventions.[27-33] The approach also reflects findings by Zehetmair *et al.*, highlighting guided imagery as a valuable self-regulation tool in trauma-exposed populations.[20] Similar benefits have also been shown in meta-analyses and clinical trials of imagery rehearsal therapy for nightmares and sleep disturbances in PTSD.In addition, trauma-focused guided imagery and music interventions have shown promise in both refugees and military veterans.[34,35.]

The pre-test scores confirmed that both groups were comparable before interventions, ensuring the validity of comparisons. The experimental group showed significantly greater improvements in sleep quality, validating the effectiveness of guided imagery for addressing trauma-related sleep difficulties in young adults with post-traumatic stress disorder. The significant difference in post-test scores between the experimental and control groups highlights the superiority of guided imagery over conventional occupational therapy.

The findings emphasize the need for guided imagery to be integrated into occupational therapy programs for young adults with post-traumatic stress disorder. Such interventions address persistent sleep disturbances, enhance relaxation, and support overall well-being. Therapists should consider incorporating multisensory, client-centred imagery practices into treatment to optimize outcomes.

Limitations

The scale of intervention was limited because of the small sample size and restricted generalizability, as the study was conducted within a single clinical setting. Individual psychosocial factors such as coping abilities, resilience, adaptability, and personal stress-management skills were not independently assessed and may have influenced participants' responses to the intervention. Additionally, the duration of the intervention was limited and long-term follow-up measurements were not conducted. Future studies should include larger sample sizes, longer intervention periods, follow-up assessments, and consideration of individual difference variables to strengthen outcome interpretation.

Conclusion

This study demonstrates that guided imagery significantly improves sleep quality in young adults with post-traumatic stress disorder. The findings underscore the importance of using safe, non-pharmacological, and engaging interventions to address trauma-related sleep problems, contributing to better rest, emotional stability, and overall quality of life.

Ethical approval

The study was approved by the Institutional Ethics Committee of Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai, India.(Approval No.: SCOT/ISRB/040/2025).

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare no conflict of interest.

CRedit authorship contribution statement

Vikraman G: conceptualization, methodology, supervision, writing: review and edit; **Benitta Paulin Giftlin R**: investigation, data curation, formal analysis, writing: original draft, review and edit.

Acknowledgements

The authors would like to thank Saveetha College of Occupational Therapy and Saveetha Institute of Medical and Technical Sciences (SIMATS) for their academic support. The authors also express their gratitude to all the participants for their cooperation and participation in this study.

References

[1] American Psychiatric Association. Posttraumatic stress disorder (PTSD) [Internet]. 2023 [cited 2025 Sep 11]. Available from: https://www.psychiatry.org/patients-families/ptsd.

[2] Capaldi VF 2nd, Guerrero ML, Killgore WDS. Sleep disruptions among returning combat veterans from Iraq and Afghanistan. Mil Med. 2011; 176(8): 879-88. doi: 10.7205/milmed-d-10-00440.

[3] World Federation of Occupational Therapists. Definition of occupational therapy [Internet]. 2012 [cited 2025 Sep 11]. Available from: https://wfot.org/about/about-occupational-therapy.

[4] Fisher AG. Occupation-centred, occupation-based, occupation-focused: Same, same or different? Scand J Occup Ther. 2013; 20(3): 162-73. doi: 10.3109/11038128.2012.754492.

[5] Petersen KS, Bjørkedal STB, Torsting AM, Eplov LF. Occupational therapy interventions in mental health: A scoping review of recent evidence. Int J Ther Rehabil. 2017; 24(1): 30-40. doi: 10.12968/ijtr.2016.0070.

[6] Eklund M. Mental health – An area of high priority in occupational therapy. Br J Occup Ther. 2021; 84(12): 733-5. doi: 10.1177/03080226211037864.

[7] Freeman D, Sheaves B, Waite F, Harvey AG, Harrison PJ. Sleep disturbance and psychiatric disorders. Lancet Psychiatry. 2020; 7(7): 628-37. doi: 10.1016/S2215-0366(20)30136-X.

[8] Lucchesi LM, Pradella-Hallinan M, Lucchesi M, Moraes WA dos S. Sleep in psychiatric disorders. Rev Bras Psiquiatr. 2005; 27(5): 333-9. doi: 10.1590/s1516-44462005000500006.

[9] Yehuda R, Hoge CW, McFarlane AC, Vermetten E, Lanius RA, Nievergelt CM, et al. Post-traumatic stress disorder. Nat Rev Dis Primers. 2015; 1: 15057. doi: 10.1038/nrdp.2015.57.

[10] Koenen KC, Ratanatharathorn A, Ng L, McLaughlin KA, Bromet EJ, Stein DJ, et al. Posttraumatic stress disorder in the world mental health surveys. Psychol Med. 2017; 47(13): 2260-74. doi: 10.1017/s0033291717000708.

[11] Shalev A, Liberzon I, Marmar C. Post-traumatic stress disorder. N Engl J Med. 2017; 376(25): 2459-69. doi: 10.1056/nejmra1612499.

[12] Javidi H, Yadollahie M. Post-traumatic stress disorder. Int J Occup Environ Med. 2012; 3(1): 2-9. Available from: https://pubmed.ncbi.nlm.nih.gov/23022845/

[13] Maher MJ, Rego SA, Asnis GM. Sleep disturbances in patients with post-traumatic stress disorder: Epidemiology, impact and approaches to management. CNS Drugs. 2006; 20(7): 567-90. doi: 10.2165/00023210-200620070-00003.

[14] Miller KE, Brownlow JA, Gehrman PR. Sleep in PTSD: Treatment approaches and outcomes. Curr Opin Psychol. 2019; 34: 12-7. doi: 10.1016/j.copsyc.2019.08.017.

[15] Belleville G, Guay S, Marchand A. Impact of sleep disturbances on PTSD symptoms and perceived health. J Nerv Ment Dis. 2009; 197(2): 126-32. doi: 10.1097/NMD.0b013e3181961d8e.

[16] Germain A. Sleep disturbances as the hallmark of PTSD: Where are we now? Am J Psychiatry. 2013; 170(4): 372-82. doi: 10.1176/appi.ajp.2012.12040432.

[17] Krau SD. The multiple uses of guided imagery. Nurs Clin North Am. 2020; 55(4): 467-74. doi: 10.1016/j.cnur.2020.06.013.

[18] Hart J. Guided imagery. Altern Complement Ther. 2008; 14(6): 295-9. doi: 10.1089/act.2008.14604.

[19] Srivastava A, Chattoraj T. Exploring guided imagery, mindfulness meditation and cognitive restructuring as therapeutic techniques for treating post-traumatic stress disorder. In: The Asian Conference on Psychology & the Behavioral Sciences; 2023 Apr; Tokyo, Japan. doi: 10.22492/issn.2187-4743.2023.9.

[20] Zehetmair C, Tegeler I, Kaufmann C, Klippel A, Reddemann L, Junne F, et al. Stabilizing techniques and guided imagery for traumatized male refugees in a German state registration and reception center: A qualitative study on a psychotherapeutic group intervention. J Clin Med. 2019; 8(6): 894. doi: 10.3390/jcm8060894.

[21] Jain S, McMahon GF, Hasen P, Kozub MP, Porter V, King R, et al. Healing touch with guided imagery for PTSD in returning active duty military: A randomized controlled trial. Mil Med. 2012; 177(9): 1015-21. doi: 10.7205/MILMED-D-11-00290.

[22] Weathers FW, Bovin MJ, Lee DJ, Sloan DM, Schnurr PP, Kaloupek DG, et al. The clinician-administered PTSD scale for DSM–5 (CAPS-5): Development and initial psychometric evaluation in military veterans. Psychol Assess. 2018; 30(3): 383-95. doi: 10.1037/pas0000486.

[23] Germain A, Hall MH, Krakow B, Shear MK, Buysse DJ. A brief sleep scale for posttraumatic stress disorder: Pittsburgh Sleep Quality Index addendum for PTSD. J Anxiety Disord. 2005; 19(2): 233-44. doi: 10.1016/j.janxdis.2004.02.001.

[24] Ho ECM, Siu AMH. Evaluation of an occupation-based sleep programme for people with insomnia. Hong Kong J Occup Ther. 2022; 35(2): 168-79. doi.org/10.1177/15691861221136261.

[25] Schaffer L, Jallo N, Howland L, James K, Glaser D, Arnell K. Guided imagery: An innovative approach to improving maternal sleep quality. J Perinat Neonatal Nurs. 2013; 27(2): 151-9. doi: 10.1097/JPN.0b013e3182870426.

[26] Afshar M, Mohsenzadeh A, Gilasi H, Sadeghi-Gandomani H. The effects of guided imagery on state and trait anxiety and sleep quality among patients receiving hemodialysis: A randomized controlled trial. Complement Ther Med. 2018; 40: 116-21. doi: 10.1016/j.ctim.2018.07.006.

[27] Seda G, Sanchez-Ortuno MM, Welsh CH, Halbower AC, Edinger JD. Comparative meta-analysis of prazosin and imagery rehearsal therapy for nightmare frequency, sleep quality, and posttraumatic stress. J Clin Sleep Med. 2015; 11(1): 11-22. doi: 10.5664/jcsm.4354.

[28] Casement MD, Swanson LM. A meta-analysis of imagery rehearsal for post-trauma nightmares: Effects on nightmare frequency, sleep quality, and posttraumatic stress. Clin Psychol Rev. 2012; 32(6): 566-74. doi: 10.1016/j.cpr.2012.06.002.

[29] Cook JM, Harb GC, Gehrman PR, Cary MS, Gamble GM, Forbes D, et al. Imagery rehearsal for posttraumatic nightmares: A randomized controlled trial. J Trauma Stress. 2010; 23(5): 553-63. doi: 10.1002/jts.20569.

[30] Harb GC, Cook JM, Phelps AJ, Gehrman PR, Forbes D, Localio R, et al. Randomized controlled trial of imagery rehearsal for posttraumatic nightmares in combat veterans. J Clin Sleep Med. 2019; 15(5): 757-67. doi: 10.5664/jcsm.7770.

[31] Lancee J, Effting M, Kunze AE. Telephone-guided imagery rehearsal therapy for nightmares: Efficacy and mediator of change. J Sleep Res. 2021; 30(3): e13123. doi: 10.1111/jsr.13123.

[32] Krakow B, Hollifield M, Johnston L, Koss M, Schrader R, Warner TD, et al. Imagery rehearsal

therapy for chronic nightmares in sexual assault survivors with posttraumatic stress disorder: A randomized controlled trial. JAMA. 2001; 286(5): 537-45. doi: 10.1001/jama.286.5.537.

[33] Sandahl H, Jennum P, Baandrup L, Mortensen EL, Carlsson J. Imagery rehearsal therapy and/or mianserin in treatment of refugees diagnosed with PTSD: Results from a randomized controlled trial. J Sleep Res. 2021; 30(4): e13276. doi: 10.1111/jsr.13276.

[34] Beck BD, Messel C, Meyer SL, Cordtz TO, Søgaard U, Simonsen E, et al. Feasibility of trauma-focused guided imagery and music with adult refugees diagnosed with PTSD: A pilot study. Nord J Music Ther. 2018; 27(1): 67-86. doi: 10.1080/08098131.2017.1286368.

[35] Story KM, Beck BD. Guided imagery and music with female military veterans: An intervention development study. Arts Psychother. 2017; 55: 93-102. doi: 10.1016/j.aip.2017.05.003.

Appendix A

Guided imagery therapy for experimental group

Session 1:
- Pre-test examined on the basis of Pittsburgh Sleep Quality Index (PSQI).

Session 2:
- Theme: Safe Place Visualization.
- Materials: Pre-recorded guided imagery audio.
- Practice deep breathing and visualize a personally safe, calm location.

Session 3:
- Theme: Nature Scene Relaxation.
- Materials: Guided imagery script with soft instrumental background music.
- Imagine lying in a peaceful meadow, focusing on sensory details.

Session 4:
- Theme: Ocean Wave Breathing.
- Materials: Audio track with rhythmic ocean sounds.
- Match breathing to the sound of gentle waves.

Session 5:
- Theme: Sunset Visualization.
- Materials: Script describing a slow-setting sun.
- Focus on body relaxation as the sky changes colours.

Session 6:
- Theme: Forest Walk.
- Materials: Guided imagery script with nature sounds.
- Imagine walking slowly through a forest, feeling grounded.

Session 7:
- Theme: Mountain Top Calm.
- Materials: Script and soft wind sound effects.
- Visualize reaching a mountain top and feeling peace.

Session 8:
- Theme: Healing Light.
- Materials: Audio describing warm golden light.
- Imagine light gently flowing through the body, easing tension.

Session 9:
- Theme: River Flow Relaxation.
- Materials: Script with flowing water sounds.
- Imagine floating along a calm river.

Session 10:
- Theme: Starry Night Sky.
- Materials: Audio with gentle ambient tones.
- Visualize lying under a clear night sky, focusing on stillness.

Session 11:
- Theme: Garden of Peace.
- Materials: Script with bird song sounds.
- Imagine walking in a flower garden, smelling the scents.

Session 12:
- Theme: Cloud Drift.
- Materials: Audio with slow, soft music
- Imagine lying on clouds drifting slowly across the sky.

Session 13:
- Theme: Rainfall Relaxation.
- Materials: Script with gentle rain sound effects.
- Imagine sitting by a window watching soft rain.

Session 14:
- Theme: Cozy Cabin.
- Materials: Guided imagery describing a warm cabin fireplace.
- Focus on warmth, comfort, and safety.

Session 15:
- Theme: Morning Sunrise Energy.
- Materials: Audio with soft morning bird sounds.
- Visualize a sunrise bringing gentle warmth and optimism.

Session 16:
- Theme: Floating on Water.
- Materials: Script with soft wave background.
- Imagine floating effortlessly on calm water.

Session 17:
- Theme: Blossoming Tree.
- Materials: Guided imagery describing a tree growing and blooming.
- Focus on personal growth and healing.

Session 18:
- Theme: Peaceful Countryside.
- Materials: Script with distant farm and meadow sounds.
- Imagine walking through open fields with fresh air.

Session 19:
- Theme: Candlelight Calm.
- Materials: Audio describing soft flickering candlelight.
- Focus on slow breathing and visualizing the flame.

Session 20:
- Theme: Gentle Stream.
- Materials: Script with soft running water.
- Imagine dipping hands and feet into a cool stream.

Session 21:
- Theme: Desert Dusk.
- Materials: Audio with light wind sounds.
- Visualize a sunset over a peaceful desert landscape.

Session 22:
- Theme: Butterfly Meadow.
- Materials: Script with soft nature sounds.
- Imagine butterflies moving gently through flowers.

Session 23:
- Theme: Snowy Silence.
- Materials: Audio describing a snowy, quiet forest.
- Focus on stillness and calm.

Session 24:
- Theme: Horizon Journey.
- Materials: Guided imagery describing walking toward a peaceful horizon.
- Focus on letting go of stress with each step.

Session 25:
- Theme: Island Retreat.
- Materials: Audio with ocean and bird sounds.
- Imagine resting in a hammock on a quiet island.

Session 26:
- Theme: Spring Renewal.
- Materials: Script with soft spring rain and birdsong.
- Visualize nature coming alive after winter.

Session 27:
- Theme: Gentle Breeze.
- Materials: Audio with light wind.
- Imagine feeling a cool breeze passing over the skin.

Session 28:
- Theme: Harbor Rest.
- Materials: Script with water and distant seagull sounds.
- Imagine sitting safely by a quiet harbour.

Session 29:
- Theme: Crystal Cave Calm.
- Materials: Guided imagery describing a glowing cave.
- Visualize shimmering crystals reflecting peaceful light.

Session 30:
- Post-test assessed using the **Pittsburgh Sleep Quality Index (PSQI)**.

Wind-up Activities

Involved gentle grounding, slow stretching, and positive affirmations to close the session calmly.[25]

Journal of Associated Medical Sciences 2026; 59(2): 163-177

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Assessment of imaging radiation dose from tomotherapy systems

Pratchayakan Hompeng[1,2], Wannapha Nobnop[3], Tinnagorn Donmoon[4], Komkrit Krongkietlearts[5], Kananan Utitsarn[5], Sumalee Yabsantia[1,6], Thunyarat Chusin[1,6*]

[1]Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.
[2]Regional Medical Sciences Center 7 Khonkaen, Khonkaen Province, Thailand.
[3]Division of Radiation Oncology, Faculty of Medicine, Chiang Mai University, ChiangMai Province, Thailand.
[4]Maha Vajiralongkorn Thanyaburi Hospital, Department of Medical Services, Ministry of Public Health, Pathum Thani Province, Thailand.
[5]Department of Radiotherapy, Lopburi Cancer Hospital, Lopburi Province, Thailand.
[6]Interdisciplinary Health and Data Sciences Research Unit, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.

ARTICLE INFO

Article history:
Received 19 October 2025
Accepted as revised 30 January 2026
Available online 10 February 2026

Keywords:
Tomotherapy systems, imaging radiation dose, kVCT, MVCT.

ABSTRACT

Background: Tomotherapy systems employ image-guided radiotherapy (IGRT) using either megavoltage (MVCT) or kilovoltage (kVCT) computed tomography for daily patient positioning verification. However, these imaging procedures can deliver a non-negligible radiation dose to the patient. According to the American Association of Physicists in Medicine (AAPM) Task Group 180, this dose should be incorporated into the prescribed treatment dose if it exceeds 5%.

Objectives: This study aimed to assess IGRT imaging doses and determine whether the cumulative dose exceeded 5% of the prescribed treatment dose.

Materials and methods: Imaging doses from IGRT were evaluated under clinical scan protocols across four tomotherapy machines at three centers, including two MVCT systems and two kVCT systems. Doses were measured as Weighted Multiple Slice Average Dose ($MSAD_w$) using a CTDI phantom. Cumulative imaging dose was calculated by multiplying $MSAD_w$ by the number of imaging sessions, then compared to the prescribed dose for each patient.

Results: The mean $MSAD_w$ values for MVCT ranged from 0.85 to 0.86 cGy per scan, while for kVCT it ranged from 0.19 to 2.65 cGy per scan. Over the treatment course, cumulative imaging doses accounted for 0.34-0.44% of the prescribed dose for MVCT and 0.28-0.87% for kVCT. Notable inter-center variations were observed, with differences of up to 51.16% in breast cancer cases and 35.71% in prostate cancer cases.

Conclusion: Cumulative imaging doses from both MVCT and kVCT across all centers remained below the 5% threshold, indicating no need to adjust the prescribed dose in clinical practice. Nevertheless, further studies are required to identify the causes of inter-center variations and to develop optimized imaging protocols aimed at minimizing patient dose while maintaining adequate image quality.

Introduction

Image-guided radiation therapy (IGRT) is an advanced technique designed to improve treatment accuracy by using daily pre-treatment imaging to verify patient positioning. The acquired images are registered with the reference images through alignment of key anatomical landmarks, thereby ensuring precise patient setup and enhancing tumor localization.[1]

*** Corresponding contributor.**
Author's Address: *Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*
E-mail address: thunyaratc@nu.ac.th
doi: 10.12982/JAMS.2026.048
E-ISSN: 2539-6056

Tomotherapy systems incorporate an IGRT platform capable of generating cross-sectional 3D images using either megavoltage computed tomography (MVCT) or kilovoltage computed tomography (kVCT) for daily position verification. Historically, MVCT was the standard modality. MVCT images are produced using the same X-ray source as the treatment beam, operated at a reduced energy of 3.5 MV. However, this approach is limited by relatively low image quality and prolonged acquisition times. To address these limitations, newer-generation tomotherapy units may be equipped with kVCT, which provides high-resolution images comparable to diagnostic CT and enables shorter scan times. Thus, the choice between MVCT and kVCT is determined by clinical requirements and system availability.[2]

Since cross-sectional images are used to verify and adjust patient positioning, sufficient image resolution is essential for visualizing target structures such as bone and soft tissue. Even though higher imaging doses can improve image quality, they also contribute additional radiation exposure. Therefore, imaging protocols should adhere to the As Low As Reasonably Achievable (ALARA) principle, ensuring adequate image quality for clinical purposes while minimizing unnecessary dose.[3,4] Although radiation therapy patients already receive high localized treatment doses, the additional exposure from imaging still carries risk and must be kept as low as reasonably possible.[5]

According to the American Association of Physicists in Medicine (AAPM) Task Group 180 report, the imaging dose should be included in the total prescribed dose if it exceeds 5% of the therapeutic dose.[6] This recommendation is supported by a study by Dische *et al.*,[7] which found that dose variations as small as 5%, based on clinical data from the continuous hyperfractionated accelerated radiotherapy (CHART) pilot study, may influence tumor control and increase patient morbidity.[8,9]

The AAPM Task Group 148 report recommends that radiation dose measurement for MVCT be defined using the Multiple Slice Average Dose (MSAD).[10] This dose is typically measured with a calibrated ionization chamber (e.g., Exradin A1SL) within a standard cylindrical phantom (e.g., the Cheese phantom). The scan range should cover the entire phantom. For imaging in normal mode, MSAD of approximately 1-3 cGy can be expected.[10] Although a Cheese phantom is commonly used in tomotherapy centers for radiation dose measurements, it has certain limitations when used to evaluate the MSAD. These limitations include a slight displacement of the phantom's designated center, which may not perfectly align with its actual midpoint. The phantom has a diameter of 30 cm, making it mismatch for head imaging dose assessments. It is also incompatible with pencil ionization chambers, which are required for measuring imaging doses from kVCT.

Several studies have investigated imaging doses associated with IGRT across different systems and protocols. Mege *et al.*[11] and Marco *et al.*[12] reported MVCT imaging doses of approximately 1-4 cGy per scan, depending on pitch settings and anatomical site, with higher pitch values generally associated with lower doses. Trivedi *et al.*[13] reported average CBCT imaging doses of 0.24 cGy for head and neck cancer and 0.76 cGy for pelvic cancer. Imaging was performed once weekly for head and neck cancer and on three consecutive days during the first week for pelvic cancer. Although cumulative doses were not directly measured and 2D imaging doses were excluded, the authors estimated total doses of 1.44 cGy and 5.32 cGy, both remaining below 5% of the prescribed dose. Additionally, Sakai *et al.*[3] reported imaging doses per kV-CBCT and per kVCT acquisition across nine machines at seven centers, with ranges of 0.09-0.66 cGy for head and 0.15-2.90 cGy for pelvic protocols. While imaging doses varied among machines and centers, the maximum dose difference was sevenfold in the head protocol and up to fourteenfold in the pelvic protocol, highlighting the need to establish standardized diagnostic reference levels (DRLs) in IGRT.

A review of existing literature shows that most studies assess imaging radiation doses using phantom measurements from a single scan. However, there is limited evaluation of the cumulative doses patients receive throughout the entire treatment course across difference cancer types. Moreover, no published data from Thailand currently exist on patient imaging doses associated with IGRT using tomotherapy systems. Such data are essential not only for considering whether to include the imaging dose in the prescribed dose, but also for evaluating appropriate imaging protocols and optimizing patient dose in clinical practice, in alignment with the ALARA principle.

This study aimed to evaluate the radiation dose used for imaging with the IGRT system under clinical scan protocols for kVCT and MVCT at three radiotherapy centers. The dose was quantified as Weighted Multiple Slice Average Dose ($MSAD_w$), measured in a Computed Tomography Dose Index (CTDI) phantom. The study also assessed whether the cumulative imaging dose received by patients throughout the course of treatment on tomotherapy systems exceeded 5% of the prescribed dose, according to AAPM Task Group 180 report.

Materials and methods

Tomotherapy system

This study evaluated the imaging dose across four tomotherapy machines at three centers. The MVCT systems included a Radixact X9 at Center A (MVCT 1) and a Radixact X5 at Center B (MVCT 2). The kVCT systems consisted of a Radixact X9 at Center A (kVCT 1) and a Radixact X5 at Center C (kVCT 2). All systems operated with software version 3.0.x (Accuray Inc., Madison, WI, USA).

Megavoltage computed tomography (MVCT)

MVCT uses the same linear accelerator source as the treatment beam but operates at a reduced energy level of approximately 3.5 MV for imaging purposes. MVCT generates cross-sectional images with a 400 mm field of view (FOV) and a matrix resolution of 512×512 pixels. The jaw width is fixed at the J1 setting, corresponding to a nominal beam width of approximately 4 mm at the isocenter. The pitch setting in MVCT controls the couch movement per gantry rotation, influencing image resolution. Fine, Normal, and Coarse modes correspond to couch speeds of 4, 8, and 12 mm per rotation and reconstructed slice thicknesses of 2, 4, and 6 mm, respectively.[1,14] The scan time depends on the selected pitch; the Coarse mode requires approximately one-third of the scan time and delivers about one-third of the radiation dose compared to the Fine mode. In this study, the clinical scan protocols for MVCT typically used in all centers are detailed in Table 1.

Table 1. Clinical scan protocols for MVCT used in this study.

Mode	Voltage (MV)	Slice thickness (mm)	Pitch	FOV (mm)
Coarse	3.5	6	3	400

Kilovoltage computed tomography (kVCT)

The kV X-ray tube is mounted at 267.5 degrees relative to the treatment beam position. A CsI:Tl flat-panel detector is installed directly opposite the X-ray source, with a source-to-detector distance of 155 cm. Imaging parameters for kVCT are not manually adjustable by users. Instead, predefined scan protocols are selected based on the anatomical region and patient body size. The selected anatomical region determines the tube voltage (kVp), reconstructed slice interval, and slice thickness, while body size determines the tube current per view (mA/view). The mode selection determines the longitudinal beam width, couch speed, views per rotation, scan time, and pitch. The pitch is fixed for 440 mm and 500 mm fields of view (FOV), but adjustable for a 270 mm FOV. FOV selection also determines the type of filtration: a 0.5 mm copper filter is used for 270 mm FOV, while a bowtie filter is used for 440 mm and 500 mm FOVs.[1,14] These clinical scan protocols for kVCT used in this study are detailed in Table 2.

Table 2. Clinical scan protocols for kVCT were used in this study.

Protocol	Voltage (kVp)	Body size	mAs	Slice interval (mm)	Slice thickness (mm)	Mode	Pitch	FOV (mm)
Head	100	Small	240	1.2	2.4	Fine	0.86	270
		Medium	375					
		Small	192			Normal	1	
		Medium	300					
		Small	144			Coarse	1.4	
		Medium	225					
Thorax	120	Small	240	1.8	3.6	Fine	0.75	440
		Medium	375					
		Large	480					
		Small	192			Normal		
		Medium	300					
		Large	384					
		Small	144			Coarse		
		Medium	225					
		Large	288					

Table 2. Clinical scan protocols for kVCT were used in this study (continued).

Protocol	Voltage (kVp)	Body size	mAs	Slice interval (mm)	Slice thickness (mm)	Mode	Pitch	FOV (mm)
Pelvis	140	Small	240	1.8	3.6	Fine	0.75	440
		Medium	375					
		Large	480					
		Small	192			Normal		
		Medium	300					
		Large	384					
		Small	144			Coarse		
		Medium	225					
		Large	288					

MSAD measurement

Dose assessment of MVCT imaging

The AAPM Task Group 148 report recommends that radiation dose measurements for MVCT be defined using the MSAD, which is measured with a calibrated ionization chamber placed within a standard cylindrical phantom. The report specifies that the scan range should cover the entire length of the phantom to ensure accurate integration of dose contributions from multiple slices.[10]

The Exradin A1SL ionization chamber (Standard Imaging Inc., Middleton, WI, USA) was used for dose measurements at both Center A and Center B. The chamber used at Center A was calibrated in water by the Bureau of Radiation and Medical Devices, Department of Medical Sciences, in November 2022, while the chamber used at Center B was calibrated in water by the Department of Medical Physics, University of Wisconsin-Madison, in February 2022. Dose measurements were performed with the chamber positioned in a CTDI phantom (GAMMEX Inc., Middleton, WI, USA). Both chambers were calibrated within a two-year interval, which is considered acceptable under national dosimetry practice, provided that calibration is traceable to a national standards laboratory and detector stability is maintained.

The Exradin A1SL ionization chamber has a sensitive volume of 0.053 cm^3 and an outer diameter of 6.35 mm. The sensitive (collecting) volume has an axial length of approximately 4.4 mm. Because the chamber diameter is smaller than the standard CTDI phantom insert hole, a custom-designed PMMA rod was fabricated to ensure a secure fit, as illustrated in Figure 1.

The custom-designed PMMA rod has an effective length of 140 mm, excluding the 10-mm head portion, and contains a central hole with a length of 74.1 mm. This hole length was determined as the sum of the distance from the rod end to the center of the phantom (70 mm) and the distance from the center of the centroid of the collecting volume to the end of the A1SL chamber wall (4.1 mm), thereby ensuring accurate centering of the sensitive volume within the CTDI phantom.

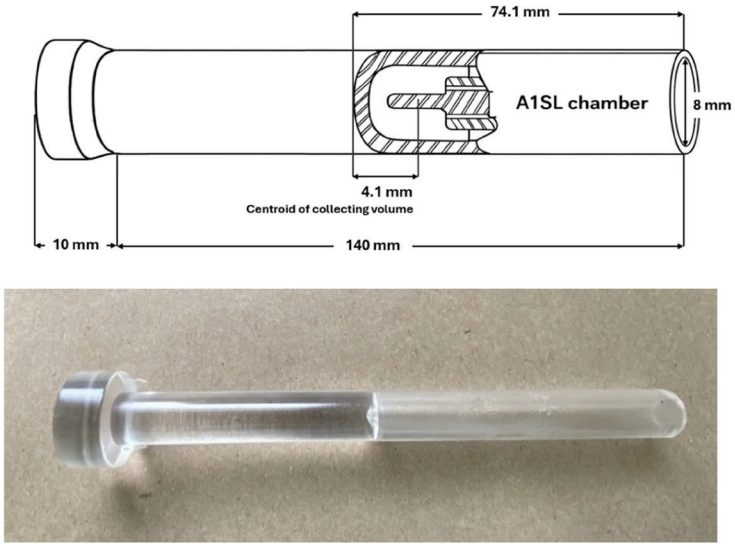


Figure 1. *Custom-designed PMMA rod for holding the Exradin A1SL ionization chamber.*

The CTDI phantom measured 16 cm in diameter was used for the head scan protocol and 32 cm for the body scan protocol, with a uniform length of 15 cm. The phantom was placed on the table, with its center aligned to the isocenter of the tomotherapy machines as shown in Figure 2.

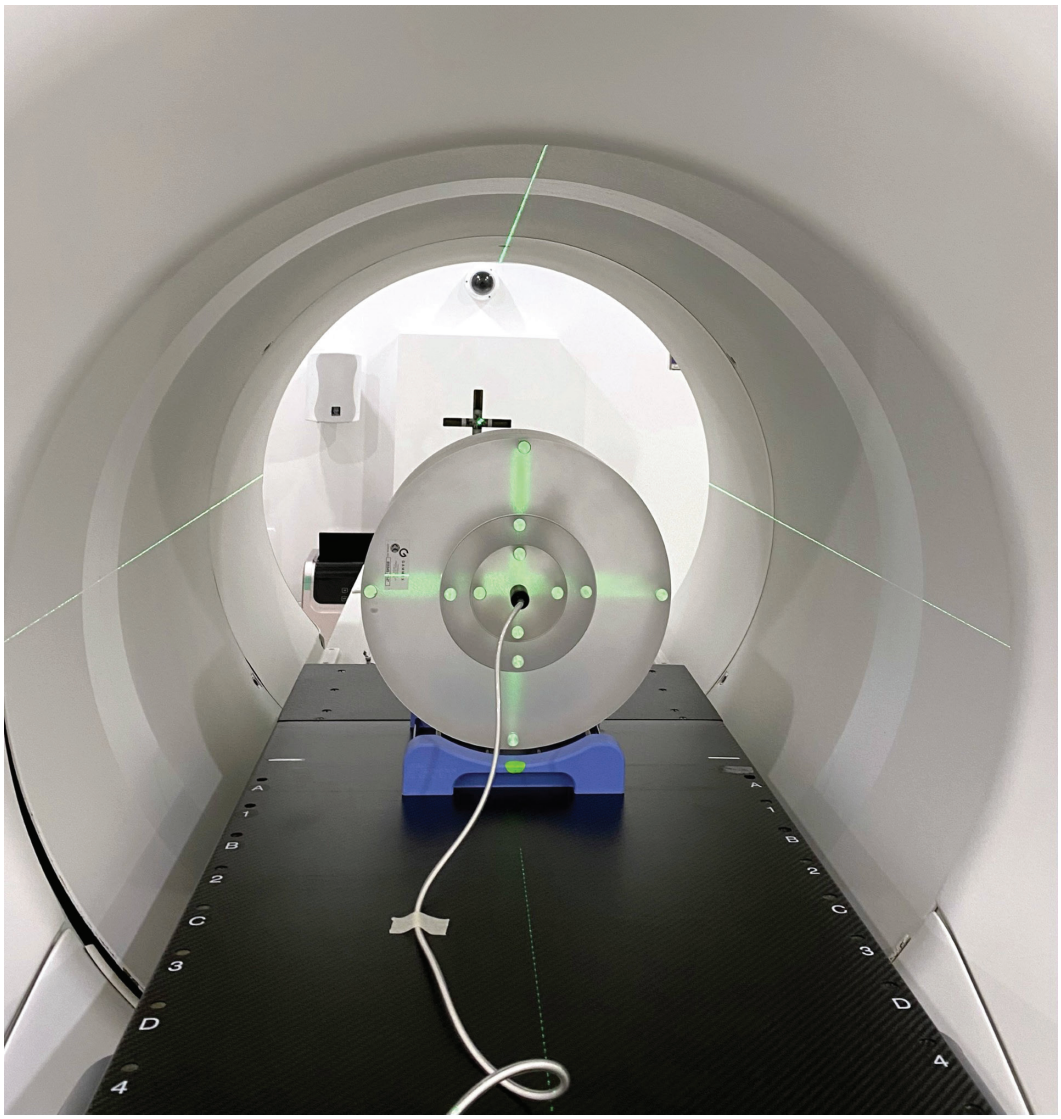


Figure 2. *The equipment setup for body CTDI phantom dose measurement.*

All scans were performed in Coarse mode using a 3.5 MV beam, with a pitch of 3 and a slice thickness of 6 mm (Table 1). A scan length of 15 cm was selected to cover the entire phantom. Each measurement was repeated twice. The MSAD and per scan were calculated according to Equations (1) and (2), respectively.

$$MSAD = M \times N_{D,w,Q_o} \quad (1)$$

where M is the meter reading (nC), and N_{D,w,Q_o} is the absorbed dose-to-water calibration factor (cGy/nC). The values of N_{D,w,Q_o} were 59.58 cGy/nC for Center A (MVCT 1) and 57.87 cGy/nC for Center B (MVCT 2).

$$MSAD_w^{MVCT} = \frac{1}{3}MSAD_{center} + \frac{2}{3}MSAD_{periphery} \quad (2)$$

where $MSAD_{center}$ is the dose measured at the center of the phantom, and $MSAD_{periphery}$ is the average dose measured at the four peripheral positions. Further details of the calculation can be found in the AAPM Task Group 75 report.[15]

Dose measurements were performed according to the recommendations of AAPM Task Group 148 report. Accordingly, no correction was applied for spectral differences between imaging and treatment beams. Pitch correction was also unnecessary, as meter readings were acquired using the clinical pitch settings.[16]

Dose assessment of kVCT imaging

Since axial scanning is unavailable for kVCT imaging on the Radixact system, an MSAD-based methodology consistent with MVCT was used, reflecting the similar helical acquisition principles of both modalities.[17]

A 100-mm pencil ionization chamber (RaySafe X2 CT, Unfors RaySafe AB, Billdal, Sweden) was employed for dose measurements. The chamber was calibrated in air by the Bureau of Radiation and Medical Devices, Department of Medical Sciences, in April 2024. All scans followed the clinical protocols summarized in Table 2, with a fixed scan length of 10 cm, as described by Leon *et al.*[16] MSAD per scan were calculated using Equations (3) and (4), respectively.

$$MSAD = M \times N_{K,Q_0} \times f \quad (3)$$

where M is the meter reading, N_{K,Q_0} is the air kerma calibration factor, and f is a correction factor of 1.06 used to convert air kerma to absorbed dose in tissue for typical diagnostic CT beam qualities, corresponding to an effective photon energy of approximately 70 keV (~120 kVp).[18] Although the air-to-tissue conversion factor can exhibit some energy dependence, variations across the kVp range used in this study are expected to be small, on the order of approximately 2% across 100-140 kVp. The values of N_{K,Q_0} were 0.994, 0.995, and 0.997 for beam qualities of 100 kVp, 120 kVp, and 140 kVp, respectively, for both Center A and Center C.

$$MSAD_w^{kVCT} = \frac{1}{3}MSAD_{center} + \frac{2}{3}MSAD_{periphery} \quad (4)$$

Patient data

The study was approved by the Ethics Committee of Naresuan University, Thailand (IRB No. P1-0127/2024; approval date: 3 September 2024). Patient data were retrospectively collected between July 2023 and September 2024 from those who underwent IGRT for positioning verification. Inclusion criteria included patients aged ≥18 years with a body weight of 60±15 kg during treatment, representing the standard Thai adult population in accordance with the National Diagnostic Reference Levels in Thailand 2023.[19] These criteria were applied to ensure consistency in patient anatomy and imaging protocols and to minimize body size–related bias. Patients were also required to have completed the full prescribed treatment course and to have received IGRT scans encompassing the planning target volume (PTV). Patient data were categorized into four categories: 1) sex, age, and body weight, 2) prescribed radiation dose, 3) imaging protocol, and 4) number of imaging sessions. For each type of cancer, 30 patients were randomly selected.[20] If fewer patients were available, all eligible cases from July 2023 to September 2024 were included.

At Center A, MVCT 1 was used for breast, prostate, and cervical cancer patients, while kVCT 1 was used for brain, head and neck, breast, prostate, and cervical cancer patients. Centers B and C contributed to breast and prostate cancer patients.

Cumulative dose

To estimate the cumulative imaging dose, the $MSAD_w$ measured in a phantom for each scan protocol was multiplied by the total number of imaging sessions (n) each patient received during treatment. The calculation methods for MVCT and kVCT are shown in Equations (5) and (6).

$$\text{Cumulative imaging dose}_{MVCT} = MSAD_w^{MVCT} \times n \quad (5)$$

$$\text{Cumulative imaging dose}_{kVCT} = MSAD_w^{kVCT} \times n \quad (6)$$

The cumulative imaging dose throughout the treatment course was expressed as a percentage of the mean prescribed dose, as shown in Equation (7).

$$\%\text{Imaging dose} = \left(\frac{\text{Cumulative imaging dose throughout the treatment}}{\text{Mean prescribed dose}}\right) \times 100 \quad (7)$$

The percentage of imaging dose throughout the treatment course was used to determine the proportion of patients whose cumulative imaging dose exceeded the 5% threshold, according to AAPM Task Group 180 report.[6]

Data analysis

An MSAD value was obtained for each measurement condition at each center based on twice-repeated phantom measurements. Owing to the limited number of independent MSAD values, inter-center comparisons were performed using descriptive metrics, specifically percentage differences, rather than formal statistical testing.

Results

Imaging dose measurement in the CTDI phantom

MVCT imaging

The mean $MSAD_w^{MVCT}$ values were 0.85±0.01 cGy at Center A and 0.86±0.00 cGy at Center B. The percentage difference between the two centers was 1.17%, as shown in Figure 3.

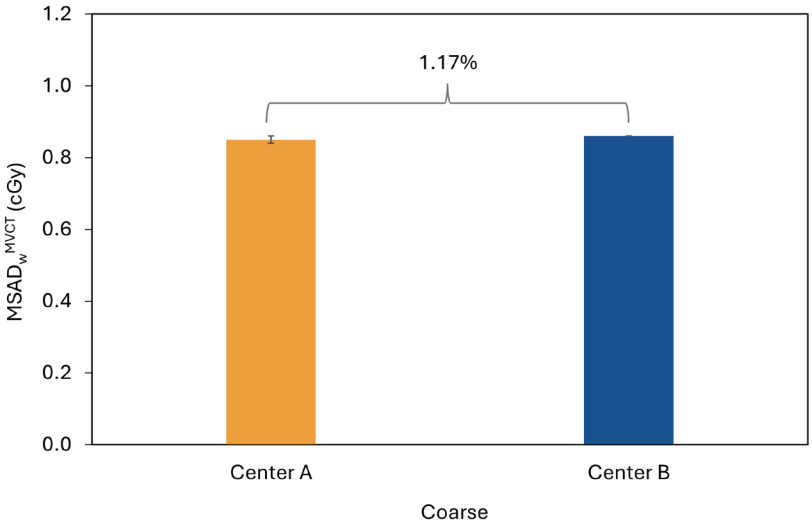


Figure 3. *The MSAD$_w^{MVCT}$ values for Coarse mode obtained from Center A and Center B.*

kVCT imaging

Overall, the MSAD$_w^{kVCT}$ values demonstrate notable variations across different scan modes and body sizes. The Fine mode yielded the highest MSAD$_w^{kVCT}$ values, whereas the Coarse mode resulted in the lowest values. MSAD$_w^{kVCT}$ values increased by a range of 1.5-1.7 fold when comparing small to medium body sizes.

The mean MSAD$_w^{kVCT}$ values for the head protocol at Center A across the three modes (fine, normal, and coarse) ranged from 0.19±0.01 to 0.49±0.00 cGy and 0.26±0.02 to 0.82±0.00 cGy for small and medium body sizes, respectively. At Center C, the corresponding ranges were 0.22±0.01 to 0.47±0.01 and 0.33±0.00 to 0.83±0.01 cGy, respectively. The percentage difference between the two centers was within 20% for all modes, except for the Coarse mode with the medium body size, which showed a difference of 23.73%, as shown in Figure 4.

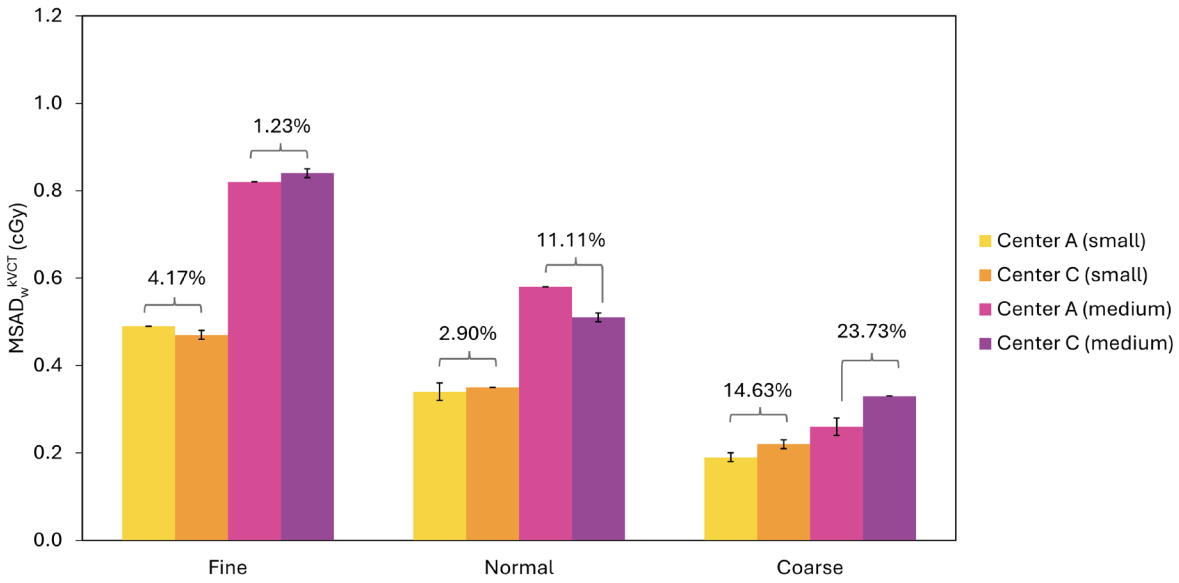


Figure 4. *The MSAD$_w^{kVCT}$ values for head protocol obtained from Center A and Center C.*

The mean MSAD$_w^{kVCT}$ values for the thorax protocol at Center A across the three modes (fine, normal coarse) ranged from 0.49±0.02 to 0.86±0.00 cGy, 0.78±0.03 to 1.31±0.01 cGy, and 1.11±0.01 to 1.68±0.00 cGy for small, medium, and large body sizes, respectively. At Center C, the corresponding ranges were 0.55±0.01 to 0.97±0.06 cGy, 0.83±0.02 to 1.46±0.01 cGy, and 1.12±0.05 to 1.96±0.04 cGy, respectively. The percentage difference between the two centers was within 20% for all modes, except for the Normal mode with small body size, which showed a difference of 21.33%, as shown in Figure 5.

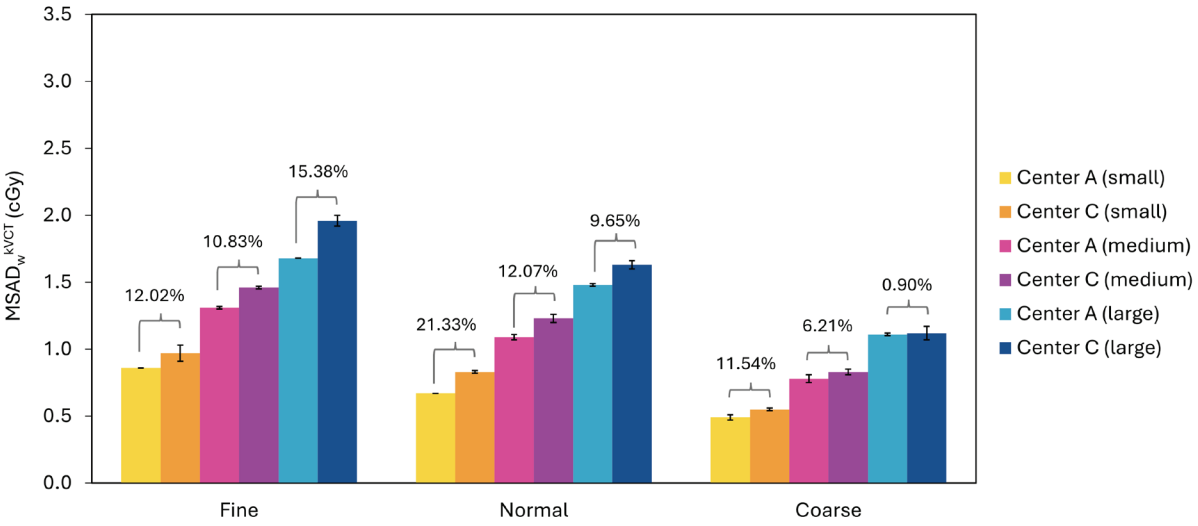


Figure 5. *The MSAD$_w^{kVCT}$ values measured using the thorax protocol from Center A and Center C.*

The mean MSAD$_w^{kVCT}$ values for the pelvis protocol at Center A across the three modes (fine, normal, and coarse) ranged from 0.74±0.01 to 1.19±0.00 cGy, 1.13±0.00 to 1.93±0.06 cGy, and 1.53±0.02 to 2.65±0.05 cGy for small, medium, and large body sizes, respectively. At Center C, the corresponding ranges were 0.79 ± 0.01 to 1.42±0.02 cGy, 1.27±0.01 to 2.22±0.02 cGy, and 1.58±0.05 to 2.61±0.04 cGy, respectively. The percentage difference between the two centers was within 20% for all modes, except for the Normal mode with the small body size, which showed a difference of 21.00%, as shown in Figure 6.

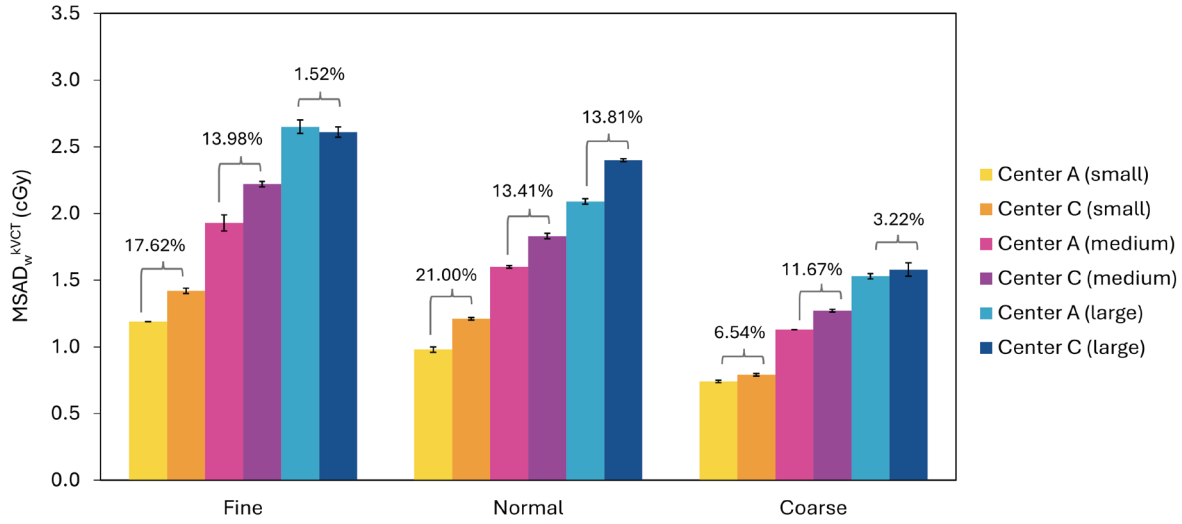


Figure 6. *The MSAD$_w^{kVCT}$ values measured using the pelvis protocol from Center A and Center C.*

The cumulative imaging dose received by the patient throughout the treatment course

MVCT imaging

Patient data were collected from MVCT imaging throughout the treatment course at Center A (MVCT 1) for patients with breast, prostate, and cervical cancers and at Center B (MVCT 2) for patients with breast and prostate cancers, as summarized in Table 3. The mean age differences between centers were 3.88% for breast cancer and 2.59% for prostate cancer patients. Mean body weight differences for the same groups were 0.30% and 6.78%, respectively.

Table 3. Patient data from MVCT imaging at Centers A and B.

Center	Cancer type	Sex	Number	Age (years)	Weight (kg)
A (MVCT 1)	Breast	Female	30	53.57±11.60	60.20±9.81
	Prostate	Male	30	72.73±6.83	69.87±13.42
	Cervix	Female	30	53.53±13.29	56.37±8.74
B (MVCT 2)	Breast	Female	30	51.53±11.87	60.38±10.25
	Prostate	Male	30	70.87±7.63	65.29±9.62

Note: *data presented as Mean±SD*

Prescribed dose data collected from MVCT imaging are summarized in Table 4. The mean prescribed dose for breast cancer patients at Center A was 3.06% higher than at Center B. In contrast, for prostate cancer patients, the mean prescribed dose of Center B was 1.21% higher than that at Center A. Additionally, the average number of fractions at Center B was 31.58% higher for breast cancer patients and 3.92% higher for prostate cancer patients compared to Center A.

Table 4. Prescribed dose data from MVCT imaging at Centers A and B.

Center	Cancer type	Prescribed dose (PD)	PD (cGy)	Fraction
A (MVCT 1)	Breast	4,240 cGy/16 fraction-4,960 cGy/16 fraction	4,840±272.92	16±0.00
	Prostate	5,500 cGy/20 fraction-7,560 cGy/28 fraction	6,564±589.40	25±4.12
	Cervix	4,600 cGy/23 fraction-5,700 cGy/25 fraction	4,852±399.46	23±0.81
B (MVCT 2)	Breast	4,256 cGy/16 fraction-5,000 cGy/25 fraction	4,694±366.58	22±4.61
	Prostate	6,000 cGy/20 fraction-7,600 cGy/38 fraction	6,644±624.78	26±6.70

Note: *data presented as Mean±SD*

The average number of MVCT imaging sessions per patient throughout the treatment course was 13.95% higher at Center B than at Center A for breast cancer. For prostate cancer patients, both centers reported the same average number of sessions, as shown in Table 5. Additionally, most patients at both centers received one MVCT scan per treatment fraction.

Table 5. Number of MVCT imaging sessions per patient throughout the treatment course.

Cancer type	Center A MVCT 1 (sessions)	Center B MVCT 2 (sessions)
Breast	20±4.9 (16-30)	23±4.6 (16-28)
Prostate	27±5.0 (20-37)	27±6.7 (20-40)
Cervix	25±2.5 (23-34)	-

Note: *data presented as Mean±SD*

The cumulative MVCT imaging dose throughout the treatment course at Center A was 17.40±4.15 cGy for breast cancer, 22.61±4.26 cGy for prostate cancer, and 21.34±2.12 cGy for cervical cancer patients. At Center B, doses were 20.56±4.68 cGy for breast cancer and 24.02±6.85 cGy for prostate cancer, as shown in Figure 7. The mean dose at Center B was 16.65% higher than at Center A for breast cancer and 6.05% higher for prostate cancer.

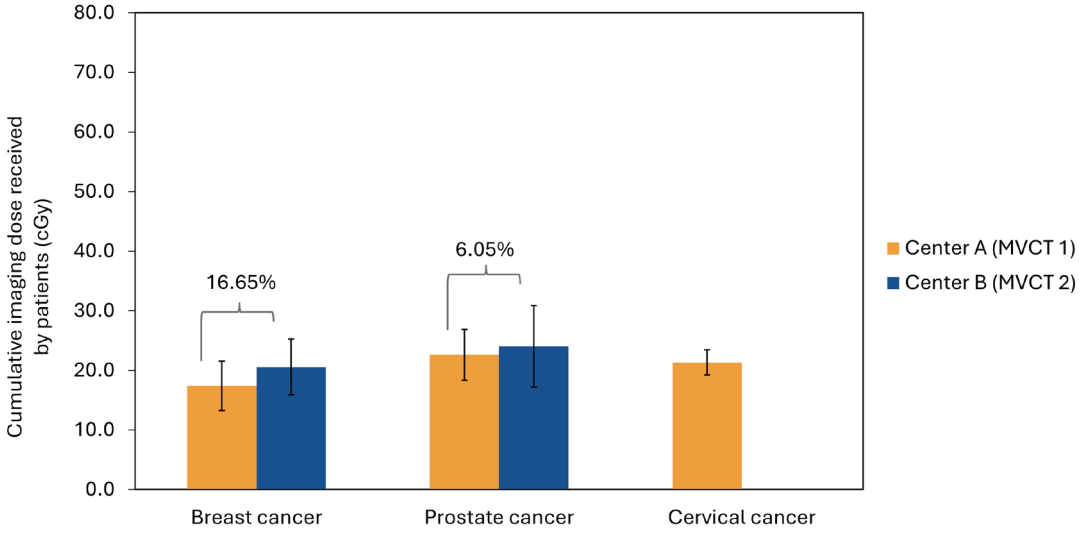


Figure 7. *Cumulative MVCT imaging dose received by patients throughout the treatment course.*

The percentage of cumulative MVCT imaging dose relative to the prescribed dose at Center A was 0.36±0.11 for breast cancer, 0.34±0.04 for prostate cancer, and 0.44±0.03 for cervical cancer. At Center B, the corresponding values were 0.43±0.08 for breast cancer and 0.36±0.07 for prostate cancer. Nevertheless, the cumulative imaging dose remained below 5% of the prescribed dose for all cancer types at both centers.

kVCT imaging

Patient data were collected from kVCT imaging throughout the treatment course at Center A (kVCT 1) for patients with brain, head and neck, breast, prostate, and cervical cancers and at Center C (kVCT 2) for patients with breast and prostate cancers, as summarized in Table 6. The mean age differences between centers were 7.03% for breast cancer and 0.50% for prostate cancer patients. Mean body weight differences for the same groups were 6.36% and 0.55%, respectively.

Center A included a wider variety of cancer types because the kVCT 1 system was used across multiple disease sites, while some brain and head and neck cancer patients were treated with other radiotherapy machines. In contrast, the kVCT 2 system at Center C was used exclusively for breast and prostate cancer patients.

Table 6. Patient data from kVCT imaging at Centers A and Center C.

Center	Cancer type	Sex	Number	Age (years)	Weight (kg)
A (kVCT 1)	Brain	Male	5	45.56±17.62	60.88±12.99
		Female	11		
	Head and neck	Male	17	60.20±10.69	53.60±8.24
		Female	13		
	Breast	Female	30	56.97±10.69	60.52±12.30
	Prostate	Male	30	69.97±5.40	67.00±13.23
	Cervix	Female	30	52.40±13.28	59.69±11.82
C (kVCT 2)	Breast	Female	30	61.12±11.13	56.79±9.37
	Prostate	Male	30	69.62±6.25	67.37±13.06

Note: *data presented as Mean±SD*

Prescribed dose data collected from kVCT imaging are summarized in Table 7. The prescribed dose for breast cancer patients at Center A was 2.44% higher than at Center C. In contrast, for prostate cancer patients, the prescribed dose of Center C was 1.61% higher than that at Center A. Additionally, the average number of fractions at Center A was 6.06% higher for breast cancer patients and 4.44% higher for prostate cancer patients compared to Center C.

Table 7. Prescribed dose data from kVCT imaging at Centers A and Center C.

Center	Cancer type	Prescribed dose (PD)	PD (cGy)	Fraction
A kVCT 1)	Brain	5,400 cGy/27 fraction -7,000 cGy/33 fraction	5,925±640.31	30±1.90
	Head and neck	6,000 cGy/30 fraction -7,000 cGy/33 fraction	6,880±275.93	33±0.76
	Breast	4,240cGy/16 fraction -5,000 cGy/25 fraction	4,558±350.58	17±1.72
	Prostate	5,240 cGy/20 fraction -7,000 cGy/28 fraction	6,049±727.12	23 ± 3.84
	Cervix	4,600 cGy/23 fraction -5,500 cGy/25 fraction	4,848 ± 392.14	23±0.37
C (kVCT 2)	Breast	4,240 cGy/16 fraction -5,000 cGy/25 fraction	4,448±426.26	16±2.63
	Prostate	6,000 cGy/20 fraction -7,800 cGy/39 fraction	6,147±406.39	22±5.33

Note: *data presented as Mean±SD*

The average number of kVCT imaging sessions per patient throughout the treatment course at Center C was 28.57% higher than at Center A for breast cancer and 21.43% higher for prostate cancer, as shown in Table 8.

Table 8. Number of kVCT imaging sessions per patient throughout the treatment course.

Cancer type	Center A kVCT 1 (sessions)	Center C kVCT 2 (sessions)
Brain cancer	32±2.8 (27-38)	-
Head and neck	37±3.5 (31-45)	-
Breast cancer	21±3.8 (17-32)	28±6.5 (15-45)
Prostate cancer	25±4.9 (20-34)	31±7.4 (23-58)
Cervical cancer	25±3.1(23-38)	-

Note: *data presented as Mean±SD*

The cumulative kVCT imaging dose delivered over the treatment course at Center A was 16.20±6.82 cGy for brain cancer, 19.73±5.29 cGy for head and neck cancer, 19.17±3.79 cGy for breast cancer, 37.38±8.53 cGy for prostate cancer, and 38.80±7.72 cGy for cervical cancer. In comparison, at Center C, the corresponding doses were 32.35±8.23 cGy for breast cancer and 53.63±16.02 cGy for prostate cancer, as illustrated in Figure 8. Notably, the mean dose at Center C was 51.16% higher than at Center A for breast cancer, and 35.71% higher for prostate cancer.

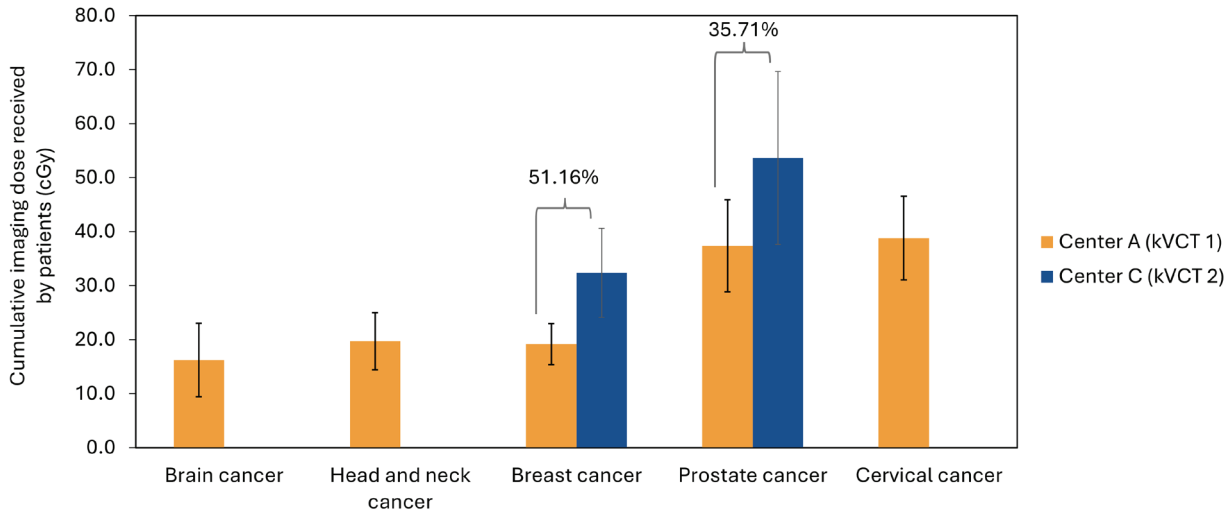


Figure 8. *Cumulative kVCT imaging dose received by patients throughout the treatment course.*

The percentage of cumulative kVCT imaging dose relative to the prescribed dose at Center A was 0.28±0.12 for brain cancer, 0.29±0.08 for head and neck cancer, 0.42±0.10 for breast cancer, 0.62±0.10 for prostate cancer, and 0.80±0.16 for cervical cancer. At Center C, the corresponding values were 0.73±0.19 for breast cancer and 0.87±0.26 for prostate cancer. Nevertheless, the cumulative imaging dose remained below 5% of the prescribed dose for all cancer types at both centers.

In all centers and for all cancer types, the cumulative imaging doses from both MVCT and kVCT did not exceed 5% of the prescribed dose, as summarized in Table 9.

Table 9. The percentage of cumulative MVCT and kVCT imaging dose relative to the prescribed dose.

Imaging type	Center	Cancer type	%Cumulative imaging dose
MVCT	A	Breast	0.36±0.11
		Prostate	0.34±0.04
		Cervix	0.44±0.03
	B	Breast	0.43±0.08
		Prostate	0.36±0.07
kVCT	A	Brain	0.28±0.12
		Head and neck	0.29±0.08
		Breast	0.42±0.10
		Prostate	0.62±0.10
		Cervix	0.80±0.16
	C	Breast	0.73±0.19
		Prostate	0.87±0.26

Note: *data presented as Mean±SD*

Discussion

This study evaluated the radiation dose associated with imaging using the IGRT system under clinically scanned protocols. The dose was quantified as the $MSAD_w$, measured in a CTDI phantom. Additionally, the study assessed whether the cumulative imaging dose received by patients throughout the treatment course with tomotherapy systems exceeded 5% of the prescribed therapeutic dose.

The measured $MSAD_w^{MVCT}$ in the CTDI phantom with a 15 cm scan length was 0.85 cGy at Center A and 0.86 cGy at Center B. These values are consistent with previous findings by Devi R *et al.*[21] (0.82 cGy) and the value reported in the manufacturer's manual[14] (0.84 cGy), both of which were obtained using a Cheese phantom with an 18 cm scan length provided by the manufacturer. The dose differences between the method used in this study and that described in the manufacturer's manual were within 5%. Importantly, all measured values remained below the 1 cGy threshold specified by the manufacturer for MVCT imaging in Coarse mode.[14] The $MSAD_w^{MVCT}$ was evaluated only in Coarse mode, as this scan mode is preferred at both Center A and Center B due to its shorter acquisition time compared with the Fine and Normal modes.

An evaluation of kVCT imaging doses revealed that radiation exposure generally increased with the anatomical region being scanned, reflecting the tube potential (kVp) applied. Additionally, dose levels also rose with patient body size, corresponding to the tube current–time product (mAs). Imaging dose also varied depending on the imaging mode, which is influenced by pitch settings. Lower pitch values yielded higher image quality but were associated with increased radiation doses. Among the modes, Fine provided the highest image quality, followed by Normal and Coarse.[14] When pelvic and thoracic imaging were compared using the same FOV size (440 mm), pelvic protocols consistently resulted in higher radiation doses across all imaging modes and body sizes. This difference was primarily attributable to the higher tube potential used for pelvic imaging (140 kVp) compared to thoracic imaging (120 kVp).

A study by Ehler *et al.*[17] reported kVCT imaging doses measured using an 18 cm scan length without applying a correction factor to convert air kerma to absorbed dose in tissue (f factor), with reported dose ranges of 0.18-0.91 cGy (head), 0.45-1.67 cGy (thorax), and 0.65-2.38 cGy (pelvis), and percentage differences between measured doses and displayed $CTDI_{vol}$ of -3.25% to 12.50%. In comparison, the present study observed comparable dose ranges when doses were calculated without applying the f factor, with reported dose ranges of 0.18-0.79 cGy (head), 0.46-

1.85 cGy (thorax), and 0.70-2.51 cGy (pelvis), and percentage differences between measured doses and displayed $CTDI_{vol}$ of -8.00% to 16.67%. Additional context is provided by the study of Leon *et al.*[16] which evaluated imaging dose in diagnostic CT systems using a 10 cm scan length and a measurement methodology comparable to that employed in the present study. Leon *et al*.[16] reported percentage differences between measured and displayed $CTDI_{vol}$ ranging from -9.3% to 8.2%. It is acknowledged that the use of a 10 cm scan length with a 100-mm pencil ionization chamber may underestimate the absolute kVCT dose due to incomplete integration of scatter. However, despite this potential underestimation, the thorax and pelvis dose ranges were comparable to, and in some cases higher than, those reported by Ehler *et al.*[17] using a longer scan length (18 cm), indicating that scan length alone does not fully explain the observed differences. The remaining discrepancies are likely due to methodological differences and system-specific factors, including variations in hardware configuration, detector design, and system optimization that can influence mAs requirements and effective imaging output.

This study found that Center A used kVCT imaging protocols (Normal and Coarse modes) that delivered lower radiation doses compared to Center C (Fine and Normal modes). For breast cancer, Normal medium (50.72%) and Coarse small (30.72%) modes were most common at Center A, while Center C used a higher proportion of Fine medium (71.55%) and Normal medium (1.42%). A similar trend was seen for prostate cancer imaging, where Center A used mainly Normal medium (72.67%) and Coarse medium (10.28%), while Center C used more Fine small (5.54%), Fine medium (14.02%), and Normal medium (59.89%). As a result, cumulative imaging doses differed by 51.16% for breast cancer and 35.71% for prostate cancer.

The cumulative imaging dose delivered to breast and prostate cancer patients was higher with kVCT (19.17-53.63 cGy) compared to MVCT (17.40-24.02 cGy). This was primarily due to the greater number of imaging sessions and the use of higher-dose imaging protocols in kVCT. Center C showed a 28.57% higher imaging session count for breast cancer and 21.43% for prostate cancer compared to Center A. This increase may be attributed to differences in patient setup practices, with Center C using support devices rather than rigid immobilization systems, leading to more positional variability. When target alignment could not be corrected by couch shifts alone, patient repositioning was required, resulting in repeated IGRT acquisitions.

Patient size can influence imaging dose estimates derived from phantom-based $MSAD_w$, as standardized PMMA phantoms may not fully represent individual patient anatomy. Smaller patients may receive higher absorbed doses due to reduced x-ray attenuation, whereas larger patients may receive lower doses as a result of increased attenuation and scatter. To more accurately evaluate the impact of patient size on dose estimation, size-specific dose estimates (SSDE) can be used, as they account for differences between patient dimensions and standardized phantom sizes and enable more patient-specific dose assessment.

In this study, the cumulative imaging dose throughout the treatment course did not exceed 5% of the prescribed dose for any cancer type at any center, remaining within clinically acceptable limits. Where clinically feasible, the use of lower-dose imaging protocols, such as the Coarse mode instead of Fine or Normal modes, may further reduce cumulative radiation exposure without compromising image quality. This is supported by Marco *et al.*,[12] who reported negligible differences in image quality among Fine, Normal, and Coarse modes, with no clinically relevant impact on IGRT patient positioning. These findings support dose optimization strategies and may inform the development of institutional policies for tomotherapy-based IGRT. Furthermore, AAPM Task Group 142 report recommends performing annual evaluations of imaging doses, as well as reassessments after any replacement of imaging-related equipment, to ensure that patient doses remain within safe and clinically acceptable limits.[22]

Limitations

This study has several limitations. Firstly, dose measurements were performed twice for each imaging condition, which may underestimate random uncertainty; additional repetitions would be required for a more robust characterization of measurement variability and for supporting statistically significant analysis. Secondly, kVCT doses were measured using a fixed 10 cm scan length with a 100 mm pencil ionization chamber, which may underestimate the absolute dose due to incomplete scatter integration. Thirdly, the imaging protocols evaluated were limited to those used at three centers, which may not fully represent the range of clinical practices. Fourthly, dose measurements were based on a phantom-derived $MSAD_w$, which may not reflect patient-specific scan lengths and body sizes in clinical practice, affecting dose estimations. Finally, all measurements were performed at the isocenter, which may not accurately reflect patient positioning in clinical settings and could lead to variations in delivered dose.

Conclusion

The mean $MSAD_w^{MVCT}$ values per scan at Center A and B were 0.85±0.01 cGy and 0.86±0.00 cGy, respectively, with a percentage difference of 1.17%. The mean $MSAD_w^{kVCT}$ values per scan at Center A and C ranged from 0.19-0.83 cGy, 0.49-1.96 cGy, and 0.74-2.65 cGy for the head, thorax, and pelvis protocols, respectively. The imaging doses between the two centers differed by less than 20% in most cases, except for the head protocol

in Coarse mode for medium size (23.73%), the thorax protocol in Normal mode for small size (21.33%), and the pelvis protocol in Normal mode for small size (21.00%). The cumulative imaging doses from MVCT and kVCT over the treatment course accounted for 0.34-0.44% and 0.28-0.87% of the prescribed dose, respectively. In all centers and for all cancer types, the cumulative imaging dose did not exceed 5% of the prescribed dose. However, a substantial variation in cumulative imaging dose was observed between centers, with differences of 51.16% in breast cancer cases and 35.71% in prostate cancer cases. These findings highlight the need to investigate the underlying causes of this variation and to establish optimized imaging protocols in future study.

Ethical approval

This research was approved by the ethics committee of Naresuan University, Thailand (IRB No. P1-0127/2024).

Funding

This work was supported by Naresuan University (NU), and National Science, Research and Innovation Fund (NSRF) [Grant number R2569B107].

Conflict of interest

The authors declare no potential conflicts of interest.

CRediT authorship contribution statement

Wannapha Nobnop, Sumalee Yabsantia, and **Thunyarat Chusin**: supervision; **Pratchayakan Hompeng**: manuscript drafting; **all authors**: study conception, design, material preparation, data collection, and analysis, writing: reviewed and approved the final version.

Acknowledgements

The authors gratefully acknowledge the medical physicists at the Radiation Therapy Departments of Maharaj Nakorn Chiang Mai Hospital, Lopburi Cancer Hospital, and Maha Vajiralongkorn Thanyaburi Hospital for their kind cooperation, technical assistance, and facilitation during data collection.

References

[1] Tegtmeier RC, Ferris WS, Bayouth JE, Miller JR, Culberson WS. Characterization of imaging performance of a novel helical kVCT for use in image-guided and adaptive radiotherapy. J Appl Clin Med Phys. 2022; 23: e13648. doi.org/10.1002/acm2.13648.

[2] Yang B, Geng H, Chang TYA, Tse MY, Lam WW, Huang CY, et al. Clinical implementation of kVCT-guided tomotherapy with ClearRT. Phys Eng Sci Med. 2022; 45: 915-24. doi.org/10.1007/s13246-022-01162-y.

[3] Sakai Y, Monzen H, Takei Y, Kosaka H, Nakamura K, Yanagi Y, et al. Evaluation of in-room volumetric imaging doses for image-guided radiotherapy: A multi-institutional study. J Med Phys. 2023; 48: 189-94. doi.org/10.4103/jmp.jmp_109_22.

[4] Tomita T, Isobe T, Furuyama Y, Takei H, Kobayashi D, Mori Y, et al. Evaluation of dose distribution and normal tissue complication probability of a combined dose of cone-beam computed tomography imaging with treatment in prostate intensity-modulated radiation therapy. J Med Phys. 2020; 45: 78-87. doi.org/10.4103/jmp.JMP_4_20.

[5] Brenner DJ. Induced second cancers after prostate-cancer radiotherapy: No cause for concern? Int J Radiat Oncol Biol Phys. 2006; 65: 637-9. doi.org/10.1016/j.ijrobp.2006.02.044.

[6] Ding GX, Alaei P, Curran B, Flynn R, Gossman M, Mackie TR, et al. Image guidance doses delivered during radiotherapy: Quantification, management, and reduction: Report of the AAPM Therapy Physics Committee Task Group 180. Med Phys. 2018; 45: e84-9. doi.org/10.1002/mp.12824.

[7] Dische S, Saunders MI, Williams C, Hopkins A, Aird E. Precision in reporting the dose given in a course of radiotherapy. Rad Onc.1993; 29: 287-93. doi.org/10.1016/0167-8140(93)90146-Y.

[8] Saunders MI, Dische S, Grosch EJ, Fermont DC, Ashford RFU, Maher EJ, et al. Experience with CHART. Int J Radiat Oncol Biol Phys. 1991; 21: 871-8. doi.org/10.1016/0360-3016(91)90722-G.

[9] Morgan-Fletcher SL. Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU Report 50), ICRU Report 62. ICRU, pp. ix+52, 1999 (ICRU Bethesda, MD) $65.00 ISBN 0-913394-61-0. Br J Radiol. 2001; 74: 294. doi.org/10.1259/bjr.74.879.740294.

[10] Langen KM, Papanikolaou N, Balog J, Crilly R, Followill D, Goddu SM, et al. QA for helical tomotherapy: Report of the AAPM Task Group 148. Med Phys. 2010; 37: 4817-53. doi.org/10.1118/1.3462971.

[11] Mege JP, Wenzhao S, Veres A, Auzac G, Diallo I, Lefkopoulos D. Evaluation of MVCT imaging dose levels during helical IGRT: Comparison between ion chamber, TLD, and EBT3 films. J Appl Clin Med Phys. 2016; 17: 143-57. doi.org/10.1120/jacmp.v17i1.5774.

[12] De Marco P, Abdi Osman I, Castellini F, Ricotti R, Leonardi MC, Miglietta E, et al. Image quality and dose evaluation of MVCT TomoTherapy acquisitions: A phantom study. Physica Medica. 2019; 57: 200-6. doi.org/10.1016/j.ejmp.2019.01.009.

[13] Trivedi G, Dixit C, Oinam A, Kapoor R, Bahl A. Kilovoltage cone-beam computed tomography imaging dose estimation and optimization: Need of daily cone-beam computed tomography. J Cancer Res Ther. 2019; 15: 470-4. doi.org/10.4103/jcrt.JCRT_949_17.

[14] Radixact physics essentials guide., Version 3.0.x ed.

Published online: Accuray Incorporated; 2023.

[15] Murphy MJ, Balter J, Balter S, Bencomo JA, Das IJ, Jiang SB, et al. The management of imaging dose during image-guided radiotherapy: Report of the AAPM Task Group 75. Med Phys. 2007; 34:4041-63. doi.org/10.1118/1.2775667.

[16] Leon SM, Kobistek RJ, Olguin EA, Zhang Z, Barreto IL, Schwarz BC. The helically-acquired $CTDI_{vol}$ as an alternative to traditional methodology. J Appl Clin Med Phys .2020; 21: 263-71. doi.org/10.1002/acm2.12944.

[17] Ehler ED, Alaei P. Imaging dose and image quality of kilovoltage imaging implemented on a helical tomotherapy unit. Z Med Phys. 2025.; in press. doi.org/10.1016/j.zemedi.2024.12.003.

[18] McCollough CH, Cody D, Edyvean S, Geise R, Gould B, Keat N, et al. The measurement, reporting, and management of radiation dose in CT. College Park, MD, USA: AAPM. 2008; Report No. 96: pp 1-68. doi.org/10.37206/97

[19] Department of Medical Sciences, Ministry of Public Health. National diagnostic reference levels in Thailand 2023. Nonthaburi: Department of Medical Sciences, Ministry of Public Health; 2023: pp 42, ISBN 978-616-11-5048-8.

[20] Vañó E, Miller DL, Martin CJ, Rehani MM, Kang K, Rosenstein M, et al. ICRP Publication 135: Diagnostic reference levels in medical imaging. Ann ICRP. 2017; 46(1): 1-144. doi:10.1177/0146645317717209.

[21] Devi R, Singh B, Basandrai D, Passi K, Singh V. Estimation of weighted computed tomography dose index ($CTDI_w$) in megavoltage computed tomography (MVCT). Egypt J Radiol Nucl Med. 2024; 55:205:1- 8. doi:10.1186/s43055-024-01374-0.

[22] Klein EE, Hanley J, Bayouth J, Yin FF, Simon W, Dresser S, et al. Task group 142 report: Quality assurance of medical acceleratorsa. Med Phys. 2009; 36: 4197-212. doi.org/10.1118/1.3190392.

Journal of Associated Medical Sciences 2026; 59(2): 178-186

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Serum Klotho as a feasible diagnostic biomarker for metabolic syndrome in Iraqi adults aged over 50 years

Haitham Ibrahim Azrak* **and Entedhar Rifaat Sarhat**

Department of Clinical Biochemistry, College of Medicine, University of Tikrit, Iraq.

ARTICLE INFO

Article history:
Received 25 October 2025
Accepted as revised 28 January 2026
Available online 12 February 2026

Keywords:
Metabolic syndrome, Klotho Beta, insulin

ABSTRACT

Background: Metabolic Syndrome (MetS) is a cluster of cardiometabolic risk factors primarily driven by insulin resistance, with a high prevalence in older adult populations.

Objectives: This study aimed to evaluate the diagnostic accuracy of serum Klotho Beta (KLB) and insulin as novel biomarkers for MetS in Iraqi adults over 50 years old.

Materials and methods: This case-control study included 60 MetS patients and 30 healthy controls, recruited from Iraqi hospitals. MetS was diagnosed using harmonized IDF/AHA/NCEP ATP III criteria. Serum Klotho Beta and insulin levels were measured by ELISA.

Results: Serum Klotho Beta levels were significantly lower in the MetS group (25.54±4.41 nmol/L) compared to controls (59.00±11.05 nmol/L, $p<0.001$). Klotho Beta was inversely correlated with HOMA-IR (r=-0.59, $p<0.001$), BMI, HbA1c, and triglycerides, and positively with HDL-C. ROC analysis showed an AUC of 0.997 for Klotho Beta (sensitivity 95%, specificity 100% at ≤35.248 nmol/L) and an AUC of 0.998 for insulin. Logistic regression confirmed Klotho Beta as a strong independent predictor of MetS (OR=1.85 per 1 nmol/L decrease).

Conclusion: Serum Klotho Beta and insulin are highly sensitive and specific biomarkers for MetS in older adults. Klotho Beta deficiency, coupled with hyperinsulinemia, reflects underlying insulin resistance and metabolic dysregulation, highlighting its potential for risk stratification and early diagnosis in high-risk populations.

Introduction

Metabolic syndrome (MetS) is a cluster of interrelated cardiovascular risk factors, including abdominal obesity, dyslipidemia, impaired glucose tolerance, and hypertension, and is driven largely by insulin resistance.[1,2] The global prevalence of MetS has surged owing to rising obesity rates and sedentary lifestyles, posing a significant public health challenge.[3] Diagnostically, MetS is defined by the harmonized criteria established by the International Diabetes Federation (IDF), American Heart Association (AHA), and other major organizations, requiring the presence of at least three of the following: elevated waist circumference (≥90 cm for men and ≥85 cm for women), fasting glucose ≥100 mg/dL, triglycerides ≥150 mg/dL, reduced HDL cholesterol (<40 mg/dL for men and <50 mg/dL for women), or blood pressure ≥130/85 mmHg.[4]

* **Corresponding contributor.**
Author's Address: *Department of Clinical Biochemistry, College of Medicine, University of Tikrit, Iraq.*
E-mail address: Hi240043me@st.tu.edu.iq
doi: 10.12982/JAMS.2025.049
E-ISSN: 2539-6056

While earlier definitions (e.g., WHO) emphasized insulin resistance as a prerequisite, modern frameworks, such as the NCEP ATP III criteria, prioritize clinical applicability, reflecting the multifactorial nature.[5]

Geographic and demographic disparities in the prevalence of MetS are conspicuous. Urbanized regions and ethnic groups such as Mexican Americans exhibit considerably higher rates, with prevalence in Middle Eastern and Gulf nations frequently exceeding 35-40% among adults.[6,7] Latin American countries, including Mexico, Chile, and Brazil, report prevalence estimates between 30% and 35% due to shifting dietary patterns towards ultra-processed foods and reduced physical activity.[8] The United States has a prevalence of approximately 34%.[9,10] Rates in South and Southeast Asia, including India, Malaysia, and Thailand, remain low (20-25%) but are rapidly escalating.[11] Europe has a prevalence range of 20-30%, with higher rates in Eastern Europe than in Western regions, reflecting variances in healthcare access and lifestyle. Historically, low-prevalence settings such as sub-Saharan Africa are witnessing increases linked to urbanization and the adoption of Westernized dietary habits.[9]

Given the substantial morbidity associated with MetS, early identification and intervention are paramount. Traditional diagnostic parameters, while valuable, may lack the sensitivity to detect subclinical diseases, particularly in at-risk older adults. As a result, the search for reliable biomarkers that reflect the underlying pathophysiology and predict disease progression has intensified.[12,13] Klotho Beta has emerged as a promising candidate in this context. KLB is a type I transmembrane protein that is structurally related to the anti-aging factor α-Klotho and functions as a crucial co-receptor for the endocrine fibroblast growth factors FGF19 and FGF21.[14] It is highly expressed in metabolic organs such as the liver, pancreas, and adipose tissue and is integral to the regulation of lipid and carbohydrate metabolism.[15] Insulin, a peptide hormone produced by the pancreas, is pivotal for glucose uptake and metabolic regulation.[16,17] Dysfunction, manifested by insulin resistance, is a core driver of MetS and its related complications.[18] Despite the recognition of MetS burden, especially among elderly individuals, there remains a gap in validated, sensitive biomarkers capable of stratifying risk and reflecting disease mechanisms.[19]

This study specifically investigated the predictive utility of serum β-Klotho and insulin as direct indices of metabolic regulation and insulin resistance in Iraqi adults aged ≥ 50 years with MetS.

Materials and methods

Study design and participants

This case-control study was conducted from January 1st to April 30th, 2025. Participants were recruited from the outpatient clinics of several public hospitals in the Diyala Governorate, Iraq. The study comprised 90 adults aged over 50 years: 60 patients diagnosed with MetS and 30 age- and sex-matched healthy controls.

Inclusion and exclusion criteria

MetS was diagnosed according to the harmonized criteria, which mandates at least three of five components: 1) elevated waist circumference (≥90 cm in men, ≥85 cm in women), 2) elevated fasting blood glucose (≥100 mg/dL) or pharmacologic treatment for hyperglycemia, 3) hypertriglyceridemia (≥150 mg/dL) or lipid-lowering drug use, 4) low HDL-C (<40 mg/dL in men, <50 mg/dL in women), or 5) elevated blood pressure (systolic ≥130 mmHg, diastolic ≥85 mmHg) or antihypertensive medication use.

Participants in the control group were confirmed to be free of MetS based on the same criteria. Exclusion criteria for both groups included a history of endocrine disorders (e.g., Cushing's syndrome, thyroid dysfunction), chronic organ failure (eGFR <60 mL/min/1.73 m^2; severe hepatic or pulmonary disease), cancer, rheumatoid arthritis, cerebrovascular disease, or current use of bone-modifying drugs (e.g., glucocorticoids, bisphosphonates, thiazolidinediones, SGLT-2 inhibitors, vitamin D/calcium supplements, estrogen).

Sample size justification

The sample size was determined as a convenience sample based on patient availability and logistical constraints during the study period, rather than by a formal a priori power calculation.

Biological sample collection

Fasting venous blood samples (5 mL) were collected from all participants in the morning (between 8:00 and 10:00 AM) after an overnight fast of 10-12 hours. Blood was drawn into two vacuum tubes: 1 mL in a K2EDTA tube and 4 mL in a serum separator gel tube. The serum tubes were allowed to clot for 30 minutes at room temperature before centrifugation at 3,000 rpm for 10 minutes. The resulting serum was aliquoted into cryovials and stored immediately at -20 °C until analysis. Hemolyzed samples were discarded. All samples were analyzed within one month of collection, undergoing a single freeze-thaw cycle.

Laboratory analysis

The MetS parameters, including fasting glucose and lipid profiles (total cholesterol, triglycerides, LDL, VLDL, and HDL), were assessed using an automated system. All measurements were performed as a single determination for each participant. Other biomarkers include liver enzymes (ALT) and (AST), uric acid, and glycated hemoglobin (HbA1c) using Cobas C111-Roche-Germany. Human Insulin was measured using a kit (Cat. No. E0010Hu, BT Laboratory, China) with a minimum detection limit of 0.11 mIU/L. The reported intra-assay CV was <8% and the inter-assay CV was <10%. Human Klotho Beta was measured using a

kit (Cat. No. E2782Hu, BT Laboratory, China) with a minimum detection limit of 0.29 nmol/L. The reported intra-assay CV was <6% and the inter-assay CV was <9%.

Statistical analysis

The data analysis was performed utilizing IBM SPSS Statistics software, Version 26 (IBM Corp., USA). Descriptive statistics for continuous measures are reported as means with their standard deviations. To compare these measures between the MetS and control cohorts, independent samples t-tests were employed. We evaluated the biomarker performance of serum klotho Beta, insulin, and HOMA-IR by constructing receiver operating characteristic (ROC) curves; the discriminatory power of each biomarker was quantified by the area under the curve (AUC). A p-value below 0.05 was established as the threshold for statistical significance.

Results

The demographic and clinical characteristics of the study participants are summarized in Table 1. The groups were well-matched for age and sex. As expected, the MetS group exhibited significant alterations in classic metabolic parameters, including higher BMI, waist circumference, fasting glucose, HbA1c, insulin, HOMA-IR, and triglycerides, alongside lower HDL-cholesterol (all $p<0.001$). Crucially, the novel biomarker serum Klotho Beta was profoundly lower in the MetS group (25.54±4.41 nmol/L) compared to healthy controls (59.00±11.05 nmol/L; $p<0.001$).

Table 1. Demographic, clinical, and biochemical characteristics of the study participants.

Variable		Control group (N=30)	MetS Group (N=60)	*p* value
Demographics				
Age (years)		58.33±4.84	60.28±7.62	0.20
Gender (N, %)	Male	12 (40%)	21 (35 %)	0.34
	Female	18 (60%)	39 (65 %)	
Anthropometrics				
BMI (kg/m^2)		24.67±2.55	29.18±2.81	<0.001
Waist circumference (cm)		84.6±2.55	98.8±6.7	<0.001
Blood pressure (mmHg)				
Systolic		120.5±8.2	135.2±10.5	<0.001
Diastolic		78.9±5.6	88.1±7.3	<0.001
Glucose metabolism				
Fasting blood sugar (mg/dL)		80.90±11.03	124.23±44.11	<0.001
HbA1c (%)		5.22±0.52	7.68±2.15	<0.001
Insulin (mIU/L)		3.31±0.56	7.07±1.21	<0.001
HOMA-IR		0.63±0.14	2.16±0.82	<0.001
Lipid profile (mg/dL)				
Total cholesterol		172.07±32.37	171.15±41.11	0.92
Triglycerides		87.17±30.10	164.85±71.07	<0.001
HDL-cholesterol		51.13±7.84	40.37±9.72	<0.001
LDL-cholesterol		103.63±27.77	99.15±37.94	0.57
VLDL-cholesterol		17.37±5.90	33.00±14.54	<0.001
Other biomarkers				
Serum uric acid (mg/dL)		4.14±1.46	4.39±1.48	0.46
ALT (U/L)		20.43±5.68	20.60±6.15	0.90
AST (U/L)		21.53±6.19	19.37±6.74	0.14
Proposed biomarker				
Klotho Beta (nmol/L)		59.00±11.05	25.54±4.41	<0.001

Note: Data are presented as Mean±SD or counts, *p* values from independent samples t-tests (or Chi-square for gender). Statistically significant values ($p<0.05$).

Table 2 presents the association between serum Klotho Beta levels and the presence of individual components of metabolic syndrome, as defined by IDF criteria. The results demonstrate that Klotho Beta concentrations were significantly lower in participants with central obesity (24.8±5.1 nmol/L vs 31.5±12.1 nmol/L, $p<0.001$), hypertriglyceridemia (23.9±3.0 nmol/L vs 27.5±5.1 nmol/L, $p=0.0012$), low HDL-C (25.1±5.3 nmol/L vs 29.8±11.5 nmol/L, $p=0.002$), and impaired fasting glucose (24.6±3.5 nmol/L vs 27.5±5.2 nmol/L, $p=0.014$) compared to those without these components.

Table 2. Klotho Beta levels according to the presence of metabolic syndrome components (IDF criteria).

Metabolic syndrome component	Absent	Present	*p* value
Central obesity (waist circumference)	31.5±12.1 (N=12, 20%)	24.8±5.1 (N=48, 80%)	<0.001
Hypertriglyceridemia (≥150 mg/dL)	27.5±5.1 (N=25, 41.7%)	23.9±3.0 (N=35, 58.3%)	0.0012
Low HDL-C (<50 mg/dL F, <40 mg/dL M)	29.8±11.5 (N=13, 21.7%)	25.1±5.3 (N=47, 78.3%)	0.002
Impaired fasting glucose (≥100 mg/dL)	27.5±5.2 (N=20, 33%)	24.6±3.5 (N=40, 67%)	0.014

Note: Klotho Beta (nmol/L) is presented as Mean±SD.

Analysis revealed that serum Klotho β levels were significantly inversely correlated with key indicators of metabolic dysregulation, including HOMA-IR (r=-0.59, $p<0.001$) as shown in Figure 1, triglycerides (r=-0.44, $p<0.001$), HbA1c (r=-0.41, $p<0.001$), BMI (r=-0.32, $p<0.01$), and WC (r=-0.28, $p=0.03$). In contrast, Klotho Beta showed a favorable positive correlation with HDL-C (r=0.38, $p<0.01$). We have also specified that Pearson correlation was used for this analysis.

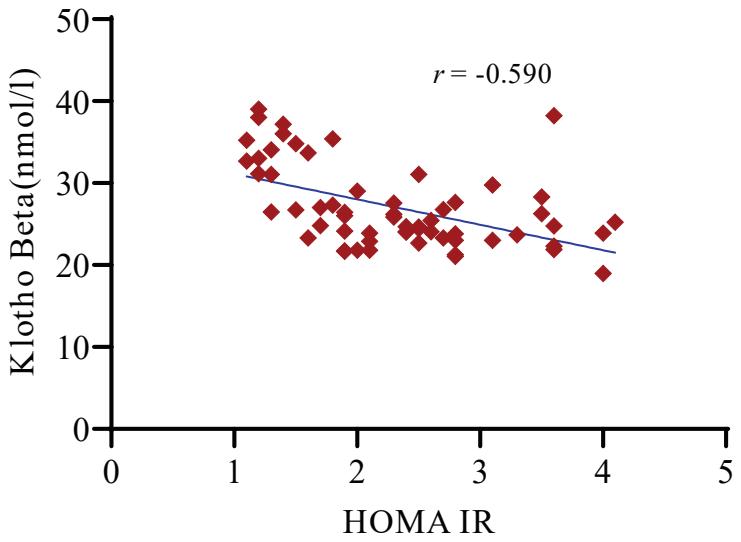


Figure 1. *Scatter plot illustrating the inverse correlation between serum Klotho Beta levels and HOMA-IR.*

The diagnostic performance of serum Klotho Beta and insulin for distinguishing individuals with MetS from healthy controls was evaluated using ROC curve analysis. Klotho Beta exhibited exceptional discriminatory ability, with an AUC of 0.997 (95% CI: 0.953-1.000), achieving a sensitivity of 95.00% and a specificity of 100% at an optimal cutoff value of ≤35.248 nmol/L as show in Figure 2. Similarly, insulin demonstrated near-perfect discrimination, with an AUC of 0.998 (95% CI: 0.955-1.000), a sensitivity of 96.67%, and a specificity of 100% at a cutoff value of >4.755 mIU/L.

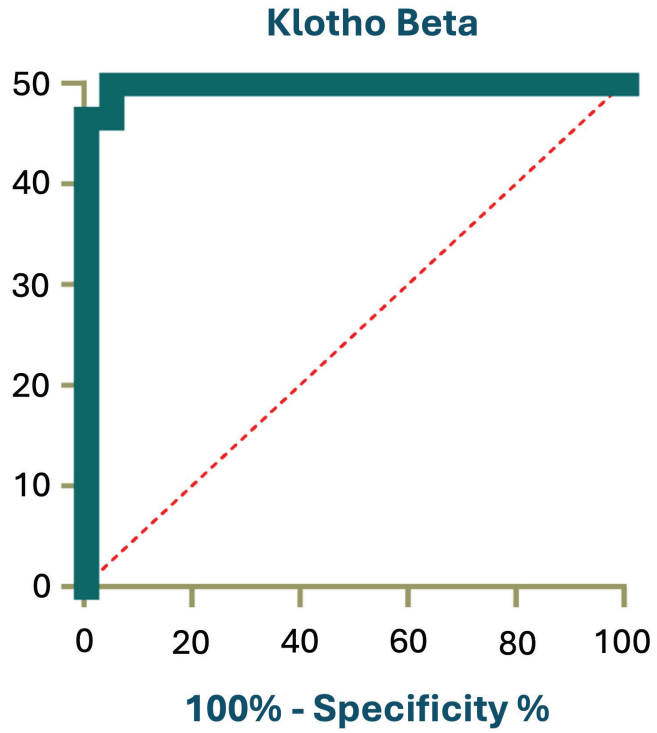


Figure 2. *Receiver operating characteristic (ROC) curve for Klotho Beta in classifying patients from controls.*

A binary logistic regression model (Table 3) was constructed to distinguish patients with MetS from controls based on key variables: Klotho Beta, BMI, HbA1c, gender, and age. The analysis identified a significant inverse relationship between serum Klotho Beta and MetS, where each 1 nmol/L decrement corresponded to an 85% increase in the odds of having MetS (OR=1.85, 95% CI: 1.52-2.25, $p<0.001$). Positive associations were also observed for BMI (OR=1.62, 95% CI: 1.28-2.05, $p<0.001$) and HbA1c (OR=8.45, 95% CI: 3.92-18.20, $p<0.001$). The model further indicated that being male was a protective factor (OR=0.24, 95% CI: 0.07-0.79, $p=0.019$), but age did not contribute significantly to predicting MetS status ($p=0.610$). Both the Hosmer-Lemeshow test ($p=0.451$) and a Nagelkerke R^2 value of 0.89 confirmed the model robust fit and explanatory power.

Table 3. Binary logistic regression for predicting patient status (patients vs controls).

Predictor	**Odds Ratio** (OR)	**95% CI for OR**	***p* value**
Klotho Beta (per 1 nmol/L decrease)	1.85	1.52-2.25	<0.001
BMI (per 1 kg/m^2 increase)	1.62	1.28-2.05	<0.001
HbA1c (per 1% increase)	8.45	3.92-18.20	<0.001
Gender (male vs female)	0.24	0.07-0.79	0.019
Age (per 1 year increase)	1.02	0.94-1.11	0.610

Note: Hosmer-Lemeshow test at $p=0.451$ indicates a good model fit). Nagelkerke $R^2=0.89$.

Discussion

The present study provides strong evidence supporting the utility of serum klotho Beta and insulin as highly sensitive and specific biomarkers for diagnosing MetS in Iraqi adults over 50 years of age. A key finding was the exceptional diagnostic performance of both biomarkers: Klotho Beta achieved an AUC of 0.997 with 95% sensitivity and 100% specificity, while insulin demonstrated near-perfect discrimination with an AUC of 0.998, 96.67% sensitivity, and 100% specificity.

Patients with MetS exhibited significantly lower serum Klotho Beta concentrations than healthy controls, a finding consistent with previous research suggesting its role as a protective factor in metabolic homeostasis.[20,21] The observed reduction may indicate compromised anti-aging and metabolic regulatory pathways, as Klotho deficiency has been implicated in impaired insulin signaling, diminished antioxidant capacity, and increased inflammation all pivotal mechanisms in metabolic dysfunction.[20,21] This substantial difference in Klotho Beta levels between groups supports its clinical relevance for risk stratification in aging populations.[20,22] The pronounced reduction suggests accelerated metabolic aging and a loss of protective mechanisms against oxidative damage and metabolic stress. Klotho Beta deficiency may contribute to insulin resistance by impairing insulin signaling pathways, reducing antioxidant capacity, and enhancing inflammatory responses.[23,24] This finding suggests that strategies aimed at restoring Klotho Beta function could represent a novel therapeutic target for addressing the underlying pathophysiology of metabolic dysfunction, aligning with the work of Wang *et al.*[25] and others who propose that reduced KLB function disrupts FGF21-mediated metabolic regulation.[26]

Parallel to the alterations in the FGF21-KLB pathway, our study confirmed the central role of insulin resistance in MetS, as evidenced by significantly elevated insulin levels and HOMA-IR values in the patient group. Hyperinsulinemia is initially a compensatory response to maintain euglycemia in the face of developing insulin resistance. However, chronic hyperinsulinemia itself contributes to the worsening of insulin resistance by promoting the downregulation of insulin receptors and impairing post-receptor signaling pathways.[27,28] This creates a vicious cycle that perpetuates metabolic dysfunction. The strong correlation between insulin and HOMA-IR in our study reinforces their utility as reliable measures of insulin resistance.[29,30] The elevated HOMA-IR values observed in our patients are consistent with established thresholds for increased risk of type 2 diabetes and cardiovascular complications, confirming a significant impairment in insulin-mediated glucose uptake and metabolism.[31]

This compensatory hyperinsulinemia has broader metabolic consequences, including enhanced lipogenesis, reduced lipolysis, and promotion of weight gain, creating a vicious cycle that perpetuates and worsens insulin resistance over time.[32] The HOMA-IR values in the patient group approached the threshold typically used to define insulin resistance (≥2.5), confirming a significant impairment in insulin-mediated glucose uptake and metabolism in peripheral tissues.[33] The higher variability in patients suggests varying degrees of insulin resistance severity, which is associated with an increased risk of T2DM progression, cardiovascular disease, and MetS.[34,35]

These biomarker alterations occur within the broader context of insulin resistance, the core pathophysiological mechanism of MetS, as reflected by the significant differences in HbA1c and fasting blood glucose between groups.[36,37] Focusing on an older cohort is clinically relevant, as MetS in this population often manifests with more severe complications and higher cardiovascular risk. The high diagnostic accuracy of these biomarkers in individuals over 50 points to their potential utility in targeted screening initiatives for this high-risk demographic, where early diagnosis is essential for implementing interventions to prevent adverse outcomes.[38,39]

The high diagnostic accuracy of both Klotho Beta and insulin in our study underscores their clinical relevance for risk stratification and early-stage screening of MetS in older adults.[12] Traditional diagnostic criteria for MetS identify individuals who already have a cluster of metabolic abnormalities. In contrast, Klotho Beta and insulin are biomarkers that reflect the underlying pathophysiology of the disease. A low Klotho Beta level may indicate a state of FGF21 resistance and a predisposition to metabolic dysfunction, while elevated insulin is a direct measure of insulin resistance. Therefore, these markers could potentially identify individuals at high risk for developing MetS even before they meet the full diagnostic criteria.[12] This would allow for earlier intervention with lifestyle modifications or pharmacological therapies to prevent or delay the onset of overt disease and its associated cardiovascular complications. Furthermore, monitoring these biomarkers could provide a more dynamic assessment of an individual's metabolic health and their response to treatment.

Limitations

The main strengths of this study are its exacting statistical methodology and the remarkable diagnostic results produced by both biomarkers. Strong proof of its clinical value is provided by ROC analysis combined with reliable AUC computations. The focus on persons over 50 targets a high-risk demographic where early diagnosis is especially beneficial, and the age-matched methodology reduces confounding by age-related metabolic changes. These large effect sizes are consistent with findings from prior research, which lends further support to the robustness of our results despite the modest sample size. Nevertheless, future multicenter studies with larger, more diverse populations and formal power calculations are warranted

to confirm these findings and enhance generalizability. Additional limitations include the study's single-center, cross-sectional design, which precludes causal inferences and may limit the external validity of our conclusions to other populations.

Conclusion

This study demonstrates that serum Beta-Klotho and insulin serve as highly sensitive and specific biomarkers for metabolic syndrome in older Iraqi adults. The exceptional diagnostic accuracy achieved supports the clinical utility of these markers for early diagnosis and risk stratification in aging populations. These findings contribute to the growing evidence base supporting Klotho-related pathways as key regulators of metabolic health and potential therapeutic targets for age-related metabolic dysfunction.

Ethical approval

An agreement to conduct this study was signed before sample collection began. The Tikrit University College of Medicine collaborated in this study. This study was approved by the research committee (No. 25/11/2024) in accordance with the Declaration of Helsinki. All participants provided written informed consent to participate in the study.

Funding

This research received no specific funding.

Availability of data and materials

The datasets generated and analyzed during the current study are not publicly available because of participant privacy and confidentiality concerns but are available from the corresponding author upon reasonable request.

Conflict of interests

The authors declare no competing interests, financial or otherwise.

CRedit authorship contribution statement

Haitham Ibrahim Azrak: conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, visualization, project administration, funding acquisition, writing: original draft, review and edit; **Entedhar Rifaat Sarhat**: supervision, project administration, conceptualization, validation, writing: review and editing.

Acknowledgements

The authors sincerely thank the medical staff and management of the participating hospitals in the Diyala Governorate for their invaluable assistance in facilitating patient enrollment and sample collection. We are especially grateful to the Tikrit University College of Medicine for its resources and institutional support. We would also like to express our gratitude to everyone who volunteered for this study.

References

[1] Chew NWS, Ng CH, Tan DJH, et al. The global burden of metabolic disease: Data from 2000 to 2019. Cell Metab. 2023; 35(3): 414-28.e3. doi:10.1016/j.cmet.2023.02.003.

[2] Neeland IJ, Lim S, Tchernof A, et al. Metabolic syndrome. Nat Rev Dis Prim. 2024; 10(1): 60-77. doi:10.1038/s41572-024-00563-5.

[3] Nwankwo M, Okamkpa CJ, Danborno B. Comparison of diagnostic criteria and prevalence of metabolic syndrome using WHO, NCEP-ATP III, IDF and harmonized criteria: A case study from urban southeast Nigeria. Diabetes Metab Syndr Clin Res Rev. 2022; 16(12): 1-34. doi: 10.1016/j.dsx.2022.102665.

[4] Fahed G, Aoun L, Bou Zerdan M, et al. Metabolic syndrome: updates on pathophysiology and management in 2021. Int J Mol Sci. 2022; 23(2): 786. doi: 10.3390/ijms23020786.

[5] Liu L, Cai XC, Sun XY, et al. Global prevalence of metabolic syndrome in patients with psoriasis in the past two decades: current evidence. J Eur Acad Dermatol Venereol. 2022; 36(11): 1969-79. doi:10.1111/jdv.18296.

[6] Jamali Z, Ayoobi F, Jalali Z, et al. Metabolic syndrome: a population-based study of prevalence and risk factors. Sci Rep. 2024; 14(1): 3987. doi:10.1038/s41598-024-54367-4.

[7] Al-Rubeaan K, Bawazeer N, Al Farsi Y, et al. Prevalence of metabolic syndrome in Saudi Arabia-a cross sectional study. BMC Endocr Disord. 2018; 18(1): 1-9. doi: 10.1186/s12902-018-0244-4.

[8] Liang X, Or B, Tsoi MF, Cheung CL, Cheung BMY. Prevalence of metabolic syndrome in the United States National Health and Nutrition Examination Survey 2011-18. Postgrad Med J. 2023; 99(1175): 985-92. doi:10.1093/postmj/qgad008.

[9] Pigeot I, Ahrens W. Epidemiology of metabolic syndrome. Pflugers Arch. 2025; 477(5): 669-80. doi:10.1007/s00424-024-03051-7.

[10] Bowo-Ngandji A, Kenmoe S, Ebogo-Belobo JT, et al. Prevalence of the metabolic syndrome in African populations: A systematic review and meta-analysis. PLoS One. 2023; 18(7): e0289155. doi:10.1371/journal.pone.0289155.

[11] Noubiap JJ, Nansseu JR, Lontchi-Yimagou E, et al. Geographic distribution of metabolic syndrome and its components in the general adult population: A meta-analysis of global data from 28 million individuals. Diabetes Res Clin Pract. 2022; 188: 109924. doi: 10.1016/j.diabres.2022.109924.

[12] Srikanthan K, Feyh A, Visweshwar H, Shapiro JI, Sodhi K. Systematic review of metabolic syndrome biomarkers: A panel for early detection, management, and risk stratification in the West Virginian population. Int J Med Sci. 2016; 13(1): 25-38. doi:10.7150/ijms.13800.

[13] Warner JD, Blake GM, Garrett JW, et al. Correlation

of HbA1c levels with CT-based body composition biomarkers in diabetes mellitus and metabolic syndrome. Sci Rep. 2024; 14(1): 21875. doi:10.1038/s41598-024-72702-7.

[14] Geng L, Yi X, Lin Y, et al. Site-specific analysis and functional characterization of N-linked glycosylation for β-Klotho protein. Int J Biol Macromol. 2025; 289: 138846. doi: 10.1016/j.ijbiomac.2024.138846.

[15] Yang M, Liu C, Jiang N, et al. Fibroblast growth factor 21 in metabolic syndrome. Front Endocrinol (Lausanne). 2023; 14: 1220426. doi: 10.3389/fendo.2023.1220426.

[16] Wu SE, Chen WL. Soluble klotho as an effective biomarker to characterize inflammatory states. Ann Med. 2022; 54(1): 1520-9. doi:10.1080/07853890.2022.2077428.

[17] Zhou J, Ben S, Xu T, Xu L, Yao X. Serum β-klotho is a potential biomarker in the prediction of clinical outcomes among patients with NSCLC. J Thorac Dis. 2021; 13(5): 3137-50. doi:10.21037/jtd-21-798.

[18] ElSayed NA, Aleppo G, Aroda VR, et al. Introduction and methodology: standards of care in diabetes—2023. Diabetes Care. 2023; 46(Suppl_1): S1-S4. doi: 10.2337/dc23-Sint.

[19] Li S, Wang P, Wang J, Zhao J, Wang X, Liu T. Effect of mind-body exercise on risk factors for metabolic syndrome including insulin resistance: a meta-analysis. Front Endocrinol (Lausanne). 2024; 15: 1289254. doi: 10.3389/fendo.2024.1289254.

[20] Yuguang L, Chang Y, Chen N, et al. Serum klotho as a novel biomarker for metabolic syndrome: findings from a large national cohort. Front Endocrinol (Lausanne). 2024; 15: 1-8. https://www.frontiersin.org/journals/endocrinology/articles/10.3389/fendo.2024.1295927.

[21] Kuro-o M. The Klotho proteins in health and disease. Nat Rev Nephrol. 2019; 15(1): 27-44. doi: 10.1038/s41581-018-0078-3.

[22] Liang Y, Liu Y, Tan Q, Zhou K, Wu Y, Yu L. Systemic immune-inflammation mediates the association between Klotho protein and metabolic syndrome: findings from a large-scale population-based study. Lipids Health Dis. 2024; 23(1): 360. doi:10.1186/s12944-024-02339-y.

[23] Lin Y, Chen J, Sun Z. Antiaging gene Klotho deficiency promoted high-fat diet–induced arterial stiffening via inactivation of AMP-activated protein kinase. Hypertension. 2016; 67(3): 564-73. doi: 10.1161/HYPERTENSIONAHA.115.06825.

[24] Hu MC, Shi M, Zhang J, et al. Klotho deficiency causes vascular calcification in chronic kidney disease. J Am Soc Nephrol. 2011; 22(1): 124-36. doi: 10.1681/ASN.2009121311%0A.

[25] Jiang S, Wang Y, Wang Z, Zhang L, Jin F, Li B. The association of serum Klotho concentrations with hyperlipidemia prevalence and lipid levels among US adults: a cross-sectional study. BMC Public Health. 2023; 23(1): 1-12. doi: 10.1186/s12889-023-16566-y.

[26] Keuper M, Häring HU, Staiger H. Circulating FGF21 levels in human health and metabolic disease. Exp Clin Endocrinol Diabetes. 2020; 128(11): 752-70. doi: 10.1055/a-0879-2968.

[27] Bagchi DP, Li Z, Corsa CA, et al. Wntless regulates lipogenic gene expression in adipocytes and protects against diet-induced metabolic dysfunction. Mol Metab. 2020; 39: 100992. doi: 10.1016/j.molmet.2020.100992.

[28] Wyble K. Implementing the American Diabetes Association Glycemic Assessment Clinical Practice Guidelines in the care of patients diagnosed with type 2 diabetes in a single rural healthcare clinic in Southwest Louisiana. University of Louisiana at Lafayette; 2023.

[29] Tsai SF, Yang CT, Liu WJ, Lee CL. Development and validation of an insulin resistance model for a population without diabetes mellitus and its clinical implication: a prospective cohort study. EClinicalMedicine. 2023; 58: 101934. doi:10. 1016/j.eclinm.2023.101934.

[30] Kawada T. Insulin-related biomarkers to predict the risk of metabolic syndrome. Int J Endocrinol Metab. 2013; 11(4): e10418. doi:10.5812/ijem.10418.

[31] Gao K, Lin Y. Insulin resistance mediates the association between physical activity and mortality in US adults with metabolic syndrome. Sci Rep. 2025; 15(1): 18872. doi:10.1038/s41598-025-02921-z

[32] Nolan CJ, Prentki M. Insulin resistance and insulin hypersecretion in the metabolic syndrome and type 2 diabetes: Time for a conceptual framework shift. Diabetes Vasc Dis Res. 2019; 16(2): 118-27. doi: 10.1177/1479164119827611.

[33] Kim S, Song K, Lee M, et al. Trends in HOMA-IR values among South Korean adolescents from 2007–2010 to 2019–2020: a sex-, age-, and weight status-specific analysis. Int J Obes. 2023; 47(9): 865-72. doi: 10.1177/1479164119827611.

[34] Petersen MC, Shulman GI. Mechanisms of insulin action and insulin resistance. Physiol Rev. 2018; 98(4).1-91 doi: 10.1152/physrev.00063.2017.

[35] Meroni M, Dongiovanni P, Tiano F, Piciotti R, Alisi A, Panera N. β-Klotho as novel therapeutic target in Metabolic dysfunction-associated steatotic liver disease (MASLD): A narrative review. Biomed Pharmacother. 2024; 180: 117608. doi: 10.1016/j.biopha.2024.117608.

[36] Islam MS, Wei P, Suzauddula M, et al. The interplay of factors in metabolic syndrome: understanding its roots and complexity. Mol Med. 2024; 30(1): 1-21. doi:10.1186/s10020-024-01019-y.

[37] Bovolini A, Garcia J, Andrade MA, Duarte JA. Metabolic syndrome pathophysiology and predisposing Factors. Int J Sports Med. 2021;

42(3): 199-214. doi:10.1055/a-1263-0898.

[38] Hooshmand E, Akbarzade I, Delbari D, Niroomand M, Ghavidel F, Saatchi M. Prevalence of metabolic syndrome among adults aged 50 and above and associated factors: A cross-sectional study from Ardakan cohort study on aging (ACSA). Heal Sci Reports. 2025; 8(3):1-7. doi: 10.1002/hsr2.70508.

[39] Holmannova D, Borsky P, Andrys C, et al. The influence of metabolic syndrome on potential aging biomarkers in participants with metabolic syndrome compared to healthy controls. Biomedicines. 2024; 12(1): 1-15. doi:10.3390/biomedicines12010242.

Journal of Associated Medical Sciences 2026; 59(2): 187-195

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Towards precision hematology: Machine learning-driven comparative feature importance analysis of hematologic parameters in newly diagnosed chronic myeloid leukemia vs healthy controls

Hossein Ayatollahi[1,2], Arefeh Mazhari[1,2], Seyed Mohammad Tabatabaei[3,4], Samaneh Talebi[5], Mahnaz Dehghan[2], Maryam Sheikhi[2], Yasamin Ayatollahi[2], Zahra Darvish Khalilabadi[1,2*]

[1]Department of Hematology and Blood Banking, Mashhad University of Medical Sciences, Mashhad, Iran.
[2]Cancer Molecular Pathology Research Center, Mashhad University of Medical Sciences, Mashhad, Iran.
[3]Department of Medical Informatics, Mashhad University of Medical Sciences, Mashhad, Iran.
[4]Applied Biomedical Research Center, Basic Sciences Research Institute, Mashhad University of Medical Sciences, Mashhad, Iran.
[5]Center for Healthcare Data Modeling, Department of Biostatistics and Epidemiology, School of Public Health, Shahid Sadoughi University of Medical Sciences, Yazd, Iran.

ARTICLE INFO

Article history:
Received 5 October 2025
Accepted as revised 2 February 2026
Available online 16 February 2026

Keywords:
Chronic myeloid leukemia, CML, machine learning, Random Forest, simulated annealing, NLR, BLR, dNLR.

ABSTRACT

Background: Chronic Myeloid Leukemia (CML) is a myeloproliferative neoplasm characterized by uncontrolled granulocytic proliferation and is often initially suspected based on peripheral blood smear findings. Utilizing machine learning to analyze routinely available Complete Blood Count (CBC) parameters and derived inflammatory indices may facilitate early identification of CML at the time of initial diagnosis.

Objectives: The objective of this study is to evaluate the diagnostic value of routinely available Complete Blood Count (CBC) parameters and derived inflammatory indices for the early identification of chronic myeloid leukemia at the time of initial diagnosis, using machine learning–based models.

Materials and methods: This study was conducted on 295 newly diagnosed cases of CML and 340 normal control samples. A total of 49 factors were subjected to study, including the following: variables included in the CBC of patients, inflammatory indices, and demographic data. Logistic regression analysis was performed to identify relevant variables, resulting in the selection of 22 features. Subsequently, multiple machine learning algorithms, including Random Forest (RF), Recursive Feature Elimination (RFE), Simulated Annealing (SA), Decision Tree (DT), K-Nearest Neighbor (KNN), and Xg-Boost (XGB), were applied to evaluate the diagnostic performance pf the selected features.

Results: The findings of this study indicate that the factors most pertinent in initial diagnosis in comparison with normal control include the WBC count, the relative percentage of cells including neutrophils, monocytes, and lymphocytes, and a series of indicators related to RBC such as RBC count and RDW-CV, as well as the index of inflammatory NLR, and BLR and PDW.

Conclusion: This study demonstrates that machine learning models based solely on routinely available CBC parameters and derived inflammatory indices can support the early identification of CML at the time of initial diagnosis. In addition to leukocyte-related variables, RBC-related and inflammatory indices provided complementary diagnostic information, highlighting their potential value in early-stage CML detection. The application of machine learning techniques could prioritize the development of more user-friendly dashboards to facilitate the diagnosis of CML at initial diagnosis.

* **Corresponding contributor.**
Author's Address: *Cancer Molecular Pathology Research Center, Mashhad University of Medical Sciences, Mashhad, Iran.*
E-mail address: zdarvish99@gmail.com
doi: 10.12982/JAMS.2026.050
E-ISSN: 2539-6056

Introduction

Chronic myeloid leukemia (CML) is a myeloproliferative neoplasm, defined as a disorder characterized by uncontrolled granulocyte proliferation. It is estimated that CML accounts for approximately 15% of all adult leukemias.[1] The disease predominantly affects older age groups; however, it can manifest in any age demographic. Approximately 50% of cases of this disease are identified incidentally during a routine medical examination. In instances where symptoms are present, the most observed manifestations are fatigue, weight loss, and fever.[2] The initial stage of diagnosis is the examination of a peripheral blood smear. Subsequent molecular testing is then performed to confirm the BCR-ABL fusion.

Machine learning is a subfield of artificial intelligence that employs sophisticated algorithms to analyze intricate data sets, such as those pertaining to disease prevalence or disease outcomes. The application of machine learning enables the development of models capable of accurate estimation based on the analysis of raw data.[1]

The application of machine learning techniques has been limited within the context of CML.[3] In spite of the use of Tyrosine kinase inhibitors (TKIs), some patients require alternative treatments in accordance with their specific conditions. These treatments may be identified through the use of machine learning algorithms, enabling the selection of the optimal treatment.[1] In a study on CML, machine learning models have been employed to distinguish CML-related mature neutrophils from normal mature neutrophils, with a sensitivity of ≤95.30% and ≤95.80% specificity.[4] In addition, models have been constructed to predict treatment responses within a period of 6 to 18 months, utilizing data obtained from cyTOF.[5] Attempts have been made to apply machine learning in the field of diagnosis, particularly in the context of CML and other leukemias. These efforts have largely focused on the digitization of microscopic smear images. The training group employed blood smears with a definitive diagnosis of CML. Subsequently, the models created were subjected to a verification process with the test group.[6-8] A study was conducted on cell blood counts from CML patients just five years before the definitive diagnosis. The results indicated that the BCR-ABL test result can be predicted with a high degree of accuracy using blood cell count results.[9]

This study aims to examine the factors that can be identified at the initial stage of diagnosis through the utilization of a straightforward and readily accessible CBC (complete blood count) test, rather than the recommendation of a digital blood smear, which has been the focus of investigation in previous studies. The aforementioned factors have been identified as essential in the diagnostic process, particularly in the initial diagnosis of CML patients in comparison to normal controls. A diagnosis of CML can be reached with greater efficiency by investigating these factors, which can then be followed by the appropriate CML confirmation tests, if necessary.

Material and method

Patients and data collection

This retrospective study was conducted between 2010 and 2023 on 325 newly diagnosed cases of CML and 341 normal control samples at the Cancer Molecular Pathology Research Center, Mashhad University of Medical Sciences, Mashhad, Iran. The demographic data of the patients was extracted from the electronic medical records. The information encompasses the complete set of variables included in the CBC of patients and demographic data such as age and gender. The data obtained from CBC were utilized to calculate the inflammatory ratios, including neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), monocyte-to-lymphocyte ratio (MLR), platelet-to-hemoglobin ratio (PHR), basophil-to-lymphocyte ratio (BLR), eosinophil-to-lymphocyte ratio (ELR), derived neutrophil-to-lymphocyte ratio (DNLR), neutrophil-to-lymphocyte*platelet ratio (NLPR), systemic inflammation index (SII), and systemic inflammation response Index (SIR-I). Furthermore, a comprehensive examination of the peripheral blood smear was conducted to ascertain the morphology of the blood cells, particularly in view of the presence of immature myeloid cells, and to evaluate the cells' differentiation. All the results are confirmed by a pathologist and hematologists. A total of 49 factors were subjected to study, including the following: age, gender, RBC count, hemoglobin, hematocrit, MCV, MCH, MCHC, RDW-CV, platelet count, PDW, MPV, WBC count, the absolute and relative blast count, absolute and relative promyelocyte count, absolute and relative myelocyte count, absolute and relative Metamyelocyte count, absolute and relative band cell count, absolute and relative neutrophil count, absolute and relative eosinophil count, absolute and relative basophil count, absolute and relative monocyte count, absolute and relative lymphocyte count, absolute and relative nRBC count, NLR, PLR, MLR, PHR, BLR, ELR, DNLR, NLPR, SIRI, and SII. Methods were carried out in accordance with relevant guidelines and regulations, including the Declaration of Helsinki.

Statistical analysis

In the initial phase of the study, 325 patients and 341 controls were considered, with those exhibiting deficiencies in information being excluded. Ultimately, 295 patients and 340 controls were included in the study. A total of 49 factors were investigated in the patient group in comparison to the controls. A logistic regression analysis was conducted between the control group and the patient utilizing R 4.4.2 and RStudio 2024.09.1, resulting in the incorporation of a total of 22 factors into the model. Furthermore, patients were divided into two groups—the training (80%) and the testing (20%) groups—for the purpose of developing

machine learning models. The following section of the analysis employs a variety of machine learning algorithms, including Random Forest (RF), Recursive Feature Elimination (RFE), Simulated Annealing (SA), Decision Tree (DT), K-Nearest Neighbor (KNN), and Xg-Boost (XGB). Subsequently, the importance of the factors in new cases was evaluated in comparison to the control group. Subsequently, to evaluate the predictive performance of the models, a range of metrics were employed, including sensitivity, specificity, accuracy, and the area under the ROC curve (AUC). To conclude, an overall assessment was conducted in order to ascertain the relative importance of each factor, with assistance from the derived importance of each factor from each model. Figure 1 depicts the sequence of steps undertaken during the study. Statistical analysis was performed using the R 4.4.2 and RStudio 2024.09.1 and packages including "Random Forest", "caret", "e1071", "XGBoost", "rminer" and "rpart" was employed for feature importance analysis with a significance level of 0.05.

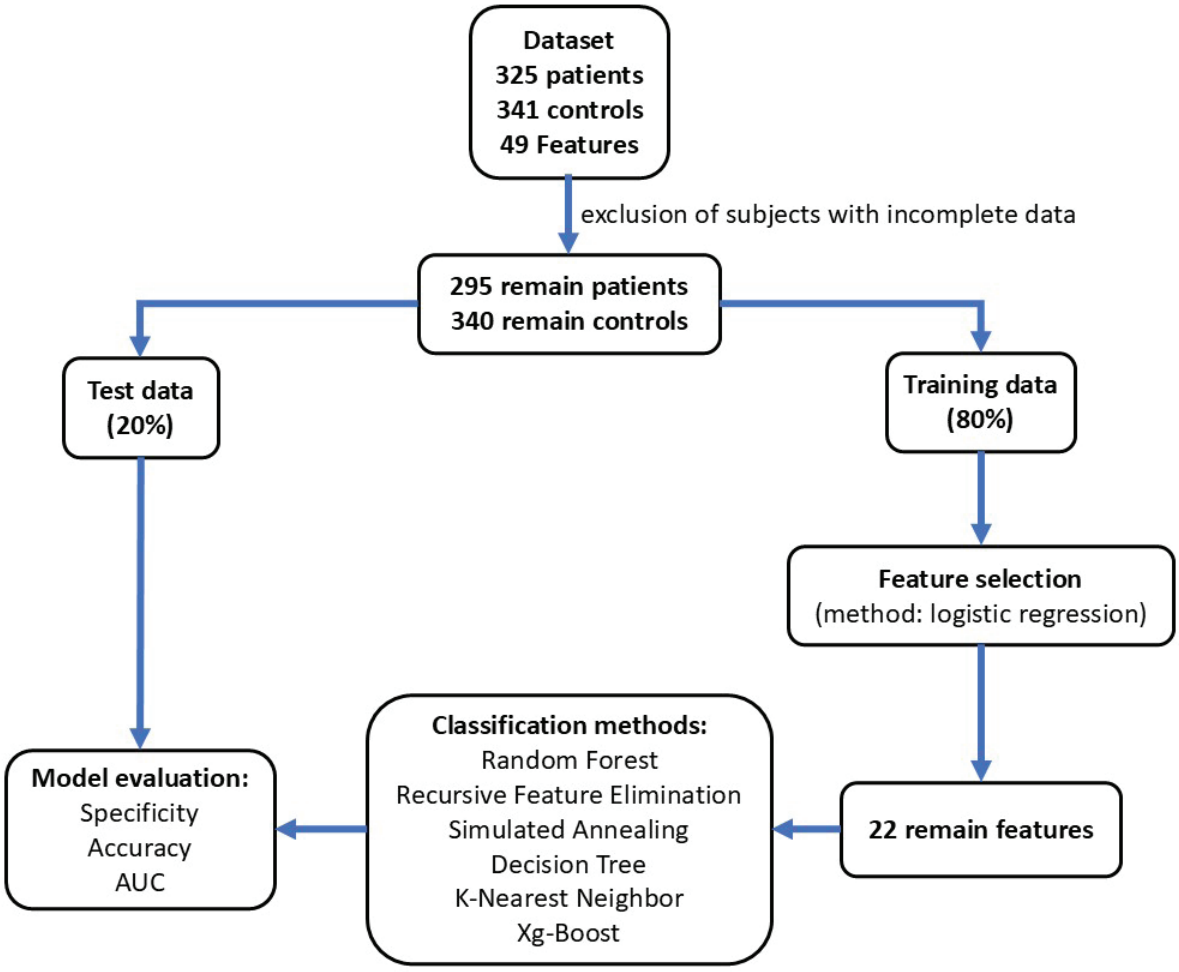


Figure 1. *The framework of the study design.*

Result

In total, 295 patients were included in the study, with 53% men (N=155) and 47% women (N=140) representing the sample population. An investigation was carried out into the 49 factors derived from the initial diagnoses of the subjects.

At the outset of the investigation, a logistic regression analysis was conducted between the two study groups, resulting in the identification of 22 statistically significant factors. The aforementioned factors encompass age, gender, RBC count, hematocrit, MCV, RDW-CV, PLT, PDW, WBC, neutrophil relative count, eosinophil relative count, basophil relative count, lymphocyte relative count, monocyte relative count, NLR, PLR, PHR, ELR, BLR, dNLR, and NLPR.

As shown in Table 1, there was a statistically significant difference between CML patients and healthy controls in all hematologic parameters. Notably, the median WBC count was significantly higher in the CML group (p<0.001).

Table 1. Comparison of laboratory parameters between normal control and CML patients.

Parameters	Control normal (N=340)	Patients (N=295)	*p* value
Age (year)	51 (39, 64)	44 (32, 56)	<0.001
RBC (x10^6/µL)	4.7 (4.4, 4.9)	3.7 (3.1, 4.3)	<0.001
HCT (%)	40 (37.9, 42)	33.6 (28.9, 38.7)	<0.001
MCV (fL)	85.8 (83.5, 88.4)	89.4 (85.2, 93.5)	<0.001
RDW-CV (%)	13.2 (12.7, 13.8)	17.3 (16,18.7)	<0.001
PLT (x10^3/µL)	246 (207, 287)	322 (224, 516)	<0.001
PDW (fL)	15.9 (12,16.9)	12.5 (11.2,14.4)	<0.001
MPV (fL)	9.5 (8.4,10.4)	9.8 (9,10.5)	<0.001
WBC (x10^3/µL)	6.7 (5.7, 7.7)	92.2 (33.1,156.7)	<0.001
Neutrophils (%)	56.4 (50.9, 62.8)	43 (32, 56)	<0.001
Lymphocytes (%)	32.9 (28, 38.4)	6 (3,11)	<0.001
Monocytes (%)	6.2 (4.7, 8.3)	0 (0, 2)	<0.001
NLR	1.7 (1.3, 2.2)	7.3 (4.1,15)	<0.001
PLR	114.5 (91,142.3)	99.2 (41.8, 215.3)	0.023
PHR	1.8 (1.4, 2.1)	2.9 (2, 5)	<0.001
BLR	0.004 (0.002, 0.008)	0.00 (0.00, 0.00)	<0.001
ELR	0.07 (0.05, 0.11)	0.14 (0.00, 0.97)	0.008
DNLR	1.29 (1.04,1.69)	0.75 (0.47, 1.27)	<0.001
NLPR	0.67 (0.52, 0.96)	2.35 (0.99, 4.8)	<0.001

In the subsequent phase, the aforementioned factors were employed within the RF, RFE, SA, DT, KNN, and XGB models to ascertain the feature importance within the context of each model. The results of this step are presented in Figure 2.

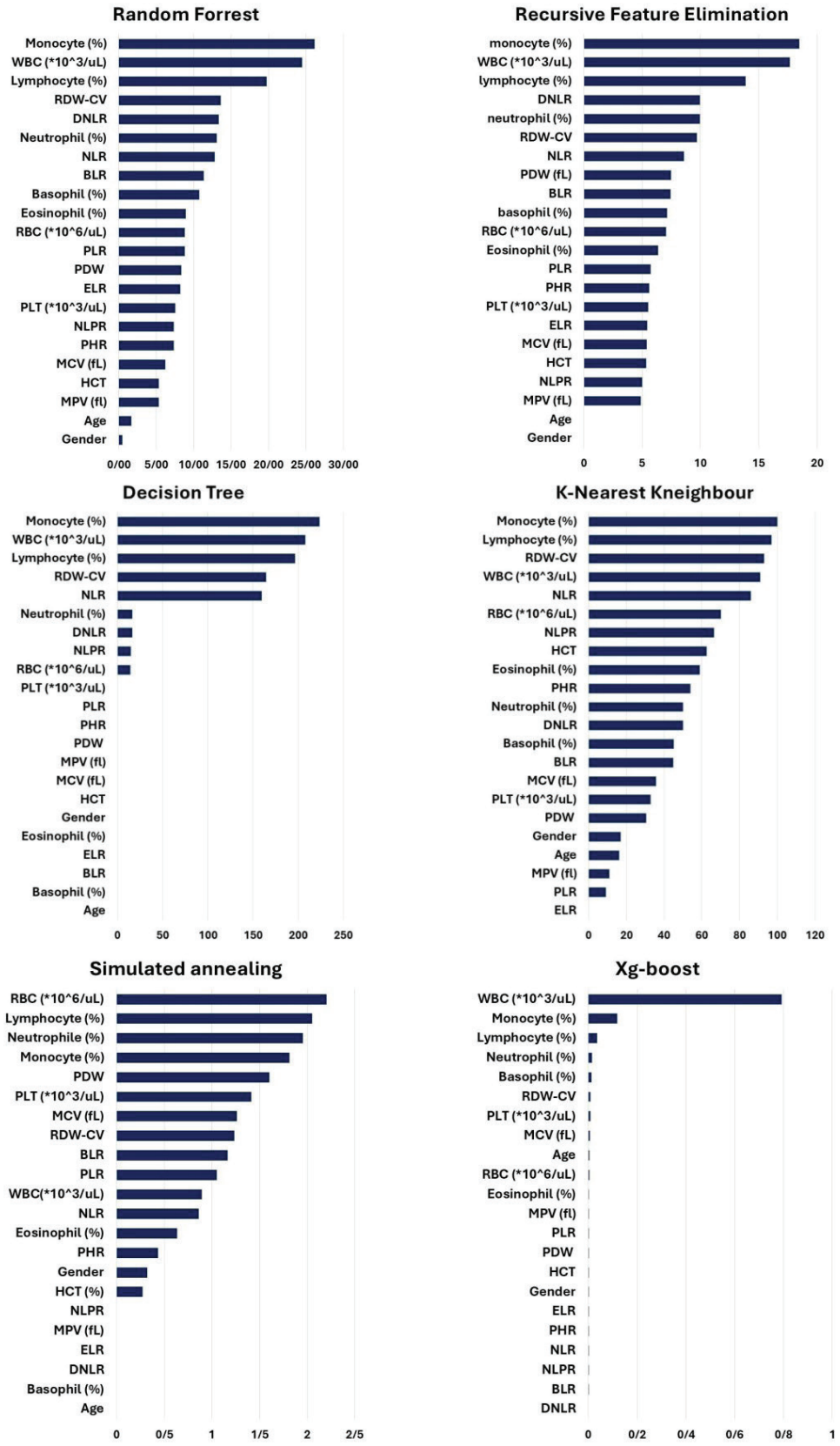


Figure 2. *Importance index of variables.*

The findings of this study indicate that the factors most pertinent in initial diagnosis in comparison with normal control include the WBC count, the relative percentage of cells including neutrophils, monocytes, and lymphocytes, and a series of indicators related to RBC such as RBC count and RDW-CV, as well as the index of inflammatory NLR, and BLR and PDW.In order to assess the robustness of the proposed models, a ROC curve has been constructed, the result of which is illustrated in Figure 3. The ROC curve analysis indicates that the RF and RFE models with an AUC level of 99 exhibit the greatest predictive efficacy, while other models with an AUC level of 98 demonstrate the second-highest level of performance.

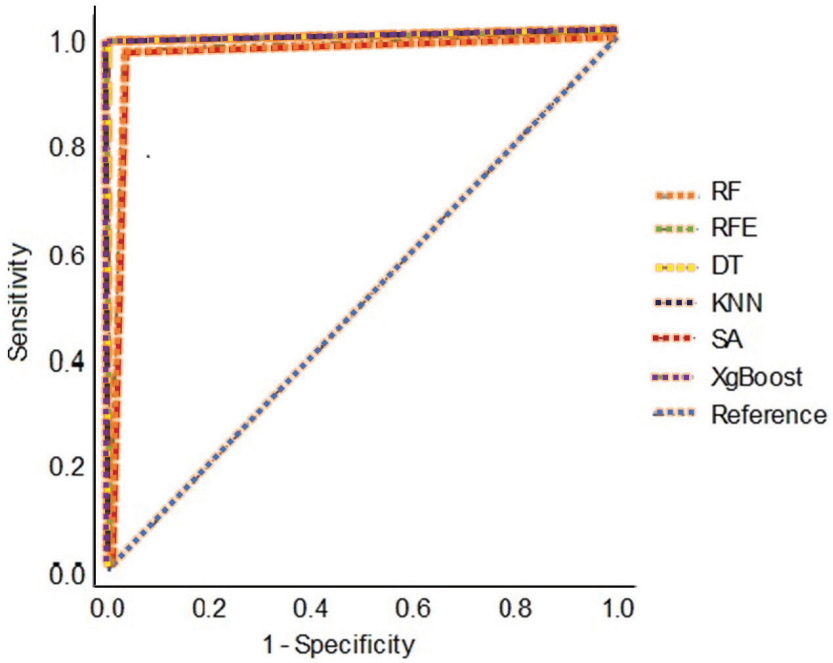


Figure 3. *Receiver operating characteristic (ROC) curve for different machine learning models.*

The subsequent phase entails the computation of sensitivity, specificity, and area under the curve (AUC) indices, which are utilized to assess the performance of these models, as illustrated in Table 2. Among the used models, the RF and RFE models exhibit the greatest sensitivity (equating to a value of 1) and the highest specificity (equating to a value of 0.98), respectively. These models indicate that WBC, the relative count of monocytes, neutrophils, and lymphocytes, as well as indices such as DNLR and NLR, as well as RDW, may be of diagnostic importance.

Table 2. Comparison of predictive power indices for models.

Model	**Sensitivity**	**Specificity**	**Accuracy** (0.95% CI)	**AUC** (0.95% CI)
Random Forest	1.00	0.98	0.99 (0.95, 0.99)	0.992 (0.97, 1)
Recursive Feature Alimination	0.97	0.96	0.97 (0.92, 0.99)	0.97 (0.93, 1)
Simulated Annealing	1.00	0.98	0.99 (0.95, 0.99)	0.992 (0.97, 1)
Decision Tree	0.97	0.96	0.97 (0.92, 0.99)	0.97 (0.93, 1)
K-Nearest Neighbor	0.97	0.96	0.97 (0.92, 0.99)	0.97 (0.93, 1)
Xg-Boost	0.97	0.96	0.97 (0.92, 0.99)	0.97 (0.93, 1)

To enhance the clinical utility of our findings, ROC curve analysis was conducted for the important variables based on the RF model. The optimal cutoff value WBC count was 10.35 (AUC: 0.97, Sensitivity: 90.8%, Specificity: 99.7%). Detailed thresholds for other parameters are summarized in Table 3.

Table 3. Cut-off points for the important variables based on the RF model.

Variable	Cut-off point	Sensitivity	Specificity	AUC (0.95% CI)
WBC (×10^3/uL)	>10.35	90.8	99.7	0.97 (0.95,0.98)
Lymphocyte (%)	<21.03	94.4	94.6	0.94 (0.92,0.96)
RDW-CV (%)	>15.15	87.5	93.1	0.93 (0.90,0.96)
Monocyte (%)	<3.05	97.1	91.2	0.98 (0.96,0.99)
DNLR	<0.82	85.9	85.7	0.91 (0.89,0.95)
Neutrophil (%)	<45.03	83.9	87.2	0.89 (0.87,0.91)

Discussion

As previously indicated, recent research findings have demonstrated that machine learning is a highly prevalent method in the field of medical diagnostics and prognostication. While this method has received less attention in the context of CML. The present study evaluated multiple machine learning models to identify key predictors of CML at the initial stage of diagnosis using CBC parameters, derived inflammatory indices, and demographic data.

The application of multiple machine learning algorithms in this study was intended to enhance the robustness and reliability of the identified predictors for early CML detection. Different machine learning methods capture distinct patterns within the data, and agreement across models increases confidence in the observed findings.[10]

Despite methodological differences among the applied algorithms, a high degree of consistency was observed in the ranking of key features. Variables such as WBC count, relative proportions of neutrophils, lymphocytes, and monocytes, as well as RDW-CV and inflammatory indices including NLR and DNLR, were repeatedly identified as important predictors across Random Forest, Recursive Feature Elimination, Simulated Annealing, and XGBoost models. Such convergence across independent modeling strategies supports the biological plausibility and clinical relevance of these CBC-derived parameters in the early identification of CML.[11,12]

Tree-based models, particularly Random Forest and Decision Tree algorithms, demonstrated strong diagnostic performance and relative interpretability, making them well-suited for clinical screening applications. Random Forest has been shown to perform robustly in biomedical datasets characterized by nonlinearity and complex feature interactions, while also providing stable feature importance measures.[13] Feature selection techniques such as Recursive Feature Elimination and Simulated Annealing further strengthened the analysis by reducing dimensionality and mitigating overfitting, thereby enhancing model generalizability.[14,15]

Other algorithms, including K-Nearest Neighbor and XGBoost, contributed complementary perspectives but present specific limitations. KNN performance is sensitive to feature scaling and class imbalance, while XGBoost requires careful hyperparameter tuning to avoid overfitting, particularly in single-center retrospective datasets.[16,17]

Machine learning has been increasingly applied to leukemia diagnosis and management, but most studies focus on prognosis or subtype classification rather than early clinical diagnosis of CML. Previous Machine learning studies using CBC or molecular markers reported moderate to high accuracy but did not systematically integrate inflammatory or RBC-related indices. The integration of low-cost, interpretable biomarkers into machine learning models provides a potentially useful tool for early screening, especially where molecular assays are unavailable. Logistic regression identified several significant predictors, including WBC count, neutrophil and lymphocyte percentages, NLR, BLR, RDW, and PDW. In addition, the Random Forest and Simulated Annealing models consistently ranked these variables among the most important features, contributing to the strong diagnostic performance (sensitivity =1.00, specificity =0.98) observed in our analysis.

The high feature importance of WBC aligns with established clinical guidelines, where leukocytosis is a hallmark of CML. In the context of CML, WBC count is a critical parameter, ranking amongst the most significant factors in the clinical management of the disease. A study of 56 CML patients revealed that individuals presenting with elevated WBC levels, abdominal discomfort, and splenomegaly should undergo evaluation for CML.[18] Furthermore, another research has indicated that a WBC count exceeding 20,000/μL in conjunction with basophilia may serve as a useful indicator for the identification of CML.[19]

Relative percentages of neutrophils, lymphocytes, and monocytes provide additional information beyond absolute counts, reflecting early disease-related shifts in white cell distributions. relative percentage of monocytes and lymphocytes was identified as a

significant factor in most models. Despite findings from some studies indicating that absolute cell count is not associated with molecular response, there has been minimal discussion in the literature regarding its significance during the initial diagnosis.[20] In some studies, an absolute lymphocyte count (ALC) above 4000/uL within 3-5 months after treatment has been reported to be associated with a decreased overall survival (OS) rate.[21] The relative percentage of neutrophil cells is another significant factor that must be considered. Neutrophilia is a common initial finding in CML patients.[1]

Inflammatory indices such as NLR and BLR also emerged as significant predictors. Although their role in CML diagnostics has been less explored, their accessibility and low cost make them attractive for early screening and prioritization.

Similarly, red blood cell distribution width (RDW) and platelet distribution width (PDW) were among the top predictors identified. While RDW has been recognized as a prognostic marker in other hematologic disorders, its application in early CML diagnosis has not been extensively evaluated.[22] PDW remains understudied in the context of CML, but its association with platelet variability could reflect underlying pathophysiological changes in hematopoiesis.[21] The identification of accessible, cost-effective predictors highlights the potential for developing decision support tools that flag high-risk individuals for confirmatory testing. Future studies should validate these models across multi-center cohorts and explore the integration of additional clinical variables, such as symptomatology and physical examination findings. Furthermore, future research could prioritize the development of more user-friendly dashboards to facilitate the diagnosis of CML.

Limitations

One potential limitation of this study is the relatively small sample size. This study is limited by its retrospective design and single-center data, which may affect generalizability. External validation and prospective studies would strengthen the evidence for clinical implementation.

Conclusion

CML is a myeloproliferative disorder, defined as a disorder characterized by uncontrolled granulocyte proliferation. While several models performed exceptionally well, we recommend the Random Forest model for high-precision diagnostic tasks and the Decision Tree model for its ease of visual interpretation in clinical settings. This study aims to examine the factors that can be identified at the initial stage of diagnosis through the utilization of a straightforward and readily accessible CBC test, rather than the recommendation of a digital blood smear, which has been the focus of investigation in previous studies. Key discriminative features included WBC count, relative leukocyte proportions, RBC-related indices (RBC count and RDW-CV), and inflammatory markers such as NLR, BLR, and PDW. These findings highlight the potential of integrating CBC-derived features with machine learning approaches to enable efficient early screening and support clinical decision-making. Furthermore, future research could prioritize the development of more user-friendly dashboards to facilitate the diagnosis of CML.

Funding

This article has not been supported or received any grant or fund.

Ethical approval

This study received approval from the Ethics Committee of MUMS, and all participants provided informed consent (Ethics code: IR.MUMS.MEDICAL.REC.1402.525).

Conflict of interest

Authors declared no conflict of interest.

Data availability

All the data have been included in the manuscript and will be made available upon publication of the manuscript.

CRediT authorship contribution statement

Hossein Ayatollahi: conceptualization, data curation, investigation, methodology, visualization, project administration, validation, writing: original draft, review and edit; **Arefeh Mazhari**: data curation, methodology, writing: original draft, review and edit; **Seyed Mohammad Tabatabaei**: formal analysis, methodology, writing: original draft, review and edit; **Samaneh Talebi**: formal analysis, methodology, writing : original draft, review and edit; **Mahnaz Dehghan**: data curation, writing: review and edit; **Maryam Sheikhi**: methodology, writing: review and edit; **Yasamin Ayatollahi**: data curation, writing: review and edit; **Zahra Darvish Khalilabadi**: conceptualization, data curation, investigation, methodology, visualization, writing: original draft, review and edit.

Acknowledgements

The authors wish to thank all colleagues at Mashhad University of Medical Sciences.

References

[1] Elhadary M, Elsabagh AA, Ferih K, Elsayed B, Elshoeibi AM, Kaddoura R, et al. Applications of machine learning in chronic myeloid leukemia. Diagnostics. 2023;13.

[2] Pandey MK, Pal S. Evaluation of chronic myelogenous leukemia (CML) as the chronic phase of disease using machine learning techniques. Int J Mech Eng Educ. 2022; 7: 1-12.

[3] Shanbehzadeh M, Afrash MR, Mirani N, Kazemi-

Arpanahi H. Comparing machine learning algorithms to predict 5-year survival in patients with chronic myeloid leukemia. BMC Med Inform Decis Mak. 2022; 22(1): 236. doi:10.1186/s12911-022-01980-w.

[4] Ni W, Tong X, Qian W, Jin J, Zhao H. Discrimination of malignant neutrophils of chronic myelogenous leukemia from normal neutrophils by support vector machine. Comput Biol Med. 2013; 43(9): 1192-5. doi:10.1016/j.compbiomed.2013.06.004

[5] Bergenheim MB. Towards prediction of CML treatment outcomes with machine learning and CyTOF data: The University of Bergen; 2024.

[6] Cerrato TR. Use of artificial intelligence to improve access to initial leukemia diagnosis in low-and middle-income countries. American Society of Clinical Oncology; 2020.

[7] Dese K, Raj H, Ayana G, Yemane T, Adissu W, Krishnamoorthy J, et al. Accurate machine-learning-based classification of leukemia from blood smear images. Clin Lymphoma Myeloma Leuk. 2021; 21(11): e903-e14. doi: 10.1016/j.clml. 2021. 06.025.

[8] Salah HT, Muhsen IN, Salama ME, Owaidah T, Hashmi SK. Machine learning applications in the diagnosis of leukemia: Current trends and future directions. Int J Lab Hematol. 2019; 41(6): 717-25. doi:10.1111/ijlh.13089

[9] Hauser RG, Esserman D, Beste LA, Ong SY, Colomb DG, Bhargava A, et al. A aachine learning model to successfully predict future diagnosis of chronic myelogenous leukemia with retrospective electronic health records data. Am J Clin Pathol. 2021; 156(6): 1142-8. doi:10.1093/ajcp/aqab086

[10] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. Springer series in statistics New-York; 2009.

[11] Guyon I, Elisseeff A. An introduction to variable and feature selection. J Mach Learn Res. 2003;(3): 1157-82.

[12] Saeys Y, Inza I, Larranaga P. A review of feature selection techniques in bioinformatics. Bioinformatics. 2007; 23(19): 2507-17. doi:10.1093/bioinformatics/btm344.

[13] Couronné R, Probst P, Boulesteix A-L. Random forest versus logistic regression: a large-scale benchmark experiment. BMC Bioinformatics. 2018; 19(1): 270. doi: 10.1186/s12859-018-2264-5.

[14] Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. Machine learning. 2002; 46(1): 389-422.

[15] Kirkpatrick S, Gelatt Jr CD, Vecchi MP. Optimization by simulated annealing. Science. 1983; 220(4598): 671-80.

[16] Chen T. XGBoost: A scalable tree boosting system. Cornell University. 2016.

[17] Kotsiantis SB, Zaharakis I, Pintelas P. Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering. 2007; 160(1): 3-24.

[18] Algahtani FH, Alqahtany FS. Evaluation and characterisation of Chronic myeloid leukemia and various treatments in Saudi Arabia: A retrospective study. J Infect Public Health. 2020; 13(2): 295-8. doi:10.1016/j.jiph.2019.12.006

[19] Fenu E, O'Neill SS, Insuasti-Beltran G. BCR-ABL1 p210 screening for chronic myeloid leukemia in patients with peripheral blood cytoses. Int J Lab Hematol. 2021; 43(6): 1458-64. doi:10.1111/ijlh.13635

[20] Pepedil-Tanrikulu F, Buyukkurt N, Korur A, Sariturk C, Aytan P, Boga C, et al. Significance of lymphocyte count, monocyte count, and lymphocyte-to-monocyte ratio in predicting molecular response in patients with chronic myeloid leukemia: a single-centre experience. Clin Lab. 2020; 66(3). doi:10.7754/Clin.Lab.2019.190628

[21] Mulas O, Mola B, Madeddu C, Caocci G, Macciò A, Nasa GL. Prognostic role of cell blood count in chronic myeloid neoplasm and acute myeloid leukemia and its possible implications in hematopoietic stem cell transplantation. Diagnostics. 2022; 12(10): 2493. doi.org/10.3390/diagnostics12102493

[22] Iriyama N, Hatta Y, Kobayashi S, Uchino Y, Miura K, Kurita D, et al. Higher red blood cell distribution width is an adverse prognostic factor in chronic-phase chronic myeloid leukemia patients treated with tyrosine kinase inhibitors. Anticancer Res. 2015; 35(10): 5473-8. PMID: 26408711

Journal of Associated Medical Sciences 2026; 59(2): 196-207

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



The usefulness of endometrial receptivity biomarkers as predictors of IVF outcome

Zainab A. Aljarah[1], Rana M. Hameed[1,2*], Wasan G. Al Safi[3]

[1]Department of Chemistry and Biochemistry, College of Medicine, University of Kerbala, Karbala, Iraq.
[2]Department of Biochemistry, Medical College, Al-Subtain University of Medical Sciences/International Branch of Tehran University of Medical Sciences, Karbala, Iraq.
[3]Department of Obstetrics and Gynecology, College of Medicine, University of Kerbala, Karbala, Iraq.

ARTICLE INFO

Article history:
Received 2 August 2025
Accepted as revised 28 January 2026
Available online 16 February 2026

Keywords:
Assisted reproductive techniques, biomarkers, in vitro fertilization, ovarian function tests, pregnancy.

ABSTRACT

Background: Despite the progress in assisted reproductive technology, predicting *in vitro* fertilization (IVF) success remains challenging.

Objectives: This study aimed to investigate the relationship between endometrial receptivity biomarkers and Ovarian Sensitivity Index (OSI) as predictors of IVF outcomes.

Materials and methods: A total of 72 women undergoing IVF treatment, between December 2023 and July 2024, participated in this study. Serum levels of MUC1, GDF-9, and HLA were measured, OSI was calculated, and pregnancy outcomes (β-hCG) were recorded. Participants were divided into two groups based on OSI (≤4 and >4), and biomarker levels were compared. Correlation analysis, ROC curves, and logistic regression models were used to assess predictive capabilities.

Results: OSI showed strong predictive value (AUC=0.751, 95% CI: 0.613-0.859, $p<0.05$) with a threshold of >8 (sensitivity=52.63%, specificity=94.12%). Endometrial biomarkers showed limited individual predictive capacity (AUCs: MUC1=0.625, GDF-9=0.507, HLA=0.589). OSI retained significant association with pregnancy outcomes across all regression models (OR=1.19-1.38, $p<0.05$). Combining hormonal parameters with OSI significantly improved model performance (AUC=0.867, $R^2=0.509$). Significant correlations were observed between OSI and hormonal parameters, particularly AMH (r=0.66, $p<0.05$) and E2 (r=0.49, $p<0.05$).

Conclusion: While endometrial receptivity biomarkers showed limited individual predictive value, the incorporation of OSI with hormonal parameters provides better predictive value of IVF outcomes, suggesting the importance of a comprehensive assessment in assisted reproduction.

Introduction

Yet, after more than four decades of IVF practice, predicting treatment success remains one of the most difficult aspects of fertility treatment.[1] The success rates are influenced by multiple factors, including psychological such as anxiety and depression,[2] environmental and occupational exposures,[3] and lifestyle ones including smoking, alcohol consumption, and dietary habits.[4] The measurement of IVF success itself is a complex challenge, requiring the addition of clinical, financial, and patient-centered perspectives.[5] Furthermore, even when pregnancy is achieved, IVF pregnancies are associated with increased risks of

* **Corresponding contributor.**
Author's Address: *Department of Chemistry and Biochemistry, College of Medicine, University of Kerbala, Karbala, Iraq.*
E-mail address: rana.m@uokerbala.edu.iq
doi: 10.12982/JAMS.2026.051
E-ISSN: 2539-6056

adverse obstetric and perinatal outcomes, including higher rates of congenital anomalies. [6]

Endometrial receptivity, a dynamic state that determines the endometrium's ability to support embryo implantation, is fundamental to successful IVF outcomes. This process is characterized by a precisely timed "window of implantation" (WOI) during which the endometrium exhibits optimal conditions for embryo attachment. [7] Moreover, this complex biological process involves complex molecular interactions with various factors, including adhesion molecules, cytokines, and growth factors.[8] In recent years, significant advances in transcriptomics and proteomics have reformed the understanding of endometrial receptivity, leading to the development of sophisticated diagnostic tools such as the Endometrial Receptivity Array (ERA) (9). Building upon this understanding, researchers have found several promising biomarkers, particularly MUC1 and GDF-9, which play important roles in the implantation process. Particularly, MUC1, a trans-membrane glycoprotein, functions as an important mediator of embryo attachment and implantation, with its expression carefully regulated throughout the menstrual cycle.[10] Also, GDF-9, a member of the transforming growth factor-β superfamily, has been reported as an essential factor in early ovarian development and oocyte survival. [11,12] Furthermore, recent studies have found that these biomarkers not only indicate endometrial receptivity but also show potential as predictive tools for IVF success.[13-15]

In addition, the role of immune factors, particularly HLA molecules, has gained increasing attention in the context of IVF outcomes. Specifically, HLA molecules are fundamental in starting and maintaining maternal-fetal immune tolerance, a critical aspect of successful implantation.[16] Of particular interest is the non-classical HLA class Ib, which shows a vital role in protecting the developing embryo from maternal immune responses.[17] Consequently, the assessment of HLA expression patterns, alongside other established biomarkers, offers a more comprehensive approach to predicting implantation success and overall IVF outcomes.

The concept of ovarian responsiveness, as measured by the Ovarian Sensitivity Index (OSI), has appeared as a useful tool in predicting IVF outcomes. This index, calculated as the ratio between retrieved oocytes and total gonadotropin dose, offers important information into individual patient responses to hormonal stimulation.[18] OSI has showed stronger predictive value for clinical pregnancy than conventional measures such as total oocyte count, principally in patients with normal expected ovarian response.[19] However, despite the growing understanding of both endometrial receptivity biomarkers and ovarian sensitivity, there remains a significant knowledge gap regarding the potential relationship between these two essential aspects of IVF success. Specifically, the interaction between endometrial receptivity markers (MUC1, GDF-9, and HLA) and OSI has not been thoroughly investigated. Therefore, the current study sought to examine the relationship between endometrial receptivity biomarkers (MUC1, GDF-9, HLA) and the OSI in women undergoing IVF treatment, and to determine whether these markers, independently or in combination, could help as predictors of IVF success.

Materials and methods

Study participants

This study was conducted at the Fertility Center of Al-Kafeel super-specialty hospital in Karbala, Iraq. A total of 72 women with primary or secondary infertility who were undergoing IVF treatment were recruited between December 2023 and July 2024. Women aged 21-42 years, who were diagnosed with primary or secondary infertility, undergoing IVF treatment with FSH stimulation, and having had serum AMH levels evaluated before treatment were included. The exclusion criteria were females with endocrine disorders, thyroid diseases, hypogonadotropic hypogonadism, severe endometriosis, previous ovarian surgery, radiation or chemotherapy and premature ovarian failure.

Sample size was determined by the available patient population meeting our strict inclusion criteria during the study period. A formal a priori power calculation was not performed, which represents a limitation. Post-hoc power analysis indicated that with our sample size of 72 participants, we had approximately 80% power to detect the observed effect size for OSI.

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki. The research protocol was reviewed and approved by the Institutional Review Board of the College of Medicine, University of Kerbala D 6/259 in 15/1/2024. In addition, all participants were well informed about the study's purpose and procedures. Informed consent was obtained from all participants prior to their enrollment in the study.

Data collection

At initial consultation, comprehensive demographic data were collected from each participant, including age, height, and weight. In addition, a detailed reproductive history was obtained, including the duration and type of infertility (primary or secondary), previous pregnancies, and prior fertility treatments. Along with these baseline characteristics, the gynecological history included menstrual cycle characteristics, reproductive disorders, and any relevant medical conditions. Following the initial documentation, all participants underwent a comprehensive physical examination complemented by transvaginal ultrasonography. Specifically, the ultrasound assessment evaluated ovarian reserve parameters, including antral follicle count (AFC), while simultaneously screening for any anatomical

abnormalities. As part of this comprehensive evaluation, baseline measurements of ovarian dimensions and endometrial thickness were carefully recorded. In conjunction with the physical assessments, laboratory analyses were performed according to standardized protocols. Male partner semen analysis was conducted in accordance with World Health Organization guidelines (2010), evaluating key parameters including volume, sperm concentration, motility, and morphology.[20] Moreover, on day 2 of the menstrual cycle, baseline hormonal assessments were performed, measuring follicle-stimulating hormone (FSH), luteinizing hormone (LH), estradiol (E2), prolactin, and anti-Müllerian hormone (AMH) levels.

Throughout the subsequent IVF cycle, a detailed monitoring protocol was implemented to track treatment progress. This comprehensive monitoring included documentation of stimulation protocol specifications, total gonadotropin dose administered, duration of ovarian stimulation, number of developing follicles, quantity of oocytes retrieved, fertilization rates, and embryo quality assessment results. Finally, following embryo transfer, pregnancy outcomes were systematically monitored through serial β-hCG measurements and ultrasound confirmation of clinical pregnancy.

Biochemical analysis

Venous blood samples were collected from all participants on day 2 of their menstrual cycle for biochemical analysis. All measurements, including both hormonal parameters (E2, LH, FSH, AMH) and endometrial receptivity markers (MUC1, GDF-9, HLA), were performed using serum samples collected at this single time point. The uniform timing was chosen to standardize measurements across all participants. Blood samples were centrifuged at 1000×g for 15 minutes at room temperature. The separated serum was aliquoted into 1.5 mL Eppendorf tubes and stored at -20°C until analysis. Baseline hormonal assessments, including E2, LH, FSH, and AMH were performed using the MiniVidas automated immunoassay analyzer (Biomerieux, France). Although AMH levels are known to fluctuate during the menstrual cycle, it was measured concurrently with other hormones on day 2 due to its established value as a predictor of ovarian reserve. Moreover, MUC1, GDF-9, and HLA levels were quantified using commercially available enzyme-linked immunosorbent assay (ELISA) kits (Elabscience, USA) according to the manufacturer's instructions. The assay used a quantitative sandwich enzyme immunoassay technique. Briefly, microplates pre-coated with specific monoclonal antibodies against MUC1, GDF-9, and HLA were used. Standards and serum samples were pipetted into the wells, and any present MUC1, GDF-9, and HLA were bound by the immobilized antibodies. After washing to remove unbound substances, enzyme-linked polyclonal antibodies specific to each biomarker were added to the wells. Following a wash to remove unbound antibody-enzyme reagent, a substrate solution was added to develop color proportional to the amount of biomarker bound in the initial step. The color intensity was measured photometrically to determine the concentration of each biomarker.

Calculation of OSI

The Ovarian Sensitivity Index (OSI) was calculated for each participant as a measure of ovarian response to gonadotropin stimulation during the IVF cycle (21). The formula used was:

$$OSI=\left(\frac{\text{No. of retrived oocytes}}{\text{total gonadotropin dose}}\right)\times 1000$$

The number of retrieved oocytes was determined during the oocyte pickup procedure, while the total gonadotropin dose was calculated as the cumulative amount of FSH administered during ovarian stimulation, measured in international units (IU). To analysis, participants were categorized into two groups based on their OSI values: those with OSI <4 and those with OSI ≥4. This cut-off value was determined according to a previous study by Abbara et al.[21]

Statistical analysis

Data analysis was performed using R (version 4.3.2). Descriptive statistics were calculated with categorical variables presented as frequencies and percentages, while continuous variables were expressed as mean±standard deviation after confirming normality using the Shapiro-Wilk test (or as median with interquartile range (IQR) for non-normally distributed data). For two-group comparisons, independent t-tests were used for normally distributed continuous variables, while Mann-Whitney U tests were applied for non-normally distributed data. Categorical variables were compared using Fisher's exact test. Correlations between variables were assessed using Spearman's (or point-biserial) correlation coefficient and visualized through a heatmap. The predictive performance of individual biomarkers was evaluated using ROC curve analysis, with calculation of AUC, sensitivity, specificity, and optimal cutoff values using Youden's index. Binary logistic regression models were constructed to assess the predictive value of variables for pregnancy outcomes, with four sequential models of increasing complexity: Model 1 (OSI alone), Model 2 (OSI plus biomarkers), Model 3 (addition of hormonal parameters), and Model 4 (inclusion of demographic confounders). Model performance was evaluated using Nagelkerke's R^2, AUC, and AIC values. Cases with missing data for any variable in each model were excluded from that specific model using listwise deletion. Statistical significance was set at $p<0.05$ (two-sided) for all analyses.

Results

Comparison of baseline characteristics and clinical outcomes between OSI groups revealed several significant differences (Table 1). Regarding continuous variables, patients with OSI >4 had significantly higher AMH levels compared to those with OSI ≤4 (3.89±3.67 vs 2.28±1.99, $p<0.05$), while other parameters including age (31.76±4.89 vs 32.01±5.12 years, $p>0.05$), BMI (24.44±2.89 vs 24.87±3.11 kg/m^2, $p>0.05$), FSH (5.27±1.77 vs 6.85±5.02, $p>0.05$), E2 (2396.18±1464.85 vs 2213.50±1289.47, $p>0.05$), and total gonadotropin dose (2771.32±1284.95 vs 3102.63±1477.98 IU, $p>0.05$) showed no significant differences between groups.

For categorical variables, the distribution of age groups was comparable between OSI groups (p>0.05), with most patients in both groups aged ≤35 years (79.4% vs 76.3%). BMI categories showed similar distributions (p>0.05), with normal weight being most prevalent in both groups (55.9% vs 52.6%). The type of infertility was predominantly primary in both groups (79.4% vs 73.7%, $p>0.05$). However, significant differences were observed in the cause of infertility ($p<0.05$), with male factor being more prevalent in the OSI >4 group (55.9% vs 21.1%), while female factor cases were exclusively found in the OSI ≤4 group (21.1% vs 0%). Most patients in both groups underwent only one IVF attempt (82.4% vs 89.5%, $p>0.05$). Especially, pregnancy outcomes differed significantly between groups ($p<0.05$), with higher positive β-HCG rates in the OSI >4 group compared to OSI ≤4 (58.3% vs 17.2%)

Table 1. Baseline characteristics and clinical outcomes by ovarian sensitivity index.

Variable	OSI ≤4 (N=38)	OSI >4 (N=34)	*p* value
Age (years)	32.01±5.12	31.76±4.89	0.829
BMI (kg/m^2)	24.87±3.11	24.44±2.89	0.535
AMH	2.28±1.99	3.89±3.67	0.019*
FSH	6.85±5.02	5.27±1.77	0.082
E2	2213.50±1289.47	2396.18±1464.85	0.575
Gonadotropin dose (IU)	3102.63±1477.98	2771.32±1284.95	0.315
Age groups, N (%)			
≤30 years	14 (36.8%)	15 (44.1%)	0.630
31-35 years	15 (39.5%)	12 (35.3%)	
36-40 years	5 (13.2%)	6 (17.6%)	
>40 years	4 (10.5%)	1 (2.9%)	
BMI groups, N (%)			
Normal weight	20 (52.6%)	19 (55.9%)	0.535
Overweight	14 (36.8%)	14 (41.2%)	
Obese	4 (10.5%)	1 (2.9%)	
Infertility type, nN(%)			
Primary	28 (73.7%)	27 (79.4%)	0.593
Secondary	10 (26.3%)	7 (20.6%)	
Cause of infertility, N (%)			
Male factor	8 (21.1%)	19 (55.9%)	0.001*
Female factor	8 (21.1%)	0 (0.0%)	
Combined	16 (42.1%)	9 (26.5%)	
Unexplained	6 (15.8%)	6 (17.6%)	
IVF attempts, N (%)			
One	34 (89.5%)	28 (82.4%)	0.501
More than one	4 (10.5%)	6 (17.6%)	
β-HCG, N (%)			
Negative	24 (82.8%)	10 (41.7%)	0.004*
Positive	5 (17.2%)	14 (58.3%)	

Note: **Statistically significant at p<0.05 (tests were independent sample t test or Mann Whitney's U test, and Fisher's exact test). For continuous variables, data is presented as Mean±SD if it normally distributed. Categorical variables were presented as frequencies and %.*

The significant difference in infertility causes between OSI groups is noteworthy. The higher prevalence of male factor infertility in the OSI >4 group (55.9% vs 21.1%) likely reflects the fact that women with male factor infertility typically have normal ovarian function and thus respond better to gonadotropin stimulation, resulting in higher OSI values. On the contrary, female factor infertility was exclusively found in the OSI ≤4 group, suggesting that underlying female reproductive issues may be associated with reduced ovarian responsiveness.

The analysis of endometrial receptivity biomarkers revealed no significant differences between OSI groups. Figure 1 shows MUC1 levels were comparable between OSI ≤4 and OSI >4 groups (mean±SD: 7.34±2.06 ng/mL vs 7.22±1.88 ng/mL, respectively, p=0.928). Also, GDF-9 concentrations showed no significant difference between groups (3.41±2.90 ng/mL vs 3.05±2.87 ng/mL, p=0.465; Figure 2). However, HLA levels were slightly higher in the OSI >4 group but did not reach statistical significance (19.70±10.17 ng/mL vs 17.85±8.78 ng/mL, p=0.160; Figure 3).

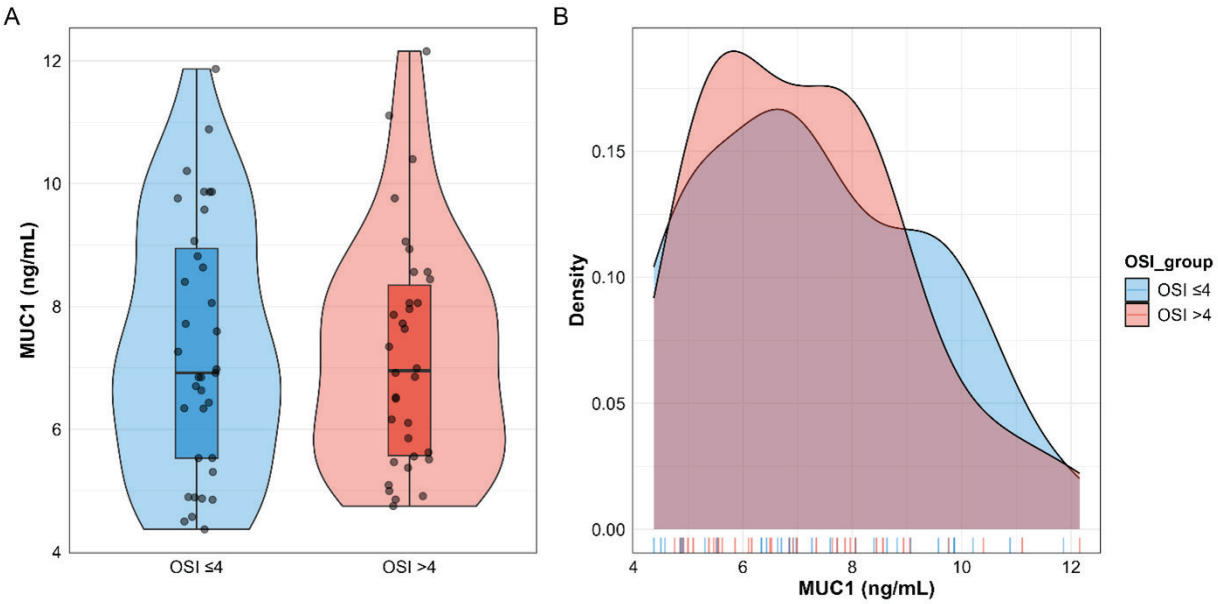


Figure 1. *MUC1 expression levels by OSI groups. A: box and violin plots comparing MUC1 (ng/mL) distribution between OSI ≤4 (N=38) and OSI >4 (N=4) groups, B: density plot showing MUC1 distribution patterns (blue: OSI ≤4; red: OSI >4).*

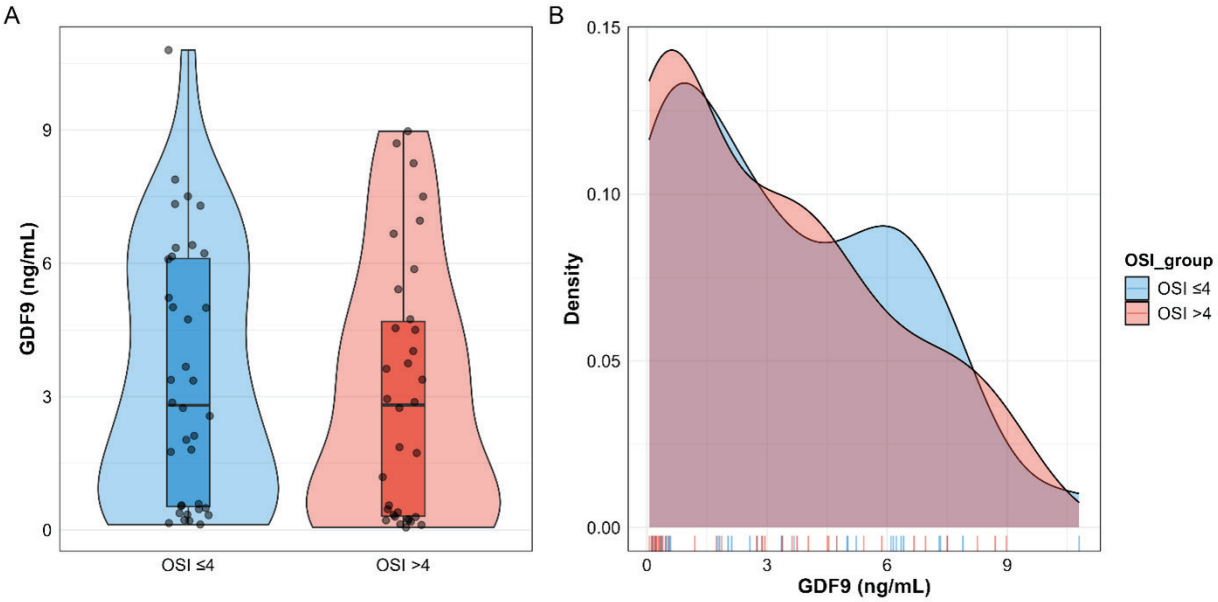


Figure 2. *GDF-9 expression levels by OSI groups. A: box and violin plots comparing GDF-9 (ng/mL) distribution between OSI ≤4 (N=38) and OSI >4 (N=34) groups, B: density plot showing GDF-9 distribution patterns (blue: OSI ≤4;* red: OSI >4).

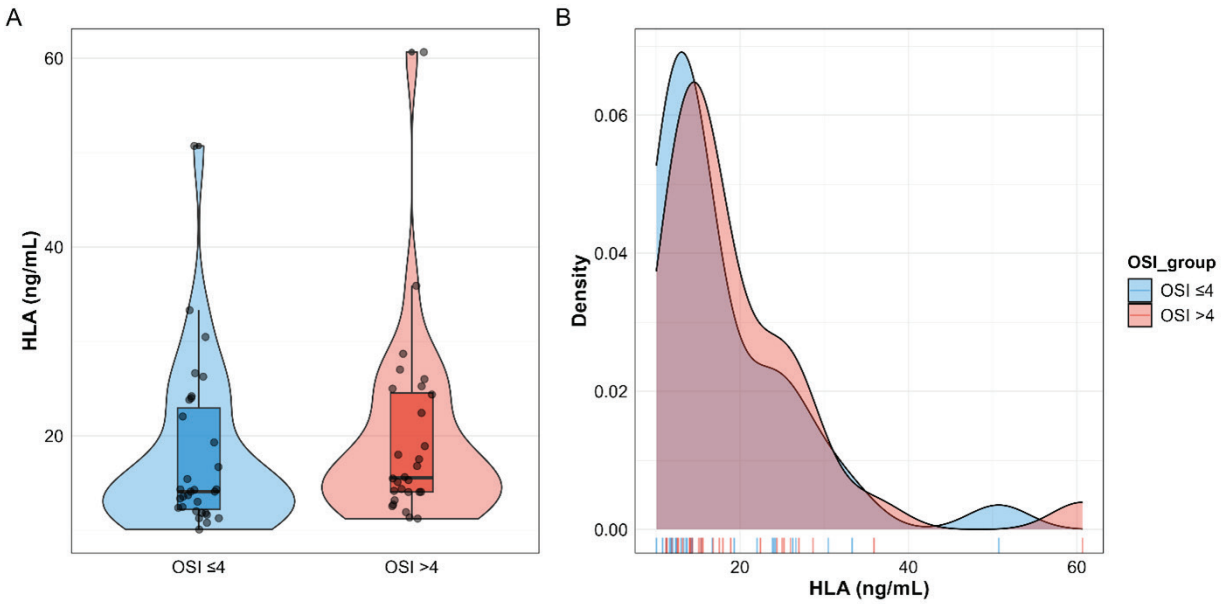


Figure 3. *HLA expression levels by OSI groups. A: box and violin plots comparing HLA (ng/mL) distribution between OSI ≤4 (N=38) and OSI >4 (N=34) groups, B: density plot showing HLA distribution patterns (blue: OSI ≤4; red: OSI >4).*

Correlation analysis shown several significant relationships among the studied variables (Figure 4). Most notably, OSI showed strong positive correlations with AMH (r=0.66, p<0.001), E2 (r=0.49, p<0.001), and β-HCG (r=0.41, p=0.002). AMH had additional significant positive correlations with E2 (r=0.53, p<0.001). A strong positive correlation was observed between HLA and LH (r=0.46, p<0.001), while E2 showed a significant negative correlation with FSH (r=-0.24, p=0.048). The biomarkers MUC1 and GDF-9 did not show significant correlations with other variables, though MUC1 showed a trend toward negative correlation with βHCG (r=-0.24, p=0.090) and positive correlation with E2 (r=0.23, p=0.076). Age showed a negative trend with OSI (r=-0.21, p=0.078) and E2 (r=-0.22, p=0.070), while BMI did not demonstrate any significant correlations with the measured parameters. These findings suggest a complex relationship between ovarian response markers, hormonal parameters, and pregnancy outcomes, with OSI, AMH, and E2 forming a closely related cluster of variables.

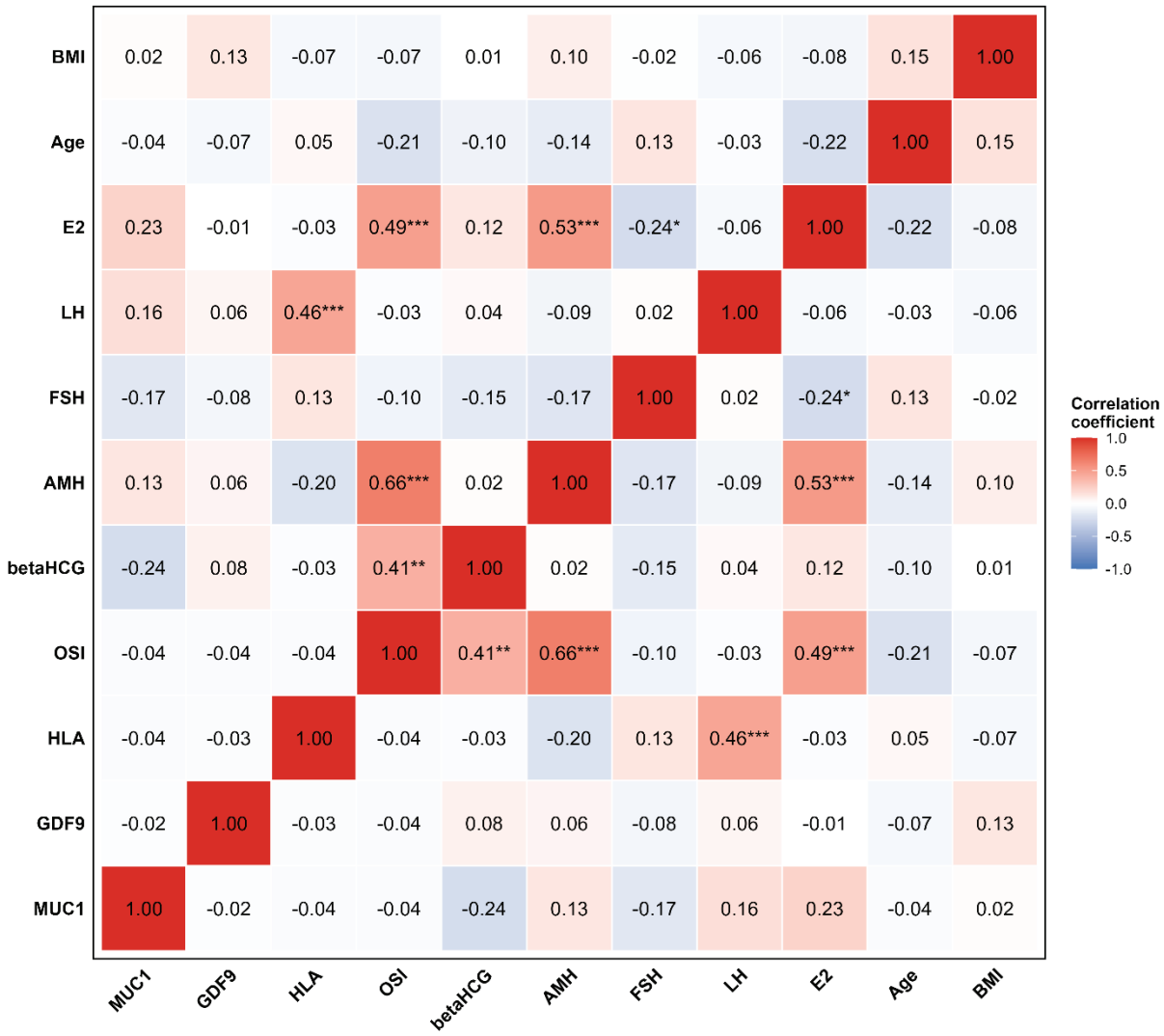


Figure 4. *Correlation heatmap of demographic parameters and biomarkers and IVF outcomes. Color intensity denotes correlation strength (blue: negative; red: positive) and numbers indicate correlation coefficients. Statistical significance: *p<0.05, **p<0.01, ***p<0.001.*

The ROC analysis showed varying predictive capabilities among the assessed parameters for IVF pregnancy outcomes (Figure 5). Firstly, OSI showed the strongest predictive value with the highest AUC of 0.751 (95% CI: 0.613-0.859, *p*=0.001), achieving optimal performance at a cutoff value of >8, yielding a sensitivity of 52.63% (95% CI: 28.9-75.6%) and specificity of 94.12% (95% CI: 80.3-99.3%). Secondly, MUC1 showed moderate discriminative ability with an AUC of 0.625 (95% CI: 0.477-0.758, *p*=0.116), with optimal cutoff at ≤8.819, providing high sensitivity of 94.74% (95% CI: 74.0-99.9%) but lower specificity of 38.71% (95% CI: 21.8-57.8%). Thirdly, HLA exhibited limited predictive capability with an AUC of 0.589 (95% CI: 0.431-0.735, *p*=0.302), showing optimal performance at >14.091 with sensitivity of 75.00% (95% CI: 47.6-92.7%) and specificity of 57.14% (95% CI: 37.2-75.5%). And lastly, GDF-9 demonstrated the lowest predictive value with an AUC of 0.507 (95% CI: 0.363-0.649, *p*=0.943), with optimal cutoff at >3.749, yielding sensitivity of 47.37% (95% CI: 24.4-71.1%) and specificity of 75.00% (95% CI: 56.6-88.5%). Among all parameters, only OSI showed statistically significant discriminative ability (*p*<0.05).

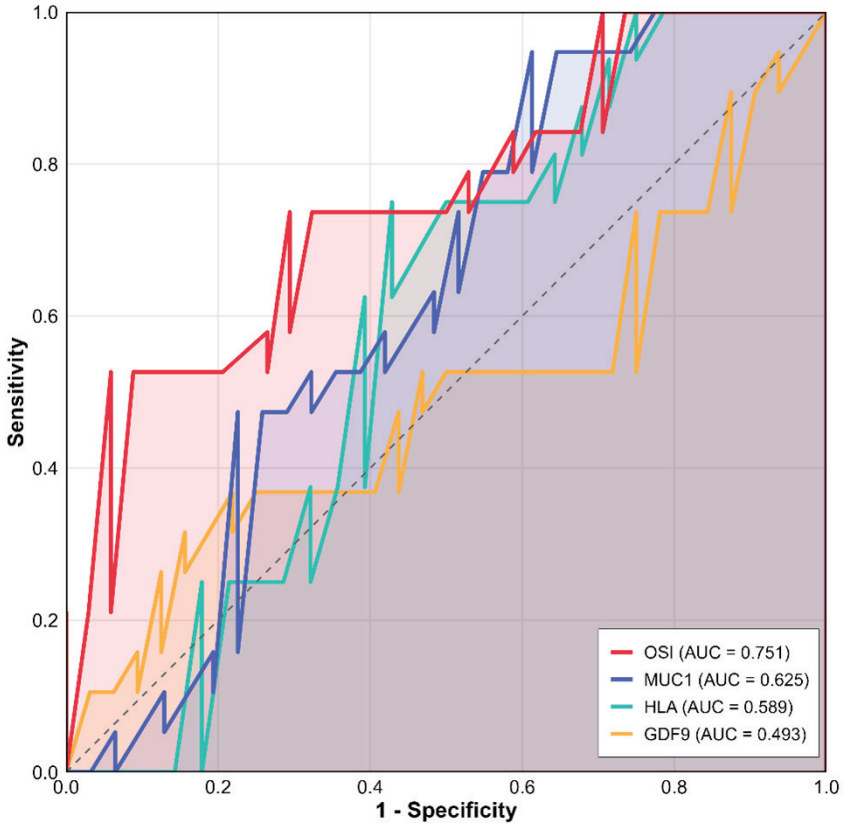


Figure 5. *ROC curves comparing the predictive performance of OSI and biomarkers (MUC1, GDF-9, HLA) for pregnancy outcomes. The diagonal dashed line represents random prediction (AUC= 0.5). Areas under curves are shaded with corresponding colors.*

The logistic regression analysis showed a progressive improvement in predictive performance throughout four increasingly complex models for pregnancy outcomes in IVF treatment (Table 2, Figure 6). The baseline model (model 1, N=53) showed that OSI was a significant predictor of pregnancy outcomes (OR= 1.19, 95% CI: 1.05-1.36, p<0.01), with moderate differential ability (AUC=0.751) and explaining 25.5% of outcome variance (Nagelkerke R^2). The addition of biomarkers in model 2 (N=41) maintained OSI's significant predictive value (OR=1.18, 95% CI: 1.01-1.38, p<0.05) while modestly improving model performance (AUC= 0.782, R^2=0.311), although individual biomarkers; MUC1 (OR= 0.77, 95% CI: 0.52-1.15), GDF-9 (OR= 1.22, 95% CI: 0.91-1.65), and HLA (OR=1.02, 95% CI: 0.91-1.15), did not reach statistical significance. Model 3 (N=37), including hormonal parameters, significantly enhanced predictive performance (AUC=0.867, R^2=0.509), with OSI showing stronger association (OR=1.38, 95% CI: 1.04-1.84, p<0.05). While hormonal parameters including AMH (OR=0.56, 95% CI: 0.25-1.25), FSH (OR=0.68, 95% CI: 0.41-1.11), LH (OR=0.94, 95% CI: 0.61-1.46), and E2 (OR=1.00, 95% CI: 1.00-1.00) did not achieve statistical significance, their inclusion improved overall model fit, as evidenced by reduced AIC (50.465) and residual deviance (32.465). The final model (model 4, N=37) maintained similar differential ability (AUC=0.867, R^2=0.509) with OSI remaining a significant predictor (OR=1.38, 95% CI: 1.01-1.89, p<0.05), while additional confounders age (OR=1.00, 95% CI: 0.83-1.21) and BMI (OR=0.99, 95% CI: 0.77-1.28) showed no significant associations.

The consistent significance of OSI across all models, coupled with considerable improvements in model performance upon addition of hormonal parameters, suggests that integrating these markers provides optimal predictive capability for pregnancy outcomes in IVF treatment. The sample size decreased across regression models due to missing data for certain variables. Model 1 (n=53) included all participants with complete OSI and pregnancy outcome data. Model 2 (N=41) was restricted to participants with complete biomarker measurements (MUC1, GDF-9, HLA). Models 3 and 4 (N=37) included only participants with complete hormonal parameter data (AMH, FSH, LH, E2). We used complete case analysis rather than imputation to maintain data integrity.

Table 2. Multivariate logistic regression analysis of factors associated with pregnancy outcomes.

Variables	Model 1 (N=53)	Model 2 (N=41)	Model 3 (N=37)	Model 4 (N=37)
Primary factor				
OSI	1.19 (1.05-1.36)**	1.18 (1.01-1.38)*	1.38 (1.04-1.84)*	1.38 (1.01-1.89)*
Biomarkers				
MUC1	-	0.77 (0.52-1.15)	0.92 (0.51-1.67)	0.92 (0.50-1.71)
GDF-9	-	1.22 (0.91-1.65)	1.39 (0.92-2.12)	1.40 (0.92-2.12)
HLA	-	1.02 (0.91-1.15)	1.07 (0.90-1.27)	1.07 (0.90-1.27)
Hormonal parameters				
AMH	-	-	0.56 (0.25-1.25)	0.56 (0.25-1.27)
FSH	-	-	0.68 (0.41-1.11)	0.68 (0.41-1.12)
LH	-	-	0.94 (0.61-1.46)	0.94 (0.60-1.46)
E2	-	-	1.00 (1.00-1.00)	1.00 (1.00-1.00)
Confounding factors				
Age	-	-	-	1.00 (0.83-1.21)
BMI	-	-	-	0.99 (0.77-1.28)
Model performance				
Null deviance	69.170	54.846	49.961	49.961
Residual deviance	58.288	44.148	32.465	32.461
R^2	0.255	0.311	0.509	0.509
AUC	0.751	0.782	0.867	0.867
AIC	62.288	54.148	50.465	54.461

Note: *Data are presented as odds ratio (95% confidence interval). Model 1: OSI only, model 2: OSI + biomarkers (MUC1, GDF-9, HLA), model 3: OSI + biomarkers (MUC1, GDF-9, HLA) + hormonal parameters (AMH, FSH, LH, E2), model 4: full model with confounders (age, BMI), *p<0.05, **p<0.01. AUC: area under the curve; R^2: Nagelkerke's R^2, AIC: Akaike information criterion, N: sample size, (-): variable not included in model.*

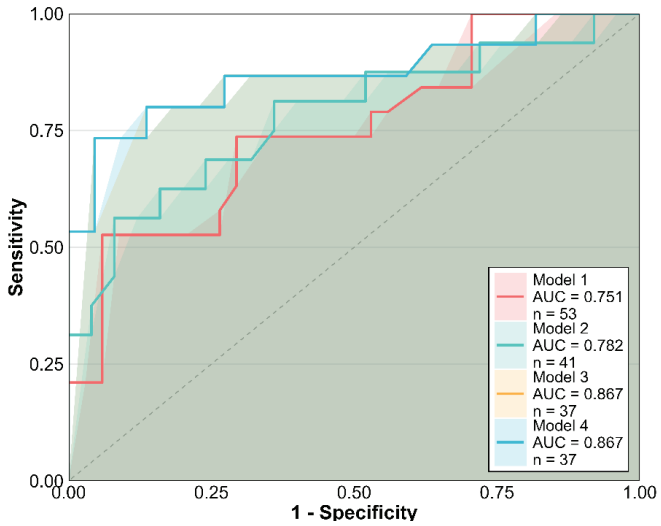


Figure 6. *ROC curves comparing predictive models for IVF outcomes. Blue: Model 1 (OSI only, AUC=0.751), Red: model 2 (add biomarkers MUC1, GDF-9, and HLA (AUC=0.782), Green: model 3 (incorporates hormonal parameters AMH, FSH, and E2, AUC=0.857), Orange: model 4 (include additional confounders age, BMI, and infertility duration, AUC=0.857). Shaded areas under each curve represent the area under the curve (AUC). The diagonal dashed line indicates random prediction (AUC=0.5).*

Discussion

The present study aimed to investigate the relationship between endometrial receptivity biomarkers (MUC1, GDF-9, HLA) and the Ovarian Sensitivity Index (OSI), and to evaluate their ability as predictors of IVF outcomes. Our findings exhibited that OSI was the strongest independent predictor of pregnancy outcomes, with a significant AUC of 0.751 (95% CI: 0.613-0.859) and maintained its predictive value across all logistic regression models (OR=1.19-1.38). While the endometrial biomarkers showed limited individual predictive capacity, with AUCs ranging from 0.507 to 0.625, the combination of hormonal parameters with OSI largely improved the model's predictive performance (AUC=0.867, R^2=0.509). Particularly, correlation analysis showed significant associations between OSI and key hormonal parameters, particularly AMH (r=0.66, p<0.001) and E2 (r=0.49, p<0.001), which suggests a complex relationship between ovarian response markers and IVF success. These results draw attention to the possible value of combining OSI with hormonal parameters for better prediction of IVF outcomes.

Our findings display that OSI is a robust predictor of IVF outcomes, as it shows significant predictive value (AUC=0.751) with an optimal threshold of >8, yielding high specificity (94.12%). This observation aligns with previous studies, particularly Revelli et al. who reported that OSI outperforms total oocyte count in predicting clinical pregnancy.[18] Similarly, Weghofer et al. found OSI to be a significant predictor of live birth rates, [22] while Hu *et al.* displayed its superior sensitivity compared to other ovarian response markers in patients with normal expected response.[19] The strong positive correlations we observed between OSI and key hormonal parameters (AMH: r=0.66; E2: r=0.49) support findings by Li *et al.* regarding OSI's reliability in measuring ovarian responsiveness.[23] Furthermore, our logistic regression analysis revealed that OSI maintained its significant predictive value even after adjusting for multiple confounders (OR=1.38, 95% CI: 1.01-1.89). Together, this suggests its probable clinical utility as an independent predictor. The association between cause of infertility and OSI in our study may support that OSI primarily reflects ovarian function rather than overall fertility potential. Women with isolated male factor infertility and normal ovarian reserve expectedly showed higher OSI values.

In this study, the endometrial receptivity biomarkers individually showed limited predictive value for IVF outcomes. While MUC1 showed moderate discriminative ability (AUC=0.625), both GDF-9 (AUC=0.507) and HLA (AUC=0.589) failed to achieve significant predictive power. These findings align with Chakravarty and Kamal who reported that endometrial MUC1 expression could not reliably predict IVF outcomes.[13] The lack of significant associations in our study might be attributed to several factors. Firstly, the complex temporal dynamics of these markers during the implantation window, as suggested by Baştu et al. Secondly, the possible influence of various confounding factors including hormonal milieu and immune responses; and third, the possibility that local endometrial expression may not be accurately reflected in systemic levels.[24] Our correlation analysis revealed interesting trends, such as MUC1's negative correlation with βHCG (r=-0.24), which partially supports Wu et al.'s findings linking decreased MUC1 expression with implantation outcomes.[25] As well, the strong correlation between HLA and LH (r=0.46) suggests potential immune-endocrine interactions that need further investigation, as previously reported by Hviid *et al.* in their study of HLA-G genotype associations with IVF success.[26]

Our study used biochemical pregnancy (positive β-hCG) as the primary outcome measure rather than clinical pregnancy or live birth. This is also an important limitation, as biochemical pregnancies do not always progress to viable pregnancies. The predictive relationships we observed, particularly for OSI, might differ if clinical pregnancy or live birth outcomes were examined. Biochemical pregnancy losses may occur due to factors unrelated to the biomarkers we measured, possibly affecting our results. Future studies should evaluate these biomarkers in relation to sustained clinical pregnancies and live birth rates to provide more clinically relevant information for patients and clinicians. Moreover, the measurement of endometrial receptivity biomarkers in serum rather than endometrial tissue is a limitation. Local endometrial expression of MUC1, GDF-9, and HLA may not correlate well with systemic serum levels, as these markers may undergo local regulation and metabolism within the endometrium. The poor predictive performance of these serum biomarkers in our study may partly reflect this discordance between local tissue expression and circulating levels. Direct assessment of endometrial tissue biomarkers through biopsy during the window of implantation might provide more accurate predictive information.

Limitations

Overall, and although the current study comes with important findings, several other limitations should be acknowledged. Firstly, the single-center design limits the generalizability of our findings to other populations and clinical settings with potentially different patient characteristics and treatment protocols. Secondly, the relatively small sample size may have limited our statistical power to detect smaller effect sizes and increases the risk of chance findings. Thirdly, as discussed above, we used biochemical pregnancy (positive β-hCG) rather than clinical pregnancy or live birth as our outcome measure, which may not fully reflect the clinical success that matters most to patients. Fourth, biomarker measurements were performed at a single time point (day 2 of menstrual cycle), which may not capture the dynamic changes during the window of implantation.

Finally, we did not assess embryo quality parameters in detail, which are known to significantly influence IVF outcomes. Future multi-center studies with larger sample sizes, serial biomarker measurements, comprehensive confounding variable assessment, and live birth as the primary outcome are needed to validate and extend our findings.

Conclusion

In conclusion, this study supports that OSI is a robust predictor of IVF success, particularly when integrated with hormonal parameters. While the endometrial receptivity biomarkers (MUC1, GDF-9, and HLA) showed limited individual predictive value, their assessment alongside OSI and hormonal markers offers a more comprehensive understanding of factors influencing IVF outcomes. The strong associations between OSI and hormonal parameters, particularly AMH and E2, draw attention to the importance of considering multiple factors in predicting IVF success. These findings together suggest that a combined assessment approach, primarily focusing on OSI and hormonal parameters, may offer more reliable predictive value of pregnancy outcomes in IVF treatment, possibly contributing to better treatment strategies in assisted reproduction.

Funding

None.

Ethical approval

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki. The research protocol was reviewed and approved by the Institutional Review Board of the College of Medicine, University of Kerbala D 6/259 in 15/1/2024. In addition, all participants were well informed about the study's purpose and procedures.

Conflict of interest

The authors declare that they have no competing of interest.

CRediT authorship contribution statement

Zainab A. Aljarah and **Rana M. Hameed**: designed the experiments, writing the manuscript; **Zainab A. Aljarah, Rana M. Hameed,** and **Wasan G. Al Safi**: performed experiments, collected data, writing: first draft. All authors read and approved the final manuscript.

Acknowledgements

None

References

[1] Shingshetty L, Cameron NJ, Mclernon DJ, Bhattacharya S. Predictors of success after in vitro fertilization. Fertil Steril. 2024; 121: 742-51. doi: 10.1016/j.fertnstert.2024.03.003.

[2] Yurci A. The impact of psychological factors on in vitro fertilization (IVF) outcomes: a comprehensive review. J Compr Surg. 2024; 2: 35-8. doi: 10.51271/JOCS-0031

[3] Younglai E V., Holloway AC, Foster WG. Environmental and occupational factors affecting fertility and IVF success. Hum Reprod Update. 2005; 11: 43-57. doi: 10.1093/humupd/dmh055.

[4] Hornstein MD. Lifestyle and IVF outcomes. Reprod Sci. 2016; 23: 1626-9. doi: 10.1177/1933719116667226.

[5] Rienzi L, Cimadomo D, Vaiarelli A, Gennarelli G, Holte J, Livi C, et al. Measuring success in IVF is a complex multidisciplinary task: time for a consensus? Reprod Biomed Online. 2021; 43: 775-8. doi: 10.1016/j.rbmo.2021.08.012.

[6] Sullivan-Pyke CS, Senapati S, Mainigi MA, Barnhart KT. In Vitro fertilization and adverse obstetric and perinatal outcomes. Semin Perinatol. 2017; 41: 345-53. doi: 10.1053/j.semperi.2017.07.001.

[7] López-Luna A, Hernández-Melchor D, Ramírez-Martínez L, López-Bayghen E. The genetic and biochemical blueprint of endometrial receptivity: Past, present, and future factors involved in embryo implantation success. Modern medical genetics and genomics. IntechOpen; 2019.; ID: 7904292.

[8] Achache H, Revel A. Endometrial receptivity markers, the journey to successful embryo implantation. Hum Reprod Update. 2006; 12: 731-46. doi: 10.1093/humupd/dml004.

[9] Miravet-Valenciano JA, Rincon-Bertolin A, Vilella F, Simon C. Understanding and improving endometrial receptivity. Curr Opin Obstet Gynecol. 2015; 27: 187-92. doi:10.1097/GCO.0000000000000173.

[10] Brayman M, Thathiah A, Carson DD. MUC1: A multifunctional cell surface component of reproductive tissue epithelia. Reproductive Biology and Endocrinology. 2004; 2: No.4.

[11] Hreinsson JG, Scott JE, Rasmussen C, Swahn ML, Hsueh AJW, Hovatta O. Growth differentiation factor-9 promotes the growth, development, and survival of human ovarian follicles in organ culture. J Clin Endocrinol Metab. 2002; 87: 316-21. doi:10.1210/jc.87.1.316.

[12] Fountas S, Petinaki E, Bolaris S, Kargakou M, Dafopoulos S, Zikopoulos A, et al. The roles of GDF-9, BMP-15, BMP-4 and EMMPRIN in folliculo-genesis and in vitro fertilization. J Clin Med. 2024; 13: 3775. doi.org/10.3390/cm13133775.

[13] Chakravarty S, Kamal M. MUC1 and E-cadherin immunohistochemistry of endometrium cannot predict the outcome of in vitro fertilization: A case-control study. F1000Res. 2019; 8: 162.

[14] Huang T, Chen F, Zhang Y, Chen S, Long F, Wei J, et al. Decreased GDF9 and BMP15 in follicle fluid and granulosa cells and outcomes of IVF-ET

among young patients with low prognosis. J Assist Reprod Genet. 2023; 40: 567-76. doi: 10.1007/s10815-023-02723-0.

[15] Li Y, Li R-Q, Ou S-B, Zhang N-F, Ren L, Wei L-N, et al. Increased GDF9 and BMP15 mRNA levels in cumulus granulosa cells correlate with oocyte maturation, fertilization, and embryo quality in humans. Reprod Biol Endocrinol. 2014; 12: 81. doi: 10.1186/1477-7827-12-81.

[16] Peng L, Zhao W, Yin T, Xu C, Wang G, Du M. The unique expression pattern of human leukocyte antigen in trophoblasts potentially explains the key mechanism of maternal-fetal tolerance and successful pregnancy. J Reprod Immunol. 2023; 158: 103980. doi:10.1016/j.jri.2023.103980.

[17] Zhuang B, Shang J, Yao Y. HLA-G: An important mediator of maternal-fetal immune-tolerance. Front Immunol. 2021 ; 12: 744324. doi.org/10.3389/fimmu.2021.744324

[18] Revelli A, Gennarelli G, Biasoni V, Chiadò A, Carosso A, Evangelista F, et al. The ovarian sensitivity index (OSI) significantly correlates with ovarian reserve biomarkers, is more predictive of clinical pregnancy than the total number of oocytes and is consistent in consecutive IVF cycles. J Clin Med. 2020; 9: 1914. PMCID: PMC6486569. PMID: 31058052

[19] Hu F, Wang X, Ren H, Lv Y, Li H, Liu S, et al. Ovarian sensitivity index can be used as a more sensitive indicator than follicular output rate to predict IVF/ICSI outcomes in patients of normal expected ovarian response stimulated with GnRH antagonist protocol. Hum Fertil. 2023; 26: 1264-70. doi: 10.1080/14647273.2023.2164869.

[20] World Health Organization. WHO laboratory manual for the examination and processing of human semen. Geneva, Switzerland: WHO Press; 2010.

[21] Biasoni V, Patriarca A, Dalmasso P, Bertagna A, Manieri C, Benedetto C, et al. Ovarian sensitivity index is strongly related to circulating AMH and may be used to predict ovarian response to exogenous gonadotropins in IVF. Reprod Biol Endocrinol. 2011; 9: 112. doi: 10.1186/1477-7827-9-112.

[22] Weghofer A, Barad DH, Darmon SK, Kushnir VA, Albertini DF, Gleicher N. The ovarian sensitivity index is predictive of live birth chances after IVF in infertile patients. Hum Reprod Open. 2020; 2020(4): hoaa049. doi: 10.1093/hropen/hoaa049

[23] Li HWR, Lee VCY, Ho PC, Ng EHY. Ovarian sensitivity index is a better measure of ovarian responsiveness to gonadotrophin stimulation than the number of oocytes during in-vitro fertilization treatment. J Assist Reprod Genet. 2014; 31: 199-203. doi: 10.1007/s10815-013-0144-5.

[24] Bastu E, Mutlu MF, Yasa C, Dural O, Nehir Aytan A, Celik C, et al. Role of Mucin 1 and Glycodelin A in recurrent implantation failure. Fertil Steril. 2015; 103: 1059-1064.e2. doi: 10.1016/j.fertnstert.2015.01.025.

[25] Wu F, Chen X, Liu Y, Liang B, Xu H, Li TC, et al. Decreased MUC1 in endometrium is an independent receptivity marker in recurrent implantation failure during implantation window. Reproductive Biology and Endocrinology. 2018; 16: 60. doi: 10.1186/s12958-018-0379-1.

[26] Hviid TVF, Hylenius S, Lindhard A, Christiansen OB. Association between human leukocyte antigen-G genotype and success of in vitro fertilization and pregnancy outcome. Tissue Antigens. 2004; 64: 66-9. doi: 10.1111/j.1399-0039.2004.00239.x.

Journal of Associated Medical Sciences 2026; 59(2): 208-219

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Examination of the validity and reliability of the Perceive, Recall, Plan, and Perform-long COVID (PRPP-Long COVID) self-report assessment for Thai adults with long COVID

Sutinun Juntorn[1], Peeradech Thichanpiang[1], Pawared Ontong[2], Thitiya Wangkawan[1], Peeraya Munkhetvit[3*]

[1]Division of Occupational Therapy, Faculty of Physical Therapy, Mahidol University, Nakhon Pathom, Thailand.
[2]Department of Community Medical Technology, Faculty of Medical Technology, Mahidol University, Nakhon Pathom, Thailand.
[3]Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 7 November 2025
Accepted as revised 6 February 2026
Available online 17 February 2026

Keywords:
Long COVID, PRPP System of Task Analysis, information processing, cognitive impairment, self-report, performance-based assessment

ABSTRACT

Background: Long COVID harms multiple aspects of occupational performance, affecting daily routines, academic pursuits, and work efficiency due to compromised physical, mental, and cognitive abilities. Notably, patients with long COVID commonly report persistent cognitive impairments.

Objectives: This study aims to adapt and evaluate the validity and reliability of the self-report assessment; the PRPP-Long COVID in Thai adults with long COVID**.**

Materials and methods: The Perceive, Recall, Plan, and Perform-Long COVID (PRPP-Long COVID) was adapted from the perceive, recall, plan, and perform (PRPP) System of Task Analysis: Thai version. This self-report assessment is used to assess the deficits of information processing capacity in individuals recovering from COVID-19. The psychometric properties of the PRPP-Long COVID were investigated, including content validity, internal consistency, and intra-rater reliability. Thirty individuals aged between 21 and 59 years had previously tested positive for COVID-19 by RT-PCR or Antigen Test Kit (ATK) and had recovered at least three months before participating in this study. All of them had an education level of at least secondary school. The PRPP-Long COVID was sent to three experts to examine for the Index of Item Objective Congruence (IOC). After that, the revised version of the PRPP-Long COVID was distributed to the participants. They were given the Google form link submission in 2 rounds, seven days apart, to test intra-rater reliability. For data analysis, descriptive statistics were utilized to calculate demographic characteristics, while the Cronbach alpha coefficient and the Intra-class Correlation Coefficients (ICCs) were used to estimate the internal consistency and intra-rater reliability of the PRPP-Long COVID, respectively.

Results: The IOC values for each question demonstrated a high level of agreement (IOC=0.67-1.00). The internal consistency of the 34 questions in the PRPP-Long COVID indicated a high level of reliability, with α=0.93. The intra-rater reliability of the PRPP-Long COVID was moderate to good (ICC=0.55-0.88).

Conclusion: The PRPP-Long COVID self-rated assessment has been determined to have acceptable validity and reliability.

* **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: peeraya.m@cmu.ac.th
doi: 10.12982/JAMS.2026.052
E-ISSN: 2539-6056

Introduction

Covid-19 is an infectious disease caused by the SARS-CoV-2 coronavirus. Recent data shows many

confirmed COVID-19 cases totaling 770,437,327, including 6,956,900 deaths. Therefore, the spread of SARS-CoV-2 has significantly impacted the world.[1] Although there are many changes in virus strains, vaccination can reduce the severity of COVID-19 symptoms. However, several studies have reported that a large number of people have persistent abnormal symptoms after COVID-19 infection, a condition known as "Long COVID".[2] Long COVID refers to a condition that can affect individuals who have previously been infected with the SARS-Co V-2 virus, have been treated, and have generally recovered for at least three months but still have persistent abnormal symptoms such as fatigue, shortness of breath, and cognitive dysfunction. These symptoms may persist for over two months and significantly interfere with daily life. The prevalence of Long COVID is more common among females than males, and individuals aged 35-49 have a higher likelihood of experiencing Long COVID.[3]

Long COVID-19 has demonstrated a multifaceted impact on various aspects of occupational performance, including activities of daily living, academic, and work efficiency, due to impairments in multiple domains such as physical, mental, and information processing.[4] Patients with Long COVID report prolonged effects on their cognitive abilities. A previous study indicated that individuals who recovered from COVID-19 at 3 and 6 months demonstrated lower attention and information processing speed scores than healthy people.[5] Another study showed that problematic information processing, verbal fluency, and executive functions are usually found during 1-3 months after experiencing COVID-19.[6,7] Moreover, after seven months, many patients show impaired memory, perception, and cognition. They are still unable to function at their prior capacities and persist with an overwhelming number of symptoms. These problems affect work performance [8] and reduce overall quality of life.

An essential role of occupational therapists in health promotion and prevention is providing services to people with long-term COVID-19 and evaluating information processing skills. This evaluation aims to plan rehabilitation programs to improve information processing abilities. In doing so, developing self-assessment tools for information processing deficits is essential for individuals with long COVID who may have problems with information processing.

At present, there are many assessment tools for evaluating cognitive strategies deficits which developed by occupational therapists such as, the assessment of motor and process skills (AMPS),[9] the executive function performance test (EFPT),[10] the cognitive performance test (CPT),[11] and the Perceive, Recall, Plan, and Perform (PRPP) system of task analysis.[12] PRPP system of task analysis is a client-centered assessment tool for individuals with cognitive difficulties. It identifies deficits in information processing strategies during task performance and guides targeted intervention.[13] The process begins with task analysis and focuses on problematic components of performance. Its strengths include adaptability to clients' preferences, systematic task breakdown, and clear identification of functional skills and deficits. Consequently, the PRPP system of task analysis is valuable for detecting performance changes and informing treatment planning.

In Thailand, the PRPP system of task analysis had been translated into a Thai-language known as the "PRPP System of Task Analysis: Thai version"[14] and then it was evaluated in various groups including patients with stroke,[15] schizophrenias,[16] and children with learning disabilities.[17] The PRPP system of task analysis was originally developed as a performance-based assessment and later adapted into questionnaire formats, such as the PRPP@SCHOOL-1 (TQ & PQ), for parent or teacher assessment of children with learning disabilities.[18] However, when relevant behaviors are not directly observable, performance-based evaluations may be limited. Self-report assessments provide direct insight into individuals' cognitive processes that may not be evident through observation alone.[19] At present, there remains a lack of a self-report assessment based on the PRPP system of task analysis for adults to determine deficits in information processing. Utilizing self-reported assessment is essential for detecting information processing difficulties among people with Long COVID, as it is highly prevalent and affects populations worldwide. The major advantage of self-report assessment is its ability to efficiently gather a vast amount of quantitative data from a large sample size without requiring significant time or financial investment. Consequently, self-report assessments are widely employed as screening instruments in epidemiological research. Moreover, information provided directly by individuals is often more accurate than proxy reports. As a result, certain problems may not be readily observable through behavior alone, yet individuals may perceive subtle changes or unusual experiences that are not immediately apparent to others.[20]

For this reason, the self-report assessment is suitable for identifying the information processing strategies dysfunction in individuals with Long COVID. Based on the points discussed earlier, the objective of this study was to adapt and evaluate the validity and reliability of the self-report assessment; the PRPP-Long COVID in Thai adults with long COVID.

Materials and methods

Study design and participants

The population includes individuals who have previously tested positive for COVID-19 by RT-PCR or Antigen Test Kit (ATK) and had recovered after at least three months. The study focuses on Thai people aged 18 to 59 years who live in Thailand and have an education level of at least secondary school. The sample consisted of 35 individuals who met the above criteria. Five individuals were in the pilot use phase, while 30 were in the internal consistency and intra-rater

reliability examination phases. The inclusion criteria were individuals who previously tested positive for COVID-19 by real-time polymerase chain reaction (RT-PCR) or Antigen Test Kit (ATK) and had recovered for at least three months, age 18 to 59 years, and possessed the ability to read the Thai language. As for exclusion criteria, these were individuals who were diagnosed with central nervous system disorders that result in impaired cognition and poor comprehension, such as cerebrovascular disease, dementia, or traumatic brain injury. Furthermore, this includes those who were diagnosed with psychiatric disorders such as mood disorders, psychotic disorders, or other mental illnesses, and were diagnosed with a substance use disorder that affects the nervous system.

Instrument

The Perceive, Recall, Plan and Perform-Long COVID (PRPP-Long COVID) was adapted from the Perceive, Recall, Plan, and Perform (PRPP) System of Task Analysis: Thai version,[14] a criterion-referenced, performance-based, client-centered assessment designed to assess the effectiveness of information processing strategies patients used in their daily activities. It consists of two stage of analysis. Stage One utilizes standard behavioral task analysis to determine the individual's mastery in specific occupations, and Stage Two Analysis is founded on a preliminary model of human behavior information processing established by Romiszowski.[12] This is necessary to explain the cognitive processes that underlie task performance and the cognitive strategies employed in complex situations. The PRPP system of task analysis comprises four quadrants - perceive, recall, plan, and perform (Figure 1) - each subdivided into three sub-quadrants and 34 information processing strategies ("descriptors"). Observers rate performance using a three-point Likert scale to identify abilities and deficits based on effective task execution.[12] The assessment employs standardized terminology, evaluation procedures, and scoring, with established validity and reliability across diverse populations,[14,21-27] including individuals with traumatic brain injury,[14,23,26] schizophrenia,[21] early psychosis,[27] HIV-associated neurocognitive disorder,[28] and learning disabilities.[22] The Thai version of the PRPP System of Task Analysis was translated from the original instrument and demonstrated high test-retest and inter-rater reliability in Thai patients with acquired brain injury,[14] and has since been applied in patients with stroke,[15] schizophrenia,[16] and children with learning disabilities.[17]

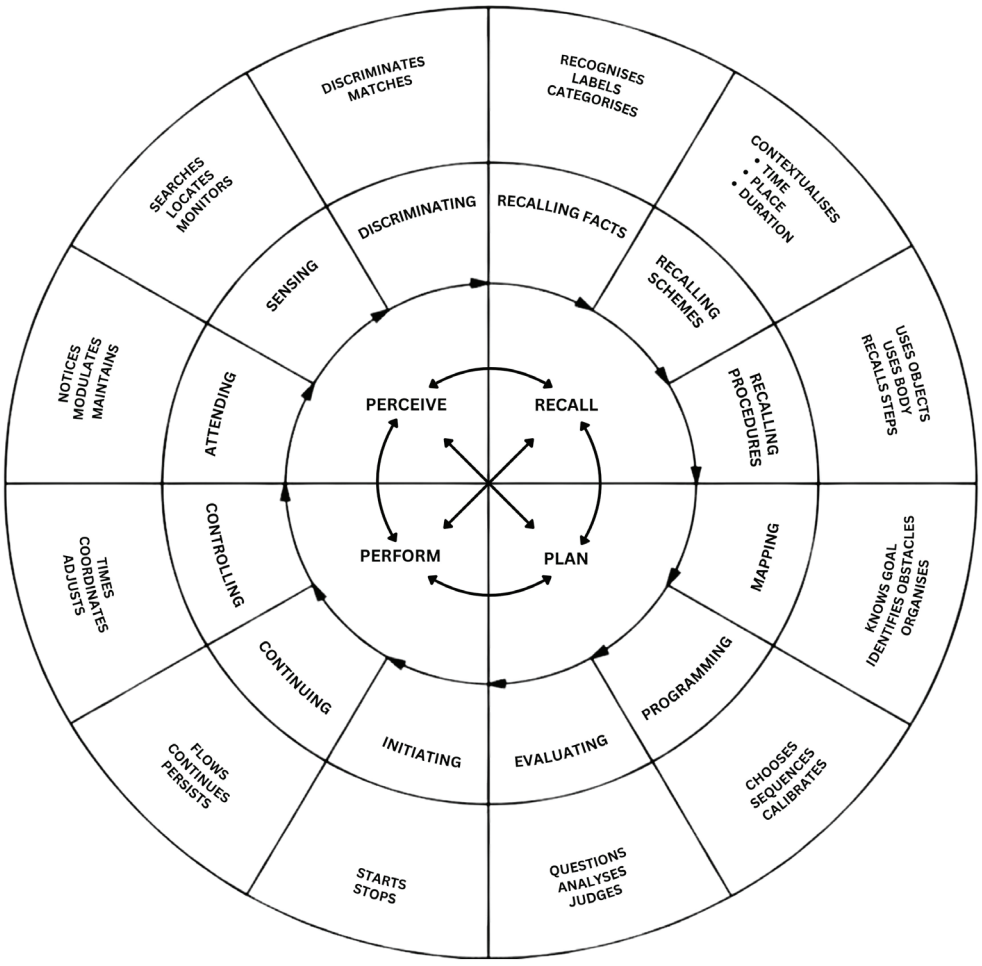


Figure 1*. The PRPP system of task analysis (Stage Two) conceptual model.*[12]

Phase 1: Modification and content validity

Modification of Perceive, Recall, Plan and Perform-long COVID (PRPP-Long COVID)

The modification entailed changing the wording to conform to the self-rating structure based on the PRPP System of Task Analysis's content, as reviewed by two experts who are lecturers in the field of occupational therapy. Additionally, the PRPP-Long COVID utilizes a Likert scale to assess the frequency of behavioral problems, which is rated on a scale of 0 to 4 (0 = never, 1 = rarely, 2 = sometimes, 3 = often, 4 = very often). The Likert scale is a reliable and valid method for quantifying abstract thoughts, feelings, and actions.[29] The frequency of problematic behavior was chosen by the researcher as the measuring scale to represent the degree of difficulty. A higher total score indicates a more self-perceived information processing dysfunction.

Content validity and pilot use

The researcher evaluated the content validity of the PRPP-Long COVID through the Index of Item Objective Congruence (IOC) method with the agreement of three experts who possess extensive experience in utilizing the PRPP system of task analysis. These experts considered it unambiguous and consistent with the intended objectives of the assessment. After that, 5 participants were asked to rate themselves using the PRPP-Long COVID to evaluate comprehension and clarity in the questions. Consequently, the final version of the PRPP-Long COVID was obtained.

Phase 2: Examination of internal consistency and Intra-rater reliability

Participants

Thirty eligible individuals with long COVID participated in the internal consistency and intra-rater reliability phases. The sample size was calculated using the R statistical package to estimate the intra-class correlation coefficient (ICC) with a 95% confidence interval width (CIW) of 0.05 and an estimate of 0.85. [30]

Internal consistency

The final version of the PRPP-Long COVID was used to assess 30 participants via the Google form channel. The scores of the self-rating assessment were determined by the internal consistency using Cronbach's alpha coefficient statistic.

Intra-rater reliability

Thirty participants were assessed by the PRPP-Long COVID twice via a Google form. The time interval for rating scores between the first and second times was seven days to reason the effect of recall bias.[31] The scores from the first and second assessments were analyzed to determine intra-rater reliability.

Statistical analysis

SPSS for Windows (version 17.0; SPSS Inc, Chicago, IL) was used for all statistical analyses. Descriptive statistics were used to calculate demographic characteristics, and the Cronbach alpha coefficient was used to estimate the internal consistency. Intra-class correlation coefficients (ICCs) using the ICC (3,1) model (two-way mixed-effects, single measurement) were utilized to determine the intra-rater reliability of the PRPP-Long COVID.

Ethical considerations

This research project has been sanctioned by the Human Research Ethics Committee of Mahidol University, as evidenced by document number COE No.MU-CIRB 2023/035.2002, granted in February 2023. The researcher ensured the protection of the rights of the participants, and the data collected from the sample were kept confidential. They were not utilized to identify individual respondents specifically, as all answers were incorporated into the overall research analysis.

Results

Content validity and pilot use

IOC values for each question demonstrate a high level of agreement (IOC=0.67-1.00). This indicates that every question in the self-assessment tool for cognitive impairment in adults with long COVID aligns well with the intended objectives of the evaluation. The results are shown in Table 1. Moreover, the pilot use process showed that the sample group demonstrated high satisfaction with the clarity and comprehensibility of the questions and the scoring criteria.

Table 1. Index of Item Objective Congruence (IOC) of the PRPP-Long COVID (N=30).

PRPP-Long COVID	Item	IOC	PRPP-Long COVID	Item	IOC
Perceive quadrant			*Recall quadrant*		
Attending sub-quadrant			Recalling Facts sub-quadrant		
Notices	Item 1	1	Recognizes	Item 9	1
Modulates	Item 2	1*	Labels	Item 10	1
Maintains	Item 3	1	Categories	Item 11	1
Sensing sub-quadrant			Recalling schemes sub-quadrant		
Searches	Item 4	1	Contextualizes to time	Item 12	1
Locates	Item 5	0.67	Contextualizes to place	Item 13	1
Monitors	Item 6	1*	Contextualizes to duration	Item 14	1
Discriminating sub-quadrant			Recalling procedures sub-quadrant		
Discrimination	Item 7		Uses objects	Item 15	0.67
Matches	Item 8		Uses body	Item 16	1
			Recalls steps	Item 17	1*
Plan quadrant			*Perform quadrant*		
Mapping sub-quadrant			Initiating sub-quadrant		
Knows goal	Item 18	1	Starts	Item 27	0.67
Identifies obstacles	Item 19	1	Stops	Item 28	1
Organizes	Item 20	0.67			
Programming sub-quadrant			Continuing sub-quadrant		
Chooses	Item 21	1	Flows	Item 29	1
Sequences	Item 22	0.67	Continues	Item 30	1
Calibrates	Item 23	1	Persists	Item 31	1
Evaluating sub-quadrant			Controlling sub-quadrant		
Questions	Item 24	1*	Times	Item 32	0.67
Analyzes	Item 25	0.67	Coordinates	Item 33	1*
Judges	Item 26	1*	Adjusts	Item 34	1

Note: *The IOC value has been adjusted according to experts' recommendations.

The characteristics of the participants are presented in Table 2. The participants were 30 individuals who were infected with COVID-19. Of these, 7 were male (23.30%) and 23 were female (76.70%). The mean age of the sample was 29.30 (SD=9.85, range=21-59). The mean period after infection with COVID-19 was 3.30 months (SD=0.54, range=3-5). Most participants held a bachelor's degree (80.00%).

Table 2. Demographic characteristics (N=30).

Characteristics	Participants	Percentage
Sex		
Female	23	76.70
Male	7	23.30
Education		
Secondary Education	2	6.70
Bachelor's Degree	24	80.00
Master's Degree	4	13.30
	Mean (SD)	**Range**
Age (years)	29.30 (9.85)	21-59
The period after being infected with COVID-19 (months)	3.30 (0.54)	3-5

The internal consistency of the 34 items of the PRPP-Long COVID was evaluated using Cronbach's alpha. The results indicated a high level of reliability, with α=0.93. Furthermore, Cronbach's alpha values for each quadrant (Perceive, Recall, Plan, and Perform) were found to be at a moderate to high level of reliability,[32] ranging from 0.70 to 0.86. These findings are presented in Table 3.

Table 3. Internal consistencies of the PRPP-Long COVID (N=30).

Internal consistencies	Cronbach's alpha
Perceive	0.70
Recall	0.78
Plan	0.86
Perform	0.80
All Items	**0.93**

The intra-rater reliability of the PRPP-Long COVID was determined using the intraclass correlation coefficients (ICC), and the results indicate moderate to good reliability for quadrant (ICC=0.75-0.86), moderate to good reliability for sub-quadrant (ICC=0.63-0.84), and moderate to good reliability for items (ICC=0.55-0.88). Specifically, items 27 which are flow descriptors showed the highest reliability (ICC=0.88), while items 1, 24 exhibited showed the lowest reliability (ICC=0.55). Overall, the findings suggest that the PRPP-Long COVID has acceptable intra-rater reliability as shown in Table 4.

Table 4. Intra-rater reliability of the PRPP-Long COVID (N=30).

ITEMS	ICC ICC (2,1) (95% CI)	95% CI	ITEMS	ICC ICC (2,1) (95% CI)	95% CI
Perceive quadrant	0.86	0.70-0.93	***Recall quadrant***	0.82	0.62-0.92
Attending sub-quadrant	0.84	0.66-0.92	Recalling facts sub-quadrant	0.81	0.61-0.91
Item 1	0.55	0.06-0.79	Item 9	0.64	0.24-0.83
Item 2	0.73	0.44-0.87	Item 10	0.87	0.73-0.94
Item 3	0.86	0.70-0.93	Item 11	0.79	0.55-0.90
Sensing sub-quadrant	0.77	0.52-0.90	Recalling schemes sub-quadrant	0.76	0.49-0.89
Item 4	0.77	0.50-0.89	Item 12	0.68	0.32-0.85
Item 5	0.81	0.61-0.91	Item 13	0.60	0.15-0.81
Item 6	0.58	0.14-0.80	Item 14	0.65	0.26-0.83
Discriminating sub-quadrant	0.81	0.59-0.91	Recalling procedures sub-quadrant	0.83	0.65-0.92
Item 7	0.66	0.29-0.84	Item1 5	0.62	0.18-0.82
Item 8	0.85	0.69-0.93	Item 16	0.75	0.48-0.88
			Item 17	0.71	0.39-0.86

Table 4. Intra-rater reliability of the PRPP-Long COVID (N=30) (continued).

ITEMS	ICC ICC (2,1) (95% CI)	95% CI	ITEMS	ICC ICC (2,1) (95% CI)	95% CI
Plan quadrant	0.77	0.52-0.89	***Perform quadrant***	0.75	0.47-0.88
Mapping sub-quadrant	0.80	0.58-0.91	Initiating sub-quadrant	0.83	0.65-0.92
Item 18	0.64	0.23-0.83	Item 27	0.88	0.74-0.94
Item 19	0.75	0.47-0.88	Item 28	0.75	0.47-0.88
Item 20	0.82	0.62-0.91	Continuing sub-quadrant	0.76	0.50-0.88
Programming sub-quadrant	0.75	0.48-0.88	Item 29	0.68	0.33-0.85
Item 21	0.70	0.37-0.86	Item 30	0.73	0.42-0.87
Item 22	0.66	0.29-0.83	Item 31	0.67	0.30-0.84
Item 23	0.57	0.09-0.80	Controlling sub-quadrant	0.63	0.22-0.83
Evaluating sub-quadrant	0.67	0.33-0.84	Item 32	0.58	0.10-0.80
Item 24	0.55	0.08-0.78	Item 33	0.64	0.23-0.83
Item 25	0.62	0.22-0.82	Item 34	0.66	0.29-0.84
Item 26	0.68	0.34-0.85			

Discussion

In this study, we focused on examining the psychometric properties of the PRPP-Long COVID, which was modified from the perceive, recall, plan, and perform (PRPP) system of task analysis: Thai version. The results revealed that the PRPP-Long COVID had acceptable validity and reliability. As for content validity, the index of item objective congruence (IOC) from 3 experts showed that all items had a high value of IOC (Table 1). This indicates that the PRPP-Long COVID can identify various cognitive strategic problems in individuals with Long COVID and also reveals difficulties regarding principles of information processing theory and the fundamental strategy application model developed for the PRPP System of Task Analysis.[12] This study placed special emphasis on identifying the information processing strategies needed for effectively carrying out daily activities in adulthood, particularly with regard to education or work. A previous study of the effects of Long COVID on cognitive function found that long COVID influenced memory loss, especially short-term memory. Additionally, it can lead to difficulties in attention, overall thinking, executive functioning, and problem-solving. Furthermore, it can affect the speed of thought, cause confusion, and even cause agnosia. In adulthood, there are levels of symptoms ranging from mild to severe in equal proportions. This condition has the greatest impact on work function. In addition, it affects other lifestyle activities that require advanced and complex processing, including making serious decisions, communicating thoughts, having conversations with others, and driving.[8]

In terms of internal consistency investigation, the Cronbach's alpha coefficients among the 34 items in each quadrant were high (Table 3). This finding aligns with previous studies, which examined the internal consistency of the PRPP System of Task Analysis. For example, a study by Van Keulen-Rouweler *et al*.[33] assessed the ability to perform occupation and information processing skills in elderly Parkinson's patients or students with learning difficulties using the methods of Lowe.[18] This study demonstrated a high level of internal consistency in the PRPP assessment. Therefore, the results confirm the homogeneity of items in the PRPP quadrants and sub-quadrants that involve the problematic informational processing strategies in people with Long COVID.

For intra-rater reliability examination, the results revealed a significantly moderate to high correlation (See Table 4). Several factors have contributed to the stability of the PRPP-Long COVID. Firstly, there is the clarity and simplicity in the item wording, eliminating any potential confusion for the respondents. Secondly, the contents of PRPP-Long COVID were created based on symptoms commonly associated with impaired information processing in individuals who had COVID-19, drawing from a thorough analysis of past research.[8] As a result, all inquiries about items are related to the level of difficulty in information processing during daily activities in adulthood. These are supported by Streiner *et al*.[34] who argue that the reliability of the questionnaire or self-rating assessment depends on the interface between the instrument, the specific group of respondents, and the situation at hand. The PRPP-Long COVID had a high degree of interface with respondents (individuals who had Long COVID) and the situation (strategies used to perform complex and advanced skilled activities in adulthood).[34] Therefore, this contributed to acceptable reliability in the participants in this study. Finally, the intra-rater interval of 7 days was deemed suitable, as many researchers have selected a range of 2 days to 2 weeks.

It is generally accepted that this duration provides an adequate compromise between unexpected clinical change and recall bias.[35]

These findings align with the study by Aubin *et al*. which employed the PRPP assessment to examine the information processing deficits experienced by schizophrenia patients during both simple and complex cooking tasks.[36] They assessed the intra-rater reliability of the PRPP assessment by evaluating the participants' performance on the task. The results revealed that the total score of the PRPP assessment demonstrated a high degree of intra-rater reliability. However, the intra-rater reliability of each quadrant of the assessment showed a moderate value. Likewise, in the study by Lindenschot *et al*., the intra-rater reliability of the PRPP assessment was investigated in children with mitochondrial disease, revealing a high reliability.[37] Furthermore, the result of the present study also utilized the PRPP Questionnaire, which applied the PRPP@SCHOOL-1 to children with learning difficulties.[18] The contents of the PRPP@SCHOOL-1 are consistent with the problematic information processing identified in children with learning disabilities when they participate in naturalistic contexts, including home and school. In this study, the investigators examined the test-retest reliability of the PRPP@SCHOOL-1 parent questionnaire version. Therefore, we confirm that the PRPP-Long COVID is reliable and consistent for application for adult Thais with Long COVID. In clinical practice, it is crucial to recognize that when information processing deficits are identified, the information obtained can be utilized to receive occupational therapy rehabilitation services that involve analyzing and synthesizing activities aimed at restoring cognitive skills that are appropriate and consistent with the individual's information processing impairments, with the ultimate goal of regaining the ability to perform practical daily activities. In addition, the PRPP-Long COVID can be modified and utilized as a multidisciplinary professional guide for evaluation and intervention planning.

Limitations

This study has some limitations. The first is the distribution of the participants. In this study, most participants were female in early adulthood. Differences in biological mechanisms between males and females, as well as variations in health perception and symptom reporting, may have influenced the study outcomes. Future studies should involve participants with a more balanced proportion of genders and ages. The second limitation is the period after recovery from COVID-19 infection. Previous research has found that Long COVID can occur up to 6-9 months after recovery from COVID-19 infection. However, in the present study, the period after recovery from COVID-19 was found to be within 3-5 months. Therefore, future studies should involve participants who cover the period after recovery from coronavirus infection at 3-9 months to gain a more comprehensive understanding of information processing problems in people with Long COVID. Lastly, the lack of information regarding the symptoms and severity of COVID-19 infections, as well as the use of medication and treatment modalities, may significantly impact the degree of information processing disorder. Therefore, future studies should consider these factors more thoroughly to better understand the relationship between COVID-19 and information processing disorders.

Conclusion

The findings from this study demonstrated that the Perceive, Recall, Plan, and Perform (PRPP) in adults with Long COVID (PRPP-Long COVID) has acceptable psychometric properties, including content validity, internal consistency, and intra-rater reliability. This suggests that the PRPP-Long COVID is appropriate for evaluating information processing impairment in individuals with Long-COVID in the Thai context. If the self-report assessment reveals problematic information processing, it can help individuals recognize their abilities and facilitate access to specialized treatment. In addition, the assessment can serve as an outcome measure for planning occupational therapy interventions aimed at improving occupational performance.

Future direction and clinical implementation

Future research should further examine the PRPP-Long COVID across larger and more diverse populations who experience information processing impairments. Longitudinal studies could explore its predictive value for functional recovery and quality of life among individuals with Long COVID. Clinically, the PRPP-Long COVID can be integrated into occupational therapy practice as both a screening and outcome measure for cognitive information-processing deficits. Its use can guide personalized intervention planning, monitor progress, and enhance communication among multidisciplinary rehabilitation teams in managing cognitive sequelae of Long COVID.

Ethical approval

The study adhered to the ethical standards of the Declaration of Helsinki and complied with all relevant institutional and national guidelines for human research. Ethical approval was obtained from the Human Research Ethics Committee of Mahidol University (COE No.MU-CIRB 2023/035.2002). All participants were informed of the study's objectives and procedures, provided written consent prior to participation, and were free to withdraw at any time without any consequences.

Funding

This research project was supported by Faculty of Physical Therapy, Mahidol University.

Conflict of interest

The authors declare no conflicts of interest related to the research, authorship, or publication of this article.

CRediT authorship contribution statement

Sutinun Juntorn: conceptualization, methodology, data curation, formal analysis, writing: original draft, supervision, project administration; **Peeraya Munkhetvit:** conceptualization, methodology, data curation, formal analysis, writing: original draft, supervision, project administration, correspondence; **Peeradech Thichanpiang:** writing: review and edit; **Thitiya Wangkawan:** data curation; **Pawared Ontong:** data curation.

Acknowledgements

This research project was supported by Faculty of Physical Therapy, Mahidol University. We are truly grateful for the research participants for their time and effort. We also would like to thank our occupational therapy students including Ms. Jongrak Aungsuwirun, Ms. Khwandao Khamlao and Ms. Passorn Bundhukul for their kind help.

References

[1] World Health Organization. WHO Coronavirus (COVID-19) Dasboard. 2023. Available from: https://dashboards-dev.sprinklr.com.

[2] Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV; WHO Clinical Case Definition Working Group on Post-COVID-19 Condition. A clinical case definition of post-COVID-19 condition by a Delphi consensus. Lancet Infect Dis. 2022; 22(4): e102-e7. doi: 10.1016/S1473-3099(21)00703-9.

[3] Crook H, Raza S, Nowell J, Young M, Edison P. Long covid-mechanisms, risk factors, and management. BMJ. 2021; 374: n1648. doi: 10.1136/bmj.n1648.

[4] Taquet M, Dercon Q, Luciano S, Geddes JR, Husain M, Harrison PJ. Incidence, co-occurrence, and evolution of long-COVID features: A 6-month retrospective cohort study of 273,618 survivors of COVID-19. PLoS Med. 2021; 18: e1003773. doi: 10.1371/journal.pmed.1003773.

[5] Poletti S, Palladini M, Mazza MG, De Lorenzo R, COVID-19 BioB Outpatient Clinic Study group, Furlan R, *et al*. Long-term consequences of COVID-19 on cognitive functioning up to 6 months after discharge: role of depression and impact on quality of life. Eur Arch Psychiatry Clin Neurosci. 2022; 272: 773-82. doi: 10.1007/s00406-021-01346-9.

[6] Almeria M, Cejudo JC, Sotoca J, Deus J, Krupinski J. Cognitive profile following COVID-19 infection: Clinical predictors leading to neuropsychological impairment. Brain Behav Immun Health. 2020; 9: 100163. doi:10.1016/j.bbih.2020.100163.

[7] Méndez R, Balanzá-Martínez V, Luperdi SC, Estrada I, Latorre A, González-Jiménez P,*et al*. Short-term neuropsychiatric outcomes and quality of life in COVID-19 survivors. J Intern Med.2021; 290: 621-31. doi: 10.1111/joim.13262.

[8] Davis HE, Assaf GS, McCorkell L, Wei H, Low RJ, Re'em Y, *et al*. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. ECM. 2021; 38: 101019. doi: 10.1016/j.eclinm.2021.101019.

[9] Fisher AG, Jones KB. Assessment of motor and process skills: Development, standardization, and administration manual. United States: Three Star Press; 2012 .

[10] Baum CM, Connor LT, Morrison T, Hahn M, Dromerick AW, Edwards DF. Reliability, validity, and clinical utility of the executive function performance test: a measure of executive function in a sample of people with stroke. Am J Occup Ther. 2008; 62: 446-55. doi: 10.5014/ajot.62.4.446.

[11] Burns T, Mortimer JA, Merchak P. Cognitive Performance Test: a new approach to functional assessment in Alzheimer's disease. J Geriatr Psychiatry Neurol.1994; 7: 46-54. doi: 10.1177/089198879400700109

[12] Chapparo C, Ranka J. The PRPP system of task analysis: User's training manual: Available from the Discipline of Occupational Therap. Faculty of Health Sciences, University of Sydney, Lidcombe, Australia. 2005

[13] Nott MT, Chapparo C, Heard R. Effective occupational therapy intervention with adults demonstrating agitation during post-traumatic amnesia. Brain Inj.2008; 22: 669-83. doi: 10.1080/02699050802227170.

[14] Munkhetvit P. Cognitive assessment in patients with acquired brain injury in Thailand [Unpublished thesis]. University of Sydney, NSW, Australia. 2005

[15] Munkhetvit P. Types of cognitive deficits during the performance of ADL and IADL in persons with CVA. JOTAT. 2008; 11: 19-30.

[16] Bunyachatakul S, Munkhetvit P, Srikamjak T, Sarakam A. Cognitive deficits during the performance of activities of daily living in Schizophrenic patients. Bulletin of Suan Prung. 2010; 26: 15-24.

[17] Juntorn S, Sriphetcharawut S, Munkhetvit P. Effectiveness of information processing strategy training on academic task performance in children with learning disabilities: A pilot study. Occup Ther Int. 2017: 6237689.doi:10.1155/2017/6237689.

[18] Lowe S. Cognitive strategies and school participation for students with learning difficulties [Dissertation on the Internet]. University of Sydney, NSW, Australia; 2010. Available from: https://ses.library.usyd.edu.au/handle/2123/7170.

[19] Montag C, Dagum P, Hall B, Elhai J. Do we still need psychological self-report questionnaires in the age of the Internet of Things? Discov Psychol. 2021; 2(1). doi: 10.1007/s44202-021-00012-4.

[20] Demetriou A, Kazi S. Self-awareness in g (with processing efficiency and reasoning). Intell. 2006;

34: 297-317. doi: 10.1016/j.intell.2005.10.002.

[21] Aubin G, Stip E, Gelinas I, Rainville C, Chapparo C. Assessment of information-processing skills during daily task performance in persons with schizophrenia: usefulness of the Perceive, Recall, Plan, and Perform (PRPP) System of Task Analysis. In Proceedings of the 15th World Conference of the Federation of Occupational Therapists. 2010.

[22] Boland K. Assessing information processing in the context of pretend play. University of Sydney, NSW, Australia. 2004

[23] Chapparo C,Ranka J. Development of an information processing task analysis to identify errors in self-maintenance performance of brain injured clients: Final Category B Grants Research Report. 1992.

[24] Fordham M. Perceive, Recall, Plan, Perform (PRPP) rating scales (parent and teacher): reliability and validity [Unpublished thesis]. University of Sydney, NSW, Australia. 2001.

[25] Lohri J. Information and sensory processing abilities of young children with autism spectrum disorder [Unpublished thesis]. University of Sydney, NSW, Australia. 2005.

[26] Nott MT, Chapparo C. Measuring information processing in a client with extreme agitation following traumatic brain injury using the Perceive, Recall, Plan and Perform system of task analysis. Aust Occup Ther J. 2008; 55: 188-98. doi: 10.1111/j.1440-1630.2007.00685.x.

[27] Still M, Beltran R, Catts S, Chapparo C. Cognitive and functional assessment of patients with early psychosis. Acta Neurol Scand Suppl. 2002; 413: 40-41. Published abstracts from 3rd International Conference on Early Psychosis, Copenhagen, Denmark, September.

[28] Ranka JL, Chapparo CJ. Assessment of productivity performance in men with HIV associated neuro-cognitive disorder (HAND). Work. 2010; 36: 193-206. doi: 10.3233/wor-2010-1020 .

[29] Joshi A, Kale S, Chandel S,Pal D. Likert Scale: Explored and explained. Br J Appl Sci Technol. 2015; 7: 396-403. doi: 10.9734/BJAST/2015/14975.

[30] Wolak ME, Fairbairn DJ, Paulsen YR. Guidelines for estimating repeatability. Methods Ecol Evol. 2012; 3: 129-37. doi: 10.1111/j.2041-210X.2011.00125.x.

[31] Marinus J, Visser M, Stiggelbout AM, Rabey JM, Martínez-Martín P, Bonuccelli U, *et al.* A short scale for the assessment of motor impairments and disabilities in Parkinson's disease: the SPES/SCOPA. J Neurol Neurosurg Psychiatry.2004; 75: 388-95. doi: 10.1136/jnnp.2003.017509.

[32] Ahdika A. Improvement of quality, interest, critical, and analytical thinking ability of students through the application of research-based learning (RBL) in introduction to stochastic processes subject. Int Elect J Math Ed.2017; 12(2): 167-91. doi: 10.29333/iejme/608.

[33] Van Keulen-Rouweler BJ, Sturkenboom IH, Kottorp A, Graff MJ, Nijhuis-Van der Sanden MW, Steultjens EM. The Perceive, Recall, Plan and Perform (PRPP) system for persons with Parkinson's disease: a psychometric study. Scand J Occup Ther. 2017; 24: 65-73. doi: 10.1080/11038128.2016.1233291.

[34] Streiner DL, Norman GR, Cairney J. Health measurement scales: A practical guide to their development and use. New York: Oxford University Press; 2014.

[35] Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. comparison of two time intervals for test-retest reliability of health status instruments. J Clin Epidemiol. 2003; 56(8): 730-35. doi: 10.1016/s0895-4356(03)00084-2.

[36] Aubin G, Chapparo C, Gélinas I, Stip E, Rainville C. Use of the Perceive, Recall, Plan and Perform System of Task Analysis for persons with schizophrenia: a preliminary study. Aust Occup Ther J. 2009; 56: 189-99. doi: 10.1111/j.1440-1630.2007. 00725.x.

[37] Lindenschot M, Groot IJM, Nijhuis-van der Sanden MWG, Steultjens EMJ, Graff MJL, Koene S. Perceive, Recall, Plan and Perform (PRPP)-assessment based on parent-provided videos of children with mitochondrial disorder: Action design research on implementation challenges. Phys Occup Ther Pediatr. 2023; 43: 74-92. doi: 10.1080/01942638.2022.2086024.

Appendix

PRPP-Long COVID	Item
Perceive quadrant	
Notices	1. Do you neglect incoming stimuli and/or react more slowly than usual to incoming stimuli, for example, by not noticing lights flashing or reacting more slowly to different warning sounds?
Modulates	2. Do you have difficulty to shift your attention from one point to another?
Maintains	3. Do you have difficulty maintaining interest in completing activities?
Searches	4. Do you have difficulty finding things by sight, hearing, smell, or touch?
Locates	5. Do you have difficulty telling the location of objects or equipment that are needed during activities?
Monitors	6. Have you neglected the changes in stimuli that occur during activities, e.g., not noticing a kettle boiling?
Discrimination	7. Do you find it difficult to distinguish between similar objects, e.g., a 1 Baht coin and a 2 Baht coin?
Matches	8. Do you find it difficult to match objects, e.g., socks, chopsticks, spoons, and forks?
Recall quadrant	
Recognizes	9. Do you have difficulty remembering familiar sounds, smells, feelings, signs, and names of people or places?
Labels	10. Do you find naming familiar people, objects, and places difficult?
Categories	11. Do you find it difficult to group objects with the same function, e.g., kitchen utensils, stationery, or things in the same group, such as flowers, animals, etc.?
Contextualizes to time	12. Do you often forget when you must do that activity?
Contextualizes to place	13. Do you often forget where you must do certain activities?
Contextualizes to duration	14. Do you often spend time doing activities too quickly, causing errors, or too slowly, causing the work to be finished beyond schedule?
Uses Objects	15. Do you often make mistakes when using complex tools, e.g., electronic devices? Electrical appliances or programs in the computer?
Uses Body	16. Do you feel like you have clumsy body movements or difficulty using your body parts to work effectively, e.g., typing on the keyboard more slowly or having poor picking up things?
Recalls Steps	17. Do you not remember the steps in doing activities as well as before?

PRPP-Long COVID	Item
Plan quadrant	
Knows Goal	18. Do you feel that you have no goals in your activities or are neglecting your work goals?
Identifies Obstacles	19. Do you often ignore or are you unaware that there are obstacles to your activities?
Organizes	20. Do you have difficulty organizing things to suit the space and make it convenient for activities?
Chooses	21. Do you have difficulty choosing parts of the body or objects and equipment that are appropriate for your activities?
Sequences	22. Do you find it difficult to perform activities correctly according to the steps and continuously change actions without interruption?
Calibrates	23. Do you find it difficult to adapt, increase, or decrease the speed or force of your movements to the situation?
Questions	24. Do you neglect or not ask questions when problems arise during activities?
Analyzes	25. Do you have difficulty finding, identifying problems, and analyzing solutions to deal with obstacles while doing activities?
Judges	26. Do you have difficulty deciding on a plan or method, including using substitute materials and equipment to solve immediate problems, or have you decided to choose an inappropriate solution?
Perform quadrant	
Starts	27. Do you slow to start unfamiliar activities or hesitant to start them?
Stops	28. Can you not stop doing the activity even when you have completed all the steps, or regularly finish an activity even though it has not yet been completed?
Flows	29. Do you often perform activities inconsistently or stumble intermittently in each step?
Continues	30. Do you have difficulty completing activities consistently within the given time frame?
Persists	31. Do you often give up when obstacles arise?
Times	32. Do you find it difficult to adjust the speed of your movements to suit the activity?
Coordinates	33. Do you often have problems with movement and a lack of precision in activities that require precision, such as writing, sports, or music?
Adjusts	34. Do you have difficulty adjusting your movements to suit the activity?

Journal of Associated Medical Sciences 2026; 59(2): 220-225

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Effect of acceptance and commitment therapy on psychological well-being and quality of life among individuals living with HIV

Punitha P and Santhiya M*

Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India.

ARTICLE INFO

Article history:
Received 1 October 2025
Accepted as revised 6 February 2026
Available online 18 February 2026

Keywords:
Acceptance and commitment therapy, psychological well- being, quality of life, occupational therapy, Individuals living with HIV

ABSTRACT

Background: Psychiatric issues are notably common among individuals with HIV, heavily impacting the overall quality of life and the success of their treatment. Acceptance and Commitment Therapy (ACT) is a recent cognitive-behavioural therapy that puts great emphasis on psychological flexibility. However, while the model has been extensively studied in clinical settings, research on its application for psychological well-being and quality of life among individuals with HIV remains limited.

Objectives: This study aimed to evaluate the effect of acceptance and commitment therapy on psychological well-being and quality of life among individuals living with HIV.

Materials and methods: A total of 55 individuals with HIV met the inclusion criteria, from which 33 individuals with HIV provided consent to participate in the study. Before the randomization of the groups, 3 participants were dropped out of the study due to relocation from their living facility. The remaining 30 individuals with HIV were selected and assigned to either the experimental group (N=15) or the control group (N=15) through simple randomization using the lottery method in which participants retrieved sequentially numbered opaque sealed envelopes. The quality of life and psychological well-being were assessed using the WHOQOL HIV-BREF and Ryff's psychological well-being scale. The experimental group received Acceptance and Commitment Therapy, while the control group underwent conventional occupational therapy for three months. Data analysis was performed using SPSS version 23.0, with paired and independent t-tests applied.

Results: Post-test scores showed statistically significant improvements in the experimental group than in the control group in both quality of life (p<0.001) and psychological well-being (p<0.001). The results of this study showed that Acceptance and Commitment Therapy was effective in enhancing both quality of life and psychological well-being among individuals with HIV.

Conclusion: The findings indicate that Acceptance and Commitment Therapy effectively enhanced quality of life and psychological well-being of Individuals living with HIV positive following intervention in experimental group.

Introduction

Human immunodeficiency virus (HIV) is a long-term, life-threatening infection that targets the body's immune system by infecting CD4+T cells, which provide immune protection.[1] Psychiatric issues are notably common among individuals with HIV, heavily impacting the overall quality of life and the success of their treatment. Evidence has shown that individuals

* **Corresponding contributor.**
Author's Address: *Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India.*
E-mail address: santhiyam1311@gmail.com
doi: 10.12982/JAMS.2026.053
E-ISSN: 2539-6056

with HIV are at higher risk of developing mental illness, including depression, anxiety, and substance abuse disorders.[2] The chronic nature of HIV, in addition to its stigma, frequently associated with individuals living with HIV, could lead to social isolation and psychological distress.[3] This contributes to the emergence of the risk of poor quality of life and psychological well-being.

The World Health Organization (WHO) defines quality of life as, "an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns". Psychological well-being encompasses a sense of a person's emotional resilience, contentment with life, and coping with stress.[4]

The various interventions currently being used to improve the psychological well-being and quality of life among individuals with HIV, such as cognitive behavioural techniques (CBT), life skills training, social skills training, group therapy, mindfulness and relaxation techniques which focus on restructuring or challenging distorted thoughts, stress reduction, skill training, social and vocational support.[5-7] Despite the above-mentioned interventions, acceptance and commitment therapy (ACT), goes beyond symptom reduction and focuses on psychological flexibility. It teaches clients to accept distressing thoughts without trying to change them and align their action with values, which is especially helpful in chronic conditions, where emotional distress is ongoing.[8]

ACT is a recent cognitive-behavioral therapy that puts great emphasis on psychological flexibility. ACT places greater emphasis on the ability to tolerate unpleasant negative experiences-such as sorrow, suffering, and emotional pain-than on suppressing or confronting emotions. ACT aims to promote psychological flexibility by targeting six core competencies: acceptance, defusion, contact with the present moment, self-as-context, values, and committed action. Learning to view ideas as words or mental events rather than as absolute truths or directives that must be obeyed is known as cognitive defusion. Acceptance entails voluntarily embracing challenging feelings, ideas, and physical experiences without attempting to suppress, alter, or manage them. Present moment awareness focuses on being open and curious while fully interacting with the present moment. Self-as-context is the capacity of perspective-taking that enables people to see their experiences without letting them define them.[9-13]

ACT places a strong emphasis on defining one's personal values, or the characteristics of behavior and being that one desires to maintain. Setting objectives and acting in a way that is consistent with one's values while accepting difficulties and setbacks as a necessary part of the process are all components of committed action.[14] ACT is particularly effective in improving psychological well-being in people with chronic illness. By encouraging mindfulness and living in accordance with values, ACT allows people to cope with distress, decrease avoidance behavior, and build resilience to stigma and social pressure. Also, ACT has been shown to improve quality of life by allowing people to engage in valued activities, improve social support, and build a greater sense of purpose in life.[15]

These studies commonly lack objective measures, and adequate sample sizes, and do not assess outcomes like psychological well-being or quality of life together for individuals with HIV positive. This gap underscores the need for empirical evidence to validate the effectiveness of acceptance and commitment therapy in individuals with HIV. This study aims to examine the effectiveness of Acceptance and Commitment Therapy in improving psychological well-being and quality of life.

Materials and methods

Study design

The study employed a randomized controlled trial (RCT) to evaluate the effect of ACT on quality of life and psychological well-being among individuals with HIV. Participants were randomly assigned to experimental or control groups using the lottery method. A single blinding method was used, in which the participants were unaware of their group allocation but the therapist is aware of the group allocation.

Participants

Fifty-five samples were collected from Shelter Trust, a trust for the homeless in Tamil Nadu. Based on the inclusion and exclusion criteria, medically-diagnosed individuals with HIV-positive irrespective of gender, aged 20-40 years with HIV in early and preventable stages and without severe cognitive impairments or psychiatric disorders were identified as eligible. Using simple randomization through the lottery method in which participants retrieved sequentially numbered opaque sealed envelopes , 30 participants were selected and randomly assigned to either the experimental group (N=15) or the control group (N=15). Single blinding was implemented, wherein the participants were unaware of their group allocation.

Instruments

Ryff's psychological well-being scale

The Ryff Psychological Well-Being Scale (PWB) was developed by Carol Ryff in 1989. The 42-item version of the PWB offers insight into individuals' subjective psychological functioning. Each subscale comprises 7 items rated on a Likert scale, typically from 1 (strongly disagree) to 6 (strongly agree), with higher scores indicating greater well-being in that domain. Interpretation involves examining subscale scores individually or computing a total well-being score. Lower scores may reflect diminished functioning or psychological distress in specific areas of life. Reliability of the 42-item version is well-documented, with internal consistency coefficients (Cronbach's alpha) typically

ranging from 0.70 to 0.90 across subscales. Construct validity has been confirmed through correlations with other measures of life satisfaction, depression, and self-esteem.

WHOQOL-HIV BREF

The WHOQOL-HIV BREF is a 31-item instrument developed by the World Health Organization to assess quality of life among individuals living with HIV/AIDS. Items are rated on a 5-point Likert scale, with higher scores reflecting better perceived quality of life. Domain scores are typically transformed to a 4-20 scale for interpretation. The interpretation focuses on identifying strengths and areas of concern in each life domain. It is particularly useful in both clinical assessment and intervention planning for people living with HIV. The WHOQOL-HIV BREF demonstrates excellent reliability, with Cronbach's alpha coefficients generally above 0.70 for all domains. It has strong content, construct, and cross-cultural validity, having been validated in diverse international samples.[16]

Procedure

The experimental group (15 participants) received Acceptance and Commitment Therapy, while the control group (15 participants) underwent conventional occupational therapy. ACT aims to promote psychological flexibility by targeting six core competencies: acceptance, defusion, contact with the present moment, self-as-context, values, and committed action. Intervention includes "leaves on a stream" metaphor, "passengers on a bus" metaphor, silly voices, stigma shield, which was given with proper instruction to the individuals with HIV The final stage involves assessing the effectiveness of the intervention. The control group received conventional occupational therapy, incorporating Medication Adherence Training, Physical Activity and Exercise Therapy, Nutritional Education Workshops, Sleep Hygiene Training, Energy Conservation Training. The study duration was 3 months, three days per week, 30 sessions, and 45 minutes per session which were equally provided for the control group and the experimental group. Following that, standardized assessment procedures and validated outcome measures (WHOQOL HIV -BREF and Ryff's psychological well-being scale) were used for both the control group and the experimental group to reduce assessment bias. outcome assessments were conducted by trained independent assessor who was blinded to group allocation (Figure 1).

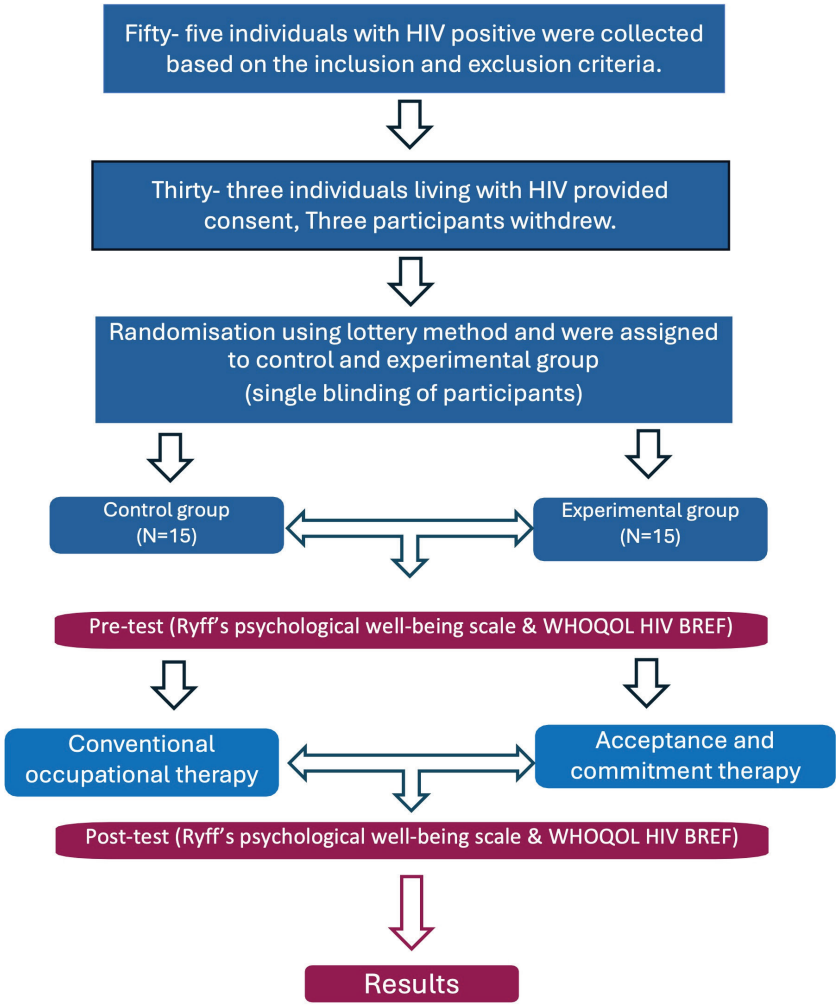


Figure 1. Consort diagram of the study.

Data analysis

Descriptive statistics (means, standard deviations, frequencies, and percentages) were used to summarize demographic and baseline data. Data were analysed according to a per-protocol approach, as no missing data or dropouts were observed during the study period Normality of the data was checked by using the Shapiro Wilk test. As the data were normally distributed, inferential statistics appropriate for a randomized control trial design is applied. Participants were randomly allocated to experimental and control groups using the lottery method. Paired T-tests were used to compare pre- and post-intervention scores within groups, and independent t-tests were used to assess differences between groups. A 95% confidence interval was used in this study to estimate the precision of the observed effects. All statistical analyses were conducted using a significance level of $p<0.05$; results with $p<0.05$ were considered statistically significant, and those with $p<0.01$ were considered highly significant. The statistical analysis was done with the help of SPSS version 23.

Results

Table 1 shows the baseline comparison between the control and experimental groups on demographic variables. All participants were individuals living with HIV. The mean age and gender distribution were comparable between the groups, with no statistically significant differences ($p>0.05$), indicating that both groups were homogeneous at the start of the study. hence, the baseline demographics were similar for control group and experimental group.

Table 1. Baseline comparison of groups on demographics.

Serial No.	Demographic variables	Class	Number of respondents in group (%)	
			Control	Experimental
1	Age (years)	20-25	6 (40.0)	6 (40.0)
		26-30	3 (20.0)	3 (20.0)
		31-35	2 (13.3)	4 (26.7)
		36-40	4 (26.7)	2 (13.3)
2	Gender	Female	10 (66.7)	10 (66.7)
		Male	5 (33.3)	5 (33.3)

Table 2 presents the statistical analysis of the Comparison of QOL and psychological well-being scores between pre-test and post-tests in control group. In the control group, the mean QoL score increased slightly from 7.26 to 9.51 (t=-27.11, $p=0.017$) and the mean psychological well-being score also increased from 19.38 to 22.75 (t=-22.18, $p=0.0014$), indicating a statistically significant improvement following conventional occupational therapy. These findings were supported by a study done by Ruby Aikat *et al*. explored how occupational therapy affects the quality of life (QOL) for individuals living with HIV, their understanding of the condition, and the knowledge possessed by occupational therapists.[17]

Table 2. Comparison of QOL and psychological Well -being scores between pre and post-tests in control group (Paired t test).

Serial No.	Variable		N	Mean	SD	T value	95% CI	Df	*p* value
1	QOL	Pre-test	15	7.26	0.548	-2.711	(-4.022)-(-0.469)	14	0.017*
		Post-test	15	9.51	3.079				
2	Well-being	Pre-test	15	19.38	1.055	-2.814	(-5.94)-(-0.802)	14	0.014*
		Post-test	15	22.75	4.422				

Note: ***significant at 1% level, *significant at 5% level.*

Table 3 presents statistical analysis of comparison of QOL and psychological Well-being scores between pre-test and post-tests in experimental group. The experimental group, which received Acceptance and Commitment Therapy alongside conventional OT, demonstrating a more pronounced improvement in QOL scores from 7.42 to 17.89 (t=-24.010, $p<0.001$) and psychological well-being scores from 19.76 to 47.94 (t=-22.902, $p<0.001$), indicated a highly significant change. This aligns with previous findings on ACT's adaptability and effectiveness across diverse populations.

Table 3. Comparison of QOL and psychological Well -being scores between pre and post-tests in experimental group (Paired t test).

Serial No.	Variable		N	Mean	SD	T value	95% CI	Df	*p* value
1	QOL	Pre-test	15	7.42	0.571	-24.010	(-11.4) - (-9.53)	14	0.000*
		Post-test	15	17.89	1.344				
2	Well-being	Pre-test	15	19.76	0.906	-22.902	(-30-818) - (-25.54)	14	0.000*
		Post-test	15	47.94	4.655				

Note: ***significant at 1% level, *significant at 5% level.*

Table 4 displays the statistical analysis of comparison of post-test level QOL and Psychological well-being scores between control and experimental groups. A statistically significant difference was observed for both QoL (t=-9.241, *p*=0.017) and psychological well-being (t=-15.195, *p*=0.014), with the experimental group performing better.

Table 4. Comparison of post-test level QOL and well-being scores between control and experimental groups (Independent t test).

Serial No.	Variable		N	Mean	SD	T value	95% CI	Df	*p* value
1	QOL	Control	15	9.87	3.079	-9.241	(-9.794) - (-6.24)	28	0.000*
		Experimental	15	17.89	1.344				
2	Well-being	Control	15	22.75	4.422	-15.195	(-28.585) - (-21.794)	28	0.000*

Note: ***significant at 1% level, *significant at 5% level.*

Discussion

ACT places greater emphasis on the ability to tolerate unpleasant negative experiences-such as sorrow, suffering, and emotional pain-than on suppressing or confronting emotions. These strategies may have improved their mental health, self-confidence, and ability to bounce back from challenges, leading to better QoL and psychological well-being.

This is supported by prior studies, Zahra Eskafi-Sabet *et al*. evaluated the effectiveness of acceptance and commitment group therapy on the severity of the disease, quality of life, and mental well-being in irritable bowel syndrome (IBS)patients.[18] Ghodsieh Ebrahimpour *et al*. gauged the impact of ACT on the psychological well-being, quality of life, and levels of depression in individuals with epilepsy.[19]

Although these studies involved different populations, they emphasize ACT's capacity to foster psychological flexibility. These findings suggest that while conventional OT is beneficial, the Acceptance and Commitment Therapy had a greater impact on psychological well-being and quality of life.

This study found that Acceptance and Commitment Therapy significantly improved both quality of life and psychological well-being among individuals with HIV with greater gains than conventional interventions. These findings highlight ACT's role in enhancing psychological flexibility, which supports better psychological well-being and overall quality of life. Future research could explore its long-term impact, adaptability across settings, and use with other palliative conditions.

Limitations

This study only targeted people living with HIV between the ages of 20 and 40 years. This reduces the generalizability of the study results to people who are younger and older. Additionally, the small sample size reduces the power of the study and the ability to generalize the results to the larger population of people living with HIV. Additionally, both outcome variables are measured using self-administered questionnaires that may be prone to response bias, social desirability bias, and individual subjective interpretation of their psychological well-being.

Conclusion

This study demonstrated that integrating the Acceptance and Commitment Therapy into occupational therapy enhanced psychological well-being and quality of life among individuals living with HIV. The positive outcomes in psychological well-being and quality of life, in the experimental group highlight ACT's effectiveness, reinforcing its value in improving psychological flexibility.

Ethical approval

This study was approved by the institution scientific review board (ISRB) of Saveetha College of Occupational Therapy with reference no. of SCOT/ ISRB/014/2025.

Funding

This study was conducted without any financial

support from any funding agency in the public, commercial or not-for-profit funding agencies.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Punitha P: conceptualization, methodology, supervision, writing: review and edit; **Santhiya M:** investigation, data collection, formal analysis, writing: original draft, review and edit.

Acknowledgements

I sincerely thank Saveetha Institute of Medical and Technical Sciences, Saveetha college of Occupational Therapy, Dr. M. Arun Kumar, MOT, PhD, for his continuous encouragement.

References

[1] Remien RH, Stirratt MJ, Nguyen N, Robbins RN, Pala AN, Mellin CA. Mental health and HIV/AIDS: the need for an integrated response. AIDS. 2019; 33(9): 1411-20. doi:10.1097/QAD.0000000000002227.

[2] Hu FH, Liu P, Jia YJ, Ge MW, Shen LT, Xia XP, et al. Prevalence of mental health problems in people living with HIV: a systematic review and meta-analysis. Psychol Health Med. 2025; 30(3): 397-413. doi:10.1080/13548506.2024.2424998.

[3] Nedelcovych MT, Manning AA, Semenova S, Gamaldo C, Haughey NJ, Slusher BS. The psychiatric impact of HIV. ACS Chem Neurosci. 2017; 8(7): 1432-4. doi:10.1021/acschemneuro.7b00169.

[4] Ryff CD, Keyes CL. The structure of psychological well-being revisited. J Pers Soc Psychol. 1995; 69(4): 719-27. doi:10.1037/0022-3514.69.4.719.

[5] Fathi M, Abdolmaleki L, Alamdari SM, Kamal SHM. Effectiveness of cognitive behavioural therapy for quality of life among women with HIV. Iran J Health Psychol. 2023; 6(15): 9-18. doi:10.30473/ijohp.2023.63309.1240.

[6] Gonzalez-Garcia M, Ferrer MJ, Borras X, Muñoz-Moreno JA, Miranda C, Puig J, et al. Effectiveness of mindfulness-based cognitive therapy on the quality of life, emotional status, and CD4 cell count of patients aging with HIV infection. AIDS Behav. 2014; 18(4): 676-85. doi:10.1007/s10461-013-0612-z.

[7] van der Heijden I, Abrahams N, Sinclair D. Psychosocial group interventions to improve psychological well-being in adults living with HIV. Cochrane Database Syst Rev. 2017; 3(3): CD010806. doi: 10.1002/14651858.CD010806.pub2.

[8] Feliu-Soler A, Montesinos F, Gutiérrez-Martínez O, Scott W, McCracken LM, Luciano JV. Current status of acceptance and commitment therapy for chronic pain: a narrative review. J Pain Res. 2018; 11: 2145-59. doi:10.2147/JPR.S144631.

[9] Rider JV, Laverdure AE. Exploring the application of acceptance and commitment therapy by occupational therapy practitioners: a scoping review protocol. JBI Evid Synth. 2025; 23(5): 10307. doi:10.11124/JBIES-24-00232.

[10] Hayes SC, Strosahl KD, Wilson KG. Acceptance and commitment therapy: The process and practice of mindful change. 2nd Ed. New York: Guilford Press; 2012.

[11] Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and commitment therapy: Model, processes and outcomes. Behav Res Ther. 2006; 44(1): 1-25. doi: 10.1016/j.brat.2005.06.006.

[12] Harris R. ACT made simple: An easy-to-read primer on acceptance and commitment therapy. Oakland (CA): New Harbinger Publ Inc; 2019.

[13] Hayes SC, Pistorello J, Levin ME. Acceptance and commitment therapy as a unified model of behaviour change. Couns Psychol. 2012; 40(7): 976-1002. doi:10.1177/0011000012460836.

[14] Wilson KG, Murrell AR, Hayes BT, Follette VM, Linehan MM. Values-centred interventions: Setting a course for behavioural treatment. In: Hayes SC, Follette VM, Linehan MM, Editors. Mindfulness and acceptance: Expanding the cognitive-behavioral tradition. New York: Guilford Press; 2004.

[15] Moitra E, Stein MD, Busch AM, Pinkston MM, Abrantes AM, Baker JV, et al. Acceptance of chronic pain in depressed patients with HIV: correlations with activity, functioning, and emotional distress. AIDS Care. 2022; 34(10): 1338-46. doi:10.1080/09540121.2021.1981819.

[16] WHOQOL-HIV Group. Preliminary development of the World Health Organization's Quality of Life HIV instrument (WHOQOL-HIV): analysis of the pilot version. Soc Sci Med. 2003; 57(7): 1259-75. doi:10.1016/S0277-9536(02)00506-3.

[17] Aikat R, Gomes O. Effect of occupational therapy intervention on the quality of life of HIV positive clients and study of knowledge about HIV on clients and occupational therapists. Int J Health Sci Res. 2015; 5(8): 369-76.

[18] Skafi Sabet Z, Ghorban Shiroodi S, Zarbakhsh-Bahri M, Aminian K. The effectiveness of acceptance and commitment group therapy on disease severity, quality of life and psychological well-being of patients with irritable bowel syndrome. Salamat Ijtimai. 2022; 9(1): 36-46. doi:10.22037/ch.v9i1.31290.

[19] Ebrahimpour G, Mirzaeian B, Hasanzade R. Effectiveness of acceptance and commitment therapy on psychological well-being, quality of life and depression in patients with epilepsy. J Shahid Sadoughi Univ Med Sci. 2019; 27(2): 1262-79 doi:10.18502/ssu.v27i2.1046. (in Persian)

Journal of Associated Medical Sciences 2026; 59(2): 226-233

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Effectiveness of reminiscence therapy to improve sleep quality among older adults with hypertension

J. Roselin Jaba Sheela[1*] and J. Abiya Evangelin Jayapriya[2]

Saveetha college of occupational therapy, Saveetha institute of medical and technical sciences (SIMATS), Saveetha Nagar, Chennai, Tamil Nadu, India.

ARTICLE INFO

Article history:
Received 10 September 2025
Accepted as revised 6 February 2026
Available online 19 February 2026

Keywords:
Reminiscence therapy, sleep quality, hypertension, older adults

ABSTRACT

Background: Sleep disturbances are highly prevalent among older adults with hypertension, further exacerbating cardiovascular risks and impairing quality of life. Reminiscence therapy, a psychosocial intervention, has shown promise in promoting psychological well-being and reducing stress, thereby improving sleep outcomes.

Objectives: To evaluate the effectiveness of reminiscence therapy In improving sleep quality among older adults with hypertension.

Materials and methods: The experimental study was conducted among 30 hypertensive older adults aged 60 years and above. Participants were allocated to the intervention (reminiscence therapy) and control groups, which received conventional occupational therapy focused on improving sleep quality. This included cognitive behavioural therapy for insomnia, sleep hygiene education, sensory modulation strategies, environmental adaptations, and activity scheduling. Sleep quality was assessed using the Pittsburgh Sleep Quality Index (PSQI).

Results: Post-intervention analysis revealed a significant improvement in PSQI scores in the reminiscence therapy group compared to the control group (p<0.05). Subdomains, including sleep latency, sleep duration, and overall sleep efficiency, showed marked improvements.

Conclusion: Reminiscence therapy is effective in enhancing sleep quality among older adults with hypertension and can be integrated as a complementary non-pharmacological intervention in geriatric care.

Introduction

Systemic arterial hypertension, also known as hypertension, is characterized by persistently high blood pressure (BP) in the systemic arteries. Blood pressure is commonly expressed as diastolic blood pressure, which measures the force the heart experiences during relaxation, and systolic blood pressure, which measures the force the blood exerts on the arterial walls during cardiac contraction. Cardiovascular disease (CVD) includes peripheral artery disease, heart failure, stroke, coronary myocardial infarction, atrial fibrillation, and coronary heart disease. The most preventable risk factor for CVD is hypertension. Globally, CVD is the primary cause of death and disability. It also results in cognitive impairment and chronic kidney disease (CKD)[1]

*** Corresponding contributor.**
Author's Address: *Saveetha college of occupational therapy, Saveetha institute of medical and technical sciences (SIMATS), Saveetha Nagar, Chennai, Tamil Nadu, India.*
E-mail address: roselinjebasheelaj.scot@gmail.com
doi: 10.12982/JAMS.2026.054
E-ISSN: 2539-6056

The risk of arterial stiffness decreased vascular responsiveness, and hemodynamic changes that can worsen arterial dysfunction increase with lifelong exposure to endothelium damaging events. The capacity to maintain normal blood pressure and blood pressure variability decline with age as the risk of vascular disease rises. This change leads to a reduced prevalence of optimal blood pressure in older people with a phenotypic pattern that frequently includes a wide pulse pressure with SBP elevations that are relatively larger than diastolic blood pressure elevations.[2] According to the report, men are more likely than women to have hypertension (24.1%), with overall prevalence in India being 22.6%. As people aged, the prevalence increased, reaching 48.4% among those aged 60 and older.[3]

Patients with hypertension commonly complain of poor sleep quality. Poor sleep quality has been noted in older adults with hypertension, whose ratio is 1.5 to 3.18. Numerous studies have shown that older adults with hypertension are more likely to have trouble sleeping.[4-6] Sleep, a vital physiological activity, occupies a large part of our lives. However, modern society's struggle with sleep loss, chronic sleep deprivation, irregular sleep patterns and declining sleep quality is a concern. Remember that sleep is not just a daily activity but an essential process for recovery and renewal. The serious consequences of sleep deprivation, including increased disease risk and impacts on metabolism, immune function, and hormone regulation, underscore its importance as a topic and should motivate us to prioritize our sleep.[7] previous study reveals that approximately 50% of older adults experience poor sleep quality.[8]

Reminiscence therapy, a non-pharmacological intervention that facilitates guided recall of meaningful past experiences, is theorized to improve sleep quality in hypertensive older adults by reducing stress, enhancing emotional regulation, and promoting parasympathetic activation. By allowing individuals to process unresolved emotions, strengthen a sense of identity, and evoke positive affect, reminiscence therapy may lower physiological arousal and nighttime rumination factors known to exacerbate both hypertension and sleep disturbances. Previous studies have demonstrated that reminiscence-based interventions can effectively enhance sleep quality among older adults, supporting its potential relevance for this population. Reminiscence therapy is the utilization of recalling past events, emotions, and thoughts to promote pleasure, enhance quality of life, and support better adjustment to present circumstances. It can be in a structured or unstructured form and can occur in one-to-one or group settings. Reminiscence becomes reminiscence therapy when there is two way or more communication and is targeted at the achievement of some objectives as per what the individual requires.[9]

Creative reminiscence revives memories from the past and can be done in various ways. Some most useful ways are: photo albums/collages, scrapbooks, Art forms(drawing, painting or clay work can prove to be an alternative for verbal expression), historical objects and meaningful objects (toys, antiques, or clothing), drama (acting out brief scenes that encourage the role playing of previous experiences), vocal and instrumental music (may result in memory recall), life story work (recorded oral accounts of childhood and early years or autobiographies), memory boxes (a three dimensional box that shows personal belongings to represent one's life and accentuate memories).[10]

Numerous studies have examined non-pharmacological interventions to improve sleep quality in older adults.[11-13] However, there is a lack of evidence specifically addressing whether reminiscence therapy is effective for older adults living with hypertension, a population uniquely vulnerable to sleep disturbances due to physiological stress, autonomic imbalance, and emotional burden associated with chronic illness. Although reminiscence therapy, which involves structured recall and reflection on meaningful experiences, has been shown to enhance psychological well-being, reduce stress, and support better sleep in the general elderly population, its targeted impact on hypertensive older adults remains largely unexplored.

This represents a clear gap in the literature as hypertension related sleep problems may require tailored psychosocial approaches beyond what has been tested in broader geriatric samples. Therefore, this study aims to address this gap by investigating the specific effects of reminiscence therapy on sleep quality among older adults with hypertension. The findings are expected to contribute evidence for incorporating reminiscence therapy into occupational therapy programs designed for this clinically vulnerable group.

Materials and methods

Study design

A true experimental study design was employed to evaluate the effectiveness of reminiscence therapy on sleep quality among older adults with hypertension. This design involved random allocation of participants into experimental and control groups to establish a cause-and-effect relationship between the intervention and the outcome.

Sampling technique

A probability sampling approach using simple random sampling was adopted.

Sample setting

The study was conducted in the Mahatma Gandhi old age home, Arcot, Ranipet district.

Sample size

A total of 30 participants were recruited, with 15 allocated to the intervention group and 15 to

the control group.

Participants were randomly allocated to intervention and control groups using a computer-generated randomization sequence with allocation concealment through sealed opaque envelopes. Outcome assessors were blinded to group allocation. Proper randomization ensured comparability of baseline characteristics, minimizing selection bias.

Inclusion criteria

The study included older adults aged 60 years and above who were clinically diagnosed with hypertension. Participants were required to be conscious and able to communicate effectively. Both male and female individuals were considered eligible. Additionally, participants had to be experiencing sleep disturbances, such as difficulty falling asleep, staying asleep, or having poor overall sleep quality.

Exclusion criteria

The study excluded older adults who were diagnosed with dementia, as well as individuals with known neurological conditions. Participants with hearing impairments were also excluded to ensure effective communication and accurate participation in the intervention.

Tools used

Pittsburgh sleep quality index (PSQI)[14]

The PSQI was designed to evaluate overall sleep quality in these clinical populations. Each of the questionnaire's 19 Self-reported items belongs to one of seven subcategories: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction.

Interpretation of PSQI

The PSQI Global Score is interpreted as follows: ≤5: shows good sleep quality; >5: suggests poor sleep quality.

Reliability

Internal Reliability (Internal Consistency): Cronbach's alpha (α)=0.83, Test-Retest Reliability: Correlation coefficient=0.85

Validity

Sensitivity: 89.6%, Specificity: 86.5%

Intervention procedure

The Study was conducted over 12 weeks, with participants attending 30 sessions (3 sessions per week), each lasting 45 minutes. The intervention was structured as follows:

- Warm-up phase (10 minutes) began with light social interaction and a simple warm-up exercise, such as guided deep breathing or gentle stretching, to create a comfortable and engaging environment for participants.
- Action phase (30 minutes) reminiscence therapy was conducted by guiding participants to recall, share, and discuss meaningful past experiences, fostering engagement and emotional expression.
- Wind down phase (5 minutes): the session concluded with calming activities such as gentle relaxation exercises.

Table 1. Reminiscence therapy program for experimental group.[15-17]

Session	Structured reminiscence topics	Memory prompts used	Activity conducted
Session 1-3	The pre-test was taken introduction to reminiscence therapy		
Session 4-6	Sharing favorite childhood memory	Old photographs	Drawing
Session 7-9	sharing memorable travel experience	Postcards and letters	Painting
Session 10-12	Sharing memorable theatre experience and their favorite cinema	Old coins or currencies	Create your movie ticket
Session 13-15	Recalling past friendships and their impact	Vintage games	Playing games Ludo/ snake and ladders
Session 16-18	sharing their favourite meal and family gatherings	Old kitchen utensils or cooking tools	Pecs book images for meal preparation activities
Session 19-21	Sharing stories about family traditions or celebrations	Religious or ritual items	Rangoli
Session 22-24	Parenting experience	Spoons or baby utensils	Role play
Session 25-27	Spiritual and religious beliefs (worship and spiritual support)	Religious jewelry or symbols	Gardening
Session 28-30	sharing music, lullabies, or poems that brings comfort or peace	Radio	Musical chair, ball passing

Table 1. Reminiscence therapy program for experimental group.[15-17](continue)

Session	Structured reminiscence topics	Memory prompts used	Activity conducted
Session 31-33	Recollecting past bedtime routines and rituals	Oil lamp or kerosene lantern	Creating memory box
Session 34-36	The post-test was taken		

Before the therapy sessions begin, therapists involved will participate the reminiscence therapy protocol. This includes learning the structured topics designed to evoke positive memories (e.g., childhood events, family traditions, spiritual beliefs). Practicing how to use various memory prompts and tools effectively to engage participants in meaning conversations. Understanding how to facilitate sessions sensitively, encouraging participants to share stories without feeling rushed or pressured. Training on how to manage any emotional responses that may arise during the sessions.

Therapists introduce one or two reminiscence topics from a pre-set list, such as sharing favorite childhood memories or memorable travel experiences. To support these conversations, therapist use carefully chosen memory prompts like old photographs, postcards, coins, vintage games and familiar items related to family or cultural traditions. Therapists also include practical activities like drawing, painting or creating "memory movie tickets" that help participants express their memories creatively. Group games like Ludo or Snakes and Ladders can be included to stimulate social interaction and enjoyment. Throughout the session, therapist encourage participants to talk freely, offering gentle prompts when needed, and recording important observations for later analysis.

Therapists use a session checklist during each meeting to confirm that all key topics and activities are covered. Feedback sessions are held to discuss any deviations from the protocol and provide additional support or refresher training if needed.

Conventional occupational therapy for the control group

The control group in this study received a structured, multi-component conventional OT program including CBT-I, sensory modulation, activity scheduling and sleep hygiene education. As these interventions are evidence- based appr0aches for improving sleep, the control condition represents an active treatment rather than standard care. This may have contributed to improvements in the control group and reduced the observable between group differences, thereby complicating the attribution of outcomes solely to reminiscence therapy.

Table 2. Conventional occupational therapy program for the control group.[18-20]

Intervention	Activities
1. Cognitive Behavioral Therapy for Insomnia (CBT-I)	• Clients were engaged in discussions to identify and challenge unhelpful sleep-related beliefs • Sleep restriction was implemented by setting consistent bedtime and wake-up schedules • Relaxation sessions were conducted using deep breathing, progressive muscle relaxation, and guided imagery.
2. Sleep hygiene education	• Clients were taught to maintain regular sleep-wake routines through visual schedules.
3. Sensory modulation strategies	• Weighted blankets were provided during rest periods to promote relaxation. • Dim lighting or blackout curtains were used during sleep practice sessions. • Calming auditory inputs, such as white noise or soft instrumental music, were introduced before bedtime.
4. Environmental adaptations	• The sleep area was evaluated and adapted for comfort and minimal distraction. • Room lighting and temperature were adjusted to soothing levels. • Ergonomic sleeping positions and supportive bedding were recommended.

Table 2. Conventional occupational therapy program for the control group.[18-20] (continue)

Intervention	Activities
5. Activity scheduling and occupational balance	• Daily routines were planned, incorporating light exercises such as walking or stretching during the day. • Work, leisure, and rest periods were balanced to prevent fatigue or inactivity. • Meaningful, calming evening activities such as reading, journaling, or meditation were included.

Data analysis

A quantitative research design was adopted, and the data were analyzed using descriptive and inferential statistics. Descriptive statistics such as mean, standard deviation, minimum, and maximum were used to summarize the data. A priori sample size calculation was conducted to ensure adequate statistical power for detecting meaningful effects with the given sample. Given the small sample size (N=30), non-parametric tests were applied. The Mann-Whitney U test and the Wilcoxon signed rank test were employed to compare pre-test and post-test scores between group A and group B. Effect sizes and corresponding confidence intervals were calculated to provide information on the magnitude and clinical significance of the observed differences. A significance level of $p<0.05$ was considered statistically significant. Statistical analysis was performed using IBM SPSS statistics version 23.0

Results

The demographic characteristics of the participants in the control and experimental groups were comparable as presented in Table 3. In term of age distribution, most participants in both groups were aged 60-65 years, with a slightly higher proportion in the control group (46.6%) compared to the experimental group (40.0%). The experimental group had a marginally higher percentage of participants aged 66-70 years (33.3%) than the control group (26.7%), while both groups had an equal proportion of participants aged 71-75 years (26.7%). Regarding gender, both groups had identical distributions, with females comprising 60% and males 40% of participants. Overall, the two groups were similar in demographic characteristics at baseline.

Table 3. Demographic variables.

Sample No.	Demographic values	Class	Control group (%)	Experimental group (%)
1	Age (years)	60-65	7 (46.6)	6 (40)
		66-70	4 (26.7)	5 (33.3)
		71-75	4 (26.7)	4 (26.7)
2	Gender	Female	9 (60.0)	9 (60.0)
		Male	6 (40.0)	6 (40.0)

In the Table 4, the *p* value (0.011) was less than 0.05 and was significant at the 5%level; hence, it could be concluded that there was a significant difference in the sleep quality index scores between the pre and post-tests in the control group.

Table 4. Comparison of Pre-test and Post-test PSQI Scores in the control group (Paired t test).

Sample No.	Test	N	Mean	SD	T Value	Df	*p* value
1	Pre test	15	12.53	1.246	2.907	14	0.011*
2	Post test	15	10.93	2.187			

Note: ***significant at 1% level , *significant at 5% level.*

Also, the mean value of the post test scores was lower than the mean value of the pretest scores, and the difference was statistically significant. Therefore, it was concluded that the PSQI scores significantly decreased in the control group as well.

As demonstrated in Table 5, the p=0.000 was highly significant at the 1% level: the mean value of the post-test scores was significantly lower than that of the pre-test scores, indicating a highly significant improvement in sleep quality. This substantial enhancement may be attributed to the high frequency and duration of reminiscence therapy sessions, which provided consistent emotional support and cognitive engagement. Furthermore, the supportive therapeutic environment and strong rapport established between therapists and participants likely enhanced motivation and active participation. Additionally, group dynamics that fostered social interaction and reduced feelings of loneliness may have contributed to these positive outcomes. Therefore, it was concluded that reminiscence therapy effectively and significantly decreased PSQI scores in the experimental group.

Table 5. Comparison of pre-test and post-test PSQI scores in the experimental group (Paired t test).

Sample No.	Test	N	Mean	SD	T Value	Df	*p* value
1	Pre test	15	12.93	1.438	20.000	14	0.000**
2	Post test	15	6.27	0.799			

Note: ***significant at 1% level , *significant at 5% level.*

In the Table 6, the p value (0.000) was less than 0.01 and was highly significant at the 1% level; hence, it was concluded that there was a highly significant difference in the post-test sleep quality index scores between the control and experimental groups. Also, the mean value of the experimental group was lower than the mean value of the control group, and the difference was highly significant. Therefore, it was concluded that the PSQI scores significantly decreased in the experimental group compared to the control group following the reminiscence therapy.

Table 6. Comparison of post-test PSQI scores between control and experimental group (Independent t test).

Sample No.	Test	N	Mean	SD	T Value	Df	*p* value
1	Control	15	10.93	2.187	7.764	28	0.000**
2	Experimental	15	6.27	0.799			

Note: ***significant at 1% level , *significant at 5% level.*

Discussion

The present study showed a significant and notably large improvement in sleep quality among older adults with hypertension after receiving reminiscence therapy, as measured by a substantial reduction in PSQI scores. Compared to Guler Duru Asiret who found a moderate improvement with weekly 25-30 minute reminiscence sessions that reduced PSQI scores but with smaller effect sizes, this study's thrice weekly 45 minute sessions over 12 weeks produced a greater magnitude of change.[21] Specifically, while Asiret reported an average PSQI score reduction of around 2 points, the current study observed a mean reduction exceeding 4 points, indicating a stronger therapeutic effect.

Similarly, Fatemah Sahragard achieved significant reductions in insomnia severity through twice-weekly group sessions over four weeks, which persisted at one-month follow-up.[22] However, the shorter intervention duration and lower total therapy time likely contributed to relatively smaller improvements compared to the current findings. The longer intervention period and higher session frequency in this study may explain the amplified benefits observed, suggesting dose-dependent effects of reminiscence therapy on sleep quality.

Beyond these quantitative comparisons, several factors may explain the large effect size. The supportive, engaging environment in which the therapy was delivered potentially enhanced participant motivation and attendance, increasing intervention efficacy. High therapist involvement and consistent rapport building likely promoted trust and openness, amplifying emotional and psychological benefits that facilitate improved sleep. The structured group format may have fostered positive social interactions and reduced loneliness, both known to influence sleep quality, as highlighted by Melike Dutkunb.[24] These contextual elements, alongside the novelty of reminiscence therapy for many participants, might have synergistically intensified outcomes beyond the direct therapeutic content.

Possible confounding influences warrant consideration. While therapists and research assistants were blinded to group allocation during assessments, their dual roles in therapy delivery and evaluation could have introduced subtle biases. Institutional routines and environmental factors specific to the study setting may have contributed to participant stability or stress levels, indirectly affecting sleep outcomes. Moreover,

group dynamics such as peer support or shared experiences might have played a role independent of the intervention itself.

Clinically, the observed improvements extend well beyond statistical significance. A mean PSQI score reduction of over 4 points corresponds to meaningful enhancements in sleep duration, efficiency, latency, and overall quality factors crucial for hypertensive older adults prone to cardiovascular risk linked to poor sleep. Improved sleep quality can promote better daytime functioning, cognitive health, mood stabilization, and adherence to hypertension management, reflecting the holistic benefits noted by Chiang *et al*.[26] Thus, reminiscence therapy offers not only a statistically effective but also a clinically valuable non-pharmacological approach to improving health outcomes in this population.

The significant improvement in sleep quality observed in the control group may be attributed to their participation in structured traditional occupational therapy programs. Regular therapist-participant interaction, consistent therapeutic attention, and engagement in purposeful activities may have promoted relaxation, emotional reassurance, and improved daily routines, all of which are known to positively influence sleep patterns. Additionally, the structured nature of conventional therapy sessions may have contributed to better sleep-wake regulation, thereby resulting in measurable improvements in sleep quality among control group participants.

In summary, the current findings demonstrate a larger and more sustained improvement in sleep quality compared to previous studies, likely due to higher treatment dosage, supportive therapeutic context, and group interaction effects. While confounding factors cannot be fully excluded, the clinically meaningful gains highlight reminiscence therapy's potential as a key intervention in older adults with hypertension. Future research should aim to isolate these variables and optimize intervention delivery for maximum benefit.

Limitations

This study had certain limitations. It was conducted with a small and homogeneous sample drawn from a single facility and over a short duration, which may limit the generalizability of the findings to the wider population of older adults with hypertension. Sleep quality was assessed using subjective self-report questionnaires only, without objective measures such as actigraphy or sleep monitoring, which may have introduced response bias. Additionally, cognitive variations among participants may have affected engagement in reminiscence therapy and influenced the effectiveness of the intervention.

Based on these limitations, it is recommended that future studies include long-term interventions and be conducted with a larger sample size to strengthen the results. Moreover, combining reminiscence therapy with other interventions, such as sleep hygiene education or relaxation techniques, may further enhance its effectiveness.

Conclusion

The findings of this study indicate that reminiscence therapy is an effective non pharmacological intervention for improving sleep quality among older adults with hypertension. Participants who received reminiscence therapy demonstrated statistically significant improvements in sleep quality, as assessed using the Pittsburgh Sleep Quality Index (PSQI). Accordingly, further research is warranted to substantiate and extend the evidence supporting reminiscence therapy as an effective non pharmacological intervention for enhancing sleep quality among this population. Nevertheless, participants who received conventional occupational therapy also exhibited significant improvement in sleep quality, underscoring the therapeutic value of structured traditional programs and the contribution of nonspecific treatment effects alongside targeted psychosocial interventions.

Ethical approval

This study has been approved by the institution's scientific review board (ISRB) of Saveetha College of Occupational Therapy with REF. NO. of SCOT/ISRB/042/2025

Funding

This research received no specific grant from any funding agency In the public, commercial, or not for profit sectors.

Conflict of interest

The authors declare no conflict of interest

CRediT authorship contribution statement

Roselin Jaba Sheela: supervision, methodology, conceptualization, writing: review and edit; **Abiya Evangelin Jayapriya**: investigation, data curation, writing: original draft, review and edit.

Acknowledgements

I sincerely thank Saveetha Institute of Medical and Technical Sciences, Saveetha College of Occupational Therapy, Dr.M. Arunkumar, MOT, Ph.D., for his continuous encouragement, and Dr. Roselin Jaba Sheela for her exemplary guidance and support throughout this research.

References

[1] Whelton PK, Carey RM, Aronow WS, Casey DE Jr, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high

blood pressure in adults: Executive summary. Hypertension. 2018; 71(6): 1269-324. doi: 10.1161/HYP.0000000000000066.

[2] Chaitoff A, Zheutlin AR. Epidemiology of hypertension in older adults. Clin Geriatr Med. 2024; 40(4): 515-28. doi.org/10.1016/j.cger.2024.04.007.

[3] Mohammad R, Bansod DW. Hypertension in India: A gender-based study of prevalence and associated risk factors. BMC Public Health [Internet]. 2024; 24(1). Article number 2681. doi.org/10.1186/s12889-024-20097-5.

[4] Liu D, Yu C, Huang K, Thomas S, Yang W, Liu S, et al. The association between hypertension and insomnia: A bidirectional meta-analysis of prospective cohort studies. Int J Hypertens. 2022; 2022: 1-11. doi.org/10.1155/2022/4476905.

[5] Budhiraja R, Roth T, Hudgel DW, Budhiraja P, Drake CL. Prevalence and polysomnographic correlates of insomnia comorbid with medical disorders. Sleep [Internet]. 2011; 34(7): 859-67. doi.org/10.5665/sleep.1114.

[6] Taylor DJ, Mallory LJ, Lichstein KL, Durrence HH, Riedel BW, Bush AJ. Comorbidity of chronic insomnia with medical problems. Sleep. 2007; 30(2): 213-8. doi.org/10.1093/sleep/30.2.213.

[7] Li L, Gan Y, Zhou X, Jiang H, Zhao Y, Tian Q, et al. Insomnia and the risk of hypertension: A meta-analysis of prospective cohort studies. Sleep Med Rev. 2020; 56: 101403. doi.org/10.1016/j.smrv.2020.101403.

[8] Kavousi P, Mali E, Seifhashemi N, Souri M, Pakravan L, Khalili F. Worldwide prevalence of poor sleep quality in older adults: A systematic review and meta-analysis. Iran J Psychiatry. 2025. doi.org/10.18502/ijps.v20i2.18207.

[9] Elias SMS, Neville C, Scott T. The effectiveness of group reminiscence therapy for loneliness, anxiety and depression in older adults in long-term care: A systematic review. Geriatr Nurs. 2015; 36(5): 372-80. doi.org/10.1016/j.gerinurse.2015.05.004.

[10] Latha KS, Bhandary PV, Tejaswini S, Sahana M. Reminiscence therapy: An overview. Middle East J Age Ageing. 2014; 11(1): 18-22. doi.org/10.5742/meaa.2014.92393.

[11] Camino M, Satorres E, Delhom I, Real E, Abella M, Meléndez JC. Mindfulness-based cognitive therapy to improve sleep quality in older adults with insomnia. Psychosoc Interv. 2022; 31(3): 159-67. doi.org/10.5093/pi2022a12.

[12] Hedayat S, Arefi M. The effect of cognitive-behavioral therapy on improving sleep quality in the elderly. Knowl Res Appl Psychol. 2015; 16(2): 60-8. http://journals.khuisf.ac.ir/jsr-p/article-1-1807-fa.pdf.

[13] Ong JC, Xia Y, Smith-Mason CE, Manber R. A randomized controlled trial of mindfulness meditation for chronic insomnia. Mindfulness. 2018; 9(6): 1702-12. doi.org/10.1007/s12671-018-0911-6.

[14] Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. Psychiatry Res. 1989; 28(2): 193-213. doi.org/10.1016/0165-1781(89)90047-4.

[15] Liu Z, Yang F, Lou Y, Zhou W, Tong F. The effectiveness of reminiscence therapy on alleviating depressive symptoms in older adults: A systematic review. Front Psychol. 2021; 12. doi.org/10.3389/fpsyg.2021.709853.

[16] Chiang K, Chu H, Chang H, Chung M, Chen C, Chiou H, et al. The effects of reminiscence therapy on psychological well-being, depression, and loneliness among the institutionalized aged. Int J Geriatr Psychiatry. 2009; 25(4): 380-8. doi.org/10.1002/gps.2350.

[17] Huang HC, Chen YT, Chen PY, Hu SHL, Liu F, Kuo YL, et al. Reminiscence therapy improves cognitive functions and reduces depressive symptoms in elderly people with dementia: A meta-analysis of randomized controlled trials. J Am Med Dir Assoc. 2015; 16(12): 1087-94. doi.org/10.1016/j.jamda.2015.07.010.

[18] Ho ECM, Siu AMH. Occupational therapy practice in sleep management: A review of conceptual models and research evidence. Occup Ther Int. 2018; 2018: 1-12. doi.org/10.1155/2018/8637498.

[19] Yoo I. A scoping review of sleep management as an occupational therapy intervention. Ir J Occup Ther. 2023; 51(2): 22-34. doi.org/10.1108/ijot-01-2023-0001.

[20] Gutman SA, Gregory KA, Sadlier-Brown MM, Schlissel MA, Schubert AM, Westover LA, et al. Comparative effectiveness of three occupational therapy sleep interventions. OTJR (Thorofare N J). 2017; 37(1): 5-13. doi.org/10.1177/1539449216673045.

[21] Aşiret GD. Effect of reminiscence therapy on the sleep quality of the elderly living in nursing homes: A randomized clinical trial. Eur J Integr Med. 2018; 20: 1-5. doi.org/10.1016/j.eujim.2018.03.007.

[22] Sahragard F, Rezaei M, Izadi-Avanji FS, Atoof F, Ahmadishad M. The effect of group reminiscence therapy on insomnia in older adults: A randomized controlled clinical trial. J Res Dev Nurs Midwifery. 2020; 17(1): 94-110. doi.org/10.29252/jgbfnm.17.1.94.

[23] Lök N, Bademli K, Selçuk-Tosun A. The effect of reminiscence therapy on cognitive functions, depression, and quality of life in Alzheimer patients: Randomized controlled trial. Int J Geriatr Psychiatry. 2019; 34(1): 47-53. doi.org/10.1002/gps.4980.

[24] Aşiret GD, Dutkun M. The effect of reminiscence therapy on the adaptation of elderly women to old age: A randomized clinical trial. Complement Ther Med. 2018; 41: 124-9. doi.org/10.1016/j.ctim.2018.09.018.

Journal of Associated Medical Sciences 2026; 59(2): 234-249

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Interprofessional education and collaboration in health professions: A narrative review of global evidence and lessons from the Philippines

Paolo Miguel P. Bulan[1*], Nikki Y. Pestaño[1], Romel V. Cabazor[1], Marelen H. Amanuence[2], Lois Vanjie G. Napalit[2], Niño Paolo M. Tan[3], Catherine S. Malaca[2], Eubina C. Estoy[4]

[1]*College of Occupational Therapy and Physical Therapy, Velez College, Philippines.*
[2]*College of Arts and Sciences, Velez College, Philippines.*
[3]*College of Medical Technology, Velez College, Philippines.*
[4]*College of Nursing, Velez College, Philippines.*

ARTICLE INFO

Article history:
Received 10 October 2025
Accepted as revised 16 February 2026
Available online 25 February 2026

Keywords:
Interprofessional education, interprofessional collaboration, education, teamwork, Philippines

ABSTRACT

Background: Interprofessional education (IPE) and interprofessional collaboration (IPC) are globally recognized as strategies for strengthening health systems and improving patient outcomes. Although frameworks and evidence have been well established in high-income countries, insights from low- and middle-income contexts remain limited. In the Philippines, studies on IPE and IPC are emerging but remain fragmented across disciplines and institutions.

Objectives: This review aimed to synthesize international and Philippine literature on IPE and IPC, identify key frameworks, experiences, and challenges, and discuss implications for education, practice, and policy development in the Philippine context.

Materials and methods: A narrative review approach was used to integrate diverse forms of evidence on IPE and IPC. Sources were identified through PubMed, Scopus, Google Scholar, and relevant organizational reports using the keywords “interprofessional education,” “interprofessional collaboration,” “health professions,” and “Philippines.” Articles published in English between 2000 and 2024 were included if they discussed definitions, frameworks, implementation strategies, challenges, or outcomes. Extracted data were synthesized thematically to operationalize the review objectives, focusing on conceptual and policy foundations of IPE and IPC, evidence of effectiveness, international experiences, the Philippine context and local evidence, and challenges, barriers, and sustainability.

Results: Global literature shows that IPE enhances teamwork, communication, and, in some cases, patient outcomes, though findings are often constrained by methodological heterogeneity. International frameworks from Australia, Canada, the United Kingdom, and the United States establish shared competency domains but highlight challenges in assessment and curriculum integration. Philippine studies demonstrate readiness among students and professionals, pilot initiatives in community and institutional settings, and recurring barriers such as faculty shortages, professional hierarchies, and limited policy support. Sustainability emerged as a major concern in both global and local contexts.

Conclusion: IPE and IPC hold significant potential to advance collaborative health professions education in the Philippines and similar low- and middle-income countries. To strengthen sustainability, contextual adaptation of global frameworks, investment in faculty development, alignment of accreditation

*Corresponding contributor.
Author’s Address: *College of Occupational Therapy and Physical Therapy, Velez College, Philippines.*
E-mail address: paolobulan@me.com
doi: 10.12982/JAMS.2026.055
E-ISSN: 2539-6056

and policy standards, and longitudinal research are essential to building enduring models that enhance interprofessional collaboration and health system outcomes.

Introduction

Interprofessional education (IPE) is defined by the World Health Organization[1] as occasions when students from two or more professions learn about, from, and with each other to enable effective collaboration and improve health outcomes. Interprofessional collaboration (IPC), in turn, refers to multiple health workers from different professional backgrounds working together with patients, families, caregivers, and communities to deliver the highest quality of care.[1-3] IPE aims to produce a collaborative practice–ready workforce, while IPC represents the application of these competencies in practice. Together, they are widely recognized as essential for strengthening health systems and improving patient outcomes.

The WHO has emphasized that IPE prepares a workforce capable of collaborative practice, which enhances access, quality, and efficiency of care.[1] In response, several countries have developed national competency frameworks to guide the integration of IPE into curricula and practice, including models from Canada, the United States, the United Kingdom, and Australia.[2-4] Systematic reviews have also demonstrated that IPE can improve teamwork, communication, and, in some cases, patient outcomes, although the evidence remains methodologically heterogeneous.[5,6]

Despite growing international attention, the literature on IPE and IPC remains uneven. Many reviews focus on high-income countries, leaving fewer insights from low- and middle-income contexts where health system challenges may be distinct.[7,8] In the Philippines, research on IPE and IPC is emerging but still limited, often descriptive, and constrained to specific professions or institutions.[9,10] These studies suggest readiness among students and professionals and demonstrate pilot initiatives, yet they also reveal barriers such as professional hierarchies, faculty shortages, and limited policy support.[11,12] To date, no comprehensive synthesis has integrated Philippine evidence with global literature to identify how international lessons might inform local development of IPE and IPC.

The purpose of this review is to synthesize literature on 1) conceptual and policy foundations of IPE and IPC, 2) evidence of effectiveness, 3) international experiences, 4) Philippine context and local evidence, and 5) challenges, barriers, and sustainability. By integrating global evidence with local perspectives, this review seeks to provide a clearer understanding of how IPE and IPC can be advanced in the Philippine context while contributing to broader discussions on collaborative practice in health professions education.

Materials and methods

This review followed a narrative approach,[13-15] chosen for its flexibility in synthesizing diverse forms of evidence and its suitability for examining broad and complex topics such as IPE and IPC. Unlike systematic or scoping reviews, the aim was not to exhaustively capture all available studies but to integrate key literature that informs the conceptual, practical, and contextual aspects of IPE and IPC.[13-15]

The scope of the review included literature that addressed definitions, frameworks, implementation strategies, challenges, and outcomes of IPE and IPC in health professions education and practice. Both international and Philippine studies were considered to highlight global trends and local experiences. Sources were identified through electronic databases (PubMed, Scopus, and Google Scholar), manual searches of reference lists, and relevant organizational reports. Search terms combined "interprofessional education," "interprofessional collaboration," "health professions," and "Philippines." Inclusion criteria were studies and reviews published in English between 2000 and 2024 that examined IPE or IPC in health professions education or practice. Exclusion criteria were articles that addressed multiprofessional education without an interprofessional focus or those outside the health professions.

Data were extracted using a standardized charting form that summarized author, year, country, study type, population, key findings, and relevance to IPE and IPC. The synthesis followed content analysis,[16] with analytical categories developed to operationalize the review objectives. Findings were grouped into 1) conceptual and policy foundations of IPE and IPC, 2) evidence of effectiveness, 3) international experiences, 4) Philippine context and local evidence, and 5) challenges, barriers, and sustainability. Reflexivity was maintained by acknowledging that the selection and interpretation of studies were shaped by the perspectives of the authors.

Results

Overview of included literature

A total of 26 source materials were included in this narrative review. The literature spanned more than two decades, with early contributions published between 1999 and 2009, followed by a steady increase in publications from 2010 onward. Most sources were published between 2010 and 2023.

In terms of study design, the included sources were predominantly reviews and descriptive studies, including systematic and umbrella reviews, cross-sectional surveys, and descriptive case reports. A smaller number of sources focused on intervention evaluation, and several papers presented competency frameworks or policy-oriented guidance rather than empirical findings. The studies represented diverse

geographic contexts, including global or multinational settings, as well as country-specific evidence from the Philippines, other parts of Asia, Europe, Africa, Australia, and North America.

The included literature addressed a range of populations and contexts, including undergraduate and postgraduate health professions students, licensed health and social care professionals, faculty members, institutional leaders, and policymakers across multiple disciplines. Key findings across sources focused on attitudes and readiness for interprofessional education, evidence of effectiveness of IPE and IPC initiatives, international and local implementation experiences, and recurring challenges related to barriers and sustainability. A summary of the characteristics of the included sources is presented in Table 1, while detailed extracted data are provided in the supplementary material.

The findings of this review are organized into five themes that reflect the study objectives: 1) conceptual and policy foundations of IPE and IPC, 2) evidence of effectiveness, 3) international experiences, 4) Philippine context and local evidence, and 5) challenges, barriers, and sustainability. These themes integrate insights from both global and Philippine literature to provide a comprehensive understanding of how IPE and IPC have been conceptualized, implemented, and evaluated across settings.

Table 1. Summary of characteristics of included source materials (N=26).

Characteristic	Summary
Publication period	Sources were published between 1999 and 2023, with the majority appearing from 2010 onward.
Geographic context	Studies were conducted in or focused on global or multinational settings, as well as country-specific contexts including the Philippines, United States, Canada, United Kingdom, Australia, Germany, South Africa, and Qatar.
Study type	Included sources comprised systematic reviews, scoping reviews, umbrella reviews, quantitative descriptive studies, qualitative studies, mixed-methods studies, descriptive case reports, and non-empirical competency or policy frameworks.
Educational and practice settings	Studies addressed academic, clinical, community-based, workplace, and policy-level contexts in health and social care.
Populations	Participants and target groups included undergraduate and postgraduate health professions students, licensed health and social care professionals, faculty members, administrators, and policymakers across multiple disciplines.
Key focus areas	Studies examined attitudes and readiness for interprofessional education, effectiveness of IPE and IPC initiatives, implementation strategies, international and local experiences, and challenges related to barriers and sustainability.
Relevance to IPE and IPC	All included sources contributed evidence or guidance relevant to the conceptualization, implementation, evaluation, or sustainability of interprofessional education and collaborative practice.

Conceptual and policy foundations of IPE and IPC

IPE and IPC have been recognized as essential strategies for strengthening health systems and improving patient outcomes. The WHO's Framework for Action emphasized that preparing a collaborative practice-ready workforce can enhance access, quality, safety, and efficiency of care, while also identifying systemic enablers such as educator training, curricular integration, and supportive policies.[1] Building on this foundation, several national and institutional models have been developed to operationalize competencies for collaborative practice.

The Canadian Interprofessional Health Collaborative (CIHC) Framework articulated six domains: role clarification, patient- and family-centered care, team functioning, collaborative leadership, communication, and conflict resolution, integrating knowledge, skills, and values for practice.[2] In the United States, the Interprofessional Education Collaborative (IPEC) Report proposed four competency domains: values and ethics, roles and responsibilities, communication, and teamwork, as the basis for curriculum and accreditation reforms.[3] Similarly, the Curtin University Interprofessional Capability Framework in Australia highlighted capabilities across developmental levels, emphasizing client-centered care, safety, and teamwork as outcomes uniquely achievable through interprofessional activity.[17]

Earlier frameworks also contributed to shaping this landscape. Walsh and colleagues developed the Interprofessional Capability Framework in the United Kingdom, identifying four domains: ethical practice,

knowledge in practice, interprofessional working, and reflection, that emphasized capability over static competence.[18] Thistlethwaite and colleagues later compared major frameworks from the UK, Canada, the US, and Australia, noting areas of convergence such as communication, teamwork, and ethics, while highlighting challenges in terminology, assessment, and curriculum integration.[4]

These frameworks collectively illustrate the global consensus that interprofessional competencies are critical for preparing health professionals for collaborative practice, while also pointing to gaps in assessment strategies and continuing debates on how best to conceptualize competence and capability. A summary of key international frameworks and their domains is provided in Table 2.

Table 2. Key international IPE frameworks.

Framework	Domains/Focus	Key features	Reference
WHO Framework for Action (2010)	Educator training, curricular integration, supportive policies	Emphasizes collaborative practice-ready workforce; identifies systemic enablers	WHO[1]
CIHC Framework (Canada)	Role clarification, patient- and family-centered care; team functioning, collaborative leadership, communication, conflict resolution	Integrates knowledge, skills, and values	CIHC[2]
IPEC Competencies (US)	Values and ethics, roles and responsibilities, communication, teamwork	Basis for accreditation and curriculum reforms	IPEC Expert Panel[3]
Curtin University Framework (Australia)	Client-centered care, safety, teamwork (across developmental levels)	Highlights capabilities as outcomes of IPE	Brewer[17]
UK Interprofessional Capability Framework	Ethical practice, knowledge in practice, interprofessional working, reflection	Focuses on *capability* rather than static competence	Walsh *et al*.[18] Thistlethwaite *et al*.[4]

Effectiveness of IPE and IPC

Evidence on the effectiveness of IPE and IPC has expanded over the past two decades. The Cochrane review update synthesized randomized controlled trials, controlled before-and-after studies, and time series analyses, showing that IPE can improve professional practice and, in some cases, patient outcomes such as diabetes management, emergency care, and prescribing practices. However, the certainty of evidence was generally low to moderate due to methodological heterogeneity and small sample sizes, underscoring the need for stronger designs.[5]

Other reviews have examined student outcomes, particularly in medicine. A systematic review of IPE interventions in medical curricula reported generally positive changes in student attitudes, with later and longer interventions demonstrating more consistent benefits. The lack of consensus on timing and the limitations of commonly used instruments such as the Readiness for Interprofessional Learning Scale (RIPLS) suggest the need for improved measures and longitudinal designs.[6]

A broader synthesis of IPE in higher education identified three key perspectives: providers, learners, and researchers. Barriers such as limited resources, heavy workloads, and lack of institutional support were commonly reported, while facilitators included student-led activities, relevance to clinical practice, and organizational commitment. Topic modeling reinforced the salience of themes such as simulation-based learning and research collaboration, but also highlighted gaps in the evidence base.[7]

Implementation-focused reviews also reveal persistent challenges. A scoping review of IPE initiatives mapped barriers at micro (silos, stereotypes), meso (institutional leadership, scheduling), and macro (policy and regulation) levels, noting sustainability as a critical issue.[8] Reviews of IPC in primary care and healthcare teams further confirmed recurring obstacles: time constraints, unclear roles, and hierarchical structures, while identifying facilitators such as co-location, strong leadership, and shared decision-making as context-specific enablers.[19,20]

Overall, the evidence shows that IPE and IPC can enhance collaboration and contribute to improved care, though outcomes remain inconsistent and methodological weaknesses are common.

International experiences in IPE and IPC

IPE has been implemented in diverse ways across international contexts, often shaped by local needs and available resources. Simulation-based learning is one widely adopted strategy. At the University of the Free State in South Africa, module leaders reported that simulation enhanced role clarification, teamwork, and empathy, though logistical barriers such as scheduling, costs, and staff attitudes limited broader implementation.[21] Similarly, in Qatar, a pre-post IPE activity on smoking cessation demonstrated significant reductions in stereotypical views among healthcare students, showing that targeted interventions can improve attitudes and role understanding.[22]

Workshops and structured training have also been tested as strategies for fostering collaboration. A Canadian study evaluating a one-day interprofessional workshop for residents, nurses, and allied health professionals reported high satisfaction and modest improvements in teamwork attitudes among residents, but limited change in other groups. This highlighted the potential of workshops for short-term gains while pointing to the need for role clarification and organizational support to sustain impact.[23] Beyond formal educational settings, practical strategies for embedding interprofessional education and collaborative practice into daily work have also been described. Guidance from Australia and the Philippines outlined twelve actionable tips for healthcare teams, including investing in champions, using shared language, and incorporating client feedback, offering adaptable approaches for day-to-day practice.[24]

Attitudes toward collaboration have also been studied among early-career professionals. In Germany, young healthcare graduates without prior IPE exposure expressed generally positive views on communication and teamwork but were neutral toward interprofessional learning and interactions, with demographic factors influencing attitudes. These findings suggest the importance of continuing education and early career reinforcement of collaborative skills.[25] Instrument development has further contributed to advancing IPE evaluation, with the Attitudes Toward Health Care Teams Scale (ATHCTS) providing a validated measure for assessing team perceptions across professions.[26]

International experiences demonstrate the adaptability of IPE strategies across contexts, from simulation and workshops to workplace integration. Results consistently highlight improvements in attitudes and collaboration, though limitations in logistics, evaluation, and long-term outcomes remain evident.

Philippine context and local evidence

IPE and IPC in the Philippines remain limited but has grown steadily in recent years, offering insights into readiness, attitudes, and implementation. A survey of healthcare students reported variation in readiness for interprofessional learning, with medical technology students showing lower teamwork and collaboration scores and physiotherapy and occupational therapy students scoring lower in role and responsibility domains. Senior students generally displayed more favorable attitudes, suggesting that readiness develops over time.[9] Similarly, a nationwide survey of licensed therapists found that those with prior IPE experience held more positive attitudes toward IPC, with community-based case discussions and didactics linked to stronger collaboration attitudes. Practice setting also influenced perceptions, with academic environments associated with more positive views.[10]

Institutional and community-based initiatives further demonstrate how IPE has been adapted to the local context. At the University of the Philippines Manila, a community-engaged IPE model integrated health and social care students into a rural development program. The initiative improved appreciation of professional roles and enhanced patient involvement but faced challenges such as professional hierarchies, faculty shortages, and resource constraints.[12] A pilot program at Angeles University Foundation also showed the potential for structured IPE implementation across multiple health science programs, though scheduling, bureaucracy, and limited resources posed barriers to sustainability.[11]

Other Philippine studies highlight system-level issues and workforce perspectives. Health and social workers caring for older adults described IPC as largely informal and ad hoc, hindered by organizational silos, staff shortages, and weak governance structures. Participants emphasized the need for integrated health and social care, leadership, and policy reforms.[27] A follow-up pilot training for these workers demonstrated that structured IPE activities improved attitudes, role understanding, and teamwork, supporting the feasibility of collaborative training even in resource-constrained settings.[28] In mental health, a baseline assessment among Filipino professionals revealed challenges in communication, role clarification, and conflict management, pointing to gaps in collaboration quality.[29] More recently, a survey of Filipino medical, nursing, and pharmacy students found generally positive attitudes toward IPE, with moral development emerging as an additional factor influencing perceptions.[30]

These studies provide an emerging picture of IPE and IPC in the Philippines, showing readiness and willingness to engage in interprofessional learning while identifying persistent challenges in faculty readiness, organizational support, and systemic silos.

Challenges, barriers, and sustainability

Despite the recognized benefits of IPE and IPC, persistent challenges have been documented across settings. Reviews consistently identify barriers at multiple levels. At the micro level, professional silos, negative stereotypes, and entrenched hierarchies hinder effective teamwork. Faculty shortages and limited training further constrain implementation.[8,19] At

the meso level, scheduling conflicts, lack of administrative support, and resource limitations are recurring issues. Macro-level barriers include insufficient policy support, unclear regulatory frameworks, and the absence of long-term funding or infrastructure to sustain initiatives.[8]

Philippine studies mirror these patterns but also highlight context-specific issues. Community- and institution-based initiatives identified faculty shortages, scheduling difficulties, and professional hierarchies as significant obstacles to program delivery.[11,12] Health and social workers emphasized systemic problems such as siloed services, organizational fragmentation, and weak leadership, which contributed to delayed care and duplication of services.[27] These concerns were echoed in mental health and geriatric care settings, where participants stressed the need for integrated governance, clearer role delineation, and policies supportive of collaboration.[28,29]

Another recurring challenge lies in the sustainability of IPE programs. While pilot projects and short-term interventions have demonstrated feasibility and short-term benefits, many lacked the infrastructure, funding, and leadership support needed to continue beyond initial phases.[8,11]

Overall, barriers to IPE and IPC are well documented and remain consistent across global and Philippine contexts. Addressing them requires systemic solutions that extend beyond isolated initiatives to ensure sustainability and integration into education and health systems. Table 3 summarizes barriers identified at micro, meso, and macro levels, comparing global patterns with Philippine findings.

Table 3. Barriers to IPE and IPC across contexts.

Level	Common barriers (global)	Barriers in the Philippines	Key references
Micro (individual/professional)	Professional silos, stereotypes, entrenched hierarchies	Hierarchical relations among professions; limited faculty training	Bogossian *et al.*[8] Rawlinson *et al.*[19] Paterno and Opina-Tan[12]
Meso (institutional/ organizational)	Scheduling conflicts, lack of administrative support, resource limitations	Faculty shortages; scheduling difficulties; limited institutional resources	Sy *et al.*[11] Paterno and Opina-Tan[12]
Macro (system/policy)	Insufficient policy support; unclear regulation; lack of funding	Siloed services; weak governance; lack of supportive policy frameworks	Bogossian *et al.*[8] Moncatar *et al.*[27] Sy *et al.*[29] Nakamura *et al.*[28]

The literature illustrates a strong international consensus on the importance of IPE and IPC, supported by global and national frameworks that define competencies and guide curricular development. Evidence from systematic and scoping reviews shows that IPE can improve collaboration and, in some cases, patient outcomes, though results are often limited by methodological weaknesses and inconsistent evaluation. International experiences demonstrate a range of strategies, including simulation, workshops, and workplace integration, that generally enhance teamwork and attitudes, but face persistent logistical and structural barriers. Philippine studies add valuable insights, showing student and professional readiness, pilot initiatives, and community-based models, while also highlighting barriers related to faculty capacity, professional hierarchies, and systemic silos. Across global and local contexts, sustainability emerges as a recurring challenge, with many initiatives constrained by limited resources, infrastructure, and long-term support.

Discussion

This review synthesized international and Philippine literature on IPE and IPC. The findings show that while there is broad consensus on the importance of IPE and IPC, and evidence of positive outcomes, significant gaps and barriers remain. Frameworks from the WHO, Australia, Canada, the UK, and the US, have established shared domains of competence and capability, and systematic reviews provide moderate evidence that IPE can improve collaboration and, in some cases, patient outcomes. International experiences further illustrate the potential of strategies such as simulation, workshops, and workplace integration, though logistical and structural constraints frequently limit their impact.

When considered alongside this global evidence, Philippine studies add valuable but still emerging contributions. Research among students and professionals indicates readiness and generally positive attitudes toward collaboration, and pilot initiatives show that IPE can be integrated into institutional and community settings. However, these studies are often descriptive, small-scale, and limited to specific professions or institutions. Barriers such as faculty shortages, scheduling conflicts, and entrenched professional hierarchies mirror international challenges but are compounded by local systemic and policy constraints. These patterns mirror global challenges but are compounded locally by systemic and policy

gaps, such as siloed organizational structures and limited regulatory guidance.

A notable difference between global and Philippine evidence lies in methodological rigor. International reviews and trials, while heterogeneous, include randomized and longitudinal designs that provide stronger evidence for the effects of IPE.[5,6] By contrast, most Philippine studies are cross-sectional surveys or descriptive case reports, which are valuable for baseline data but limited in establishing causality or long-term outcomes.[9,10,12] Measurement tools also present challenges. Instruments such as RIPLS and ATHCTS have been widely used, but concerns about subscale reliability and cultural fit remain, particularly in Philippine contexts. This indicates the need for adapting or developing validated tools that better capture interprofessional learning and collaboration in local settings.

Several implications arise from the findings. Firstly, faculty development remains a central requirement across contexts. International evidence shows that effective IPE relies on educators who can facilitate interprofessional learning and model collaborative behaviors.[8] Philippine studies echo this need, emphasizing that educators also shape student attitudes and readiness for collaboration.[10] Locally, structured programs combining workshops, mentorship, and competency-based evaluation could be developed, while globally, similar initiatives may be adapted to resource availability and cultural expectations. Monitoring through teaching evaluations and participation rates can serve as shared measures of accountability.

Secondly, curricular integration must move beyond optional or pilot activities. International frameworks have demonstrated that embedding IPE within required coursework leads to greater impact, yet translation into practice is uneven.[4] In the Philippines, community-based and institutional models have demonstrated feasibility but remain small-scale and resource-dependent.[11,12] Philippine institutions could mandate at least one simulation-based activity and one community-based project annually across health programs. Such models may also inform other low- and middle-income countries where experiential and community-oriented strategies offer cost-effective ways to normalize collaboration.

Thirdly, policy and accreditation support is critical for sustainability. Globally, competency frameworks have influenced accreditation standards, curricular reforms, and licensure examinations.[2,3] In the Philippine context, agencies such as the Commission on Higher Education, Professional Regulation Commission, and other accrediting bodies could embed interprofessional competencies into program outcomes and evaluation m echanisms, much like accrediting bodies in Canada, the UK, and Australia. For example, board exam blueprints could include interprofessional teamwork, aligning national policy with global trends toward collaborative-ready health professionals.[10]

Finally, sustainability requires institutional structures. Both global and Philippine experiences show that IPE initiatives often falter after pilot phases due to limited resources.[8,11] Establishing dedicated IPE coordination offices or committees, supported by research grants, alumni contributions, or inter-university partnerships, can ensure continuity. This model is transferable. Whether in the Philippines or elsewhere, designated leadership and sustainable funding mechanisms are essential for embedding IPE into long-term educational strategies.

Taken together, these implications demonstrate that while the Philippines shares global challenges in IPE and IPC, it also illustrates how international lessons can be adapted to local needs and, conversely, how Philippine experiences in community-based and resource-conscious strategies may inform broader global practice. By investing in faculty capacity, embedding interprofessional experiences in curricula, aligning accreditation and policy frameworks, and institutionalizing sustainable structures, both the Philippines and other countries can strengthen the preparation of a collaborative-ready health workforce.

Direction for future research

Several areas warrant further investigation. Firstly, longitudinal studies are essential to determine whether IPE experiences translate into sustained collaborative practice and measurable improvements in patient outcomes. International evidence suggests short-term gains in attitudes and teamwork,[5,6] but long-term effects remain unclear, particularly in low- and middle-income settings. Philippine studies are mostly cross-sectional and descriptive,[9,10,12] highlighting the need for designs that can assess outcomes over time.

Secondly, methodological diversification is required. While descriptive surveys and case reports provide valuable baseline insights, experimental and mixed-methods designs would strengthen the evidence base. Randomized and quasi-experimental studies conducted internationally could serve as models for Philippine and regional research, adapted to account for local constraints and contexts.[5]

Thirdly, instrument development should be prioritized. Tools such as the RIPLS and ATHCTS have been widely used internationally, yet their cultural fit and subscale reliability remain debated.[6] Developing or adapting validated instruments for the Philippine context would provide more reliable measures of attitudes, readiness, and collaboration quality, with potential applicability across other low- and middle-income countries facing similar cultural and resource considerations.

Finally, sustainability must be a focus of future inquiry. Evidence shows that many pilot initiatives, both globally and locally, do not extend beyond initial phases due to faculty shortages, resource limitations, and weak institutional leadership.[8,11] Future studies

should examine how leadership, policy frameworks, and resource allocation influence the long-term viability of IPE programs. Comparative research across different health systems may also highlight transferable strategies for embedding IPE into institutional and policy structures.

Limitations

This review followed a narrative approach, which offers flexibility in synthesizing diverse forms of evidence but does not aim to be exhaustive. While multiple databases and sources were considered, some relevant studies may not have been included due to language, indexing, or publication constraints. The review also relied on published literature, which may introduce publication bias by underrepresenting negative or inconclusive findings. Another limitation is the subjectivity inherent in narrative reviews. The selection and synthesis of studies were influenced by the perspectives of the authors, and different reviewers might interpret the same evidence differently. Finally, while efforts were made to balance global and Philippine perspectives, the relatively small number of local studies limited the depth of analysis specific to the Philippine context. These limitations highlight the need for cautious interpretation of findings and underscore the importance of continued research to expand and refine the evidence base.

Conclusion

This review synthesized global and Philippine literature on IPE and IPC. The findings reaffirm that IPE can foster collaborative competencies and, in some cases, contribute to improved health outcomes. International frameworks and reviews provide a strong conceptual and evidentiary foundation, while Philippine studies, though limited, demonstrate readiness among students and professionals and highlight feasible approaches to implementation. Persistent barriers, including professional hierarchies, faculty shortages, and resource constraints, are evident across settings and are compounded locally by systemic challenges.

Moving forward, contextual adaptation of global frameworks, investment in faculty development, and alignment of policy and accreditation standards are necessary to strengthen IPE in the Philippines. More rigorous and longitudinal research is also required to establish sustainable models and assess their long-term impact on practice and patient care. By integrating international lessons with local realities, IPE and IPC can be advanced to prepare a collaborative health workforce that is responsive to both national and global health needs.

Funding

This work is funded by Velez College as part of the faculty institutional research program.

Conflict of interest

No potential conflict of interests was reported by the author(s).

CRediT authorship contribution statement

Paolo Miguel P. Bulan: writing: the first draft of the manuscript. All authors researched literature, conceived the study, and were involved in protocol development, reviewed and edited the manuscript and approved the final version of the manuscript.

Acknowledgements

The authors would like to acknowledge the support of Velez College.

References

[1] World Health Organization. Framework for action on interprofessional education and collaborative practice [Internet]. Geneva: WHO; 2010 [cited 2024 Jan 6]. Available from: https://www.who.int/hrh/resources/framework_action/en/

[2] Canadian Interprofessional Health Collaborative. A national interprofessional competency framework. Vancouver: University of British Columbia; 2010.

[3] Interprofessional Education Collaborative Expert Panel. Core competencies for interprofessional collaborative practice: report of an expert panel. Washington (DC): Interprofessional Education Collaborative; 2011.

[4] Thistlethwaite JE, Forman D, Matthews LR, Rogers GD, Steketee C, Yassine T. Competencies and frameworks in interprofessional education: a comparative analysis. Acad Med. 2014; 89(6): 869-75. doi: 10.1097/ACM.0000000000000249.

[5] Reeves S, Pelone F, Harrison R, Goldman J, Zwarenstein M. Interprofessional collaboration to improve professional practice and healthcare outcomes. Cochrane Database Syst Rev. 2017; (6): CD000072. doi: 10.1002/14651858.CD000072.pub3.

[6] Berger-Estilita J, Fuchs A, Hahn M, Chiang H, Greif R. Attitudes towards interprofessional education in the medical curriculum: a systematic review of the literature. BMC Med Educ. 2020; 20(1): 254. doi: 10.1186/s12909-020-02120-6.

[7] Lieneck C, Wang T, Gibbs D, Russian C, Ramamonjiarivelo Z, Ari A. Interprofessional education and research in the health professions: a systematic review and supplementary topic modeling. Educ Sci. 2022; 12(12): 850. doi: 10.3390/educsci12120850.

[8] Bogossian F, New K, George K, Barr N, Dodd N, Hamilton AL, et al. The implementation of interprofessional education: a scoping review. Adv Health Sci Educ Theory Pract. 2023; 28(1): 243-77. doi: 10.1007/s10459-022-10132-4.

[9] Cervantes-Sudio MG, Ganotice FA, Navarro AT. Are Filipino students ready to collaborate?

Comparing the readiness of healthcare students for interprofessional education in the Philippines. J Interprof Care. 2021; 35(5): 718-25. doi: 10.1080/13561820.2020.1819241.

[10] Sy MP. Filipino therapists' experiences and attitudes of interprofessional education and collaboration: a cross-sectional survey. J Interprof Care. 2017; 31(6): 761-70. doi: 10.1080/13561820.2017.1345873.

[11] Sy M, Pineda RC, Sumulong RA, Navarro A. Establishing a pilot interprofessional education program in a higher education institution in the Philippines. J Health Caring Sci. 2020;2(2):180-91. doi: 10.37719/jhcs.2020.v2i2.rna004.

[12] Paterno E, Opina-Tan LA. Developing community-engaged interprofessional education in the Philippines. In: Gilbert J, Editor. Leadership development for interprofessional education and collaborative practice. London: Palgrave Macmillan; 2014: pp162-78.

[13] Sadikan MZ. A step-by-step guide to writing a review paper in medical education. Int J Transform Health Prof Educ. 2025; (1): 24-37. doi: 10.71354/ijthpe.03.01.48.

[14] Sukhera J. Narrative reviews: flexible, rigorous, and practical. J Grad Med Educ. 2022; 14(4): 414-7. doi: 10.4300/JGME-D-22-00400.1.

[15] Green BN, Johnson CD, Adams A. Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. J Chiropr Med. 2006; 5(3): 101-17.

[16] Elo S, Kääriäinen M, Kanste O, Pölkki T, Utriainen K, Kyngäs H. Qualitative content analysis. SAGE Open. 2014; 4(1): 215824401452263. doi: 10. 1177/2158244014522633.

[17] Brewer M. Curtin interprofessional capability framework [Internet]. Perth: Curtin University; 2013 [cited 2024 Jan 6]. Available from: https://www.curtin.edu.au/resources/file/faculty/hs/interprofessional_A5_broch_1-29072015.pdf

[18] Walsh CL, Gordon MF, Marshall M, Wilson F, Hunt T. Interprofessional capability: a developing framework for interprofessional education. Nurse Educ Pract. 2005; 5(4): 230-7. doi: 10.1016/j.nepr.2004.12.003.

[19] Rawlinson C, Carron T, Cohidon C, Arditi C, Hong QN, Pluye P, et al. An overview of reviews on interprofessional collaboration in primary care: barriers and facilitators. Int J Integr Care. 2021; 21(2): 32. doi: 10.5334/ijic.5589.

[20] Schot E, Tummers L, Noordegraaf M. Working on working together: a systematic review on how healthcare professionals contribute to interprofessional collaboration. J Interprof Care. 2020; 34(3): 332-42. doi: 10.1080/13561820.2019.1636007.

[21] Van Wyk R, Labuschagne MJ, Joubert G. Simulation as an educational strategy to deliver interprofessional education. Afr J Health Prof Educ. 2020; 12(2): 74. doi: 10.7196/AJHPE.2020.v12i2.1213

[22] El-Awaisi A, Awaisu A, Jaam M, El Hajj MS, Verjee MA. Does the delivery of interprofessional education have an effect on stereotypical views of healthcare students in Qatar? J Interprof Care. 2020; 34(1): 44-9. doi: 10.1080/13561820.2019.1612332.

[23] Curran VR, Heath O, Kearney A, Button P. Evaluation of an interprofessional collaboration workshop for postgraduate residents, nursing and allied health professionals. J Interprof Care. 2010; 24(3): 315-8. doi: 10.3109/13561820903163560.

[24] Ford M, Martin P, Sy M. Twelve tips to support healthcare teams to incorporate interprofessional education and collaborative practice into day-to-day workplace practices. J Pract Teach Learn. 2022; 19(3): 183-90. doi: 10.1921/jpts.v19i3.1835.

[25] Ulrich G, Homberg A, Karstens S, Mahler C. Attitudes towards interprofessional collaboration in young healthcare professionals. J Interprof Care. 2019; 33(6): 768-73. doi: 10.1080/13561820.2019.1612333.

[26] Heinemann GD, Schmitt MH, Farrell MP, Brallier SA. Development of an attitudes toward health care teams scale. Eval Health Prof. 1999; 22(1): 123-42. doi: 10.1177/01632789922034202.

[27] Moncatar TRT, Nakamura K, Siongco KLL, Seino K, Carlson R, Canila CC, et al. Interprofessional collaboration and barriers among health and social workers caring for older adults: a Philippine case study. Hum Resour Health. 2021; 19(1): 52. doi: 10.1186/s12960-021-00601-8.

[28] Nakamura K, Siongco KLL, Moncatar TRT, Tejero LMS, De La Vega SAF, Bonito SR, et al. In-service training programme for health and social care workers in the Philippines to strengthen interprofessional collaboration in caring for older adults: a mixed-methods study. Health Res Policy Syst. 2022; 20(S1): 111. doi: 10.1186/s12961-022-00914-2.

[29] Sy MP, Martinez PGV, Labung FFT, Medina MAKG, Mesina AS, Vicencio MRE, et al. Baseline assessment on the quality of interprofessional collaboration among Filipino mental health professionals. J Interprof Educ Pract. 2019; 14: 58-66. doi: 10.1016/j.xjep.2018.10.002.

[30] Causapin AKV, Sy MP, Grageda MEM. Attitudes of Filipino health profession students toward interprofessional education: a descriptive study. J Interprof Care. 2022; 36(6): 856-63. doi: 10.1080/13561820.2021.1971448.

Supplementary material
Data extraction chart

Author (year)	Country/Region	Study design/Type	Population/Context	Key findings/Themes	Relevance to IPE/IPC
Berger-Estilita J, *et al*. (2020)[6]	Multi-country (USA, Europe, Asia, Australia, New Zealand)	To systematically review IPE interventions in medical curricula and evaluate changes in medical students' attitudes, identifying the most effective timing for IPE introduction.	Systematic review of 23 studies (2011-2019); databases: PubMed, PsycINFO, EMBASE, SCOPUS, PEDro, EThOS. Inclusion: validated IPE scale, ≥35 medical students, post-2011 IPEC. Registered PROSPERO CRD42020160964. Quality appraisal with McMaster Critical Review Form.	5,231 students total, 3,229 (62%) medical students across preclinical and clinical years	About half of studies showed significant improvement in attitudes toward IPE. Interventions varied (simulation, workshops, case-based, community). RIPLS most used but problematic. Both early and late interventions effective; later longer ones more impactful. No clear optimal timing. Heterogeneity prevented meta-analysis.
Bogossian F, *et al*. (2023)[8]	International (27 studies from multiple countries)	To review how interprofessional education (IPE) has been implemented, including strategies, challenges, and lessons learned.	Scoping review of 27 studies (Arksey & O'Malley framework). Included descriptive and case studies on IPE implementation. Data mapped to micro, meso, macro levels and sustainability.	Health professions students, educators, and institutions (varied disciplines, multi-country)	Most IPE initiatives were curricular, optional, group-based, and interactive. Micro-level challenges: silos, resistance, stereotypes, unequal status, faculty gaps. Meso-level: leadership, resources, scheduling, admin barriers. Macro-level: accreditation, policy/regulation gaps, hidden power structures. Sustainability issues: lack of infrastructure, funding, long-term planning; success tied to stakeholder buy-in.
Brewer M. (2011, reprint 2015)[17]	Australia (Curtin University, Perth, WA)	To provide a structured framework for teaching and assessing interprofessional capabilities required for graduates to be collaborative practice-ready.	Framework development informed by Sheffield Hallam (2010), CIHC (2010), Barr, Hammick, Freeth, WHO (2010), and expert consensus at Curtin University.	Not empirical; intended for health science students, educators, and practitioners	Framework includes three elements: client/family/community-centred care, client safety & quality, collaborative practice. Five capabilities: communication, team function, role clarification, conflict resolution, reflection. Capabilities described across three levels: novice, intermediate, entry-to-practice. Promotes outcomes only achievable via interprofessional activity.

Author (year)	Country/Region	Study design/Type	Population/Context	Key findings/Themes	Relevance to IPE/IPC
Canadian Interprofessional Health Collaborative (CIHC), 2010[2]	Canada (national-level framework, adaptable internationally)	To develop a competency framework for interprofessional education (IPE) and collaborative practice (IPC) applicable across health professions in Canada and adaptable globally.	Literature review, synthesis of existing frameworks, stakeholder consultations, and expert group development funded by Health Canada.	Not empirical; applies to health professionals, students, educators, regulators, and organizations	Six domains: role clarification, patient/client/family/community-centred care, team functioning, collaborative leadership, interprofessional communication, conflict resolution. Competencies integrate knowledge, skills, attitudes, and values; developmental across career; underpinned by complexity, context, and quality improvement; framework guides education, practice, regulation, and organizational change.
Causapin AKV, *et al*. (2022)[30]	Philippines	To describe attitudes of health profession students toward IPE and identify variables (sex, program, year level, clinical/IPE experience, moral development) influencing these attitudes.	Quantitative descriptive study; survey of 485 medical, nursing, and pharmacy students from a private university using IPE-Attitude Scale (IPE-AS) and SPICE-R; stratified random sampling.	Undergraduate students (medicine, nursing, pharmacy)	Attitudes toward IPE were overall favorable. More positive attitudes observed among medical students, lower-year students, and those with higher moral development. Sex and prior clinical/IPE experiences not significant predictors.
Cervantes-Sudio MG, *et al*. (2021)[9]	Philippines (Angeles University Foundation, private university)	To examine readiness for interprofessional learning among Filipino healthcare students and to determine differences across program, year level, and gender.	Cross-sectional survey using RIPLS. Participants: 423 students in medical technology, pharmacy, PT, and OT. Analysis: MANOVA, ANOVA, and t-tests.	Filipino healthcare students (medical technology, pharmacy, PT, OT); pre-licensure; N=423	Program differences: MT students lower teamwork/collaboration, higher negative identity; PT/OT lower roles/responsibilities. Year-level differences: seniors more favorable attitudes. Gender differences: males higher negative identity and roles, lower positive identity. Some RIPLS subscales had low reliability.
Curran VR, *et al*. (2010)[23]	Canada (Memorial University of Newfoundland & Eastern Regional Integrated Health Authority)	To evaluate a one-day IPC workshop for postgraduate residents, nursing staff, and allied health professionals, and assess effects on satisfaction and attitudinal outcomes.	Pre- and post-test evaluation; eight workshops (2006-2007); ATHCTS and PEITS administered pre, post, and 6-8 week follow-up.	137 participants: 34 residents, 46 allied health professionals, 33 nursing staff, 20 others	88.5% satisfied. Residents showed significant positive change in teamwork attitudes (Cohen's d=0.45). Nursing and allied health: no significant attitude change but high satisfaction. No significant PEITS score changes. Need for role clarification and organizational support identified.
El-Awaisi A, *et al*. (2020)[22]	Qatar (Qatar University and Weill Cornell Medicine-Qatar)	To investigate whether healthcare students' stereotypical beliefs change before and after an IPE activity on smoking cessation.	Pre-post intervention; Student Stereotype Rating Questionnaire (SSRQ); open-ended questions; analysis with Wilcoxon Signed Rank Test and thematic content analysis.	68 healthcare students (medicine, nursing, pharmacy, public health, respiratory therapy)	Significant stereotype rating changes for medicine, pharmacy, and respiratory therapy students across all attributes. Public health students improved in competence and skills. Themes: clarified roles, patient-centered care, integrated teamwork. IPE reduced negative stereotypes.

Author (year)	Country/Region	Study design/Type	Population/Context	Key findings/Themes	Relevance to IPE/IPC
Ford M, *et al*. (2022)[24]	International (authors from Australia and the Philippines; global applicability)	To provide healthcare teams with practical strategies to implement IPECP in everyday workplace practice.	Expert-informed guidance based on literature review, experiential knowledge, and synthesis of best practices; non-empirical.	Healthcare teams and workers across professions (students, recent graduates, educators, clinical supervisors, practitioners)	Presents 12 tips: invest in champions, clarify team structures/roles, adopt shared language, peer engagement, mentoring, use resources, create opportunities, CPD with follow-up, maximize technology, incorporate client feedback, storytelling, embed IPECP in placements. Emphasizes adaptability, sustainability, bridging evidence–practice gap.
Heinemann GD, *et al.* (1999)[26]	United States (VA Western New York Healthcare System, University at Buffalo, University of Rochester)	To develop and psychometrically test the Attitudes Toward Health Care Teams Scale (ATHCTS) to measure professionals' perceptions of teamwork and physician centrality.	Instrument development and testing in 3 phases: Phase 1 (pilot, N=147, factor analysis), Phase 2 (revision, CVI=0.95, N=132), Phase 3 (national VA study, N=973, geriatric teams). Final scale: 21 items, 2 subscales (Quality of Care/Process, Physician Centrality).	Health professionals across disciplines (nurses, physicians, social workers, allied health) including geriatric care teams	Scale showed strong validity/reliability. Quality of Care/Process subscale reliable (α=0.83). Physician Centrality subscale identified differences across professions. Recommended for pre/post IPE and team training assessment.
Interprofessional Education Collaborative (IPEC), 2011[3]	United States (national-level framework)	To establish core competencies for interprofessional collaborative practice among US. health professions, guiding curricula and improving healthcare outcomes.	Consensus-building by expert panel from six US professional associations; informed by literature review, stakeholder input, and comparative analysis of frameworks.	Not empirical; applies to students, educators, practitioners, and health systems	Four competency domains: Values/Ethics, Roles/Responsibilities, Interprofessional Communication, Teams/Teamwork. Each domain includes statements on knowledge, skills, attitudes. Framework emphasizes patient-centered, community-oriented collaboration. Guides curriculum, accreditation, faculty development, and system integration.
Lieneck C, *et al*. (2022)[7]	International (multi-country review)	To systematically review IPE initiatives in health professions higher education and explore barriers, facilitators, and stakeholder perspectives, supplemented with topic modeling.	Systematic review (PRISMA 2020); databases: Medline, Scopus, CINAHL, Complementary Index (2015-2022). Inclusion: peer-reviewed, English, higher education. 41 studies. Supplementary Latent Dirichlet Allocation (LDA) topic modeling.	Health professions students, faculty, and professionals in higher education	Three perspectives: providers, learners, researchers. Providers: barriers (resources, workload, willingness); themes: patient experience, teamwork. Learners: positive attitudes; influenced by relevance, class size, patient-centered focus; student-led valued. Researchers: scarce evidence; facilitators include organizational support, mentorship, collaboration. Topic modeling confirmed constructs: medical/clinical skills, group training, simulation/learning, research support.

Author (year)	Country/Region	Study design/Type	Population/Context	Key findings/Themes	Relevance to IPE/IPC
Moncatar TRT, *et al.* (2021)[27]	Philippines (two cities, health and social workers)	To describe interprofessional collaboration practices among health and social workers in caring for older adults in the Philippines, and to explore perceived barriers and solutions.	Qualitative case study; 12 in-depth interviews and 29 FGDs (N=174 participants); thematic analysis using NVivo.	Health and social workers (nurses, physicians, midwives, barangay health workers, social workers, etc.) caring for older adults in the Philippines	IPC was largely informal, ad hoc, and administrative. Barriers: 1) personal values/beliefs—hierarchies, lack of trust, undervaluing staff, 2) organizational constraints—staff shortages, funding/ time/equipment gaps, 3) silo systems culture—fragmented structures, poor leadership, absent policies. Consequences: delayed care, duplication, inequitable services. Solutions: integrated health/social care, governance, leadership, policy development, and IPE.
Nakamura K, *et al.* (2022)[28]	Philippines (Marikina and Tagaytay)	To introduce and evaluate a pilot IPC training program for health and social care workers in geriatric care, measuring changes in attitudes, readiness, and practices.	Mixed-methods, convergent parallel design. Quantitative: pre/ post-test with ATHCTS, RIPLS, CAES. Qualitative: open-ended questionnaires, inductive content analysis (NVivo).	42 community health workers and 40 institutional health workers (nurses, physicians, nutritionists, rehab therapists, caregivers)	Quantitative: improved attitudes (ATHCTS), readiness (RIPLS), and coordinated activities (CAES) post-training. Qualitative: improved self-awareness, role understanding, patient-centred care, teamwork. Active learning (case studies, role play, group work) reinforced collaboration skills.
Paterno ER, *et al.* (2014)[12]	Philippines (UP Manila & partner rural municipality)	To document the development, implementation, and evaluation of community-engaged IPE in UP Manila's Community Health and Development Program, and to identify leadership, faculty, student, and organizational factors shaping the process.	Case study and descriptive evaluation (2007-2012). Methods: faculty workshops, community partnership agreements, IDA guidelines, evaluation surveys of students and stakeholders.	UP Manila students (medicine, nursing, public health, dentistry, pharmacy, OT, PT, speech pathology, nutrition); faculty from UP Manila & UP Diliman; rural midwives, village health workers; patients/families in rural municipality	Implementation evolved from parallel to coordinated collaboration. IDA Guidelines clarified shared care and documentation. Challenges: professional biases, faculty shortages, scheduling, records, limited resources. Faculty co-location reduced hierarchies. Student evaluation (N=39) showed improved appreciation of roles, communication, patient involvement, and comfort with collaboration. Community valued improved awareness, QoL, empowerment. Case study (child with developmental delay) illustrated success.

Author (year)	Country/Region	Study design/Type	Population/Context	Key findings/Themes	Relevance to IPE/IPC
Rawlinson C, *et al*. (2021)[19]	International (reviews synthesized from diverse countries)	To identify barriers and facilitators of IPC in primary care by synthesizing systematic reviews.	Overview of reviews (umbrella review); 9 databases searched; 29 reviews included; thematic synthesis of barriers/facilitators at system, organizational, inter-individual, and individual levels.	Healthcare professionals in primary care (physicians, nurses, pharmacists, mental health providers, specialists, allied health, community/public health)	22 barriers and 20 facilitators identified across six IPC types. Barriers: time/resource constraints, poor communication, unclear roles, identity threats, hierarchies, financial/structural issues. Facilitators: co-location, leadership, communication tools, shared vision, trust, organizational restructuring, training, incentives. Barriers more consistent across contexts; facilitators context-specific.
Reeves S, *et al*. (2017)[5]	International (15 studies across various high- and middle-income countries)	To assess the impact of interprofessional education (IPE) interventions compared to separate or no education on professional practice and healthcare outcomes.	Systematic review (update); included randomized controlled trials (RCTs), controlled before-and-after studies, and interrupted time series. Data synthesis with risk of bias assessment.	Health professionals and students across medicine, nursing, pharmacy, dentistry, social work, and allied health	15 studies included; some evidence IPE improves patient outcomes (e.g., diabetes care, emergency department management, collaborative clinical care). Positive effects on teamwork, communication, prescribing. Evidence low to moderate quality, with heterogeneity and methodological limits. Conclusion: IPE can improve collaboration and outcomes, but stronger studies needed.
Schot E, *et al*. (2020)[20]	International (64 studies from multiple countries, 2001-2017)	To synthesize empirical evidence on how healthcare professionals actively contribute to IPC in various healthcare settings.	Systematic review of 64 empirical studies (mostly qualitative); covered hospital, primary care, mental health, and cross-sectoral settings.	Healthcare professionals across multiple disciplines (physicians, nurses, allied health, social workers, etc.)	Three main contributions identified: 1) Bridging gaps (perspectives, relations, communication, tasks), 2) Negotiating overlaps (roles, responsibilities, decision-making), 3) Creating spaces (workarounds, new arrangements). Differences noted by profession, setting, sector. Effects inferred: improved teamwork, decision-making, care integration; risks in informal workarounds.
Sy MP. (2017)[10]	Philippines (nationwide survey of OTs, PTs, and SLPs)	To describe the extent of IPE experiences among Filipino therapists, examine their attitudes toward interprofessional collaboration (IPC), and compare attitudes across IPE experience, profession, years of practice, and practice setting.	Cross-sectional online survey (2015-2016); respondent-driven sampling via associations. Instrument: ATHCTS. N=189 (100 OTs, 56 PTs, 33 SLPs). Analysis: Mann–Whitney U, Kruskal–Wallis, descriptive statistics.	Licensed Filipino OTs, PTs, and SLPs practicing in the Philippines (N=189)	70.9% had prior IPE experience (mostly mandatory). Common IPE strategies: case discussions, SGD, didactics. Attitudes generally favorable (11/14 items). Prior IPE experience correlated with positive IPC attitudes (p=0.01). Didactics and community-based case discussions linked to more positive IPC attitudes. Practice setting significant (academia > others, p=0.02); profession and years not significant.

Author (year)	Country/Region	Study design/Type	Population/Context	Key findings/Themes	Relevance to IPE/IPC
Sy MP, *et al*. (2019)[29]	Philippines	To assess the quality of IPC among Filipino mental health professionals (FMHP) and determine variables affecting IPC quality.	Quantitative descriptive survey using CPAT-R; respondent-driven sampling; data analysis with descriptive statistics, Mann-Whitney U, and Kruskal-Wallis tests.	Licensed Filipino mental health professionals across nine professions (OTs, psychiatrists, nurses, psychologists, social workers, counselors, etc.)	Sample size: 51 (44 with IPC experience). Baseline IPC quality revealed issues in communication, role clarification, and conflict management. Provided insights into strengths and gaps in IPC among FMHP.
Sy MP, *et al*. (2020)[11]	Philippines (Angeles University Foundation, Central Luzon)	To develop, implement, and evaluate a pilot IPE program at Angeles University Foundation (AUF), guided by the WHO (2010) framework and adapted to the Philippine context.	Descriptive case report using a four-step action plan (2015-2018): program development, admin negotiation, pilot implementation (10 programs), evaluation with CIPP model.	Faculty and students from 10 AUF health science programs (medicine, nursing, pharmacy, medical technology, radiologic technology, PT, OT, etc.); administrators and faculty champions	Pilot program included lectures, case vignettes, and teamwork. Challenges: scheduling, bureaucracy, managerial commitment, limited resources. Facilitators: trained faculty, admin support, leadership. Evaluation (CIPP) showed: Context—need for IPE; Input—case vignettes, faculty development; Process—barriers in logistics; Product—positive attitudes, interest in IPE research, PHIPEC Network formed. Sustainability via inter-university partnerships.
Thistlethwaite JE, *et al*. (2014)[4]	International (UK, Canada, USA, Australia)	To critically appraise competency frameworks in IPE and compare four published frameworks to highlight similarities, differences, and implications for practice.	Conceptual analysis and comparative review of IPE competency frameworks (literature review and policy analysis). Frameworks analyzed: UK, Canadian, US, Curtin University.	Not empirical—focus on frameworks relevant to health professions education internationally	Competency frameworks standardize IPE outcomes but vary in terminology and domains. Common areas: teamwork, communication, roles, ethics, reflection. Confusion in terms (competence, competency, capability). Assessment strategies underdeveloped. Frameworks must complement curricula and emphasize interprofessional-only outcomes.
Ulrich G, *et al*. (2019)[25]	Germany (Academy of Health Professions, Heidelberg)	To evaluate and compare attitudes toward IPC and IPE among young healthcare professionals shortly after graduation, without prior formal IPE exposure.	Cross-sectional online survey (2013-2014); 387 invited, 129 analyzed (33.3% response). Used validated UWE-IP questionnaire (German version).	Recent graduates (12-23 months post-qualification): nursing (N=77), therapy professions (PT, SLT; N=21), diagnostics (biomedical science, radiography; N=31).	Positive attitudes in Communication & Teamwork and Interprofessional Relationships; neutral in Interprofessional Learning and Interprofessional Interactions. No significant differences across professions. Age, gender, and education level influenced attitudes (older, male, higher school education less positive). Highlights need for IPE in continuing education early in careers.

Author (year)	Country/Region	Study design/Type	Population/Context	Key findings/Themes	Relevance to IPE/IPC
van Wyk R, *et al*. (2020)[21]	South Africa (University of the Free State, Bloemfontein)	To identify current IPE methods in undergraduate health program at UFS and determine module leaders' opinions on simulation as a teaching strategy.	Quantitative cross-sectional descriptive study; structured interviews with 47 module leaders covering 66 modules across medicine, nursing, and allied health.	Module leaders in medicine, nursing, nutrition, dietetics, OT, PT, optometry	IPE present in 29/66 modules (43.9%), often coincidental. Leaders positive toward simulation. Challenges: scheduling (73.9%), logistics/cost (19.6%), staff attitudes (21.7%). Advantages: role clarification (41.3%), safe environment (32.6%), improved empathy/respect/teamwork. Simulation forms: role-play (77.3%), standardized patients (63.6%), skills training (31.8%). Proposed 3-phase longitudinal IPE: didactic, simulation, community-based.
Walsh CL, *et al*. (2005)[18]	United Kingdom (Sheffield Hallam University & University of Sheffield, CUILU project)	To develop an Interprofessional Capability Framework that articulates learning outcomes students must achieve to become capable interprofessional workers.	Framework development using grounded theory strategies applied to QAA benchmark statements across medicine, dentistry, nursing, social work, and allied health. Included coding, categorization, and mapping to align with professional competencies and NHS KSF.	Health and social care students (conceptual model development; not empirical participants)	Developed a four-domain Interprofessional Capability Framework: Ethical Practice, Knowledge in Practice, Interprofessional Working, Reflection. Identified 16 capabilities (respect for cultures, understanding team structures, communication, co-mentoring, reflection). Framework emphasizes practice-based learning, adaptability, reflection, and capability over static competence. Supports curricular integration.
WHO (2010)[1]	Global (42 countries across 6 WHO regions)	To provide policymakers with a framework and strategies for implementing interprofessional education (IPE) and collaborative practice (CP) to strengthen health systems and improve health outcomes.	Policy analysis, literature review, international environmental scan (396 respondents, 42 countries), synthesis of case studies, and expert consultations through the WHO Study Group.	Health professions students, educators, policymakers, and health workers across medicine, nursing, midwifery, allied health, and social care	IPE prepares a collaborative practice-ready workforce, improving teamwork and patient outcomes. Collaborative practice enhances access, quality, safety, and system efficiency. Mechanisms influencing IPE/CP include educator training, curricular design, governance models, culture, and legislation. Evidence shows improved patient satisfaction, reduced hospital stays, lower costs, and stronger health systems.

Journal of Associated Medical Sciences 2026; 59(2): 250-258

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Advanced YOLOv8 architecture for multi-class brain tumor detection

Kavita Singh, Rajesh Kumar*, Satyendra Singh

Department of Electronics and Communication, University of Allahabad, India.

ARTICLE INFO

Article history:
Received 25 August 2025
Accepted as revised 9 February 2026
Available online 25 February 2026

Keywords:
Brain tumor, classification, YOLO v8, tumor detection, MRI, mean average precision

ABSTRACT

Background: Accurate brain tumor detection is critical for early diagnosis and effective treatment planning in neuro-oncology. Magnetic resonance imaging (MRI) is a cornerstone for identifying and localizing brain tumors, guiding clinical interventions, and enhancing patient outcome. Advanced deep learning models, such as YOLOv8, offer promising solutions for automated and precise tumor detection in MRI.

Objectives: This study aimed to develop and evaluate a YOLOv8-based model for the accurate identification and localization of brain tumors, including pituitary tumors, meningiomas, and gliomas, using a publicly available Kaggle dataset of annotated MRI images.

Materials and methods: The YOLOv8 pretrained model was employed for brain tumor detection on a Kaggle dataset comprising annotated MRI images, including cases with pituitary, meningioma, and glioma, and no tumor. The dataset was pre-processed and split into training and validation sets for further analysis. The YOLOv8 model was fine-tuned to optimize the tumor detection and localization. Performance metrics, including precision, recall, F-1 score, and mean average precision (mAP), were calculated, and loss values were analyzed to evaluate the model's learning efficiency.

Results: The YOLOv8 model achieved a precision of 98.9%, recall of 98.9%, and accuracy of 99.5%. The mean average precision (mAP) reached 97.6%, indicating a high accuracy in detecting and localizing brain tumors. Loss value analysis demonstrated stable convergence during training, reflecting a robust model performance.

Conclusion: The YOLOv8-based approach provides a highly accurate and reliable method for detecting and localizing brain tumors in MRI. With exceptional precision, recall, and mAP, this model demonstrates significant potential for clinical applications, enabling faster and more precise neuro-oncological diagnosis and treatment planning.

Introduction

In the field of neurological disorders, brain tumors represent a significant and often life-threatening challenge. The prognosis of patients and assessment of treatment efficacy are critically dependent on early diagnosis and accurate localization.[1,2] Cancer has profoundly affected many lives, and most individuals have encountered someone who has suffered from its severe consequences.[3] Brain tumors are abnormal growths that develop within the brain or surrounding brain tissue. They can affect individuals of any age or sex, although certain risk factors may increase their likelihood. Brain tumor treatment typically involves a multidisciplinary approach, including chemotherapy,

* **Corresponding contributor.**
Author's Address: *Department of Electronics and Communication, University of Allahabad, India.*
E-mail address: rajeshkumariitbhu@gmail.com
doi: 10.12982/JAMS.2025.056
E-ISSN: 2539-6056

radiation therapy, or surgery, depending on the tumor type, location, and size. The effectiveness of treatment is contingent on early, accurate, and reliable detection, as it facilitates prompt intervention and significantly enhances the probability of a favorable prognosis.[4] Figure 1 shows a sample image of a brain tumor.

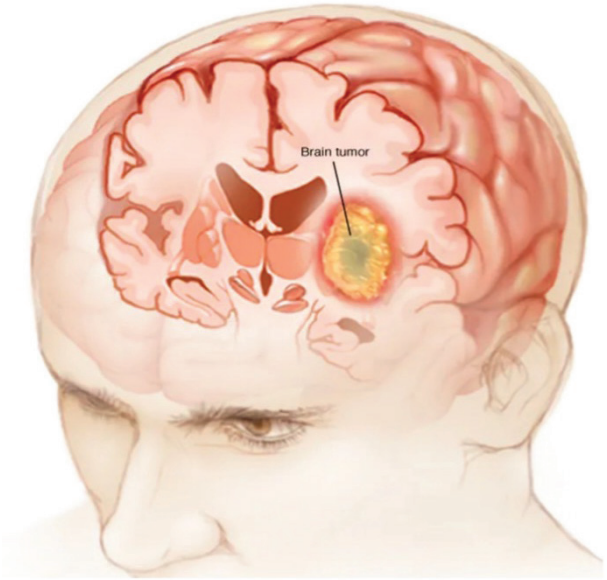


Figure 1. *Sample image of brain tumor.*

The anatomy of the human brain is highly complex, necessitating precise structural knowledge for the detection of anomalies. Various imaging techniques, such as computed tomography (CT), positron emission tomography (PET), single-photon emission computed tomography (SPECT), and magnetic resonance imaging (MRI), are employed by medical professionals to diagnose brain abnormalities.[5] Among these, MRI is the most prevalent technique because of its non-invasive nature and ability to produce high-resolution images that reveal the intricate anatomical details of the brain.[6]

The efficacy of treatment planning and improvement of patient outcomes is largely contingent upon early and accurate diagnosis. Medical imaging techniques play a crucial role in this process, particularly MRI, which is the most used because of its non-invasive nature and ability to generate high-resolution images of brain structures. Other imaging modalities include CT, PET, and SPECT.

This study focused on three primary types of brain tumors: gliomas, meningiomas, and pituitary tumors, in addition to a no-tumor category. Understanding the nature of these tumor types is essential for accurate detection and diagnosis.

Gliomas: Gliomas are tumors that originate from glial cells, which play a crucial role in supporting and protecting neurons in the brain. These tumors are often aggressive and present treatment challenges owing to their infiltrative nature. Gliomas represent the most prevalent type of primary brain tumor.

Meningioma: Meningiomas are tumors that develop in the meninges, the protective membranes covering the brain and spinal cord. Although typically benign and slow growing, meningiomas can exert significant neurological effects by exerting pressure on adjacent brain tissue.

Pituitary tumors: These tumors occur in the pituitary gland; a small but vital gland located at the base of the brain responsible for hormone regulation. Pituitary tumors can be benign or malignant and may produce a range of symptoms depending on their size and hormone secretion activity.

The state-of-the-art object detection method, You Only Look Once (YOLO),[7] is acclaimed for its ability to perform high-accuracy, real-time detection, particularly in complex scenarios. This capability makes YOLO a suitable choice for analyzing magnetic resonance imaging (MRI) scans. Moreover, YOLO and similar object detection techniques facilitate precise localization within images in addition to tumor diagnosis. In this study, we implemented YOLO v8 for brain tumor detection. The remainder of this paper is structured as follows: Section 2 presents a review of related state-of-the-artwork; Section 3 describes the proposed approach; Section 4 presents the experimental results; and Section 5 concludes with a discussion of future scope.

Literature Review

Author of this work employed YOLOv8, YOLOv11, and a custom convolutional neural network (CNN) model to accurately and efficiently identify brain malignancies, specifically pituitary, glioma, and meningioma, using artificial intelligence.[8] The models

leverage transfer learning and fine-tuning to classify magnetic resonance imaging (MRI) images into four categories, including "No Tumor." The models were trained using the publicly available CE-MRI Figshare Dataset. The custom CNN achieved an accuracy of 96.98%, whereas YOLOv11 and YOLOv8 attained remarkable accuracies of 99.56% and 99.49%, respectively. These results suggest that deep learning can substantially enhance brain tumor diagnosis and classification.

In this paper, further explored the efficacy of YOLOv8 in localizing brain tumors by analyzing the impact of model size and pretraining.[1] Experiments conducted on an NVIDIA Tesla T4 GPU using Google Colab revealed that pretrained models, particularly smaller ones, such as YOLOv8-XS, exhibited significantly higher accuracy than their non-pretrained counterparts. For example, YOLOv8-XS pretrained on the COCO dataset improved the Intersection over Union (IoU) and Dice scores by 15%. Additionally, the findings indicate that while pre-training enhances performance, its benefits diminish as the model size increases.

Work in this paper presents an enhanced YOLOv8 model for the precise identification of brain tumors in MRI data.[9] Key improvements include the integration of a vision transformer for deeper feature extraction, the use of ghost convolution to reduce computational demands, and the replacement of RT-DETR with NMS for a superior bounding box selection. When evaluated on a publicly available brain tumor dataset, the model outperformed other object detectors and standard YOLO versions, achieving a remarkable mean average precision (mAP@0.5) of 0.91, thereby demonstrating its high accuracy and efficiency.

The application of deep learning for rapid and precise malignant tumor diagnosis on mobile devices are emphasized in this work.[10] We assessed the performance of the ResNet12, DenseNet, YOLOv8, and MobileNet models using a brain tumor dataset comprising 4,489 images across 14 classes. The results indicated accuracy levels ranging from 88% to 97.3%. YOLOv8 emerged as the top performer owing to its minimal preprocessing time (0.1 ms), fastest inference time (1.8 ms), highest accuracy (97.3%), and compact size, making it ideal for real-time mobile applications.

Limitations of manual MRI analysis are addressed in this paper, by introducing an improved YOLOv8 model for brain tumor detection.[11] The model incorporates EMA for enhanced sensitivity to lesion features, CARAFE for broader contextual awareness, and C2f_DySnakeConv for improved tumor perception. The experimental results demonstrate that the revised model surpasses the original YOLOv8 and other open-source models, achieving improvements of 2.71%, 2.34%, 2.24%, and 3.73% in precision, recall, mAP@0.5, and mAP@0.5–0.95, respectively.

To enhance the detection and classification of brain tumors in challenging scenarios, this study proposes an innovative DC-YOLOv8FEN model. This model integrates advanced loss functions and dynamic attention mechanisms, along with vision transformers, self-shuffle attention, and a dual feature pyramid network, to facilitate enriched feature extraction. The model achieved an F1-score of 96.5%, accuracy of 99.5%, recall of 90.5%, and precision of 95.2%, demonstrating exceptional performance. This approach holds significant potential for assisting radiologists in real-time patient diagnosis, thereby improving precision and reducing delays.[12]

This paper also introduces a Gaussian Convolutional Neural Network (GCNN) method for multi-class brain tumor (BT) classification using MRI data. One dataset employed 3,064 images to classify tumors into pituitary, glioma, and meningioma categories, whereas another dataset utilized 516 images to classify gliomas into three grades (II, III, and IV). The proposed method achieved high accuracies of 99.8% and 97.14% on the two datasets, respectively, underscoring its efficacy in accurately identifying and categorizing various brain tumor types and grades.[13]

Furthermore, this study compares YOLOv5 and YOLOv7 for the detection and categorization of brain tumors using MRI scans of three distinct tumor types: meningiomas, gliomas, and pituitary tumors. Advanced mask alignment was employed during preprocessing to ensure the precise segmentation. Across a range of IoU thresholds, YOLOv7 exhibited slightly superior performance compared to YOLOv5 in terms of detection accuracy and mean Average Precision (mAP). Both models demonstrated their effectiveness and accuracy in medical imaging tasks, significantly outperforming traditional methods such as RCNN, Faster RCNN, and Mask RCNN.[14]

Materials and methods

Dataset

The brain tumor MRI dataset,[15] comprising 7,023 images from three sources—Figshare, SARTAJ, and Br35H—serves as an exceptional resource for classifying brain tumors into four categories: glioma, meningioma, no tumor, and pituitary. Notably, the 'no tumor' category was exclusively derived from the Br35H dataset. Additionally, the glioma class image, originally from the SARTAJ dataset, was replaced with an image from Figshare owing to misclassification errors identified in previous studies and training model results. This substitution aimed to enhance the reliability of the dataset. However, potential challenges, such as class imbalance or variations in imaging protocols across sources, may necessitate pre-processing techniques, such as normalization and augmentation, to ensure consistent model performance. Figure 2 shows sample images from the brain tumor dataset which includes glioma, meningioma, no tumor and pituitary images.

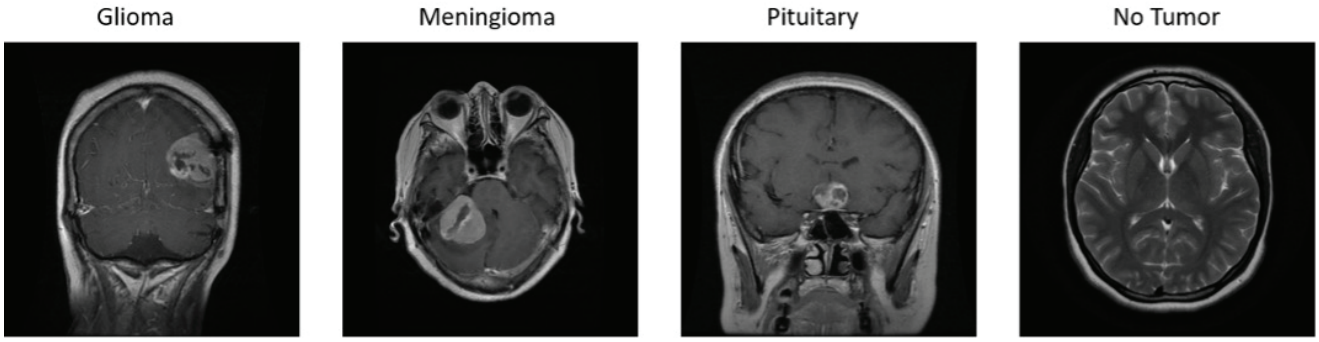


***Figure 2**. Sample images of brain tumor dataset.*

Methods

The proposed methodology for classifying brain MRI images into four categories—glioma, meningioma, no tumor, and pituitary, employs a structured pipeline based on the YOLOv8 model. As shown in Figure 3, this pipeline comprises seven stages, designed to ensure optimal model performance and rigorous scientific validation. Each stage is meticulously detailed, highlighting its technical implementation, underlying rationale, and integrative function within the overall process.

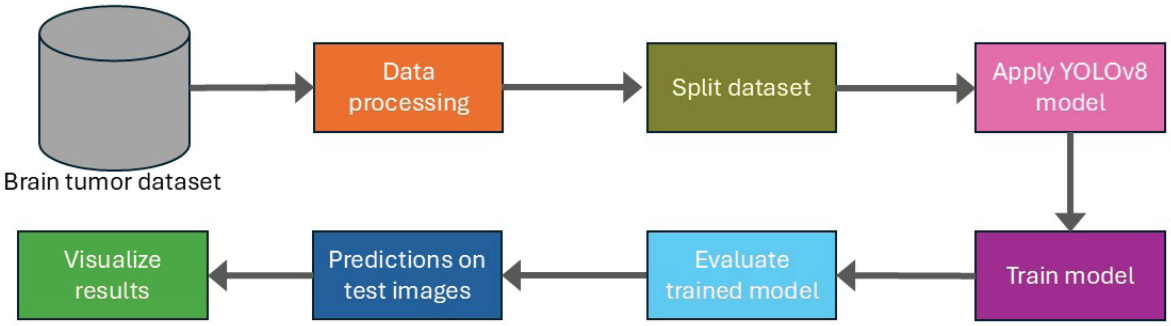


***Figure 3**. Proposed brain tumor detection framework.*

Data pre-processing

The initial phase of dataset preparation involved preprocessing, which encompassed 7,023 MRI images sourced from the Figshare, SARTAJ, and Br35H datasets. During this phase, variations in image resolution, intensity, and format are addressed using standardized transformations. The images were resized to a uniform resolution to comply with the input requirements of YOLOv8. Subsequently, the images were normalized to achieve a mean of 0 and unit variance, thereby enhancing the contrast and mitigating biases due to varying illumination. Gaussian filtering was employed to eliminate noise and other artifacts, whereas data augmentation techniques, such as horizontal flipping and brightness adjustment, were utilized to improve the model's generalization capabilities. This preprocessing is essential to ensure data reliability and coherence, which are crucial for extracting features with high fidelity in subsequent stages.

Dataset splitting

To facilitate supervised learning and ensure an unbiased evaluation of performance, the pre-processed data were partitioned into three distinct subsets: training, validation, and test data. A 70-15-15 stratified split was employed to maintain the same distribution of the four classes across the training, validation, and testing datasets. This stratification is essential, particularly given the potential imbalances in the 'no tumor' category, which is solely based on Br35H. The split was executed with a fixed seed (e.g., 42) during random sampling to ensure reproducibility. This approach allows for iterative improvement of the model and a separate assessment of its generalization performance on previously unseen observations, aligning with the best practices in machine learning evaluation.

YOLOv8 model application

To identify the most effective framework with a high success rate, we selected the state-of-the-art real-time object detection model YOLOv8 owing to its efficiency in testing speed and accuracy in medical image analysis. In this phase, the four-class classification task is fine-tuned using the pretrained YOLOv8 architecture, leveraging the weights from the ImageNet dataset. This model employs a convolutional neural network backbone to extract hierarchical

features, which are then used for bounding box regression and class prediction tasks in the detection head. The hyperparameters, including the learning rate (0.001), batch size (16) and anchor box sizes, were determined based on the image dimensions and tumor characteristics in the MRI. This step involves the application of transfer learning to expedite the convergence process and tailor the model specifically for brain MRI tumor detection.[16] Figure 4. illustrates the enhanced architecture of YOLOv8. The primary features of the YOLOv8 architecture are as follows:

Backbone: a deep convolutional neural network (CNN) that extracts hierarchical features from the input image, with CSPDarknet53 being a typical backbone.[17]

Neck: connects the backbone and head to facilitate feature flow across diverse scales. Unlike conventional feature pyramid networks (FPN), YOLOv8 employs a novel C2f module to integrate high-level features.

Head: responsible for predicting the bounding boxes and class probabilities. As an anchor-free network, YOLOv8 does not utilize predefined anchor boxes to predict the center of an object, thereby simplifying the detection process and potentially enhancing performance.

Loss function: the training of YOLOv8 involves a loss function comprising classification, localization, and abjectness losses. This multi-objective loss function aids in balancing the various components of detection tasks.

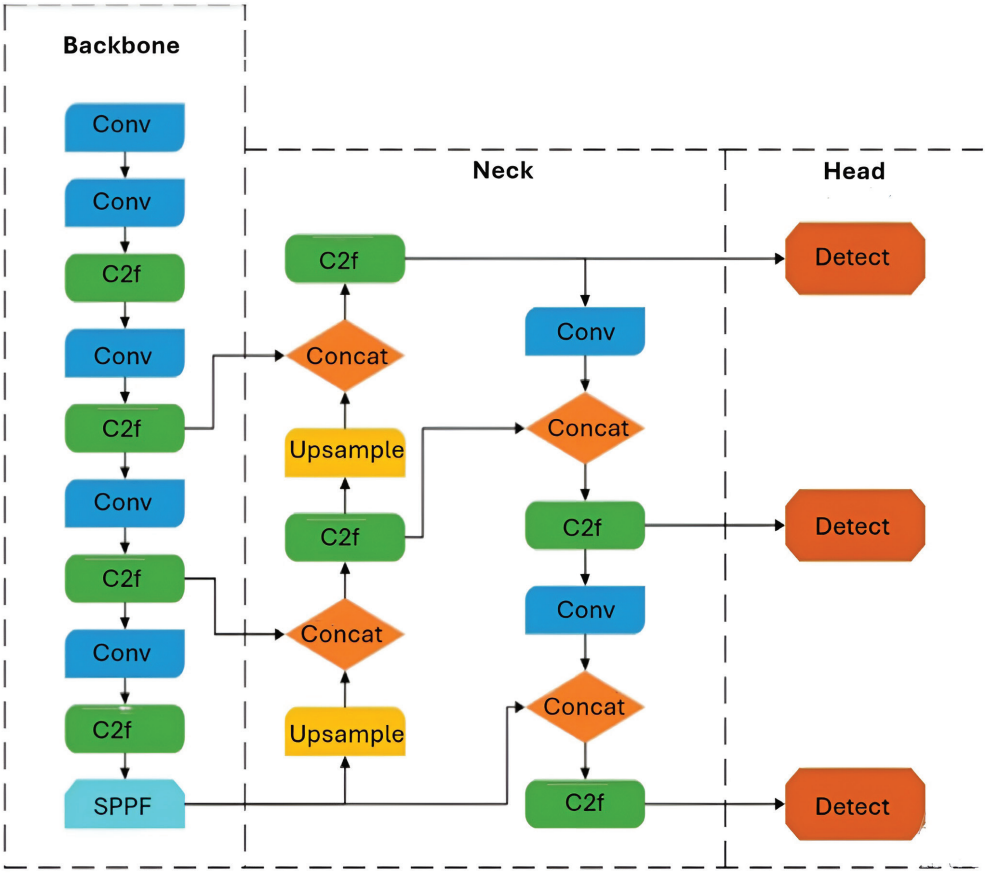


Figure 4*. The enhanced YOLOv8 architecture.*

Training model

The training subset was used to train the model, with YOLOv8 optimized through stochastic gradient descent (0.9) and a cosine-based annealing learning rate. This loss function incorporates classification loss (cross-entropy), localization loss (CIoU), and objectness loss, which are specifically designed to address the multi-class and multi-objective nature of the task to achieve equilibrium. Training was conducted until either 100 epochs were completed or convergence was achieved, with convergence being evaluated using the validation set to prevent overfitting through early stopping (patience=10 epochs). Performance metrics, including the mean Average Precision (mAP) at IoU=0.5, were monitored to assess progress. This rigorous training protocol ensures that the model effectively captures complex tumor patterns without incurring excessive computational costs.

Evaluation of the trained model

The model performance was critically evaluated after training on the independent test set. Evaluation metrics such as precision, recall, F1-score, and mean Average Precision (mAP) across all classes were considered to provide a comprehensive assessment of the accuracy and robustness of the classification.

To ascertain class-specific errors, a confusion matrix was constructed, with particular attention given to the glioma class owing to previous misclassification issues in the SARTAJ dataset. Comparisons with baseline models were conducted using statistical tests, such as the McNemar test, to ensure that the model's performance surpassed that of the baseline models. This approach serves to demonstrate the model's generalizability and clinical relevance, which are crucial for high-impact applications.

Prediction of test images

The trained YOLOv8 model was deployed on the test subset to generate predictions that included bounding boxes and class probabilities. This phase involves a forward pass through the network, followed by post-processing using non-maximum suppression (NMS) to eliminate overlapping detections, with an NMS threshold of 0.45. Ground truth labels and confidence scores were used to assess the discrepancies between the predicted labels and scores. Qualitative evaluation can be conducted by visualizing the predicted bounding boxes on the test images to identify false-positive or false-negative results. This process aims to enhance the practical applicability of the model in actual diagnostic practice.

Visualization

Visualization involves the superimposition of bounding boxes on test images to delineate the tumor regions. This process is executed through the visualize detection routine, wherein OpenCV is employed to draw rectangles and affix labels with confidence scores (e.g., the visualize routine utilizes OpenCV to render rectangles and labels with values indicating tumor detection, such as tumor: 0.97). Performance metrics were plotted using Matplotlib, with a code snippet illustrating box loss, mean Average Precision (mAP), precision, and recall on the x-axis, alongside the epoch on the y-axis, arranged in a 2×2 subplot configuration.

Evaluation metrics

Evaluation metrics are quantitative measures used to evaluate the performance of a model, algorithm, or system. We used accuracy described in Equation (1), precision represented by Equation (2), recall calculated using Equation (3), Mean Average Precision (mAP) represented by Equation (4) as performance metrics.[18]

Accuracy

This is a performance measure that indicates the percentage of correctly predicted instances (i.e., true positives and true negatives) against the total number of predictions.

$$Accuracy = \frac{TN + TP}{TP + FP + TN + FN} \quad (1)$$

Precision (P)

The percentage of accurately identified positive predictions among all positive predictions.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall (R)

The percentage of correctly predicted positive predictions among all actual positives.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Mean average precision (mAP)

The average of the Average Precision (AP) of all classes. The part of the precision-recall curve is called AP.

$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \quad (4)$$

Where N is the number of classes.

Results

In this study, the performance of the YOLOv8 model in identifying brain MRI tumors was evaluated using a dataset comprising 7,023 images categorized into four classes: glioma, meningioma, no tumor, and pituitary tumors. The YOLOv8 model employed pretrained weights through transfer learning, with training conducted at an initial learning rate of 0.001, a batch size of 16, and over 60 epochs. During test inference on 400 images with dimensions (1, 3, 640, 640), the model demonstrated processing times of 4.3, 12.8, and 3.3ms per image for preprocessing, inference, and postprocessing, respectively. Comparative performance metrics highlight the proposed model's superiority over existing approaches, as summarized in Table 1.

Table 1. Comparison of proposed model with existing models.

Author	Precision (%)	Recall (%)	Accuracy (%)
Remzan *et al.* [19]	97.71	97.71	97.71
Kaur *et al.* [20]	97.8	97.2	98.02
Rajput *et al.* [21]	99	98	99
Proposed	98.9	98.9	99.5

Notably, as shown in Figure 5, the model correctly identified a glioma in the image Tr-glTr_0006.jpg with a confidence score of 0.97. The resilience of the modified YOLOv8 architecture and its potent ability to accurately locate the glioma location inside the MRI scan are demonstrated by this detection output.

The Box Loss graph, shown in Figure 6, indicates efficient optimization and little overfitting, provides additional evidence for these findings by displaying a consistent drop in training and validation losses.

The mean average precision (mAP50-95) development over training epochs is seen in this graph shown in Figure 7. Early optimization changes are reflected in the initial oscillations. Strong convergence and high detection accuracy are then indicated by the curve stabilizing around 1.0. The trained YOLOv8 model's resilience and capacity for generalization are validated by the consistently high mAP values in subsequent epochs.

The variation in precision and recall for the different epochs is shown in Figure 8. After about 15 epochs, the precision-recall curve quickly stabilizes, and both metrics surpass 0.95, suggesting robust and steady classification accuracy throughout training.

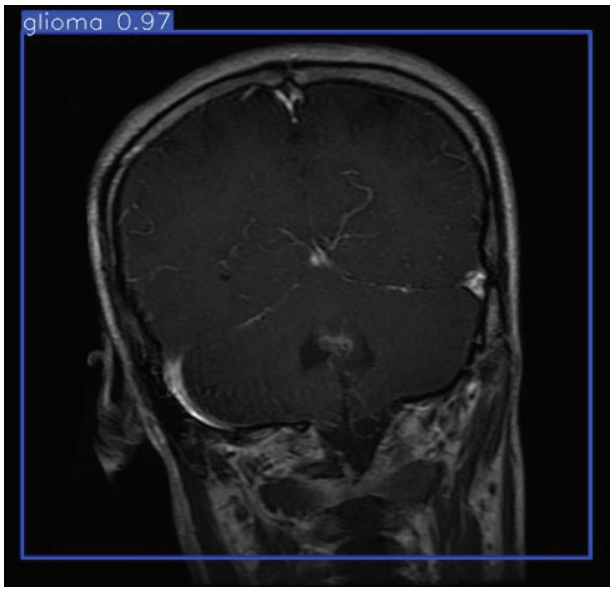


***Figure 5**. Predicted results.*

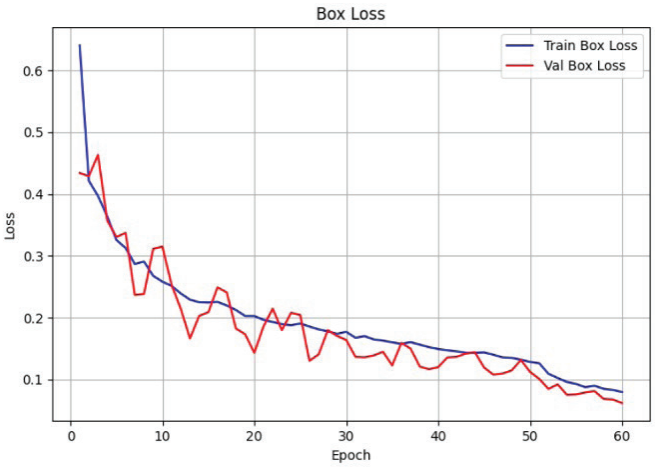


***Figure 6**. Variation in between loss and epoch.*

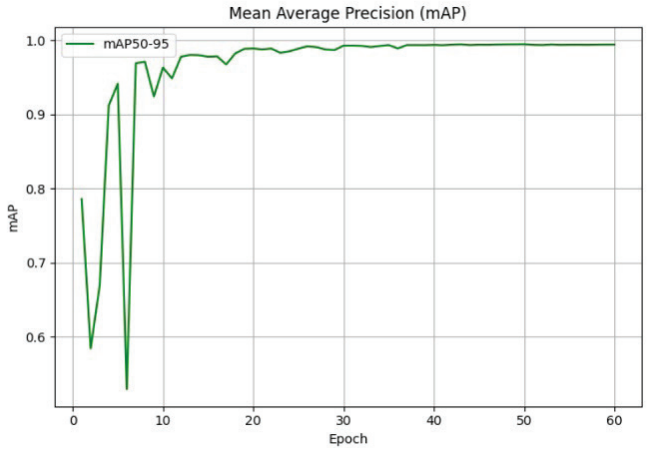


***Figure 7**. Graph plot between mean average precision and epoch.*

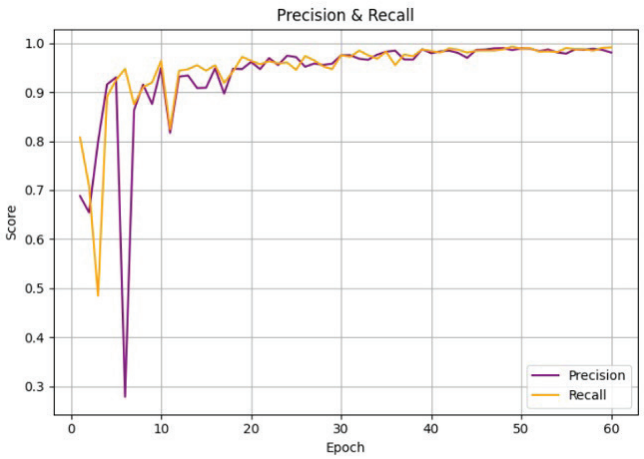


***Figure 8**. Variation of Precision and Recall between epochs.*

Discussion

By implementing and optimizing YOLOv8 for precise glioma, meningioma, and pituitary tumor detection and localization in MRI scans, this study achieves state-of-the-art performance while improving reproducibility by using a publicly available dataset. Inference testing on a subset of 400 images, each resized to a shape of (1, 3, 640, 640), revealed efficient processing times: an average of 4.3ms per image for preprocessing, 12.8ms for inference, and 3.3ms for postprocessing, resulting in a total processing time of approximately 20.4ms per image. This computational efficiency underscores the model's potential for real-time clinical applications, balancing speed with diagnostic reliability.

The model achieved a precision of 98.9%, recall of 98.9%, and overall accuracy of 99.5%, outperforming benchmarks such as Remzan *et al*. (97.71% across all metrics), Kaur *et al*. (precision 97.8%, recall 97.2%, accuracy 98.02%), and Rajput *et al*. (precision 99%, recall 98%, accuracy 99%).[19-21] These results indicate enhanced robustness in minimizing false positives and negatives, particularly in multi-class tumor classification, thereby advancing the state-of-the-art in AI-assisted brain MRI analysis. data processing and error correction.

The suggested YOLOv8-based model shows great promise as a quick and accurate method for automatically detecting brain tumors in MRI images. It can help radiologists identify and plan treatments for pituitary tumors, meningioma, and gliomas early on by increasing diagnostic efficiency and accuracy. While expert appraisal and multi-center validation are still necessary for clinical adoption, integration into clinical workflows may expedite screening and aid in decision-making.

Overall, the results suggest the model is performing well and could be useful for tumor detection, though it still requires some refinement.

Limitations

A publicly accessible Kaggle dataset was used to train and evaluate the model, which might not accurately reflect the variability of actual clinical MRI data. Furthermore, the identification performance for less-represented tumor types might have been impacted by possible class imbalance and a small sample size. The precise tumor segmentation that is essential for clinical decision-making is not addressed by the suggested YOLOv8 model, which instead concentrates on detection and localization. Additionally, the evaluation of clinical application was limited because no expert radiologist validation or comparison with established diagnostic procedures was carried out. Lastly, because no explainability techniques were used to illustrate or support the model's predictions, its interpretability is still restricted. These constraints should be addressed in future research by incorporating segmentation modules, external validation, and interpretability strategies.

Conclusion

The primary objective of this study was to employ the YOLOv8 model for the detection of tumors in brain MRI images and to evaluate its effectiveness and efficiency in comparison to existing medical object detection models, utilizing the publicly available dataset, which comprises 7,023 images categorized into glioma, meningioma, no tumor, and pituitary classes. The pre-trained YOLOv8n model, trained for more than 60 epochs, achieved a confidence level of 0.97 in detecting gliomas, with processing times of 4.3, 12.8, and 3.3ms for pre-processing, inference, and post-processing, respectively. The visualization of boxes and performance metrics, such as box loss, mAP50-95, precision, and recall, confirmed the model's capability to accurately identify tumor regions and improve over time. The experimental results indicate that the precision, recall, and accuracy values are 98.9%, 98.9%, and 99.5%, respectively. Additionally, the model provides a mean average precision (mAP) of 97.6% after comprehensive evaluations.

Future work will focus on enhancing the efficiency of the YOLOv8 model for deployment in edge devices and real-time clinical applications, as well as expanding the dataset and incorporating multimodal data to achieve improved generalization and tumor characterization.

Ethical approval

There are no humans or animals used for experimentation purpose.

Fundings

This research did not receive any specific grants from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest

There are no conflicts of interest to declare.

Data availability statement

These datasets are publicly available on the Kaggle website under the names: figshare, SARTAJ, and Br35H**.**

CRediT authorship contribution statement

Kavita Singh: developed and implemented the proposed methodology, conceptualized the study and prepared the manuscript; **Rajesh Kumar**: supervisor of this work, contributed to the conceptualization, methodological design and manuscript drafting; **Satyendra Singh**: assisted with manuscript in writing, editing and revision.

Acknowledgements

We are very thankful to Mr. Ravi Ranjan Kumar, Mr. Anand Kumar Jaiswal, Mr. Rahul Tripathi, and our supporting staff of the institute who have directly or indirectly helped to successful completion of this manuscript.

References

[1] Ranjbarzadeh R, Crane M, Bendechache M. The impact of backbone selection in YOLOv8 models on brain tumor localization. Iran J Comput Sci. 2025; 8: 939-61. doi: 10.1007/s42044-025-00258-4.

[2] Liu Z, Tong L, Chen L, Jiang Z, Zhou F, Zhang Q, et al. Deep learning based brain tumor segmentation: a survey. Complex Intell Syst. 2023; 9: 1001-26. doi: 10.1007/s40747-022-00815-5.

[3] Mahale C, Meshram S, Pakhmode A. Brain Tumor Detection with YOLOv8. 2024 4th Int. Conf. Intell. Technol. CONIT 2024, IEEE. 2024: pp 1-5. doi: 10.1109/CONIT61985.2024.10626992.

[4] Selcuk B, Serif T. Brain Tumor Detection and Localization with YOLOv8. UBMK 2023 - Proc. 8th Int. Conf. Comput. Sci. Eng., Institute of Electrical and Electronics Engineers Inc.; 2023: pp 477-81. doi: 10.1109/UBMK59864.2023.10286729.

[5] Yao Q, Zhuang D, Feng Y, Wang Y, Liu J. Accurate detection of brain tumor lesions from medical images based on improved YOLOv8 algorithm. IEEE Access. 2024; 12: 144260-79. doi: 10.1109/ACCESS.2024.3472039.

[6] Sharif MI, Khan MA, Alhussein M, Aurangzeb K, Raza M. A decision support system for multimodal brain tumor classification using deep learning. Complex Intell Syst. 2022; 8: 3007-20. doi: 10.1007/s40747-021-00321-0.

[7] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. 2016; 2016: 779-88. doi: 10.1109/CVPR.2016.91.

[8] Taha AM, Aly SA, Darwish MF. Detecting glioma, meningioma, and pituitary tumors, and normal brain tissues based on Yolov11 and Yolov8 deep learning models 2025. doi: 10.48550/arXiv.2504.00189

[9] Dulal R, Dulal R. Brain tumour identification using improved YOLOv8. Vol. 1. 2025. doi: 10.1504/ijcast.2025.10071167.

[10] Al-Shahrani A, Al-Amoudi W, Bazaraah R, Al-Sharief A, Farouquee H. An image processing-based and deep learning model to classify brain cancer. Eng Technol Appl Sci Res. 2024; 14: 15433-8. doi: 10.48084/etasr.7803.

[11] Lin Z, Lin W, Jiang F. YOLOv8-DEC: Enhancing brain tumor object detection accuracy in magnetic resonance imaging. Prog Electromagn Res M. 2024; 129: 43-52. doi: 10.2528/PIERM24061204.

[12] Abraham LA, Palanisamy G, Goutham V. Dilated convolution and YOLOv8 feature extraction network: An improved method for MRI-based brain tumor detection. IEEE Access. 2025; 13: 27238-56. doi: 10.1109/ACCESS.2025.3539924.

[13] Rizwan M, Shabbir A, Javed AR, Shabbir M, Baker T, Al-Jumeily Obe D. Brain tumor and glioma grade classification using Gaussian convolutional neural network. IEEE Access. 2022; 10: 29731-40. doi: 10.1109/ACCESS.2022.3153108.

[14] Almufareh MF, Imran M, Khan A, Humayun M, Asim M. Automated brain tumor segmentation and classification in MRI using YOLO-based deep learning. IEEE Access. 2024; 12: 16189-207. doi: 10.1109/ACCESS.2024.3359418.

[15] Msoud Nickparvar. (2021). Brain Tumor MRI Dataset [Data set]. n.d. doi: 10.34740/kaggle/dsv/14832123.

[16] Elhanashi A, Dini P, Saponara S, Zheng Q. TeleStroke: real-time stroke detection with federated learning and YOLOv8 on edge devices. J Real-Time Image Process. 2024; 21; article no. 121. doi: 10.1007/s11554-024-01500-1.

[17] Terven J, Córdova-Esparza DM, Romero-González JA. A Comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS. Mach Learn Knowl Extr. 2023; 5: 1680-716. doi: 10.3390/make5040083.

[18] Talukder MA, Islam MM, Uddin MA, Akhter A, Pramanik MAJ, Aryal S, et al. An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning. Expert Syst Appl. 2023; 230: 120534. doi: 10.1016/j.eswa.2023.120534.

[19] Remzan N, Tahiry K, Farchi A. Advancing brain tumor classification accuracy through deep learning: harnessing radimagenet pre-trained convolutional neural networks, ensemble learning, and machine learning classifiers on MRI brain images. Multimed Tools Appl. 2024; 83: 82719-47. doi: 10.1007/s11042-024-18780-1.

[20] Kaur P, Mahajan P. Detection of brain tumors using a transfer learning-based optimized ResNet152 model in MR images. Comput Biol Med. 2025; 188: 109790. doi: 10.1016/j.compbiomed.2025.109790.

[21] Rajput IS, Gupta A, Jain V, Tyagi S. A transfer learning-based brain tumor classification using magnetic resonance images. Multimed Tools Appl. 2024; 83:2 0487-506. doi: 10.1007/s11042-023-16143-w.

Journal of Associated Medical Sciences 2026; 59(2): 259-267

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Combined neutrophil-to-lymphocyte, monocyte-to-lymphocyte, and platelet-to-lymphocyte ratios as inflammatory biomarkers in tuberculosis: A retrospective study from Lampang, Thailand

Sittichok Chaiwong[1], Niphawan Jettavan[1], Yuttana Mundee[2], Pawaret Panyajai[2], Wuttichote Jansaento[2], Watcharapong Jugnam-ang[2*]

[1]Department of Medical Technology, Lampang Hospital, Lampang Province, Thailand.
[2]Department of Medical Technology, Faculty of Medical Technology, Nation University, Lampang Province, Thailand.

ARTICLE INFO

Article history:
Received 2 December 2025
Accepted as revised 17 February 2026
Available online 3 March 2026

Keywords:
Tuberculosis, neutrophil-to-lymphocyte ratio, monocyte-to-lymphocyte ratio, platelet-to-lymphocyte ratio, inflammatory biomarkers

ABSTRACT

Background: Tuberculosis (TB) remains one of the major infectious diseases causing global morbidity and mortality. Systemic inflammation plays a central role in TB pathogenesis. Hematological indices derived from complete blood counts (CBC) including neutrophil-to-lymphocyte ratio (NLR), monocyte-to-lymphocyte ratio (MLR), and platelet-to-lymphocyte ratio (PLR) have been proposed as surrogate inflammatory biomarkers.

Objectives: To investigate the role of CBC-derived ratios (NLR, MLR, and PLR) as inflammatory biomarkers in TB patients compared with healthy controls, and to assess their association with clinical characteristics and laboratory findings.

Materials and methods: This retrospective study was conducted at Lampang Hospital, Thailand. Data from 213 TB patients and 170 healthy controls were collected from medical records and hematology reports. CBC-derived ratios were calculated and analyzed using Mann-Whitney U test, multivariate logistic regression adjusted for age, gender, and body mass index (BMI), and receiver operating characteristic (ROC) curve analysis.

Results: Median NLR, MLR, and PLR values were significantly higher in TB patients compared with controls (NLR 4.21 vs 1.58; MLR 0.47 vs 0.16; PLR 223.0vs119.9; all p<0.001). In multivariate logistic regression analysis, only MLR (OR = 511.567; 95% CI: 7.384-35,439.660; p=0.004) remained independently associated with TB after adjustment, whereas NLR and PLR did not retain statistical significance. ROC analysis indicated good discrimination for NLR (AUC=0.852), MLR (AUC=0.893), and PLR (AUC=0.812). The combined model (NLR+MLR+PLR) achieved the highest accuracy (AUC=0.903; sensitivity=79.3%; specificity=92.4%).

Conclusion: Elevated NLR, MLR, and PLR values reflect systemic inflammation in TB patients. Among these markers, MLR showed the strongest independent association with TB. Although their specificity is limited, combining these ratios enhances their ability to represent the host inflammatory response and may provide additional information for clinical assessment and disease monitoring.

*** Corresponding contributor.**
Author's Address: *Department of Medical Technology, Faculty of Medical Technology, Nation University, Lampang Province, Thailand.*
E-mail address: watcharapong_jug@nation.ac.th
doi: 10.12982/JAMS.2026.057
E-ISSN: 2539-6056

Introduction

Tuberculosis (TB) remains one of the most important infectious diseases and a serious public health problem, especially in developing countries.[1] Although diagnostic

and treatment methods have improved, the global burden of TB is still high. In 2022, about 10.6 million new cases and 1.3 million deaths were reported worldwide, showing that TB continues to be a major challenge.[2] Early diagnosis and proper monitoring of disease activity are essential to improve treatment outcomes and reduce disease transmission.[3] However, advanced diagnostic tests such as nucleic acid amplification tests (NAATs) are often not available in many hospitals with limited resources.[4] Therefore, there is growing interest in using simple and low-cost inflammatory biomarkers that can be measured from routine laboratory tests, such as complete blood count (CBC), to help in TB diagnosis and disease monitoring.[5,6]

Several indices derived from CBC, including the neutrophil-to-lymphocyte ratio (NLR), monocyte-to-lymphocyte ratio (MLR), and platelet-to-lymphocyte ratio (PLR), have been proposed as potential markers of systemic inflammation.[7,8] These ratios are easy to calculate and have been reported to be associated with disease severity, bacterial load, and immune activation in many diseases, such as tuberculosis, cancer, and some chronic inflammatory disorders.[8-10] Previous studies have shown that patients with active TB usually have higher NLR and MLR compared with healthy individuals, suggesting that these ratios reflect the imbalance between innate and adaptive immune responses during *Mycobacterium tuberculosis* infection.[11,12] However, the diagnostic performance of each ratio has been inconsistent among studies and populations, meaning that no single ratio can fully distinguish between disease and healthy states.[13,14]

Although most previous studies have examined NLR, MLR, or PLR as individual parameters, evidence regarding their combined diagnostic and inflammatory significance remains scarce, particularly in real-world clinical settings. In addition, most data have come from Asian (mainly Chinese and Indian) or African populations, where diagnostic accuracy varied by study design and population characteristics.[14,15] Differences in host genetics, comorbidities, and clinical management may affect hematologic response to infection, so population-specific data are important. In Thailand, where tuberculosis remains prevalent, few studies have explored the association between hematological indices and routine clinical findings. Miyahara et al. reported that a higher neutrophil-to-lymphocyte ratio (NLR) was associated with an increased risk of tuberculosis and correlated with positive sputum smear (AFB) results under routine clinical practice.[16]

This study, therefore, aimed to evaluate the combined use of NLR, MLR, and PLR as inflammatory biomarkers in tuberculosis patients at Lampang Hospital, northern Thailand. The ratios were compared between TB patients and healthy controls, and their associations with clinical and laboratory results were analyzed. In addition, their diagnostic performance, individually and in combination, was assessed to determine their usefulness for distinguishing TB from non-TB cases. The findings may help support the application of these simple hematologic ratios as additional tools for TB monitoring, particularly in hospitals with limited diagnostic resources.

Materials and methods

Study design and population

This retrospective analytical study was conducted at Lampang Hospital, Lampang Province, Thailand. Data were collected from July 2023 to January 2025. A total of 213 patients newly diagnosed with TB (pulmonary and extrapulmonary) and 170 healthy controls who attended the hospital's health-check clinic were included. All available cases during the study period were included using a consecutive sampling approach. The study protocol was approved by the Lampang Hospital Ethics Committee (Approval No. [EC 098.1/68]). Patient confidentiality was strictly maintained.

Inclusion criteria for the TB group were adults aged ≥18 years diagnosed with TB based on clinical evaluation and/or microbiological confirmation, including acid-fast bacilli (AFB) microscopy, culture, or nucleic acid amplification tests (NAAT). The control group consisted of individuals undergoing routine annual health examinations with no history of TB, no known household exposure to TB, no clinical or radiographic evidence of active disease, and available complete blood count (CBC) results.

Exclusion criteria for both groups included autoimmune diseases, malignancies, hematologic disorders, or incomplete CBC data.

Data collection

Demographic data, body mass index (BMI), TB type, AFB status, NAAT results, and HIV coinfection were retrieved from medical records. Complete blood count (CBC) parameters were obtained at the time of diagnosis. CBC analysis was performed using an automated hematology analyzer (Sysmex XN-1000, Kobe, Japan). The inflammatory ratios were calculated as follows: NLR; Absolute neutrophil count/Absolute lymphocyte count, MLR; Absolute monocyte count/Absolute lymphocyte count, and PLR; Absolute platelet count/Absolute lymphocyte count.

Statistical analysis

Data were analyzed using SPSS software version 29 (IBM Corp., Armonk, NY, USA). Continuous variables were presented as median (interquartile range, IQR), and group differences were assessed using the Mann-Whitney U test. Multivariate logistic regression analysis was performed to assess the associations between CBC-derived inflammatory ratios and TB status after adjustment for age, gender, and BMI. Diagnostic performance was evaluated using receiver operating characteristic (ROC) curve analysis. Optimal cut-off

values were determined using Youden's index (J=sensitivity+specificity -1), with the maximum J indicating the best balance between sensitivity and specificity. Sensitivity and specificity at the selected thresholds were obtained from the ROC coordinates table. In addition, a combined diagnostic model was constructed using multivariate logistic regression, incorporating NLR, MLR, and PLR as covariates with TB diagnosis as the dependent variable. The predicted probabilities from this model were evaluated using ROC analysis. The optimal probability threshold was selected based on the highest Youden's index. A p-value<0.05 was considered statistically significant. For subgroup analyses, TB patients were further categorized into low and high groups for each inflammatory ratio (NLR, MLR, and PLR) based on the median values observed within the TB cohort. These median-based groups were used to explore associations with clinical or microbiological characteristics.

Results

Baseline demographic and hematological characteristics

A total of 213 patients with TB and 170 healthy controls were included in this study. The baseline demographic and hematological characteristics of both groups are shown in Table 1. The median age of TB patients was significantly higher than that of healthy controls (62.0±24.0 years vs. 34.0±24.0 years, $p<0.001$). The proportion of males was also much higher in the TB group compared with the control group (64.8% vs 14.7%, $p<0.001$). In addition, TB patients had a significantly lower median body mass index (BMI) than healthy controls (19.1±4.6 vs. 22.4±3.2, $p<0.001$). Regarding CBC-derived inflammatory ratios, the median values of NLR, MLR, and PLR were all significantly increased in TB patients compared with controls (NLR: 4.21 vs 1.58; MLR: 0.47 vs 0.16; PLR: 223.03 vs 119.87; all $p<0.001$). These results indicate that patients with TB had an enhanced systemic inflammatory response, as reflected by higher NLR, MLR, and PLR ratios compared with healthy individuals.

Association between inflammatory ratio groups and clinical or laboratory characteristics

The associations between inflammatory ratios (NLR, MLR, and PLR) and clinical or laboratory findings were examined by dividing each ratio based on the median cut-off values among TB patients (NLR≥4.22, MLR≥0.47, and PLR≥223.03). As presented in Table 2, sputum AFB positivity showed significant associations with all three ratios, where higher proportions of elevated NLR, MLR, and PLR were observed among AFB-positive patients ($p<0.001$, $p=0.004$, and $p=0.016$, respectively). NAAT results were also significantly related to NLR levels ($p=0.049$), but not to MLR or PLR ($p>0.05$). No significant relationships were observed between these ratios and other clinical variables, including gender, type of TB, or HIV status (all $p>0.05$). These findings suggest that increased inflammatory ratios, especially NLR, may reflect the degree of systemic inflammation and bacterial load in patients with tuberculosis.

Table 1. Baseline demographic, clinical, and hematological characteristics between healthy controls and tuberculosis patients

Parameter	Healthy control (N=170)	TB patients (N=213)	*p* value
Gender			
Male, N (%)	25 (14.71)	138 (64.79)	<0.001#
Female, N (%)	145 (85.29)	75 (35.21)	
Age (years), Median (IQR)	34.00 (4.00)	62.00 (24.00)	<0.001*
BMI, Median (IQR)	22.42 (3.17)	19.10 (4.62)	<0.001*
Level of NLR ratio, Median (IQR)	1.58 (0.85)	4.21 (6.19)	<0.001*
Level of MLR ratio, Median (IQR)	0.16 (0.07)	0.47 (0.54)	<0.001*
Level of PLR ratio, Median (IQR)	119.87 (58.78)	223.03 (240.01)	<0.001*

Note: *#Statistically significant difference between groups determined using the Chi-square test ($p<0.05$), *Statistically significant difference between groups determined using the Mann-Whitney U test ($p<0.05$).*

Table 2. Comparison of inflammatory ratios (NLR, MLR, and PLR) according to clinical and laboratory parameters among tuberculosis patients

Parameter	Low NLR (ratio<4.22)	High NLR (ratio≥4.22)	*p* value	Low MLR (ratio<0.47)	High MLR (ratio≥0.47)	*p* value	Low PLR (ratio<223.03)	High PLR (ratio≥223.03)	*p* value
Gender									
Male	64	74	0.180	43	32	0.127	43	32	0.127
Female	42	33		64	74		64	74	
TB type									
Pulmonary (P)	86	88	0.136	84	90	0.294	89	85	0.543
Extrapulmonary (EP)	18	12		19	11		14	16	
P and EP	2	7		4	5		3	6	
Sputum AFB									
Negative	63	40	<0.001*	59	44	0.004#	57	46	0.016$
Positive	23	43		23	43		24	42	
NAAT									
TB detected	37	27	0.049*	36	28	0.120	32	32	0.832
TB not detected	52	70		54	68		59	63	
HIV Status									
Positive	82	82	0.423	78	86	0.662	83	81	0.168
Negative	5	8		7	6		4	9	

Note: **statistically significant association between NLR groups and the corresponding clinical parameter (p<0.05, Chi-square test), #statistically significant association between MLR groups and the corresponding clinical parameter (p<0.05, Chi-square test), $Statistically significant association between PLR groups and the corresponding clinical parameter (p<0.05, Chi-square test), Cut-off values were defined by the median of each ratio in TB patients, AFB: acid-fast bacillus, NAAT: nucleic acid amplification test, TB: tuberculosis, EP: extrapulmonary, P&EP: both pulmonary and extrapulmonary tuberculosis, HIV: human immunodeficiency virus.*

Multivariate logistic regression analysis of inflammatory ratios associated with tuberculosis

Multivariate logistic regression analysis was performed to identify independent inflammatory predictors associated with tuberculosis after adjusting for age, gender, and BMI (Table 3). MLR remained independently associated with TB (adjusted OR=511.57; 95%CI: 7.384-35,439.660; p=0.004). In contrast, NLR and PLR were not independently associated with TB after adjustment (p=0.118 and p=0.064, respectively). Increasing age and male sex were significantly associated with higher odds of TB, whereas higher BMI was inversely associated with TB. These findings suggest that MLR demonstrated the strongest independent association with TB in this study population.

Diagnostic performance of inflammatory ratios

ROC curve analysis was conducted to evaluate the diagnostic performance of NLR, MLR, PLR, and their combined model in distinguishing TB patients from healthy controls (Table 4 and Figure 1). Among the individual inflammatory ratios, MLR showed the highest diagnostic accuracy, with an area under the curve (AUC) of 0.893, followed by NLR (AUC=0.852) and PLR (AUC=0.812). The optimal cut-off values for differentiating TB from controls were 1.562 for NLR, 0.307 for MLR, and 183.09 for PLR. These cut-offs yielded sensitivities of 82.0%, 74.2%, and 62.9%, and specificities of 76.0%, 95.3%, and 88.8%, respectively. The combined model using NLR, MLR, and PLR achieved the highest AUC of 0.903, with a sensitivity of 79.3% and a specificity of 92.4%. These results indicate that the combined use of multiple inflammatory ratios provides better discrimination between TB patients and healthy controls than any single marker. The combination model may therefore reflect the additive value of systemic inflammatory biomarkers in identifying individuals with active tuberculosis.

Table 3. Multivariable logistic regression analysis of inflammatory ratios associated with tuberculosis after adjustment for age, gender, and BMI.

Variable	B	SE	Wald	df	*p* value	Adjusted OR [Exp(B)]	95% CI for OR
NLR ratio	0.299	0.191	2.441	1	0.118	1.349	0.927-1.962
MLR Ratio	6.237	2.162	8.321	1	**0.004**	511.567	7.384-35,439.660
PLR Ratio	0.06	0.003	3.439	1	0.064	1.006	1.000-1.013
Gender	1.967	0.423	21.589	1	<0.001	7.148	3.118-16.387
Age (years)	0.078	0.014	31.495	1	<0.001	1.081	1.052-1.111
BMI	-0.175	0.048	8.433	1	<0.001	0.839	0.763-0.923

Note: *B: regression coefficient, SE: standard error, OR: odds ratio, CI: confidence interval, p<0.05 was considered statistically significant.*

Table 4. Diagnostic performance of inflammatory ratios (NLR, MLR, and PLR) and their combination in differentiating tuberculosis patients from healthy controls

Parameters	Cut-off	Sensitivity	Specificity	AUC
NLR	1.562	82%	76%	0.852
MLR	0.307	74.2%	95.3%	0.893
PLR	183.09	62.9%	88.8%	0.812
NLR+MLR+PLR	0.520	79.34%	92.35%	0.903

Note: *AUC=area under the curve; NLR=neutrophil-to-lymphocyte ratio; MLR=monocyte-to-lymphocyte ratio; PLR=platelet-to-lymphocyte ratio. AUC=area under the curve; NLR=neutrophil-to-lymphocyte ratio; MLR=monocyte-to-lymphocyte ratio; PLR=platelet-to-lymphocyte ratio. Cut-off values were determined by selecting the point with the maximum Youden's index from ROC curve analysis. Sensitivity and specificity were calculated at the selected cut-off. Cut-off value for the combined model represents the predicted probability derived from multivariant logistic regression.*

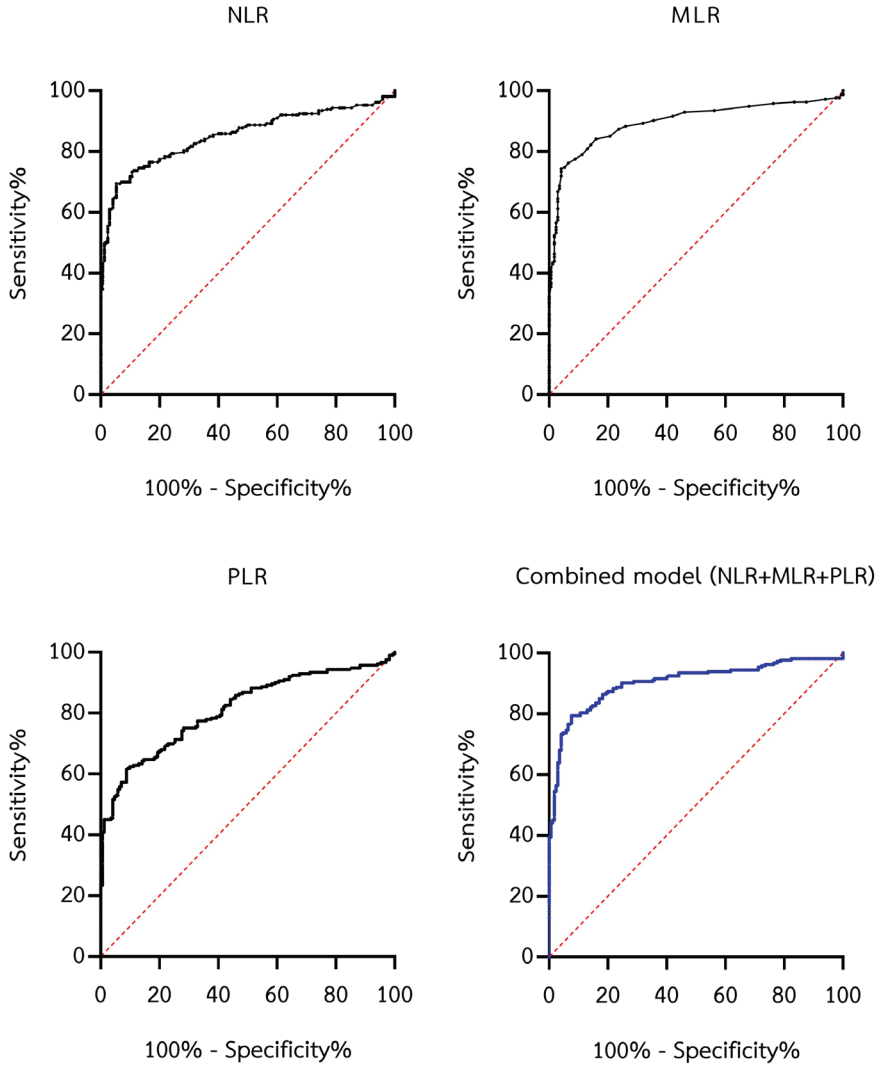


Figure 1. *Receiver operating characteristic (ROC) curves of inflammatory ratios in tuberculosis patients and healthy controls. ROC curves illustrate the diagnostic performance of neutrophil-to-lymphocyte ratio (NLR), monocyte-to-lymphocyte ratio (MLR), and platelet-to-lymphocyte ratio (PLR) for differentiating tuberculosis (TB) patients from healthy controls.*

Discussion

The present study showed that patients with TB had significantly higher NLR and MLR compared with healthy controls, while PLR did not show a significant difference. These results suggest that systemic inflammation and immune imbalance are strongly activated during *M. tuberculosis* infection.[17,18] The combination model including NLR, MLR, and PLR provided the highest diagnostic accuracy (AUC=0.903), indicating that using several inflammatory indices together can describe the host immune response more effectively than using a single marker.

The increase in NLR and MLR in TB patients can be explained by the immune-pathological mechanism of infection. Neutrophils and monocytes serve as the first line of defense by releasing cytokines and forming granulomas to limit the growth of *M. tuberculosis*.[19,20] Continuous activation of these cells leads to strong systemic inflammation, whereas lymphopenia in TB may occur due to immune exhaustion, apoptosis, or migration of lymphocytes to infection sites.[21,22] Therefore, higher NLR and MLR values indicate hyper-activation of the innate immune response and suppression of adaptive immunity, showing an imbalance between inflammatory and regulatory components.[23] In contrast, PLR may be affected by platelet turnover and hematologic conditions.[24] The absence of a significant difference in PLR in this study implies that platelet-related activity might contribute less to TB-associated inflammation than neutrophil or monocyte responses.

Our findings are consistent with previous studies by, Tu *et al.*[8], Hua *et al.*,[12] and Kim *et al.*[25] which also reported elevated NLR and MLR in active TB and their correlation with bacterial burden and disease severity. In contrast, our findings primarily support

their role as indicators of systemic inflammatory response. However, in the present multivariate analysis adjusted for age, gender, and BMI, only MLR remained independently associated with TB, whereas NLR and PLR did not retain statistical significance. Previous studies suggest that increased NLR and MLR reflect enhanced systemic inflammation and immune activation in active tuberculosis. Other recent reports also confirmed that increased inflammatory ratios, particularly NLR and MLR, can predict TB activity in different populations.[11,26] However, the diagnostic value of PLR has varied among studies, possibly due to the influence of anemia, co-infection, or platelet abnormalities.[27] These inconsistencies support the idea that a combined approach using several indices may provide a more comprehensive reflection of immune and inflammatory status in TB.

From a clinical perspective, CBC-derived ratios such as NLR and MLR are low-cost, easily obtainable, and available in most clinical laboratories. In regional hospitals where molecular diagnostics such as nucleic acid amplification tests (NAATs) are not widely accessible, these indices can serve as simple adjunctive tools for identifying patients with strong inflammatory responses, prioritizing diagnostic evaluation, and monitoring treatment outcomes.[8,27] The combined interpretation of NLR, MLR, and PLR may therefore improve diagnostic efficiency and reflect the overall inflammatory burden in TB.

Limitations

This study has several limitations. First, its retrospective design and single-center setting limit the ability to establish causality and may restrict the generalizability of the findings. Second, potential confounding factors such as nutritional status and concurrent infections were not fully controlled, which might have influenced the hematologic parameters. Third, detailed clinical indicators of systemic inflammation, disease severity, and quantitative bacterial load were not routinely available in the medical records. Microbiological data were limited to AFB smear and NAAT results, which were used primarily for diagnostic purposes rather than as quantitative measures of bacterial burden. Therefore, further multicenter prospective studies are warranted to validate these findings and to assess the prognostic value of NLR, MLR, and PLR during anti-TB therapy. In addition, this study focused on routinely available CBC-derived ratios. Advanced hematological parameters, such as neutrophil and monocyte activation indices (Neu-X, Neu-Y, Mon-X, and Mon-Y), were not included because these parameters are analyzer-dependent and were not consistently available in the retrospective dataset. Future prospective studies incorporating these advanced indices may provide additional insight into immune activation in tuberculosis.

Conclusion

In summary, patients with tuberculosis showed increased NLR, MLR, and PLR compared with healthy controls, reflecting enhanced systemic inflammation. Among them, MLR had the strongest association with TB, while the combined model of NLR, MLR, and PLR provided the best diagnostic performance. These results emphasize the potential value of integrating multiple CBC-derived ratios as affordable and accessible biomarkers for assessing systemic inflammation and disease activity, particularly in hospitals with limited laboratory resources.

Ethical approval

This study was approved by the Ethics Committee for Human Research, Lampang Hospital, Thailand (Approval No. EC 098.1/68).

Funding

This research received no external funding.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Sittichok Chaiwong: data curation, investigation, resources, writing: review and edit; **Niphawan Jettavan**: data curation, investigation, resources, writing: review and edit; **Yuttana Mundee:** validation, writing: review and edit; **Pawaret Panyajai:** validation, writing: review and edit; **Wuttichote Jansaento**: validation, writing: review and edit; **Watcharapong Jugnam-ang**: conceptualization, methodology, formal analysis, writing: original draft, review and edit. All authors have read and approved the final version of the manuscript.

Acknowledgements

The authors would like to express their sincere gratitude to the Department of Medical Technology, Lampang Hospital for providing technical support and access to clinical data. The authors also thank all medical technologists and laboratory staff for their assistance in data collection and sample analysis.

References

[1] Lv H, Wang L, Zhang X, Dang C, Liu F, Zhang X, et al. Further analysis of tuberculosis in eight high-burden countries based on the Global Burden of Disease Study 2021 data. Infect Dis Poverty. 2024; 13(1): 70. doi: 10.1186/s40249-024-01247-8.

[2] Yang H, Ruan X, Li W, Xiong J, Zheng Y. Global, regional, and national burden of tuberculosis and attributable risk factors for 204 countries and territories, 1990-2021: a systematic analysis for the Global Burden of Diseases 2021 study. BMC Public Health. 2024; 24(1): 3111. doi: 10.1186/s12889-024-20664-w.

[3] Heyckendorf J, Georghiou SB, Frahm N, Heinrich

N, Kontsevaya I, Reimann M, et al. Tuberculosis Treatment Monitoring and Outcome Measures: New Interest and New Strategies. Clin Microbiol Rev. 2022; 35(3): e0022721. doi: 10.1128/cmr.00227-21.

[4] Yayan J, Franke KJ, Berger M, Windisch W, Rasche K. Early detection of tuberculosis: a systematic review. Pneumonia (Nathan). 2024; 16(1): 11. doi: 10.1186/s41479-024-00133-z.

[5] Nyamagoud SB, Dsouza PD, Chitralu SPP, Solankure K, Swamy AHV. Evaluating hematological and inflammatory biomarkers in tuberculosis management. Monaldi Arch Chest Dis. 2025. doi: 10.4081/monaldi.2025.3433. Epub ahead of print.

[6] Leo S, Narasimhan M, Rathinam S, Banerjee A. Biomarkers in diagnosing and therapeutic monitoring of tuberculosis: a review. Ann Med. 2024; 56(1): 2386030. doi: 10.1080/07853890.2024.2386030.

[7] Adane T, Melku M, Ayalew G, Bewket G, Aynalem M, Getawa S. Accuracy of monocyte to lymphocyte ratio for tuberculosis diagnosis and its role in monitoring anti-tuberculosis treatment: Systematic review and meta-analysis. Medicine (Baltimore). 2022; 101(44): e31539. doi: 10.1097/MD.0000000000031539.

[8] Tu HZ, Lai TJ, Chen YS, Lee HS, Chen JS. Hematological Parameters as Potential Markers for Distinguishing Pulmonary Tuberculosis from Genitourinary Tuberculosis. Pathogens. 2023; 12(1): 84. doi: 10.3390/pathogens12010084.

[9] Obeagu EI. Inflammatory blood markers in breast cancer: a narrative review from early detection to therapy response. Ann Med Surg (Lond). 2025; 87(9): 5906-11. doi: 10.1097/MS9.0000000000003687.

[10] Xue F, Li Z, Yan L, Yu Q, Zhang Y. The Correlation of Blood Immune Cells with the Pathogenesis of Schizophrenia: A Meta-Analysis. Iran J Public Health. 2025; 54(8): 1563-76. doi: 10.18502/ijph.v54i8.19566.

[11] Omair M, Baig MS, Farooqui WA, Kousar S, Noori MY, Zeehan N, et al. Relationship of neutrophil lymphocyte ratio, monocyte lymphocyte ratio and neutrophil monocyte ratio with treatment response in pulmonary tuberculosis patients during intensive phase treatment. BMC Infect Dis. 2024; 24(1): 615. doi: 10.1186/s12879-024-09454-2.

[12] Hua HB, Wang HJ. Monocyte-to-lymphocyte ratio is significantly associated with positive Quanti FERON-TB Gold-In-Tube and adult survival: an observational study. Sci Rep. 2022; 12(1): 20406. doi: 10.1038/s41598-022-24376-2.

[13] Xiang ZB, Leng EL, Cao WF, Liu SM, Zhou YL, Luo CQ, et al. A systematic review and meta-analysis of the diagnostic accuracy of metagenomic next-generation sequencing for diagnosing tuberculous meningitis. Front Immunol. 2023; 14: 1223675. doi: 10.3389/fimmu.2023.1223675.

[14] Shojaan H, Kalami N, Ghasempour Alamdari M, Emami Alorizy SM, Ghaedi A, Bazrgar A, et al. Diagnostic value of the neutrophil lymphocyte ratio in discrimination between tuberculosis and bacterial community acquired pneumonia: A meta-analysis. J Clin Tuberc Other Mycobact Dis. 2023; 33: 100395. doi: 10.1016/j.jctube.2023.100395.

[15] Rahimirad S, Ghaffary MR, Rahimirad MH, Rashidi F. Association between admission neutrophil to lymphocyte ratio and outcomes in patients with acute exacerbation of chronic obstructive pulmonary disease. Tuberk Toraks. 2017; 65(1): 25-31.

[16] Miyahara R, Piyaworawong S, Naranbhai V, Prachamat P, Kriengwatanapong P, Tsuchiya N, et al. Predicting the risk of pulmonary tuberculosis based on the neutrophil-to-lymphocyte ratio at TB screening in HIV-infected individuals. BMC Infect Dis. 2019; 19(1): 667. doi: 10.1186/s12879-019-4292-9.

[17] Kaufmann SH. How can immunology contribute to the control of tuberculosis? Nat Rev Immunol. 2001;1(1):20-30. doi: 10.1038/35095558.

[18] Scriba TJ, Coussens AK, Fletcher HA. Human Immunology of Tuberculosis. Microbiol Spectr. 2017; 5(1): TBTB2-0016-2016. doi: 10.1128/microbiolspec.tbtb2-0016-2016.

[19] Hult C, Mattila JT, Gideon HP, Linderman JJ, Kirschner DE. Neutrophil Dynamics Affect Mycobacterium tuberculosis Granuloma Outcomes and Dissemination. Front Immunol. 2021; 12: 712457. doi: 10.3389/fimmu.2021.712457.

[20] Weeratunga P, Moller DR, Ho LP. Immune mechanisms of granuloma formation in sarcoidosis and tuberculosis. J Clin Invest. 2024; 134(1): e175264. doi: 10.1172/JCI175264.

[21] Chandra P, Grigsby SJ, Philips JA. Immune evasion and provocation by Mycobacterium tuberculosis. Nat Rev Microbiol. 2022; 20(12): 750-66. doi: 10.1038/s41579-022-00763-4.

[22] Li F, Chen D, Zeng Q, Du Y. Possible Mechanisms of Lymphopenia in Severe Tuberculosis. Micro organisms. 2023; 11(11): 2640. doi: 10.3390/microorganisms11112640.

[23] Xiao S, Ma X, Shi G, Zhou T, Pan J. Association between neutrophil-lymphocyte ratio and latent tuberculosis infection in the United States: a cross-sectional study from NHANES 2011-2012. BMC Infect Dis. 2025;25(1):1139. doi: 10.1186/s12879-025-11505-1.

[24] Gulgun M. Assessment of Platelet-to-Lymphocyte Ratio Can Be Affected by Several Factors. Med Princ Pract. 2017; 26(3): 299-300. doi: 10.1159/000456664.

[25] Kim MA, Park YE, Chong YP, Shim TS, Jo KW. Neutrophil-Lymphocyte Ratio and Monocyte-Lymphocyte Ratio According to the Radiologic Severity of Mycobacterium avium Complex Pulmonary Disease. J Korean Med Sci. 2022; 37(40): e292. doi: 10.3346/jkms.2022.37.e292.

[26] M Y, Vamja R, Parmar PA, Makwana N, Sundar RNS.

Haematological markers as predictive tools for tuberculosis in PLHIV: a retrospective cohort study in Gujarat, India. BMC Infect Dis. 2025; 25(1): 228. doi: 10.1186/s12879-025-10625-y.

[27] Rees CA, Pineros DB, Amour M, Munseri P, Said J, Magohe A, et al. The potential of CBC-derived ratios (monocyte-to-lymphocyte, neutrophil-to-lymphocyte, and platelet-to-lymphocyte) to predict or diagnose incident TB infection in Tanzanian adolescents. BMC Infect Dis. 2020; 20(1): 609. doi: 10.1186/s12879-020-05331-w.

Journal of Associated Medical Sciences 2026; 59(2): 268-279

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Effect of botulinum toxin-A phonophoresis on spasticity in children with hemiplegia

Ahmed F Attia[1]*, Amira M El-Tohamy[1], Gehan H El-Meniawy[1], Walaa ElNaggar[2]

[1]Department of Physical Therapy for Pediatrics, Faculty of Physical Therapy, Cairo University, Egypt.
[2]Department of Pediatrics, Faculty of Medicine, Cairo University, Egypt.

ARTICLE INFO

Article history:
Received 18 October 2025
Accepted as revised 12 February 2026
Available online 3 March 2026

Keywords:
Cerebral palsy, botulinum toxin-A phonophoresis, spasticity, Eelectromyography

ABSTRACT

Background: Spasticity is one of the most challenging problems in the management of cerebral palsy. Various trials, treatment modalities, therapeutic approaches, and exercises have been used to treat or modulate spasticity. Phonophoresis uses therapeutic ultrasound to enhance the percutaneous absorption of medication.

Objectives: To determine the effect of botulinum toxin-A phonophoresis on calf muscle spasticity and ankle range of motion in children with hemiplegia.

Materials and methods: This study was conducted on thirty children with hemiplegic cerebral palsy. Their age ranged from 5 to 10 years, their grade of ankle plantar flexors spasticity ranged from 2 to 3 according to Modified Ashworth Scale and they were able to walk alone with some limitations, classified as level II according to Gross Motor Function Classification System (GMFCS). They were randomly classified into two groups: the study group, which received botulinum toxin-A phonophoresis in addition to traditional physiotherapy program, and the control group, which received placebo ultrasonic therapy and the same traditional physiotherapy program. The Neuro-EMG-Micro (Neurosoft) was used as surface electromyography device to measure H/M ratio of calf muscles. The EasyAngle digital goniometer was used to measure ankle dorsiflexion range of motion. The SoLo Therasonic 455 ultrasonic device was used to apply botulinum toxin-A phonophoresis.

Results: Comparison within the study group revealed a significant decrease in post-treatment H/M ratio of gastrocnemius (GC) and soleus muscles compared to pre-treatment values. Additionally, there was a significant increase in post-treatment passive and active dorsiflexion range of motion (DF ROM) of ankle joint. Furthermore, there was a significant difference between the study and control groups in post-treatment H/M ratio of GC and soleus muscles, as well as in post-treatment passive and active DF ROM of ankle joint.

Conclusion: Botulinum toxin-A phonophoresis significantly decreases spasticity of calf muscle and improves ankle ROM in children with hemiplegic cerebral palsy, especially when combined with traditional physical therapy.

Introduction

Calf muscle spasticity is a common feature in children with hemiplegic cerebral palsy that can lead to equinus deformity, gait abnormalities, and functional limitations. It results from hyperexcitability of the stretch reflex. Persistent spasticity interferes with motor development, balance, and mobility, and

* **Corresponding contributor.**
Author's Address: *Department of Physical Therapy for Pediatrics, Cairo University, Egypt.*
E-mail address: ahmed.fathi@pt.cu.edu.eg
doi: 10.12982/JAMS.2026.058
E-ISSN: 2539-6056

often necessitates physical therapy, orthoses, or pharmacological interventions.[1-3]

The H-reflex to M-wave (H/M) ratio is widely used to assess spinal excitability in calf muscles. Normal H/M ratio ranges from 0.5 to 0.7 in healthy individuals. It is measured by electrical stimulation of the posterior tibial nerve in the popliteal fossa and recording commonly from the soleus and can also be recorded from gastrocnemius (GC) muscle. The maximal amplitude of the H-reflex is often obtained with low intensity and with the gradual increase of intensity, the M response increases its amplitude while the H-reflex amplitude is decreasing. An increased H/M ratio indicates enhanced alpha motor neuron excitability, commonly seen in spasticity.[4-6]

Botulinum toxin type-A (BoNT-A) reduces spasticity by inhibition of acetylcholine release at the neuromuscular junction. It cleaves synaptosomal associated protein of 25 kilodaltons (SNAP-25), a protein essential for synaptic vesicle fusion, thereby preventing muscle fiber depolarization and action potential. This leads to temporary chemodenervation, muscle relaxation, and reduced reflex hyperexcitability. Its effect typically lasts 3-6 months, providing a therapeutic window for physiotherapy and functional training in patients with spasticity.[7,8]

During the past 25 years, BoNT-A injection has become the most widely used medical intervention in children with cerebral palsy for spasticity management. Intramuscular injection of BoNT-A causes a decrease in muscle activity, usually reported as a reduction in spasticity. Injection of the GC muscle in children with hemiplegia usually increases the ankle dorsiflexion ROM.[9] Intramuscular BoNT-A injection is an effective option in spasticity treatment. However, it is limited by its cost and local discomfort.[10] Botulinum toxin-A injection is considered a complex medical intervention that may require general anesthesia for children and radiological guidance for drug delivery and requires highly skilled medical team members including a neurosurgeon, anesthesiologist, nurse and possibly a radiologist. Phonophoresis is a noninvasive method of administering molecules through the skin via ultrasonic waves that increase the transdermal absorption of drugs. Drugs such as BoNT-A can be delivered transdermally via phonophoresis or iontophoresis.[10-12] Phonophoresis can be used for delivery of BoNT-A with no adverse effects. It is safer than injection, relatively inexpensive and easy to apply.[10]

Treatment of spasticity in children with cerebral palsy by BoNT-A phonophoresis may provide a new physical therapy technique and may open a new field for physiotherapists in the treatment of cerebral palsy, which will result in improvement of physical therapy services.

Materials and methods

Study design

It was a single-blind randomized controlled study design. A preliminary power analysis was conducted for the soleus H/M ratio as the primary outcome. Using a pilot study of 10 participants (5 per group), the effect size was 1.92. Two-tailed analysis with α=0.05 and actual power (1-β)=96.8% determined a required sample size of 9 per group, which was increased by 66% to 15 per group to account for expected dropouts. Sample size calculation was performed using G*Power software (3.1.9.2).

Participants

Thirty children (17 boys and 13 girls) with hemiplegic cerebral palsy participated in this study; they were recruited from the outpatient clinic of the Faculty of Physical Therapy, Cairo University. The age range of the selected children was from 5 to 10 years old. Their ankle plantar flexors spasticity grade ranged from 2 to 3 according to Modified Ashworth Scale and they were able to walk alone with some limitations, classified as level II according to Gross Motor Function Classification System (GMFCS). They were able to follow instructions and understand commands. Children who had bone deformities in the lower leg, ankle or foot, fixed contracture of ankle plantar flexors, any metal implant in the lower leg near the area of treatment application were excluded from the study. Children who had received botulinum toxin injection or undergone surgery in gastro-soleus muscle within the 12 months prior to taking part in the study were also excluded.

Randomization

The thirty eligible children in this study were randomly classified into two equal groups through simple lottery method. Allocation concealment was ensured by sealed, opaque envelops prepared by independent person not involved in recruitment or assessment (Figure 1).

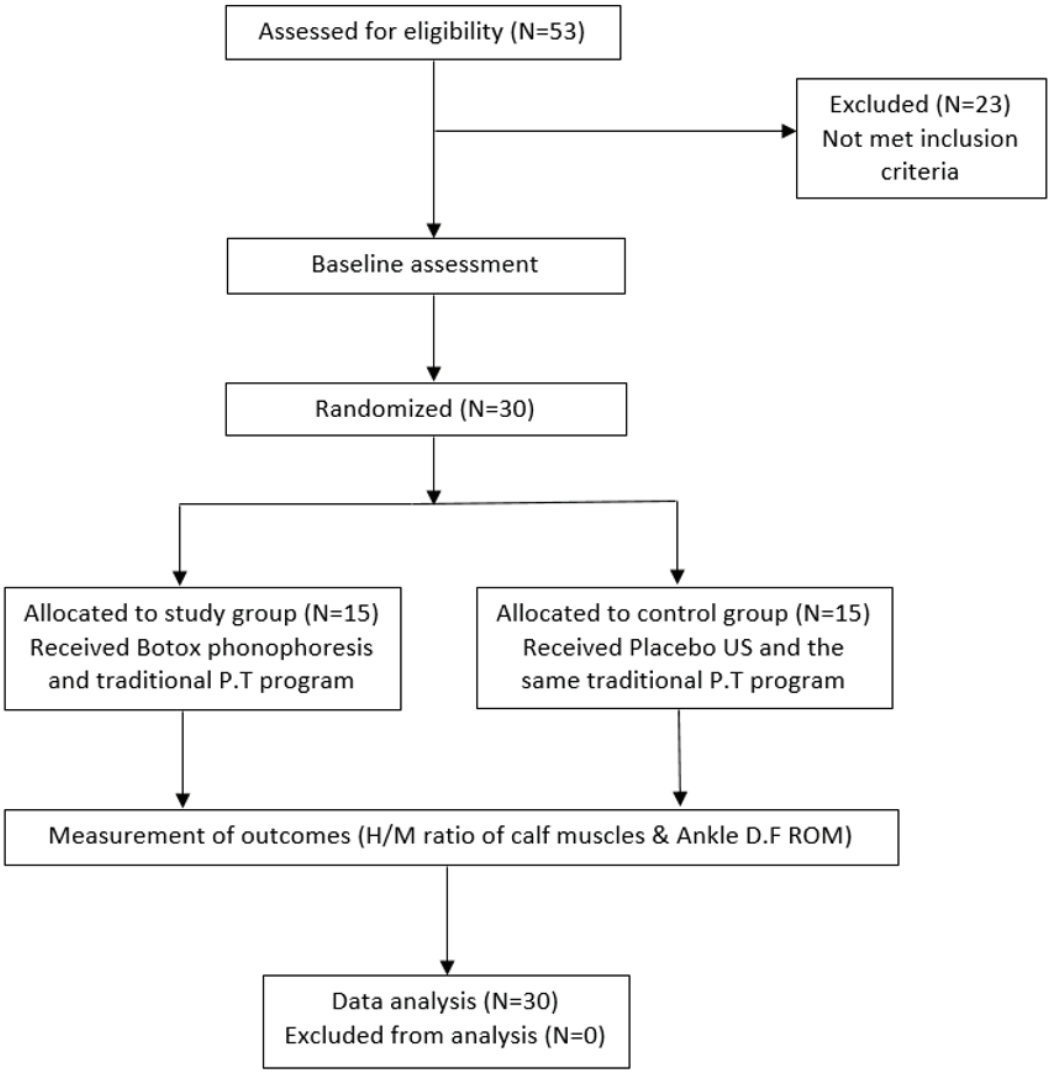


Figure 1. *Flowchart of the study.*

Study group (group A)

Botulinum toxin-A phonophoresis was applied to subjects of this group in addition to a designed traditional physical therapy program.

Control group (group B)

Subjects in this group received placebo ultrasonic therapy in addition to the same traditional physical therapy program as the study group.

Measurement of ankle dorsiflexion ROM

EasyAngle (EA) digital goniometer (Meloq AB, Stockholm, Sweden) was used to measure ankle dorsiflexion ROM (Figure 2). It is a device powered by a rechargeable battery. It has inertial measurement unit technology, with a three-axis accelerometer and a three-axis gyroscope to detect and determine its motion and position in space. It is a recent instrument for measuring joint ROM.[13] EasyAngle is supported by clinical evidence with many EA studies completed and more underway, it conforms to the highest clinical standards. With EA, the process of taking these measurements is fast, consistent, and reliable. The EA digital goniometer is a powerful tool in the hands of a physical therapist or other movement professional.[14] It is valid, reliable, easy, and fast for measuring ROM.[15]



Figure 2. *EasyAngle (EA) digital goniometer.*

The dorsi-flexion range of motion (DF ROM) was measured passively and actively before starting and at the end of treatment procedures. The child assumed supine lying position on a plinth with extended knee and full ankle plantar flexion. As shown in Figure 3, the goniometer was positioned and fixed by strap on the lateral aspect of the foot with the black line of goniometer's ruler on the lateral aspect of fifth metatarsal bone. To measure the passive DF ROM, the therapist clicked on the power button to turn on the device, then clicked on the main button to set the device to zero. The therapist clicked again on the main button taking the ankle joint toward dorsiflexion by grasping the child's foot by one hand while fixing the leg by the other hand then he clicked for the third time on the main button to record the angle once the available DF ROM was reached, defined as the first point of firm resistance perceived without eliciting pain or allowing compensatory movements. The measurement was repeated three times, and the mean value was recorded. The same procedure was applied for measurement of active DF ROM by giving instructions to the child to actively dorsiflex the ankle.

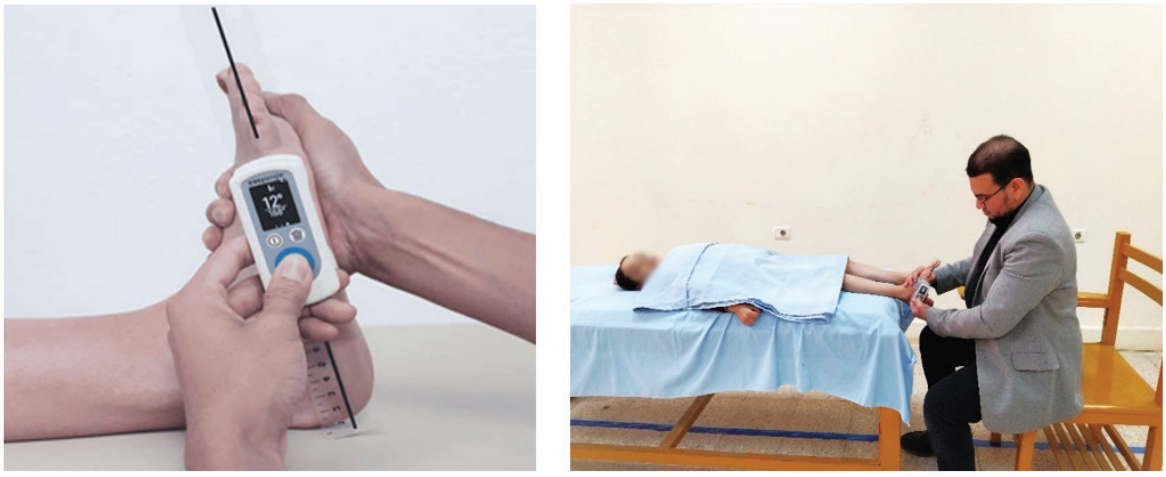


Figure 3. *ROM measurement.*

Surface electromyography

For measuring the H/M ratio, an electromyography device (Neurosoft, Ivanovo, Russia) was used for recording of myoelectric signals from calf muscles of the affected limb. The device consists of computer processing unit, screen, printer, keyboard and other components including amplifier, bipolar stimulating electrode, recording and ground electrodes and dedicated keyboard as shown in Figure 4. Surface electromyography (sEMG) is valid and reliable tool to assess the electrical activity of the muscles. It has been used in research and clinical applications for the non-invasive analysis of the myoelectric activity of the calf muscle.[16] Calf muscle electromyography was also done before starting and at the end of treatment procedures.

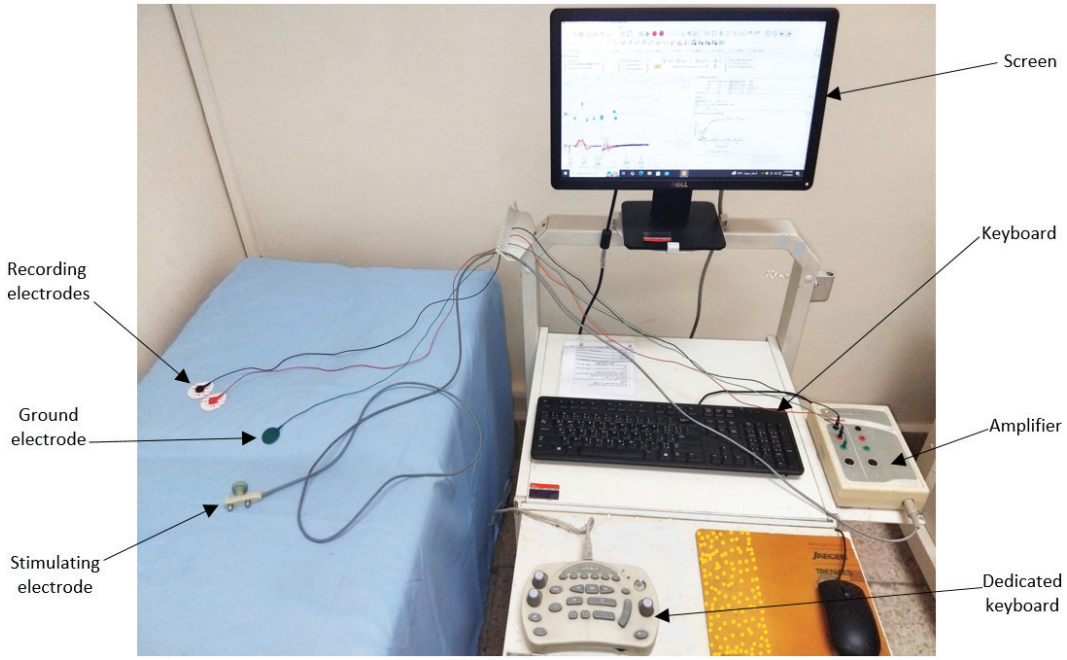


Figure 4. *Surface electromyography system.*

Before starting the electromyography procedure, all children refrained from strenuous activities and caffeine intake and stopped taking any anti-spastic drugs at least 12 hours before EMG evaluation. All procedures were applied with the patient lying in prone position and the foot outside the plinth, well positioned and relaxed for easy access to the calf muscle. The surface area of the calf muscle was cleaned with cotton wool and alcohol to reduce skin resistance.

For measuring the H/M ratio of soleus muscle as shown in Figure 5, the active recording electrode (negative, black electrode) was placed on the soleus muscle just below the contour of the gastrocnemius (GC) muscle, while the reference electrode (positive, red electrode) was placed on the heel. The ground electrode was placed midway between the negative and positive electrodes.[17,18]

For measuring the H/M ratio of GC muscle, the recording electrodes were placed on the bulkiest area of the medial head of GC muscle, and the ground electrode was placed distally on the heel (Figure 5). The stimulating electrode was used to stimulate the posterior tibial nerve at the back of knee. The stimulus duration was 1 msec. to stimulate sensory nerve fibers (Ia). It was stimulated manually at rate of once every ten seconds to reduce post-activation depression effects. The intensity of the current was gradually increased and the H-reflex and M-wave were observed on the screen. After recording, a clean towel was used to wipe the recording area.[17,18]



Figure 5. *Electrodes placement during EMG of soleus and GC muscles.*

The electromyography signals were checked on an amplitude-time display of the recordings on the screen to exclude artifacts, such as direct current artifact, motion artifact or electric noise before calculation of electromyography variables. The H_{max} to M_{max} ratio was computed via the system software by dividing the H_{max} amplitude by M_{max} amplitude.[18] The report of each muscle measurement was then printed and saved in the system software.

Intervention

Botulinum toxin-A phonophoresis session

The SoLo Therasonic 455 device (EMS Physio Ltd, Oxfordshire, England) was used for application of BoNT-A (Figure 6). It offers a broad range of uses with state-of-the-art technology in a unique design. It has the availability to use 1MHz or 3MHz frequency from the same treatment head.

Vials of BoNT-A (Dysport® 500 U/vial, Ipsen Ltd, UK) were used to prepare a mixture with carbopol gel. Each vial was dissolved in 5 ml of sterile saline solution (0.9% NaCl) and each 1 ml of the solution was gently mixed with 50 grams of carbopol gel in a sterile, capped small container. A separate labeled container was assigned for each child in the study group. All containers were stored in the refrigerator between 2 °C and 8 °C and the total amount was used for phonophoresis sessions.[10] The dosage of 100 units per child was determined based on the recommended intramuscular injection dosage (3 to 6 units/kg),[19] although the actual amount absorbed transdermally via phonophoresis remains undetermined. This dosage was selected empirically as no standardized transdermal recommendations currently exist for BoNT-A phonophoresis.

Each child in the study group received 10 phonophoresis sessions over 5 consecutive days (two sessions per day). This schedule was adopted in consideration of the stability of reconstituted BoNT-A, which does not exceed 5 days when stored in the refrigerator at 2 °C to 8 °C. A minimum interval of two hours was maintained between the two daily sessions to allow for optimal transdermal absorption and to minimize potential saturation of local tissue receptors. The ultrasonic device was applied in continuous mode, with a frequency of 1MHz, an intensity of 1W/cm^2 and a duration of 10 minutes per session.[10] Each child in the control group received 10 placebo ultrasonic sessions with the same schedule and session time.

All phonophoresis sessions as shown in Figure 6 were applied while the child was in prone position with the feet outside the plinth. The patient's skin was carefully inspected for the barriers of transcutaneous absorption of drugs such as poor circulation and skin dehydration to ensure the maximum effectiveness of phonophoresis. The surface area of the calf muscle was cleaned with cotton wool and alcohol, then an appropriate amount of gel was applied on the calf muscle and the parameters were set. The ultrasonic head was moved slowly in a circular motion over the calf muscle bulk, while maintaining a safe distance from the epiphyseal plate of growing bone. After the end of session, tissues were used to wipe the area and the skin was inspected for any signs of irritation.



Figure 6. *Phonophoresis session and SoLo Therasonic 455 device.*

Traditional physical therapy program

All participants in this study received the same traditional physiotherapy program for one month (12 sessions/month, 3 sessions/week, and 1 hr/session) after the end of all phonophresis and placebo ultrasonic therapy sessions. The program consisted of functional strengthening exercises, isolated strengthening exercises for ankle dorsi-flexors, stretching exercises for calf muscles, balance exercises and gait training. In addition, participants were instructed to wear night ankle-foot orthosis with adjustable straps and knee immobilizer at least 8 hours daily to apply positional prolonged stretching of the calf muscles.

Statistical analysis

Data were analyzed using SPSS Version 26.0. For normality test of data, Shapiro-Wilk test was performed. Unpaired t-test was used for comparison of age, pre-treatment active and passive ankle DF ROM, pre-treatment H/M ratio of GC and soleus muscle, post-treatment active and passive ankle DF ROM and post-treatment H/M ratio of GC and soleus muscle between both groups. Paired t-test was used for comparison of pre and post-treatment H/M ratio of soleus muscle within both groups and for comparison of ankle active DF ROM within group A and ankle passive DF ROM within group B. Wilcoxon signed rank test was conducted for comparison of pre and post-treatment H/M ratio of GC muscle within both groups and for comparison of ankle passive DF ROM within group A and ankle active DF ROM within group B. Chi-squared test was performed for comparison of sex and hemiplegic side distribution between the two groups. Mann-Whitney test was used for comparison of spasticity grade distribution between both groups. Descriptive statistics such as arithmetic mean and standard deviation were also performed. The level of significance for all statistical tests was set at $p<0.05$.

Results

Subject characteristics

As shown in Table (1) there was no significant difference between both groups in age ($p>0.05$). In addition, there was no significant difference in the distribution of sex, hemiplegic side, and spasticity grades between both groups ($p>0.05$).

Table 1. Basic characteristics of participants.

	Group A (N=15)	Group B (N=15)	*p* value
Age (years) mean±SD	8.20±1.70	8.07±1.91	0.841
Gender [N (%)]			
Girls	7 (46.7%)	6 (40%)	0.713
Boys	8 (53.3%)	9 (60%)	
Hemiplegic side [N (%)]			
Right	10 (66.7%)	9 (60%)	0.705
Left	5 (33.3%)	6 (40%)	
Spasticity grades [N (%)]			
MAS, Grade II	11 (73.3%)	10 (66.7%)	0.695
MAS, Grade III	4 (26.7%)	5 (33.3%)	

Note: *MAS: modified Ashworth scale*

Effect of treatment on H/M ratio and ROM

Comparison within group A

Comparison within group A revealed a significant decrease in post-treatment H/M ratio of GC muscle (p=0.001) and a significant decrease in post-treatment H/M ratio of soleus muscle (p<0.001) when compared with those of pre-treatment. In addition, there was a significant increase in post-treatment passive dorsiflexion ROM of ankle joint (p=0.001) and a significant increase in post-treatment active dorsiflexion ROM of ankle joint (p<0.001) when compared with those of pre-treatment. (Table 2).

Table 2. Comparison of pre- and post-treatment H/M ratio and ROM within group A.

Variable	Group A				
	Pre-treatment (Mean±SD)	**Post-treatment** (Mean±SD)	**MD**	**% of change**	***p* value**
GC H/M ratio, (%)	60.56±21.33	28.78±12.60	31.78	52.48%	0.001
Soleus H/M ratio, (%)	36.33±9.94	16.39±7.66	19.59	54.91%	<0.001
Passive DF ROM, (degrees)	56.53±5.85	64.47±4.21	-7.93	14.03%	0.001
Active DF ROM, (degrees)	21.53±6.93	37.80±7.05	-16.23	75.38%	<0.001

Note: *GC: gastrocnemius muscle, DF: dorsiflexion, ROM: Range of motion, MD: mean difference, H/M: Hoffman reflex/ Motor wave.*

Comparison within group B

Comparison within group B revealed a non-significant decrease in post-treatment H/M ratio of GC muscle (p=0.069) and a significant decrease in post-treatment H/M ratio of soleus muscle (p=0.009) when compared with those of pre-treatment. In addition, there was a significant increase in post-treatment passive dorsiflexion ROM of ankle joint (p<0.001) and a significant increase in post-treatment active dorsiflexion ROM of ankle joint (p=0.001) when compared with those of pre-treatment (Table 3).

Table 3. Comparison of pre- and post-treatment H/M ratio and ROM within group B.

Variable	Group B				
	Pre-treatment (Mean±SD)	**Post-treatment** (Mean±SD)	**MD**	**% of change**	***p* value**
GC H/M ratio, (%)	57.27±13.43	55.61±15.34	1.65	2.88%	0.069
Soleus H/M ratio, (%)	37.43±10.00	35.33±11.05	2.09	5.58%	0.009
Passive DF ROM, (degrees)	55.53±4.85	57.73±4.88	-2.20	3.96%	<0.001
Active DF ROM, (degrees)	21.07±6.83	23.60±7.31	-2.53	12.01%	0.001

Note: *GC: gastrocnemius muscle, DF: dorsiflexion, ROM: range of motion, MD: mean difference, H/M: Hoffman reflex/ Motor wave.*

Comparison between groups

Pre-treatment comparison

Comparison between groups revealed a non-significant difference in both pre-treatment H/M ratio of GC muscle (p=0.617) and pre-treatment H/M ratio of soleus muscle (p=0.766). In addition, there was a non-significant difference in both pre-treatment passive dorsiflexion ROM of ankle joint (p=0.615) and pre-treatment active dorsiflexion ROM of ankle joint (p=0.854). (Table 4).

Table 4. Comparison of pre-treatment H/M ratio and ROM between groups.

Variable	**Group A** (Mean±SD)	**Group B** (Mean±SD)	**MD**	***p* value**
GC H/M ratio, (%)	60.56±21.33	57.27±13.43	3.29	0.617
Soleus H/M ratio, (%)	36.33±9.94	37.43±10.00	-1.09	0.766
Passive DF ROM, (degrees)	56.53±5.85	55.53±4.85	1.00	0.615
Active DF ROM, (degrees)	21.53±6.93	21.07±6.83	0.47	0.854

Note: *GC: gastrocnemius muscle, DF: dorsiflexion, ROM: range of motion, MD: mean difference, H/M: Hoffman reflex/Motor wave.*

Post-treatment comparison

Comparison between groups revealed a significant difference in post-treatment H/M ratio of GC muscle (p<0.001) and in post-treatment H/M ratio of soleus muscle (p<0.001). In addition, there was a significant difference in post-treatment passive dorsiflexion ROM of ankle joint (p<0.001) and in post-treatment active dorsiflexion ROM of ankle joint (p<0.001) (Table 5).

Table 5. Comparison of post-treatment H/M ratio and ROM between groups

Variable	Group A (Mean±SD)	Group B (Mean±SD)	MD	*p* value
GC H/M ratio, (%)	28.78±12.60	55.61±15.34	-26.83	<0.001
Soleus H/M ratio, (%)	16.39±7.66	35.33±11.05	-18.95	<0.001
Passive DF ROM, (degrees)	64.47±4.21	57.73±4.88	6.73	<0.001
Active DF ROM, (degrees)	37.80±7.05	23.60±7.31	14.20	<0.001

Note: *GC: gastrocnemius muscle, DF: dorsiflexion, ROM: range of motion, MD: mean difference, H/M: Hoffman reflex/Motor wave.*

Discussion

This randomized controlled study investigated the efficacy of BoNT-A delivered via phonophoresis on calf muscle spasticity and ankle ROM in children with hemiplegic CP. Thirty children were divided equally into a study group, which received BoNT-A phonophoresis plus traditional physiotherapy program and a control group, which received placebo ultrasound with the same physiotherapy program. Outcomes included the H/M ratio of the soleus and gastrocnemius muscles (measured via Neurosoft EMG) and active and passive ankle dorsiflexion ROM (assessed by an EasyAngle goniometer). The study group demonstrated significant decreases in H/M ratio for both gastrocnemius and soleus muscles, along with substantial increases in active and passive dorsiflexion ROM. The control group also showed significant but smaller improvements in H/M ratio (significant only for soleus) and smaller significant ROM gains, likely arising from the physical therapy alone. Between-group comparison clearly favored the study group significantly on all primary measures, indicating a superior effect attributable to BoNT-A phonophoresis.

Potential mechanisms of BoNT-A phonophoresis

To date, BoNT-A injections are widely recognized as an effective intervention for management of spasticity in children with CP. These injections reduce muscle hyperactivity by blocking acetylcholine release at neuromuscular junctions, leading to decreased spasticity, improved ROM, and enhanced function, though often transient and requiring adjunctive interventions such as casting or orthosis to maintain benefits. In this context, several systematic reviews confirmed the role of BoNT-A as a focal spasticity management tool when combined with comprehensive rehabilitation.[9,20,21]

The present outcomes of our study, marked reductions in H/M ratio and meaningful gains in ROM are congruent with injection-based findings, suggesting that BoNT-A delivered via phonophoresis may induce neuromodulatory effects comparable to those observed following injection without implying confirmed intramuscular delivery, potentially offering a non-invasive alternative and achieving greater improvements than physiotherapy alone.

Phonophoresis employs ultrasound waves to enhance transdermal delivery of pharmacological agents. Mechanisms include thermal effects, cavitation leading to increased skin permeability, and acoustic streaming that facilitates molecular transport.[22] Despite the relatively large molecular size of BoNT-A, the significant outcomes in spasticity and ROM indicate that effective skin penetration may have occurred. In comparison to intramuscular injections, phonophoresis could offer a safer and more user-friendly option, especially for children by avoiding needle-associated discomfort, infection risk, and possible muscle damage as a non-invasive approach; however, it should not be considered equivalent to intramuscular injection in terms of pharmacokinetics or tissue distribution.

Role of traditional physical therapy

Both study and control groups received the identical physiotherapy program, including stretching, night splints, balance and gait training, functional strengthening, and dorsiflexors strengthening. While these interventions alone led to small improvements in the control group, the combined BoNT-A phonophoresis significantly enhanced outcomes. This multiplicative effect mirrors injection-based protocols where BoNT-A is most effective when paired with appropriate rehabilitative strategies.[9,21]

Phonophoresis in Spasticity: Other Agents

Although BoNT-A phonophoresis is considered a recent method in spasticity management, other pharmacologic agents delivered via phonophoresis have been studied. During our literature review, we did not find any previous studies that used BoNT-A phonophoresis in the treatment of spasticity. Most similar previous studies used other medications like lidocaine, baclofen, hydrocortisone acetate, and non-

steroidal anti-inflammatory drugs including ibuprofen, piroxicam, diclofenac sodium, and ketoprofen.[23-25]

A previous study was conducted to identify the effect of baclofen phonophoresis on spasticity in children with diplegia. It showed that baclofen delivered via phonophoresis to gastrocnemius combined with traditional physical therapy had no significant effect on decreasing spasticity but led to significant improvement in standing ability mainly because of the physical therapy program as these findings were also presented in the control group that received only physiotherapy program.[25]

Another study in children with spastic diplegic CP applied phonophoresis (using unspecified agent) combined with physical therapy and compared outcomes against traditional therapy alone. The experimental group showed significant improvements in spasticity and ankle ROM.[26] This aligns with the current findings and supports the utility of phonophoresis in CP.

Botulinum toxin phonophoresis and iontophoresis in hyperhidrosis management

Although BoNT-A is typically administered via intradermal injection for hyperhidrosis, recent explorations have examined non-invasive transdermal delivery methods, including phonophoresis and iontophoresis, to avoid the discomfort and side effects associated with needle-based therapy.

There are previous studies that used phonophoresis or iontophoresis to deliver botulinum toxin A into tissues to evaluate its effect on the treatment of palmar and axillary hyperhidrosis.[10,27]

In one of these studies, BoNT-A phonophoresis was conducted to identify its effect on palmar hyperhidrosis by using ultrasonic with continuous mode, 1MHz frequency for 10 minutes. Also in this study, BoNT-A was successfully delivered via iontophoresis by using galvanic micro current to identify it effect on palmar hyperhidrosis. Both methods revealed a significant reduction in sweating with no adverse effect over a 16-week period with higher patient acceptance due to the painless procedure.[10]

In a related investigation to compare between BoNT-A injection and iontophoresis, researchers explored BoNT-A iontophoresis for axillary hyperhidrosis, applying alternating electrical current for 30 minutes with an electrode was soaked in BoNT-A solution and fixed over the axillary region. While the sample size was small, results indicated a measurable decrease in sweat production sustained over several weeks. This study indicates that BoNT-A iontophoresis may represent a promising alternative to injections, especially for patient's intolerant to needle-based therapies or seeking a less invasive option.[27]

These findings support the feasibility of BoNT-A transdermal delivery in targeting peripheral autonomic dysfunction and provide an important conceptual precedent for its application in neuromuscular conditions like spasticity.

Collectively, the previous studies that used phonophoresis with other pharmacological agents in spasticity treatment and those that were applied for hyperhidrosis reinforce that BoNT-A delivered transdermally via ultrasound can yield clinically meaningful outcomes in spasticity management without direct evidence of muscle penetration.

Safety and practical implications

Although traditional BoNT-A injection is effective in decreasing muscle spasticity, it is invasive and carries risks of pain, muscle atrophy, systemic spread, and infection, especially concerning in pediatric populations.[9,21] Phonophoresis potentially avoids these drawbacks while delivering comparable efficacy. However, as a novel modality for delivering BoNT-A, its safety profile, long-term effects, optimal dosing, and ultrasound parameters remain to be established. Future work should include imaging and histological studies to assess potential muscle morphology changes, systemic absorption, and functional integrity over time.

Limitations

Although participants were blinded to group allocation, blinding of the outcome assessor was not possible because of the practical constraints inherent to rehabilitation-based interventions. However, the primary outcome measures (H/M ratio and ankle dorsiflexion ROM) were objective and instrument-based, which minimizes the potential risk of assessor-related detection bias and strengthens the validity of the findings. As no standardized dosage guidelines exist for BoNT-A phonophoresis, the selected dose was pragmatically based on commonly used clinical injection doses, aiming to ensure safety rather than to establish dose equivalence.

The present study did not confirm the intramuscular delivery of BoNT-A; therefore, further studies are required to investigate the pharmacokinetics and tissue distribution of BoNT-A delivered via phonophoresis.

Conclusion

Botulinum toxin-A phonophoresis significantly decreases spasticity of calf muscle and improves ankle ROM in children with hemiplegic cerebral palsy who had ankle plantar flexors spasticity grade ranged from 2 to 3 according to Modified Ashworth Scale, especially when combined with traditional physical therapy.

Ethical approval

The Research Ethics Committee, Faculty of Physical Therapy, Cairo University approved this study under approval number (P.T.REC/012/005088) and each participant's caregiver signed an informed consent form prior to take part in the study.

Funding

There was not any financial support for the research, publication process and authorship.

Conflict of interest

There were no potential conflicts of interest related to the research, authorship, and/or publication of this article.

CRediT authorship contribution statement

Ahmed F Attia: conceptualization, methodology, investigation, data curation, formal analysis, visualization, writing: original draft; **Amira M El-Tohamy:** main supervisor, project administration, writing: review and edit; **Gehan H El-Meniawy:** supervision, writing: review and edit; **Walaa ElNaggar:** supervision, writing: review and edit.

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Prof. Dr. Amira M El-Tohamy, Prof. Dr. Gehan H El-Meniawy, and Prof. Dr. Walaa ElNaggar. No words can express my appreciation for their guidance and support.

References

[1] Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damino D, et al. A report: the definition and classification of cerebral palsy April 2006. Dev Med Child Neurol Suppl. 2007; 109: 8-14. doi: 10.1111/j.1469-8749.2007.tb12610.x.

[2] Damiano DL, Alter KE, Chambers H. New clinical and research trends in lower extremity management for ambulatory children with cerebral palsy. Phys Med Rehabil Clin N Am. 2009; 20(3): 469-91. doi: 10.1016/j.pmr.2009.04.005.

[3] Zahran DAE, Bahr WM, Abd Elazim FH. Systematic review: exercise training for equinus deformity in children with cerebral palsy. Bull Fac Phys Ther. 2022; 27: 37. doi: 10.1186/s43161-022-00093-9.

[4] Zehr EP. Considerations for use of the Hoffmann reflex in exercise studies. Eur J Appl Physiol. 2002; 86(6): 455-68. doi: 10.1007/s00421-002-0577-5.

[5] Knikou M. The H-reflex as a probe: pathways and pitfalls. J Neurosci Methods. 2008; 171(1): 1-12. doi: 10.1016/j.jneumeth.2008.02.012.

[6] Elsayed SEB, Abdel Raoof NA, Elsayed OM. Effect of whole-body vibration on motor neuron excitability in healthy young men. Bull Fac Phys Ther. 2016; 21: 48-55. Doi: 10.4103/1110-6611.188027.

[7] Simpson DM, Gracies JM, Graham HK, Miyasaki JM, Naumann M, Russman B, et al. Assessment: Botulinum neurotoxin for the treatment of spasticity (an evidence-based review): report of the therapeutics and technology assessment subcommittee of the american academy of neurology. Neurology. 2008; 70(19): 1691-8. doi: 10.1212/01.wnl.0000311391.00944.c4.

[8] Dressler D, Adib Saberi F. Botulinum toxin: mechanisms of action. Eur Neurol. 2005; 53(1): 3-9. doi: 10.1159/000083259.

[9] Multani I, Manji J, Hastings-Ison T, Khot A, Graham K. Botulinum toxin in the management of children with cerebral palsy. Paediatr Drugs. 2019; 21(4): 261-81. doi: 10.1007/s40272-019-00344-8.

[10] Andrade PC, Flores GP, Uscello Jde F, Miot HA, Morsoleto MJ. Use of iontophoresis or phonophoresis for delivering botulinum toxin A in the treatment of palmar hyperhidrosis: a report on four cases. An Bras Dermatol. 2011; 86(6): 1243-6. doi: 10.1590/S0365-05962011000600037.

[11] Karatay S, Aygul R, Melikoglu MA, Yildirim K, Ugur M, Erdal A, et al. The comparison of phonophoresis, iontophoresis and local steroid injection in carpal tunnel syndrome treatment. Joint Bone Spine. 2009; 76(6): 719-21. doi: 10.1016/j.jbspin.2009.02.008.

[12] Wahba ES. Effect of calcipotriol plus betamethasone dipropionate gel phonophoresis on psoriasis: a single-blind randomized controlled trial. Bull Fac Phys Ther. 2019; 24: 57-65. Doi: 10.4103/bfpt.bfpt_23_18.

[13] Bergh A, Lauridsen NG, Hesbach AL. Concurrent validity of equine joint range of motion measurement: a novel digital goniometer versus universal goniometer. Animals (Basel). 2020; 10(12): 2436. doi: 10.3390/ani10122436.

[14] Luedtke K, Schoettker-Königer T, Hall T, Reimer C, Grassold M, Hasselhoff-Styhler P, Neulinger C, et al. Concurrent validity and reliability of measuring range of motion during the cervical flexion rotation test with a novel digital goniometer. BMC Musculoskelet Disord. 2020; 21: 535. doi: 10.1186/s12891-020-03525-6.

[15] Svensson M, Lind V, Löfgren Harringe M. Measurement of knee joint range of motion with a digital goniometer: a reliability study. Physiother Res Int. 2019; 24(2): e1765. doi: 10.1002/pri.1765.

[16] Hu B, Zhang X, Mu J, Wu M, Wang Y. Spasticity assessment based on the Hilbert–Huang transform marginal spectrum entropy and the root mean square of surface electromyography signals: a preliminary study. Biomed Eng Online. 2018; 17(1): 27. doi: 10.1186/s12938-018-0460-1.

[17] Palmieri RM, Ingersoll CD, Hoffman MA. The Hoffmann reflex: methodologic considerations and applications for use in sports medicine and athletic training research. J Athl Train. 2004; 39(3): 268-77.

[18] Tekgül H, Polat M, Tosun A, Serdaroğlu G, Gökben S. Electrophysiologic assessment of spasticity in children using H-reflex. Turk J Pediatr. 2013; 55: 519-23.

[19] Camargo CHF, Teive HAG, Zonta MB, Silva G. Botulinum toxin type A in the treatment of lower-limb spasticity in children with cerebral

palsy. Arq Neuropsiquiatr. 2009; 67(1): 62-68. doi: 10.1590/S0004-282X2009000100016

[20] Yang H, Chen S, Shen J, Chen Y, Lai M, Chen L, et al. Safety and efficacy of botulinum toxin type A injection in children with spastic cerebral palsy aged <2 years: a systematic review. J Child Neurol. 2023; 38(6-7): 454-65. doi: 10.1177/08830738231183484

[21] Papavasiliou AS, Nikaina I, Foska K, Bouros P, Mitsou G, Filiopoulos C. Safety of botulinum toxin A in children and adolescents with cerebral palsy in a pragmatic setting. Toxins (Basel). 2013; 5(3): 524-36. doi: 10.3390/toxins5030524.

[22] Vranić E. Sonophoresis - mechanisms and application. Bosn J Basic Med Sci. 2004; 4(2): 25-32. doi: 10.17305/bjbms.2004.3410.

[23] Klaiman MD, Shrader JA, Danoff JV, Hicks JE, Pesce WJ, Ferland J. Phonophoresis versus ultrasound in the treatment of common musculoskeletal conditions. Med Sci Sports Exerc. 1998; 30(9): 1349-55. doi: 10.1097/00005768-199809000-00002.

[24] Martin-Vega FJ, Lucena-Anton D, Galan-Mercant A, Perez-Cabezas V, Luque-Moreno C, Vinolo-Gil MJ, et al. Phonophoresis through nonsteroidal anti-inflammatory drugs for knee osteoarthritis treatment: systematic review and meta-analysis. Biomedicines. 2022; 10(12): 3254. doi: 10.3390/biomedicines10123254.

[25] Abd Elghany MM, El-Meniawy GH, ElNabarawi MA, Abd El Aziz HG. Effect of baclofen phonophoresis on spasticity in gastrocnemius muscle in children with spastic diplegia [master's thesis]. Faculty of Physical Therapy: Cairo University; 2022.

[26] Al-Zahrani YA, Azam AM. Efficacy of phonophoresis on spastic ankle plantar flexors in diplegic CP children. SJR Publishing [internet]. 2022; (8). Available from: https://sjr-publishing.com/wp-content/uploads/2019/03/Efficacy-of-Phonophoresis-on-Spastic-Ankle-Plantar-Flexors-Control-in-Diplegic-Cerebral-Palsy-Children-1.pdf.

[27] Montaser-Kouhsari L, Zartab H, Fanian F, Noorian N, Sadr B, Nassiri-Kashani M, et al. Comparison of intradermal injection with iontophoresis of abobotulinum toxin A for the treatment of primary axillary hyperhidrosis: a randomized controlled trial. J Dermatolog Treat. 2014; 25(4): 337-41. doi: 10.3109/09546634.2012.739679.

Journal of Associated Medical Sciences 2026; 59(2): 280-285

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Association of aerobic and anaerobic capacities with body composition in basketball players aged 13-16 years in Chiang Mai Province

Panida Prasertpong[1], Sothida Nantakool[2], Busaba Chuatrakoon[2]*

[1]Multidisciplinary and Interdisciplinary School, Chiang Mai University, Chiang Mai Province, Thailand
[2]Integrated Neuro-Musculoskeletal, Chronic Disease, and Aging Research Engagement Center (ICARE Center), Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand

ARTICLE INFO

Article history:
Received 7 May 2025
Accepted as revised 2 March 2026
Available online 11 March 2026

Keywords:
Aerobic capacity, anaerobic performance, body composition, adolescent athletes, basketball training

ABSTRACT

Background: Basketball demands both aerobic and anaerobic energy, with adolescent physiological development influencing performance capacities. While body composition impacts energy efficiency, its relationship to aerobic and anaerobic capacities in youth athletes remains underexplored.

Objectives: To examine associations between body composition and aerobic and anaerobic capacities in adolescent basketball players, and to compare these variables by gender.

Materials and methods: A cross-sectional study was conducted with 60 basketball players (33 males, 27 females) aged 13-16 years. Body composition was measured using bioelectrical impedance analysis. Aerobic capacity was assessed via the 20-meter shuttle run (VO_2max estimation), and anaerobic performance via the Running-based Anaerobic Sprint Test (RAST). Pearson's Correlation Coefficient and Independent sample t-test were used for analysis.

Results: Males had significantly higher muscle mass and VO_2max, and lower body fat than females (p<0.05). VO_2max was positively correlated with muscle mass (r=0.466, p<0.001), and negatively with BMI (r=-0.330, p=0.01) and body fat (r=-0.463, p<0.001). No significant correlations were found between anaerobic fatigue index and body composition.

Conclusion: Aerobic capacity in youth basketball players is significantly influenced by body composition, with greater muscle mass and lower fat associated with better VO_2max. However, anaerobic capacity appears independent of these morphological factors. These findings support tailored training and monitoring strategies, especially considering sex-based developmental differences.

Introduction

Basketball is a globally popular sport that demands frequent and intense physical activity.[1,2] Movements such as sprinting, changing direction, jumping, and rapid acceleration-deceleration patterns require a combination of aerobic and anaerobic energy systems.[3] Notably, anaerobic energy is dominant in basketball due to its intermittent high-intensity nature, contributing to approximately 80% of energy utilization, with the remaining 20% reliant on aerobic metabolism.[4] Basketball players frequently undergo training designed to improve performance across both

* **Corresponding contributor.**
Author's Address: *Integrated Neuro-Musculoskeletal, Chronic Disease, and Aging Research Engagement Center (ICARE Center), Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: busaba.c@cmu.ac.th
doi: 10.66285/JAMS.2026.059
E-ISSN: 2539-6056

aerobic and anaerobic domains. The anaerobic system is critical for short bursts of high-intensity efforts, such as fast breaks, quick direction changes, and jumps. Meanwhile, aerobic capacity supports recovery between these bouts and sustains performance across the full duration of the game.[4] These two energy systems work in concert to enable optimal in-game performance, particularly in young athletes who are still undergoing physiological maturation.

Previous research has reported that aerobic capacity in adolescent male basketball players typically ranges between approximately 48-60 ml/kg/min, whereas female players generally demonstrate values between 38-50 mL/kg/min, depending on age and competitive level.[5] These variations reflect differences in training exposure and biological maturation. Despite the predominance of anaerobic actions during competition, adequate aerobic fitness is essential for maintaining repeated sprint performance and facilitating recovery during intermittent play.[2,4] Similarly, anaerobic performance in youth basketball players, often assessed through repeated sprint protocols, shows considerable variability influenced by neuromuscular development and lean muscle mass.[3,5] During adolescence, progressive increases in muscle mass, particularly in males, contribute to improvements in anaerobic power output. Therefore, understanding normative performance characteristics during this developmental stage is essential for optimizing long-term athletic development.

Physiological development during adolescence leads to significant changes in body composition, including increases in muscle mass and fat mass, particularly influenced by hormonal changes.[6] During this period, sex differences become increasingly pronounced, with males typically developing greater lean muscle mass and females experiencing a relatively higher accumulation of body fat. These developmental changes may have meaningful implications for the use and efficiency of the body's energy systems.[7] Body composition is an important performance indicator in sports science, as it reflects the balance of muscle and fat mass, which in turn influences athletic potential. For instance, increased fat mass can impede movement efficiency and aerobic performance, whereas greater muscle mass enhances strength, power, and endurance capacity. Assessing how these body composition parameters relate to energy system capacities may guide training and nutritional strategies aimed at improving performance outcomes in adolescent athletes.[8]

While existing research explores energy system utilization in basketball, limited studies have examined how these capacities relate specifically to body composition in adolescent athletes. Most prior work has focused on elite or adult populations. There remains a need to understand these associations in youth athletes to inform early-stage talent development and conditioning programs.

This study aims to investigate the relationships between aerobic and anaerobic capacities with body composition variables including BMI, body fat percentage, and muscle mass in adolescent basketball players. Additionally, differences between male and female athletes will be explored.

Materials and methods

Participants

Participants were recruited through local basketball academies and schools. The inclusion criteria required that participants be actively involved in basketball training and competitions for at least one year, be between 13 and 16 years of age, and have no current musculoskeletal or cardiovascular conditions. Individuals with chronic illnesses or recent injuries were excluded from the study.

Study design

A cross-sectional study design was used with a sample of 60 adolescent basketball players (33 males, 27 females) aged 13-16 years from Chiang Mai Province. Participants were required to have at least one year of basketball experience. The study protocol was approved by the Human Ethical Review Board of the primary investigator's institution (AMSEC-67EX-039), and informed consent was secured from participants and their parents.

Outcome measures

Body composition, including body mass index (BMI), body fat percentage, and muscle mass, was assessed using a validated bioelectrical impedance analyzer (INBODY), which provides segmental analysis of lean mass and fat distribution. Aerobic capacity was evaluated by estimating VO_2max through the 20-meter shuttle run test, a widely recognized field test for assessing cardiovascular endurance. In this test, participants ran back and forth along a 20-meter track, maintaining pace with audio signals until they failed to reach the line on two consecutive occasions.[9,10] Anaerobic capacity was assessed using the Running based Anaerobic Sprint Test (RAST), in which participants completed six 35-meter sprints with 10 seconds of rest between each sprint. The time for each sprint was recorded and used to calculate the fatigue index and peak power output.[11]

Study procedure

Testing was conducted across three days. On Day 1, participants completed body composition assessments. Day 2 involved the VO_2max test. On Day 3, after 24 hours of rest, the RAST was conducted. Participants were instructed to maintain normal hydration and avoid intense exercise 24 hours before testing. Each session began with a standardized 10-minute warm-up including light jogging and dynamic stretching. All participants

completed a familiarization session at least one week prior to testing. Data collection was supervised by trained researchers, and all measurements were taken twice and averaged to ensure reliability. All performance testing took place in a climate-controlled indoor facility to ensure consistency. Ambient temperature was maintained between 22-25 °C, and tests were scheduled at similar times of day to control for circadian rhythm influences (Figure 1).

VO_2max was estimated using the Leger equation derived from the final speed attained during the 20-meter shuttle run.[10] Speed was determined by the final shuttle level reached before voluntary exhaustion.

$$VO_2max(ml/kg/min)=31.025+3.238\times speed(km/hr)-3.248\times age+0.1536\times speed\times age$$

Anaerobic performance was calculated using data from each participant's six sprint times recorded during the RAST test. Power output for each sprint was determined using a formula that incorporates body mass, sprint distance, and time. These values were then used to derive two key indicators including anaerobic capacity and fatigue index. Anaerobic capacity was represented by the average power across all six sprints, while the fatigue index reflected the decline in performance by comparing the maximum and minimum power outputs.[11]

- $Power\ (W)=(Body\ mass\times Distance^2)\div Time^3$
- *Anaerobic Capacity=Average of the six calculated power values*
- *Fatigue Index (%)=[(Maximum power-Minimum power)÷Maximum power]×100*

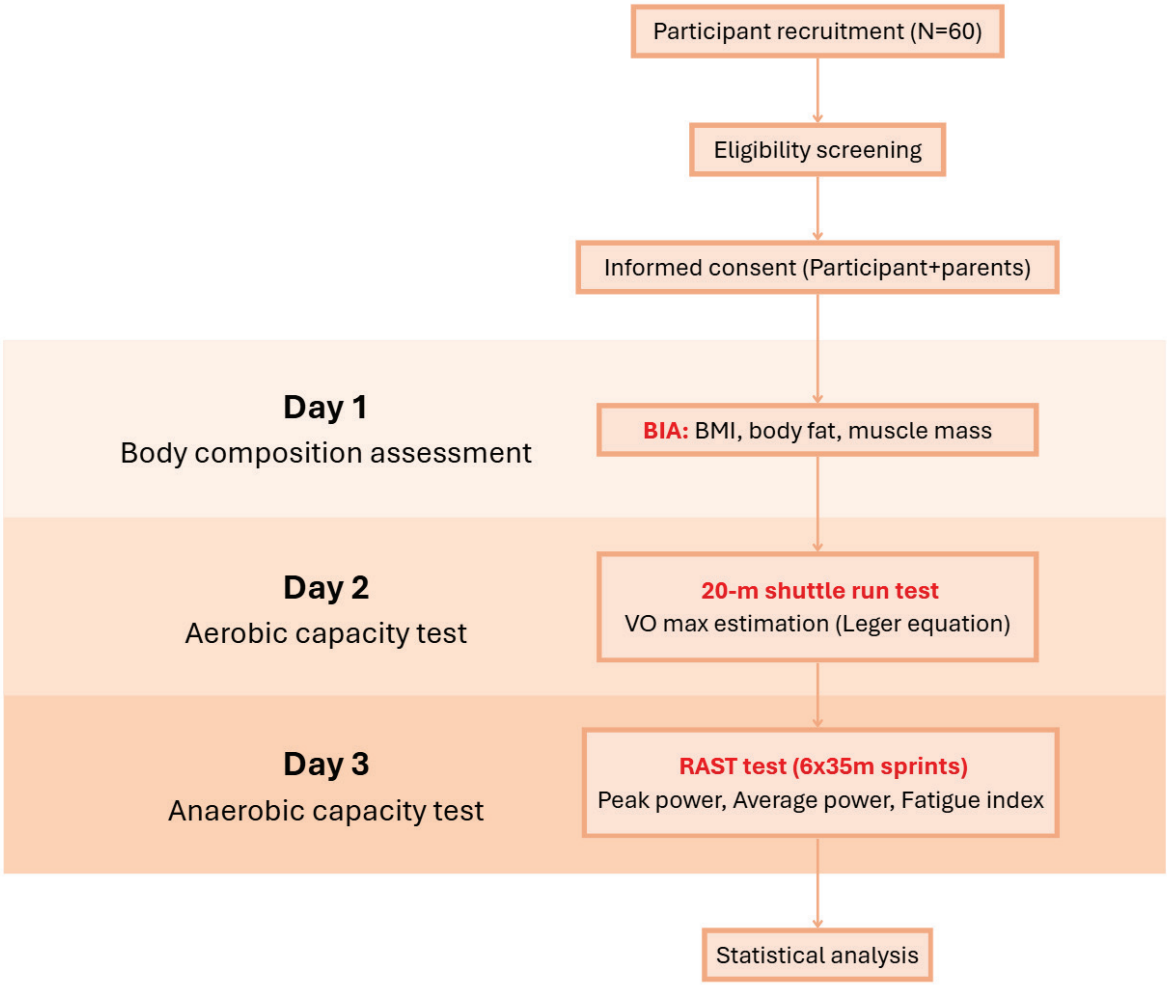


Figure 1. *Flow diagram illustrating participant recruitment, outcome measures, and study procedures.*

Data analysis

All data were analyzed using SPSS software (version 21.0, IBM Corporation, Chicago, IL). Descriptive statistics (mean±standard deviation) summarized demographic and performance data. Pearson's correlation coefficient assessed relationships between aerobic and anaerobic capacities and body composition variables. The strength of correlation was interpreted based on commonly accepted criteria as follows: values of |r|=0.00-0.19 were considered very weak, 0.20-0.39 weak, 0.40-0.59 moderate, 0.60-0.79 strong, and ≥0.80 very strong correlation.[12] Independent sample t-tests compared performance and body composition differences between gender. A significance level of $p<0.05$ was set for all analyses.

Results

The physical characteristics of participants are presented in Table 1. Male athletes demonstrated significantly higher height, weight, muscle mass, and VO_2max, and lower body fat percentage compared to females ($p<0.05$). No significant sex difference was observed in fatigue index.

Pearson's correlation analysis conducted in the overall sample (Table 2) demonstrated a moderate positive correlation between VO_2max and muscle mass (r=0.466, $p<0.001$). VO_2max also showed a moderate negative correlation with body fat percentage (r=-0.463, $p<0.001$) and a weak negative correlation with body mass index (r=-0.330, $p=0.01$). No significant correlations were observed between fatigue index and any body composition variables in the overall sample.

Table 1. Comparison of physical characteristics and energy capacities between genders.

Variables	Male (N=33)	Female (N=27)	*p* value
Height (cm)	171.79±9.28	162.30±7.26	<0.001
Weight (kg)	62.01±17.01	52.02±7.48	0.006
Body mass index (BMI)	20.83±4.02	19.68±2.38	0.196
Body fat (%)	15.35±8.01	23.21±5.96	<0.001
Muscle mass (%)	80.17±7.58	72.14±5.57	<0.001
Fatigue index (%)	38.80±12.69	40.53±12.93	0.603
VO_2max (mL/kg/min)	27.82±10.07	20.79±7.91	0.004

Note: *BMI: body mass index, cm: centimeter, kg: kilogram, min: minute, mL: milliliter.*

Table 2. Correlation between energy capacities and body composition variables.

Body composition variables	Fatigue index		VO_2max	
	r	*p* value	r	*p* value
Body mass index (BMI)	-0.068	0.608	-0.330	0.010
Body fat (%)	-0.012	0.925	-0.463	<0.001
Muscle mass (%)	0.011	0.932	0.466	<0.001

Note: *BMI: body mass index, cm: centimeter, kg: kilogram, min: minute, mL: milliliter.*

Discussion

The findings indicate a moderate correlation between aerobic capacity and body composition in adolescent basketball players. Higher VO_2max was associated with lower body fat percentage and greater muscle mass, suggesting enhanced cardiovascular and metabolic efficiency in athletes with more favorable body composition profiles. These results are consistent with previous studies that highlight the negative impact of excess body fat on aerobic performance and the supportive role of muscle mass in facilitating oxygen consumption during exercise.[2,13] The observed moderate correlations between VO_2max and both muscle mass and body fat percentage suggest that body composition plays a meaningful, although not exclusive, role in aerobic performance among adolescent basketball players. The strength of these associations indicates that while morphological characteristics contribute to aerobic fitness, other factors such as cardiovascular adaptations, training intensity, and genetic predisposition may also influence performance outcomes. The weak correlation observed between VO_2max and BMI further highlights the limitation of BMI as a general indicator, as it does not distinguish between fat and lean mass. No significant correlation was observed between anaerobic capacity and body composition. This outcome may reflect the complex and multifactorial nature of anaerobic performance,

which can be influenced by variables such as muscle fiber type, neuromuscular coordination, and sport-specific training adaptations rather than morphological characteristics alone.[6]

Sex-based differences were also evident, with male participants demonstrating higher lean muscle mass, lower body fat percentage, and superior VO_2max compared to females. These differences are likely attributable to hormonal and developmental factors, particularly the effects of puberty on muscle and fat distribution.[14] Such findings support the need for sex-specific training and nutritional interventions during adolescence to optimize physical development and performance. Moreover, the results reinforce the sensitivity of aerobic fitness to variations in body composition during adolescence, a period characterized by rapid growth and hormonal fluctuations.[15] In practical application, conditioning programs aimed at enhancing VO_2max in youth athletes should integrate strategies that simultaneously encourage muscle development and reduce body fat. This can be achieved through a balanced regimen of endurance training, resistance exercise, and targeted nutritional education tailored to the needs of adolescent athletes.

The lack of correlation between fatigue index and body composition suggests that anaerobic performance may be more dependent on neuromuscular and technical training than on morphological characteristics. Fatigue index, while a useful marker, may not capture all relevant dimensions of anaerobic performance in adolescent athletes.[16] Furthermore, the absence of a correlation between fatigue index and body composition indicates that anaerobic performance may hinge more on neuromuscular efficiency and technical skills than purely morphological characteristics. This finding emphasizes the importance of sport-specific drills and movement training to enhance anaerobic output, particularly in adolescents whose physiological systems are still maturing.[16,17]

These results also carry implications for talent identification and individualized training plans in youth basketball. Coaches may benefit from incorporating regular body composition assessments to track physical development and its relationship to energy system performance. Given that females typically show higher body fat percentages and lower muscle mass during adolescence, there may be value in designing sex-specific conditioning protocols that consider both physiological and developmental differences. In addition, the data suggests a possible threshold effect in aerobic performance where body composition must reach a certain balance before VO_2max significantly improves. Future research should aim to quantify these thresholds and investigate how long-term training interventions affect the interaction between morphology and metabolism.

Limitations

This study has some limitations that should be considered when interpreting the results. First, the relatively small sample size may limit the generalizability of the findings to broader populations of adolescent basketball players. A larger cohort would provide more statistical power and a better representation of performance variability. Second, the use of a cross-sectional design limits the ability to establish causal relationships between body composition and energy system capacities, as the data represents a single point in time rather than changes over a period. Longitudinal studies would be more appropriate for examining developmental trends and training effects. Third, the study did not control differences in training intensity, duration, or competitive experience among participants, which could have influenced both body composition and performance outcomes. Lastly, while field-based tests such as the 20-meter shuttle run and RAST are practical for large groups, they may not be as precise as laboratory-based measurements, potentially affecting the accuracy of VO_2max and anaerobic capacity estimations.

Conclusion

In the overall sample of adolescent basketball players, aerobic capacity was significantly associated with body composition, particularly muscle mass and body fat percentage. Male athletes demonstrated superior aerobic performance compared to females, likely due to more favorable body composition. No significant associations were found for anaerobic capacity. These findings support targeted training and body composition monitoring to optimize performance in youth basketball. Future research should include longitudinal data and control for training regimens to enhance interpretability.

Ethical approval

The study was performed in compliance with relevant laws and institutional guidelines and was approved by the Human Ethics Review Board of the Faculty of Associated Medical Sciences, Chiang Mai University (Approval No. AMSEC-67EX-039). Written informed consent was obtained from all participants and their parents or legal guardians prior to participation.

Funding

This research received no external funding.

Conflict of interest

The authors report no conflicts of interest in this work.

Data availability statement

The data presented in this study are available on request from the corresponding authors.

Acknowledgements

The authors would like to thank all the participants who participated in this study. Thanks also go to the Multidisciplinary and Interdisciplinary School, Chiang Mai University for supporting this research project.

CRediT authorship contribution statement

Busaba Chuatrakoon: conceptualization, methodology, investigation, data curation, writing original draft, review and editing, supervision; **Sothida Nantakool:** conceptualization, methodology, review and editing; **Panida Prasertpong:** data collection, data curation, writing original draft, review and editing.

References

[1] FIBA, CIES. International Basketball Migration Report 2019. Neuchâtel: International Centre for Sports Studies; 2019.

[2] Hoffman JR, Epstein S, Einbinder M, Weinstein Y. The influence of aerobic capacity on anaerobic performance and recovery indices in basketball players. J Strength Cond Res. 1999; 13(4): 407-11. doi: 10.1519/00124278-199911000-00018.

[3] Gottlieb R, Eliakim A, Shalom A, Dello-Iacono A, Meckel Y. Improving anaerobic fitness in young basketball players: plyometric vs. specific sprint training. J Athl Enhancement. 2014; 3(3). doi: 10.4172/2324-9080.1000148.

[4] Ben Abdelkrim N, El Fazaa S, El Ati J. Time-motion analysis and physiological data of elite under-19-year-old basketball players during competition. Br J Sports Med. 2007; 41(2): 69-75. doi: 10.1136/bjsm.2006.032318.

[5] Castagna C, Impellizzeri FM, Rampinini E, D'Ottavio S, Manzi V. The Yo-Yo intermittent recovery test in basketball players. J Sci Med Sport. 2008; 11(2): 202-8. doi: 10.1016/j.jsams.2007.02.013.

[6] Drinkwater EJ, Hopkins WG, McKenna MJ, Hunt PH, Pyne DB. Modelling age and secular differences in fitness between basketball players. J Sports Sci. 2007; 25(8): 869-78. doi: 10.1080/02640410600907870.

[7] Siervogel RM, Demerath EW, Schubert C, Remsberg KE, Chumlea WC, Sun S, et al. Puberty and body composition. Horm Res. 2003; 60(Suppl 1): 36-45. doi: 10.1159/000071224.

[8] Siervogel RM, Maynard LM, Wisemandle WA, Roche AF, Guo SS, Chumlea WC, et al. Annual changes in total body fat and fat-free mass in children from 8 to 18 years in relation to changes in body mass index. The Fels Longitudinal Study. Ann N Y Acad Sci. 2000; 904: 420-3. doi: 10.1111/j.1749-6632.2000.tb06494.x.

[9] Robinson R. 20m multistage fitness test (beep test) instructions [Internet]. Topend Sports; [cited 2025 Jan 20]. Available from: https://www.topendsports.com/testing/tests/20mshuttle.htm

[10] Léger LA, Lambert J. A maximal multistage 20-m shuttle run test to predict VO2 max. Eur J Appl Physiol Occup Physiol. 1982; 49(1): 1-12. doi: 10.1080/02640418808729800.

[11] Gobatto CA, Torres RS, Moura FA, Cunha SA, Giometti CB, Araujo GG, et al. Corresponding assessment scenarios in laboratory and on-court tests: centrality measurements by complex networks analysis in young basketball players. Sci Rep. 2020; 10(1): 8620. doi: 10.1038/s41598-020-65420-3.

[12] Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd Ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.

[13] Glaister M. Multiple sprint work: physiological responses, mechanisms of fatigue and the influence of aerobic fitness. Sports Med. 2005; 35(9): 757-77. doi: 10.2165/00007256-200535090-00003.

[14] Meckel Y, Machnai O, Eliakim A. Relationship among repeated sprint tests, aerobic fitness, and anaerobic fitness in elite adolescent soccer players. J Strength Cond Res. 2009; 23(1): 163-9. doi: 10.1519/JSC.0b013e31818b9651.

[15] França C, Martinho DV, Gouveia ÉR, Martins F, Marques A, Ribeiro T, et al. Changes in estimated body composition and physical fitness of adolescent boys after one year of soccer training. Children. 2023; 10(2): 391. doi:10.3390/children10020391.

[16] Tsoukos A, Bogdanis GC. Physiological responses and fatigue during a repeated shuttle-sprint running test in adolescent schoolchildren: a comparison between sexes and fatigue calculation methods. Children. 2023; 10(6): 1041. doi:10.3390/children10061041.

[17] Zając B, Mika A, Gaj PK, Ambroży T. Effects of anaerobic fatigue induced by sport-specific exercise on postural control in highly trained adolescent road cyclists. Appl Sci. 2023; 13(3): 1697. doi:10.3390/app13031697.

Journal of Associated Medical Sciences 2026; 59(2): 286-297

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Impact of ergonomic workstation interventions and scapular bracing on neck muscle performance, postural alignment, neck disability and work productivity in bankers with non-specific neck pain

Muhammad Awais Tassawar, Soyba Nazir, Montakarn Chaikumarn*

Department of Physical Therapy, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand

ARTICLE INFO

Article history:
Received 23 July 2025
Accepted as revised 20 February 2026
Available online 11 March 2026

Keywords:
Braces, neck pain, ergonomics, posture, perceived stress scale

ABSTRACT

Background: Musculoskeletal disorders among bankers are linked to personal traits, prolonged working posture, repetitive tasks, and workstation design. Scapular bracing has been endorsed as an intervention for improving upper-body posture; however, evidence in symptomatic office workers is inadequate.

Objective: The purpose of this study was to compare the effects of an ergonomic workstation intervention alone versus a combined scapular brace plus ergonomic workstation intervention on forward head posture, rounded shoulder posture, neck disability, neck muscle performance, work productivity, and perceived stress in bankers with nonspecific neck pain.

Materials and Methods: Fifty bankers with non-specific neck pain were randomly assigned to one of two groups: ergonomic workstation intervention (control, N=25) or scapular brace plus ergonomic workstation intervention (experimental, N=25). Craniovertebral angle (CVA), forward shoulder angle (FSA), Neck Disability Index (NDI), neck muscle performance, work productivity, and Perceived Stress Scale (PSS) were all measured at baseline and four weeks later. The data were analysed with mixed-model repeated-measures ANOVA.

Results: Significant time × group interaction effects favoured the experimental group for CVA (p=0.005), FSA (p=0.001), NDI (p=0.017), neck muscle performance (p=0.004), and work productivity (p=0.003), indicating greater postural, functional, and productivity-related improvements as opposed to the control group. There were no significant between-group differences in perceived stress.

Conclusion: The addition of a scapular brace to an ergonomic workstation intervention may provide modest short-term benefits for posture, neck disability, neck muscle function, and job productivity in bankers suffering from non-specific neck pain, although perceived stress remains unaffected.

Introduction

Computer use has grown standard in modern settings, and extended visual display unit (VDU) use is strongly linked to neck and upper quadrant musculoskeletal issues.[1] Office jobs, such as banking, require extended sitting, repetitive upper-limb activities, and static postures, which increases the risk of neck and shoulder problems. Musculoskeletal issues among bank employees have continuously been connected to individual physical characteristics, working posture,

* **Corresponding contributor.**
Author's Address: *Department of Physical Therapy, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand.*
E-mail address: Montakarn.c@chula.ac.th
doi: 10.66285/JAMS.2026.060
E-ISSN: 2539-6056

repetitive actions, and workstation design.[2] Previous surveys have shown a significant prevalence of complaints in the neck, back, shoulders, and upper extremities among bank employees, emphasising the occupational burden of computer-related musculoskeletal conditions.[3]

Ergonomic workstation approaches are typically regarded as the key technique for controlling work-related musculoskeletal conditions in office settings. Adjustments to monitor height, chair support, keyboard and mouse position, and workstation layout are common interventions for reducing biomechanical strain on the cervical spine and upper extremities. Several studies have shown that ergonomic adjustments can minimise neck and shoulder pain, improve working posture, and, in certain situations, boost job efficiency among computer users.[4-6] However, evidence suggests that while ergonomics can improve the exterior work environment,[7] its impacts on sustained postural control and neuromuscular function may be limited, especially during prolonged computer work when postural drift and muscle fatigue are present.

Neck pain is frequently accompanied by neuromuscular dysfunction, which includes decreased endurance and poor activation of the deep cervical flexor muscles. Prolonged forward head posture, which is common among VDU operators, increases cervical spine loading and contributes to deep neck flexor weakness. Concurrently, changed scapular alignment and rounded shoulder posture are linked to alterations in axioscapular muscle activity, which may increase cervical loading and pain. The scapula and cervical spine are functionally and structurally interdependent, and abnormalities in scapular control have been reported in individuals with neck pain.[8-11]

While ergonomic interventions address workstation-related risk factors, they may not effectively alleviate postural and neuromuscular deficits during task performance. As a result, supplementary techniques that provide constant posture cueing and support throughout work activities may be necessary. Scapular bracing has been proposed as one such method, with the goal of influencing scapular alignment, improving proprioceptive feedback, and facilitating more optimum muscle activation. Scapular braces, as compared to taping, may be more suitable for long-term or daily use because they prevent skin irritation and can be worn during typical occupational work tasks.[12]

Previous research has shown that scapular bracing can impact scapular posture at rest and during low ranges of motion, potentially improving scapular muscle control.[13] Short-term studies in healthy people have showed moderate improvements in rounded shoulder posture and changes in muscle activation during typing tasks when using a scapular brace.[14] However, previous research has mostly focused on asymptomatic populations, athletes, and short-duration laboratory tasks. There is also a lack of research evaluating scapular bracing as a supplement to ergonomic workplace interventions in symptomatic office workers with extended computer exposure.

In addition to pain and physical dysfunction, job-related neck disorders are linked to decreased work productivity and sickness presenteeism, which occurs when individuals remain at work despite health issues that hinder performance.[15] Musculoskeletal pain can impair concentration, endurance, and task efficiency, making productivity a clinical and economic concern.[16] Therefore, interventions that lessen musculoskeletal strain and enhance postural stability while using a computer may help worker's productivity and perceived stress in ways that go beyond just reducing symptoms.

Accordingly, the current study aimed to compare the effectiveness of ergonomic workstation changes alone with ergonomic workstation adjustments combined with scapular bracing in symptomatic bankers with extended computer use. The study examined both between-group and within-group changes over time in work productivity, neck muscle performance, head and shoulder posture, neck disability, and perceived stress. It was hypothesised that the combined intervention would lead to higher improvements in these outcomes than ergonomic adjustments alone.

Materials and methods

Participants

Inclusion criteria: The study comprised 50 volunteers aged 20 to 40 years (Males=29, Females=21) from various banks working full time in Sahiwal division, Pakistan, who had neck pain for more than three months and complained of it at least once a week for the past twelve weeks.[17] The participants had no comorbidities and any regular medication use. In this study, neck pain was defined as non-specific neck pain that usually worsens as a result of poor posture at work and can be alleviated with postural adjustment.[17] Participants' Visual Analogue Scale (VAS) scores had to be mild (1-4) to moderate (4-6).[17] A craniovertebral angle of less than 48°,[18] forward shoulder angle greater than 22°,[19] the inability to generate a spike of at least 6 mmHg from the base pressure on the Biofeedback Pressure Unit and sustain the pressure generated for 10 seconds during cervical flexion were additional eligibility requirements.[20]

Exclusion criteria: Those who had their workstations ergonomically modified, had a Neck Disability Index (NDI) score greater than 28, were enrolled in a neck treatment program for neck pain within a year of the study's start, had undergone head surgery or trauma, or had previously experienced cancer, a systemic illness process, skin sensitivity, or infection were all excluded.[6] Participants with pregnancies, certain illnesses, or inflammatory conditions were also excluded.[6]

G*Power was used to calculate sample size. Independent t-test with two tails was considered. Minimal detectable change (MDC) for craniovertebral

angle (1.1626) with the use of brace was employed to determine effect size from pilot study by Noordin NB *et al.*[21] SD set at 1.3 for both groups. Cohen's d was 0.89, test power was set at 0.80 and α=0.05. Total 42 participants were needed for this study and after including 20% dropout ratio 50 were enrolled.

Research personnel: One main researcher, a certified physiotherapist with a Doctor of Physical Therapy degree and is currently pursuing a PhD, with over three years of clinical experience. One assessor, a certified physiotherapist with a Doctor of Physical Therapy and PhD degree and more than three years of clinical experience, was blinded to group allocation.

A multistage convenience sampling strategy was used to recruit participants. Following random visits to commercial banks in the study area, bank staff members were briefed in person about the objectives and methods of the study. Enrolment was based on voluntary consent, and eligible bankers who indicated interest were then enrolled. After baseline characteristics (age, BMI score, daily computer use, gender) were recorded by assessor, participants were randomly assigned to one of two groups using the hat method: control group (N=25): ergonomic workstation intervention group, or experimental group (N=25): ergonomic workstation intervention with scapular brace group (Figure 1). Because of the nature of the interventions, the main researcher in charge of carrying out the workstation changes and scapular bracing was not blinded to group allocation. However, the outcome assessor was blinded to group assignment and did not participate in intervention delivery, lowering the possibility of assessment bias. Table 1 shows demographics of participants.

To protect participants' privacy, each was allocated a unique number that was used on all documents, data collection forms, and assessment forms. Following recruitment from all banks, participants were pooled into a single sample and randomly assigned to either the experimental or control groups using a simple random draw process, regardless of bank affiliation. This method was chosen to give the same probability of allocation for all participants. All eligible volunteers from each bank were assigned unique identification numbers, which were then placed in sealed, opaque envelopes. An impartial individual who was not engaged in recruitment or intervention delivery thoroughly mixed these envelopes and randomly assigned them to each group in alternating order. The participants were then informed, either in person or over the phone, of their random selection and group assignment. The main researcher enrolled the study participants. Subjects blinding was also impossible because some participants with braces and without braces were present in the same bank. Randomisation was done at the individual participant level, with no stratification by bank. No blocking method was used. A basic randomisation process was used to allocate groups, resulting in nearly balanced group sizes by chance.

Ethical approval letter was obtained from The Research Ethics Review Committee for Research Involving Human Research Participants, Chulalongkorn University (COA. 108/67) Dated: 15 May 2024. Written consent forms were signed by the participants before start of study explaining them the procedure and the potential use of their data in publication keeping their identities confidential.

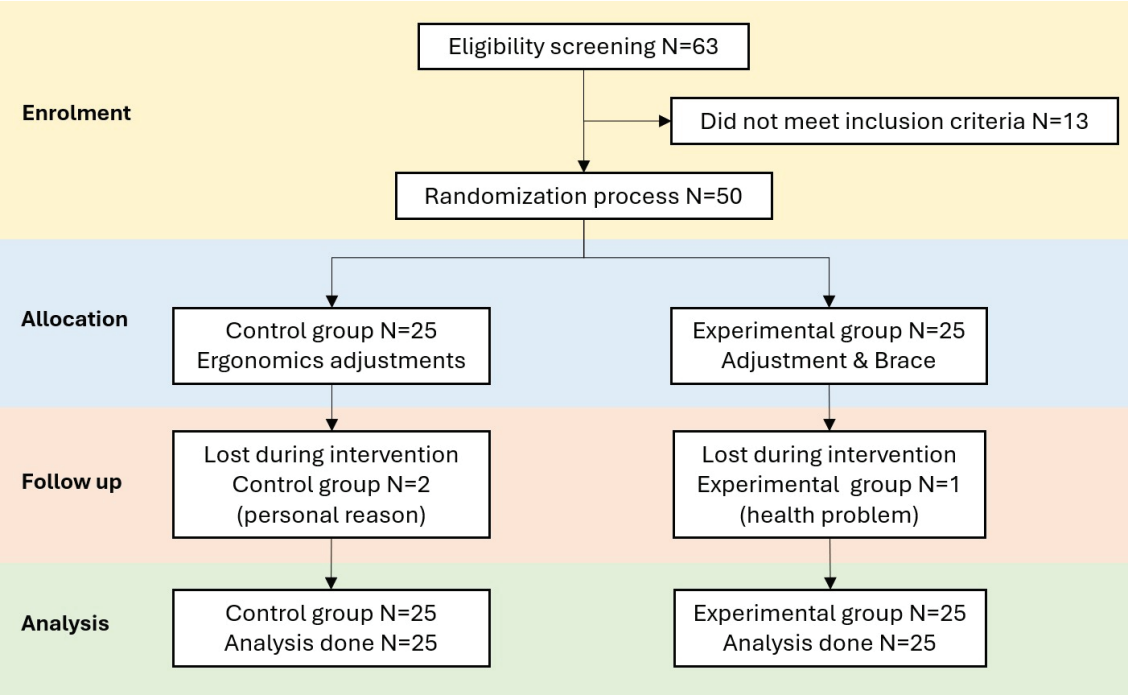


Figure 1. *CONSORT Diagram*

Interventions

Before the study started, ergonomic workstation adjustments were given to all participants based on the Canadian Standards Association's Office Ergonomics Guidelines (CSA-Z412-00) that Smith *et al.*[6] had previously employed. The adjustments were made by the main researcher and were tailored to each participant's anthropometric features and workstation configuration, while remaining consistent with the standardised guideline. Table 1 lists the modifications that participants required.

Experimental group

The 8-figure scapular brace ((Energizing Posture Support, Bell Plus company, Model Dr. Magico, Number 11180022)) was applied by the main researcher in accordance with the manufacturer's instructions see Figure 2.

Table 1. Workstation adjustments needed.

Parameters	Control N=25	Experimental N=25	Adjustments given
Chair			
• Regassing	5	7	Chair seat re-gassed
• Lumbar support	7	7	Lumbar cushion
• Armrest height correction	6	8	Armrest height corrected
• 5-8 degrees tilt	22	21	Tilted 5-8 degrees
• Footrest	3	2	Footrest provided
Monitor			
• Monitor raise	17	11	Monitor raised
• Monitor lowering	7	10	Monitor lowered
Workstation accessories			
• Flat keyboard	2	2	Flat keyboard provided
• Wrist-rest	3	5	Wrist-rest provided
• Documents holder	13	7	Document holder provided
• Headphones	2	5	Headphones provided



Figure 2. *The 8-figure scapular brace used in this study.*

The adjustable straps allow for common application of the brace by the recruited individuals in the experimental group. The brace was strapped over regular clothes on the upper back. Each subject received training on how to put on and remove the brace. It was advised that subjects wear the brace for six hours per day, with a minimum of 30 minutes in continuation during office hours, for 4 weeks, as previously implemented by Bankhele *et al*.[22] The 4-week intervention time was chosen in order to assess the interventions' immediate effects in real-world work environments. Although prior ergonomic studies have employed longer intervention periods, short-term results are still clinically relevant, especially for early symptom alleviation and feasibility assessment. Rest breaks were optional, although participants were instructed to remove the brace while standing or walking and avoid wearing on holidays. This instruction was given to reduce discomfort, avoid dependency, and limit intervention to task-specific postural support during extended desk work. However, the objective adherence metrics like wear-time sensors and daily logs were not employed. A single adjustable model of brace with adjustable straps and suitable padding was used see Figure 3.



Figure 3. *Participant wearing the scapular brace.*

The main researcher visited the participants in banks twice a week to ensure compliance and note down any complaints or issues occurred. Also, the participants were reminded of the scapular brace everyday via a phone call to the bank.

Assessment

An assessor blinded to group allocation conducted evaluations at baseline and four weeks post-intervention, when the study was completed and participants were not wearing scapular braces. Participants' VAS scores were used to measure pain, and these scores were only taken at baseline because they were a requirement for inclusion. The 10cm VAS was used by participants to report their level of pain.

Primary outcome measure (postural)

Participants' craniovertebral angles (CVA) were measured for the forward head assessment. CVA was the primary outcome measure. CVA is the point where a line from the skin above the C7 spinous process to the midline of the tragus of the ear intersects a horizontal line with the C7 spinous process (Figure 4).[23]

Secondary outcome measure (postural), forward shoulder angle FSA was measured for rounded shoulder assessment. FSA is a point where a line connecting acromion and C7 spinous processes meet a vertical line crossing C7 process see Figure 5.

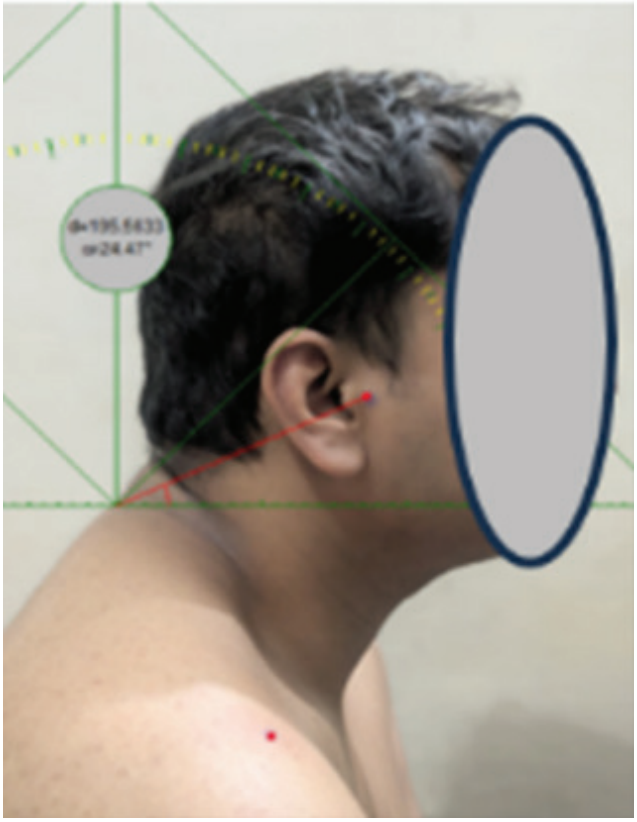


Figure 4. *CVA measured with Marcus Bader ruler program.*
(The red line connects the tragus of ear and C7 spinous process.)

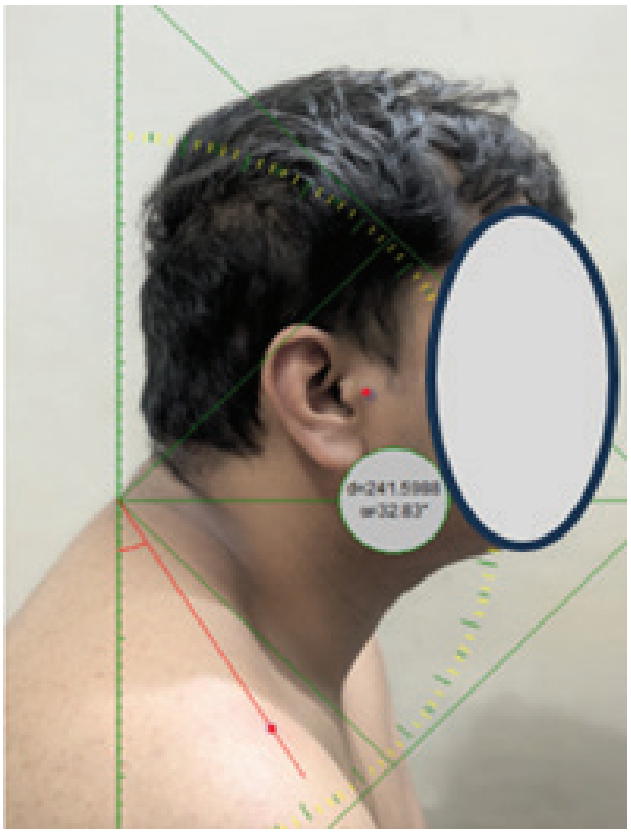


Figure 5. *FSA measured with Marcus Bader ruler program.*
(The red line connects the C7 spinous process and tip of acromion process.)

CVA and FSA were measured by photography technique in sitting position. Pictures were taken in lateral view in an empty room usually the common or break room in most of the banks, especially for the privacy of female participants. Permission from the bank manager was always taken beforehand. Prior to taking photograph in a neutral position, the participant was instructed to fully extend and flex their neck three times and then rest in the most balanced posture that feel comfortable. Pixel 6a camera was mounted on a tripod. The lens on the camera was adjusted to the level with the external auditory meatus by modifying the camera tripod's height. The camera was 1.5m away from the participant. Participant was sitting in a neutral position looking straight with back supported, hands on the thighs and feet flat. They were asked to concentrate visually on a point just ahead on the room's wall. The assessor confirmed the visual focal point when the individual had reached a comfortable seating position. This reduced the inclination of the neck to flex or extend while keeping a comfortable head position. Finally, a lateral view of each subject was photographed. One photograph per posture was obtained.[24]

Marcus Bader's MB Ruler program was used to determine the angles from the photos. After placing the ruler over the photos, the measurements were noted.[25]

Secondary outcome measure (functional)

The performance of the neck muscles was evaluated using Biofeedback Pressure Unit (BPU) readings. Participants' deep neck flexor function was evaluated using a Pressure Biofeedback Unit (BPU) while supine lying. The pressure cuff was put sub occipitally and inflated to 20 mmHg. Participants were advised to do a modest cranio-cervical flexion. DNF strength (pressure performance) was quantified as the highest-pressure level (mmHg) that the individual could achieve and sustain for 10 seconds without substitution. DNF endurance was quantified as the greatest duration (in seconds) during which the desired pressure level could be maintained without tiredness or loss of pressure control. No repetitions were averaged; only the best attempt was retained for analysis.[20]
Secondary outcome measures (self-reported):

Secondary outcome measure (self-report)

NDI was used to measure neck disability. The NDI is a self-assessment tool for the specific functional status of patients with neck pain, consisting of ten elements: pain, personal care, lifting, reading, headache, concentration, work, driving, sleeping, and leisure. Each part is scored from 0 to 5, with 0 being "painless" and 5 being "the worst pain imaginable".[26]

Work Productivity and Activity Impairment-Specific Health Problem (WPAI-SHP) questionnaire was used to measure work productivity in terms of overall work impairment (absenteeism+presenteeism). It is a validated questionnaire to evaluate work productivity loss, absenteeism, presenteeism, and activity impairment due to specific health problem over the previous 7 days.[27] The specific health problem was neck pain in this study. The WPAI:SHP results are reported as percentages of impairment, with higher numbers indicating greater impairment and lower productivity.[28]

Perceived Stress Scale questionnaire comprising of 10 questions was used to assess the stress factor in the participants. It helps participants to compare their global perceptions of stress with the epidemiology of disease.[29] No stress-specific intervention was used, as perceived stress was added as a secondary exploratory outcome to see if postural correction and reduction in musculoskeletal strain related with intervention would be accompanied by changes in total perceived stress.

Pain intensity (VAS) was measured at baseline to characterise symptom severity and confirm eligibility; however, it was not evaluated after intervention because the trial was primarily meant to investigate postural and functional changes rather than pain reduction.

Prior to data collection, the assessor was given specific training in standardised posture assessment, anatomical landmark identification, and pressure biofeedback unit testing, all in accordance with previously published protocols. To guarantee consistency, all measures were taken with the same equipment, standardised positioning, and established protocols throughout the investigation. Formal intra- or inter-rater reliability testing was not performed in this study. Previous studies have shown that the tools and processes employed have good to excellent intra- and inter-rater reliability.[8, 30-32]

All raw data, photos, and digital measurements were securely saved on password protected institutional software that could only be accessed by the assessor. Before the analysis, participant identifiers were anonymised. The data were stored in conformity with institutional research data management rules.

Statistical analysis

Data analysis was conducted using IBM SPSS Statistics version 29.0 (IBM Corporation, Armonk, NY, USA). The participant's demographic characteristics were analysed using the chi-square test for gender and job title and the independent t-test for age, BMI, and daily PC usage hours. The Shapiro-Wilk test was employed to evaluate the numerical data's normal distribution. The overall effects of time (baseline and at 4 weeks), groups (control and experimental), and time×group interaction were examined using mixed ANOVA. The Bonferroni post hoc test was used for paired comparisons. For every test, the significance threshold was set at 0.05. Baseline observation carried forward (BCOF) was used as intention-to-treat technique for dropouts.

Results

Fifty bankers were equally divided into two groups. There were no significant differences between the groups in terms of age, gender, BMI, daily computer use, or job title (p=0.194, p=0.774, p=0.753, p=0.881 and p=0.344 respectively). Table 2 shows demographics of participants.

Table 3 illustrates descriptive statistics and pairwise comparisons for every outcome measure. At baseline or at four weeks, no outcome measure showed a significant between group difference.

Table 3. The values of NDI, CVA, FSA, WPAI-SHP, BPU pressure, BPU seconds held and PSS at baseline and at 4th week.

Outcome measures	Control group (N=25)		Experimental group (N=25)	
	At baseline Mean±SD	At 4th week Mean±SD	At baseline Mean±SD	At 4th week Mean±SD
CVA (degrees)	35.68±2.19	35.74±2.16	35.22±2.21	35.58±2.20†
Neck disability index	16.32±2.62	16.20±2.90	18±2.51	16.96±2.11†
Forward shoulder angle	26.20±1.71	26.13±1.70	26.58±1.64	25.86±1.50†
BPU pressure (mmHg)	24.84±0.94	24.88±0.88	24.68±0.94	25.04±0.93†
BPU (seconds)	6.88±0.88	6.76±1.33	6.96±1.05	7.2±1.11
WPAI-SHP	36.96±6.96	36.78±6.77	40.86±7.79	36.72±5.55†
PSS	17.92±3.34	17.68±3.14	19.28±3.11	19.04±3.00

Note: *†significant difference from baseline, CVA: craniovertebral angle, FSA: forward shoulder angle, BPU: biofeedback pressure unit, WPAI-SHP: work productivity and activity impairment-specific health problem, PSS: perceived stress score.*

Table 4. Time, group and interaction main effects of all outcome measures.

Outcome measures	Main effect (group)			Main effect (Time)			Main effect (Time*group)		
	F(df)	ηp²	*p* value	F(df)	ηp²	*p* value	F(df)	ηp²	*p* value
CVA (degrees)	0.254(1)	0.005	0.617	17.34(1)	0.265	**0.001**	8.45(1)	0.150	**0.005**
NDI	3.05(1)	0.060	0.087	9.65(1)	0.167	**0.003**	6.07(1)	0.112	**0.017**
FSA (degrees)	0.013(1)	0.000	0.911	48.5(1)	0.503	**0.001**	33.27(1)	0.409	**0.001**
BPU pressure (mmHg)	0.000(1)	0.000	1.0	14.28(1)	0.229	**0.001**	9.14(1)	0.160	**0.004**
BPU (seconds)	1.385(1)	0.028	0.245	0.072(1)	0.002	0.789	0.65(1)	0.013	0.423
WPAI-SHP (overall impairment)	1.112(1)	0.023	0.297	11.60(1)	0.195	**0.001**	9.79(1)	0.169	**0.003**
PSS	2.40(1)	0.048	0.127	2.06(1)	0.041	0.158	0.000(1)	0.000	1.0

Note: *df: degrees of freedom, ηp^2: partial eta squared, CVA: craniovertebral angle, NDI: neck disability index, FSA: forward shoulder angle, BPU: biofeedback pressure unit, WPAI-SHP: work productivity and activity impairment-specific health problem, PSS: perceived stress score, bold values are significant.*

Primary outcome measure (CVA)

The craniovertebral angle (CVA) showed a significant group×time interaction (F=8.45, p=0.005, partial η^2=0.150). From 35.22°±2.21 to 35.58°±2.20, the experimental group showed an increase, with a mean change of 0.36° (95% CI: 0.18° to 0.54°). The control group showed no apparent change (Table 3 and Table 4).

Secondary outcome measures

A mixed ANOVA indicated a significant group×time interaction for neck disability (F=6.07, p=0.017, partial η^2=0.112), showing differences in change over time among groups. The experimental group's NDI scores decreased from 18.00±2.51 at baseline to 16.96±2.11 post-intervention, representing a mean drop of -1.04 points (95% CI: -1.58 to -0.50). The control group showed no significant difference (16.32±2.62 to 16.20±2.90)(Table 3 and Table 4).

A significant interaction effect was observed for the forward shoulder angle (FSA) (F=33.27, p<0.001, partial η^2=0.409). While the control group showed little change, the experimental group showed a decrease from 26.58°±1.64 to 25.86°±1.50, or a mean change of -0.72° (95% CI:-0.95° to -0.49°) (Table 3 and Table 4).

Biofeedback pressure unit (BPU) pressure measurements of deep neck flexor activation revealed a significant group×time interaction (F=9.14, p=0.004, partial η^2=0.160). The experimental group's mean

improvement was 0.36 mmHg (95% CI: 0.19 to 0.53), rising from 24.68±0.94 mmHg to 25.04±0.93 mmHg. The control group showed no discernible change (Table 3 and Table 4).

There was no significant group×time interaction for endurance (F=0.65, p=0.423). Over time, both groups showed modest and insignificant improvements. (Table 3 and Table 4).

A significant group×time interaction was observed in work productivity as measured by the WPAI-SHP (F=9.79, p=0.003, partial η^2=0.169). Overall work impairment decreased in the experimental group from 40.86±7.79 to 36.72±5.55, with a mean change of -4.14 points (95% CI: -6.10 to -2.18). The control group showed no apparent change (Table 3 and Table 4).

Experimental group had higher baseline stress scores. There was no significant group×time interaction in perceived stress, and both groups showed minimal change during the intervention period (Table 3 and Table 4).

Discussion

This study examined the short-term effects of ergonomic workstation modifications alone and in combination with scapular bracing on neck disability, posture, deep neck flexor performance, perceived stress, and job productivity in bankers who had non-specific neck pain. The study found that adding a scapular brace to ergonomic adjustments led to significant group×time interactions for several outcomes. However, the magnitude of change was modest, and findings should be interpreted as short-term within-group improvements rather than clear clinical superiority over ergonomics alone.

The neck disability index (NDI) showed a significant group × time interaction (F=6.07, p=0.017, ηp^2=0.112). Over four weeks, the experimental group's NDI scores dropped from 18.0±2.51 to 16.96±2.11, a reduction of about 1.0 point on average. Although statistically significant, this shift falls below the generally quoted minimal clinically meaningful difference levels for the NDI, implying a minor but noticeable functional improvement rather than a clinically significant change. There was no significant change in the control group.

Postural variables exhibited comparable trends. There was a significant group × time interaction for craniovertebral angle (CVA) (F=8.45, p=0.005, ηp^2=0.150). The experimental group exhibited an increase in CVA from 35.22°±2.21 to 35.58°±2.20, indicating a slight reduction in forward head posture. The experimental group improved their forward shoulder angle (FSA) from 26.58°±1.64 to 25.86°±1.50, with a significant interaction effect (F=33.27, p=0.001, ηp^2=0.409). While these variations in posture were statistically significant, their therapeutic implications should be evaluated cautiously, as very minor angular changes may not result in significant biomechanical or functional gains in the short term.

The observed postural improvements are consistent with biomechanical literature that links prolonged slumped sitting postures to higher cervical extensor loading, forward head posture, and rounded shoulders, all of which lead to neck pain and muscle fatigue.[33,34] Although scapular kinematics were not directly examined, the postural changes reported in this study may be attributed to increased postural awareness or external cueing provided by the brace.[12,22] This perspective is supported by previous work, but it remains theoretical because neuromuscular processes were not assessed in the current investigation.

Deep neck flexor performance exhibited varied changes. Biofeedback pressure unit (BPU) data revealed a significant group × time interaction for pressure (F=9.14, p=0.004, ηp^2=0.160). The experimental group showed a rise from 24.68±0.94 mmHg to 25.04±0.93 mmHg, indicating better activation or neuromuscular control. In contrast, endurance, defined as the duration of sustained contraction in seconds, did not show a significant interaction effect (F=0.65, p=0.423). The experimental group showed a tiny numerical rise from 6.96±1.05 seconds to 7.20±1.11 seconds, which was neither statistically nor clinically significant.

This distinction between power and endurance is consistent with the cervical motor control literature. Activation improvements can be achieved by improved alignment and lower inhibitory loads, but endurance adaptations often necessitate extended, low-load, task-specific training.[8,35] The lack of an organised deep neck flexor training program, along with the short intervention duration, most likely hampered endurance-related improvements. Furthermore, external postural support may have lowered muscle demand, so dampening the stimulus required for endurance development.

The WPAI-SHP revealed a significant group×time interaction on work productivity (F=9.79, p=0.003, ηp^2=0.169). The experimental group improved from 40.86±7.79 to 36.72±5.55, showing a modest reduction in perceived overall impairment. While statistically significant, productivity was assessed solely through self-reported measures, which coul d be influenced by response bias and subjective perception. Furthermore, no objective productivity measurements were provided, therefore these findings should be viewed as preliminary. To our knowledge, this is one of the first studies to look at perceived work productivity in bankers using a scapular brace; nonetheless, stronger results require objective metrics and a longer follow-up period.

There was no significant difference in perceived stress scores between the two groups, and no group×time interaction was observed. This shows that participants' work stress was more likely caused by psychological and organisational variables than by physical symptoms alone. Because no intervention addressing workplace stressors was included, the lack of change is consistent with previous research associating stress primarily to job demand, control, and organisational support.[36]

Within four weeks, the ergonomic workstation intervention did not result in significant improvements across all outcomes. This contrasts with prior research indicating effects following longer intervention periods, often eight weeks or more.[4] The absence of improvement in the control group demonstrates that short-term ergonomic changes may not be adequate to produce meaningful postural, functional, or neuromuscular adaptations. One objective of the present study was to discover whether ergonomic changes alone create short-term effects; the data imply that longer exposure may be required.

Limitation

While interpreting these results, a number of limitations should be taken into account. Despite the use of random allocation, no significant between-group differences were found, limiting inferences on the superiority of scapular bracing versus ergonomic workstation modifications. Simple randomisation using sealed envelopes was used, which may result in less allocation concealment than computer-generated approaches. The four-week intervention period may not have been long enough to capture the full effect of ergonomic changes documented in longer-term research. Participant blinding was not possible; however, outcome assessments were performed after the brace was removed to maintain assessor blindness. Regardless, performance and detection bias because of main researcher cannot be eliminated. Adherence to intervention was promoted but not objectively monitored, creating ambiguity about intervention fidelity. Several outcomes were based on self-reported questionnaires, which may be prone to recall and response bias, and pain intensity was only measured at baseline, preventing study of pain changes over time. Postural measurements were taken from single standardised pictures without any systematic intra- or inter-rater reliability testing, which could have introduced measurement error despite assessor training. And the lack of long-term follow-up and recruiting in a specific geographic region may restrict the observed effects' generalisability and durability.

Conclusion

This study examined the short-term impact of ergonomic workstation changes with and without scapular bracing in bankers complaining of non-specific neck pain. Over a four-week period, the use of a scapular brace resulted in modest short-term within-group improvements in postural alignment, neck disability, deep neck flexor activation, and work productivity. However, these benefits did not clearly outperform ergonomic intervention alone, and the magnitude of change was modest.

As a result, scapular bracing may be used in conjunction with ergonomic workstation adjustments to provide short-term postural support, rather than as a stand-alone or superior intervention. Future research should look at longer intervention periods, tailored cervical exercise programs, and follow-up examinations to further understand the role of scapular bracing in occupational neck pain management.

Funding

This study did not receive any fundings from public or commercial organizations.

Conflict of interest

No conflict of interest is present among authors.

CRediT authorship contribution statement

Montakarn Chaikumarn contributed to the title, methodology formulation, outcome measure selection and critical review of the research; **Muhammad Awais Tassawar** contributed to conception, design, the interventions implementation, interpretation and manuscript formation; **Soyba Nazir** contributed to data acquisition and statistical analysis.

Acknowledgements

We authors are thankful to the banks and participants for their cooperation and participation in the study.

References

[1] Tella B, Akodu A, Fasuba O. The prevalence of neck and upper extremity repetitive stress injury among bank workers in Surulere, Lagos, Nigeria. J Rheumatol. 2011; 6(2): 1-6. Available from: https://ispub.com/IJRH/6/2/13522.

[2] Akrouf QAS, Crawford JO, AlShatti AS, Kamel MI. Musculoskeletal disorders among bank office workers in Kuwait. East Mediterr Health J. 2010; 16(1): 94100. doi: 10.26719/2010.16.1.94.

[3] Yu IT, Wong TW. Musculoskeletal problems among VDU workers in a Hong Kong bank. Occup Med (Lond). 1996; 46(4): 27580. doi:10.1093/occmed/46.4.275.

[4] Mekhora K, Liston C, Nanthavanij S, Cole JH. The effect of ergonomic intervention on discomfort in computer users with tension neck syndrome. Int J Ind Ergon. 2000; 26(3): 36779. doi: 10.1016/S0169-8141(00)00012-3

[5] Robertson M, Amick III BC, DeRango K, Rooney T, Bazzani L, Harrist R, et al. The effects of an office ergonomics training and chair intervention on worker knowledge, behavior and musculoskeletal risk. Appl Ergon. 2009; 40(1): 12435. doi:10.1016/j.apergo.2007.12.009.

[6] Smith C. The effects of an ergonomic workstation intervention and exercise on office workers with neck pain. Faculty of Health Sciences: University of Cape Town; 2014.

[7] Esmaeilzadeh S, Ozcan E, Capan N. Effects of ergonomic intervention on workrelated upper

extremity musculoskeletal disorders among computer workers: a randomized controlled trial. Int Arch Occup Environ Health. 2014; 87(1): 7383. doi: 10.1007/s00420-012-0838-5

[8] Jull G, Sterling M, Falla D, Treleaven J, O'Leary S. Whiplash, headache, and neck pain: researchbased directions for physical therapies. 1st Edition. London, United Kingdom: Elsevier Health Sciences; 2008.

[9] Cagnie B, Dolphens M, Peeters I, Achten E, Cambier D, Danneels L. Use of muscle functional magnetic resonance imaging to compare cervical flexor activity between patients with whiplash-associated disorders and healthy controls. Phys Ther. 2010; 90(8): 115764. doi: 10.2522/ptj.20090351.

[10] Cagnie B, Struyf F, Cools A, Castelein B, Danneels L, O'Leary S. The relevance of scapular dysfunction in neck pain: a brief commentary. J Orthop Sports Phys Ther. 2014; 44(6): 4359. doi: 10.2519/jospt.2014.5038.

[11] Kim EK, Kim JS. Correlation between rounded shoulder posture, neck disability indices, and degree of forward head posture. J Phys Ther Sci. 2016; 28(10): 292932. doi: 10.1589/jpts.28.2929.

[12] Cole AK, McGrath ML, Harrington SE, Padua DA, Rucinski TJ, Prentice WE. Scapular bracing and alteration of posture and muscle activity in overhead athletes with poor posture. J Athl Train. 2013; 48(1): 1224. doi: 10.4085/1062-6050-48.1.13.

[13] Uhl T, Kibler W, Tripp B, Spigelman T, McClelland R, editors. Effects of S3 brace on scapular resting position and kinematics. Proceedings of Annual meeting of the American Orthopaedic Society for Sports Medicine; 2006.

[14] Leung M, Kan MMP, Cheng HMH, De Carvalho DE, Anwer S, Li H, et al. Effects of using a shoulder/scapular brace on the posture and muscle activity of healthy university students during prolonged typing-A randomized controlled crossover trial. Healthcare (Basel). 2023; 11(11): 1555. doi: 10.3390/healthcare11111555.

[15] Aronsson G, Gustafsson K, Dallner M. Sick but yet at work: an empirical study of sickness presenteeism. J Epidemiol Community Health. 2000; 54(7): 5029. doi: 10.1136/jech.54.7.502

[16] van Leeuwen MT, Blyth FM, March LM, Nicholas MK, Cousins MJ. Chronic pain and reduced work effectiveness: the hidden cost to Australian employers. Eur J Pain. 2006; 10(2): 1616. doi: 10.1016/j.ejpain.2005.02.007.

[17] Edmondston S, Björnsdóttir G, Pálsson T, Solgård H, Ussing K, Allison G. Endurance and fatigue characteristics of the neck flexor and extensor muscles during isometric tests in patients with postural neck pain. Man Ther. 2011; 16(4): 3328. doi: 10.1016/j.math.2010.12.005.

[18] Shaghayeghfard B, Ahmadi A, Maroufi N, Sarrafzadeh J. Evaluation of forward head posture in sitting and standing positions. Eur Spine J. 2016; 25: 357782. doi: 10.1007/s00586-015-4254-x.

[19] Thigpen CA, Padua DA, Michener LA, Guskiewicz K, Giuliani C, Keener JD, et al. Head and shoulder posture affect scapular mechanics and muscle activity in overhead tasks. J Electromyogr Kinesiol. 2010; 20(4): 7019. doi: 10.1016/j.jelekin.2009.12.003.

[20] Childs JD, Cleland JA, Elliott JM, Teyhen DS, Wainner RS, Whitman JM, et al. Neck pain: clinical practice guidelines linked to the ICF from the orthopaedic section of the APTA. J Womens Health Phys Ther. 2011; 35(2): 5790. doi: 10.2519/jospt.2008.0303.

[21] Noordin NB, Chyi LK, Ting A, Wah YC. Innovation of forward head posture orthosis with kyphosis brace. J Posit Sch Psychol. 2022: 351724.

[22] Bankhele JR, Yeradkar RS, Mulye MM. Effect of upper back brace with auditory feedback for forward head rounded shoulder posture among office workers with chronic neck pain-A prepost single group study. Indian J Occup Ther. 2020; 52(3): 859. doi: 10.4103/ijoth.ijoth_26_20

[23] Wilmarth M, Hilliard T. Measuring head posture via the craniovertebral angle. Orthop Phys Ther Pract [Internet]. 2002; 14(1): 13-15. Available from: https://sid.ir/paper/630774/en

[24] FernándezdelasPeñas C, AlonsoBlanco C, Cuadrado M, Pareja J. Forward head posture and neck mobility in chronic tensiontype headache: a blinded, controlled study. Cephalalgia. 2006; 26(3): 3149. doi: 10.1111/j.1468-2982.2005.01042.x.

[25] Kadu DV, Shetye JV. Reliability of "MBRuler Software" to measure craniovertebral angle using the photographic method. J Health Allied Sci NU. 2022; 12(3): 30711. doi: 10.1055/s-0041-1741415

[26] Cleland JA, Fritz JM, Whitman JM, Palmer JA. The reliability and construct validity of the Neck Disability Index and patient specific functional scale in patients with cervical radiculopathy. Spine (Phila Pa 1976). 2006; 31(5): 598602. doi: 10.1097/01.brs.0000201241.90914.22.

[27] Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics. 1993; 4: 35365. doi: 10.2165/00019053-199304050-00006.

[28] Blumenfeld AM, Patel AT, Turner IM, Mullin KB, Manack Adams A, Rothrock JF. Patientreported outcomes from a 1year, realworld, headtohead comparison of onabotulinumtoxinA and topiramate for headache prevention in adults with chronic migraine. J Prim Care Community Health. 2020; 11. doi: 10.1177/2150132720959936.

[29] Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav. 1983: 38596. Available from: https://doi.org/10.2307/2136404

[30] Brunton J, Brunton E, Ni Mhuiri A. Reliability of measuring natural head posture using the cranio-vertebral angle. Proceedings of Irish Ergonomics Review; 2003.
[31] Quek J, Pua YH, Clark RA, Bryant AL. Effects of thoracic kyphosis and forward head posture on cervical range of motion in older adults. Man Ther. 2013; 18(1): 6571. doi: 10.1016/j.math.2012.07.005.
[32] Singla D, Veqar Z. Methods of postural assessment used for sports persons. J Clin Diagn Res. 2014; 8(4): LE01. doi: 10.7860/JCDR/2014/6836.4266
[33] Caneiro JP, O'Sullivan P, Burnett A, Barach A, O'Neil D, Tveit O, et al. The influence of different sitting postures on head/neck posture and muscle activity. Man Ther. 2010; 15(1): 5460. doi: 10.1016/j.math.2009.06.002.
[34] Kim D, Cho M, Park Y, Yang Y. Effect of an exercise program for posture correction on musculoskeletal pain. J Phys Ther Sci. 2015; 27(6): 17914. doi: 10.1589/jpts.27.1791
[35] Falla D, Jull G, Russell T, Vicenzino B, Hodges P. Effect of neck exercise on sitting posture in patients with chronic neck pain. Phys Ther. 2007; 87(4): 40817. doi: 10.2522/ptj.20060009.
[36] Thorsteinsson EB. The relationship between workstress, psychological stress and staff health and work outcomes in office workers. Psychology. 2014. 5(10): 1301-11. doi: 10.4236/psych.2014.510141

Journal of Associated Medical Sciences 2026; 59(2): 298-306

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Comparison of carotid intima media thickness and high sensitive c-reactive protein with hemostasis model assessment for insulin resistance in prediabetes and normoglycemic patients: A case-control study

Kamani Bharath Teja[1], R Prasanna Venkatesh[2], Vignessh Raveekumaran[2]*

[1]Department of Cardiology, Amala Institute of Medical Sciences, Thrissur, Kerala, India.
[2]Department of General Medicine, Mahatma Gandhi Medical College and Research Institute, Sri Balaji Vidyapeeth (Deemed to be University), Pondicherry, India.

ARTICLE INFO

Article history:
Received 11 October 2025
Accepted as revised 16 February 2026
Available online 16 March 2026

Keywords:
Insulin resistant, carotid artery, diabetes

ABSTRACT

Background: Hyperglycaemia and insulin resistance (IR) are two significant risk factors for atherosclerotic cardiovascular disease (ASCVD). While the relationship between diabetes and atherosclerosis has been extensively studied and well established, data on prediabetes are scarce. The carotid artery intima-media thickness (CIMT) and high-sensitivity C-reactive protein (hs-CRP) are both well-established markers of subclinical atherosclerosis.

Objectives: The purpose of this study was to compare homeostasis model assessment-IR (HOMA-IR) levels in prediabetic and normoglycemic subjects, as well as to assess the relationship between HOMA-IR with hs-CRP and CIMT levels between these two groups.

Materials and methods: This case-control study enrolled 52 prediabetic patients and 52 normoglycemic controls. Glycaemic profile and hs-CRP were determined and HOMA-IR were calculated using fasting glucose and insulin values and bilateral CIMT measurements were taken. SPSS (ver_24.0) was used for analysis and based on the distribution of data analysis was performed.

Results: In both cases and controls, the mean serum fasting insulin level was 13.92±4.27 mIU/mL and 5.21±2.46 mIU/mL, respectively (p<0.001). The mean HOMA-IR values were 3.70±1.1 and 1.21±0.58, respectively (p<0.001). The mean CIMT was significantly greater in prediabetics (0.62±0.07mm) than in controls (0.45±0.08 mm) (p<0.001). Additionally, prediabetics had significantly higher serum-levels (5.70±3.35 mg/L) tan controls (1.88±1.35 mg/L) (p<0.001). A positive correlation of HOMA-IR existed between mean CIMT (p<0.001; r=0.631), and hs-CRP (p<0.001; r=0.498).

Conclusion: Prediabetic adults had higher CIMT and hs-CRP than normoglycemic controls, and HOMA-IR correlated positively with both, indicating greater subclinical atherosclerosis burden. These associations suggest IR may aid ASCVD risk stratification in prediabetes, though longitudinal studies are needed to confirm prognostic value.

Introduction

Noncommunicable diseases-especially atherosclerotic cardiovascular disease (ASCVD)-remain the leading global cause of death, and diabetes is a major accelerator of this risk. The latest International Diabetes Federation (IDF) estimates indicate a rapidly rising burden of diabetes worldwide, with substantial

* **Corresponding contributor.**
Author's Address: *Department of General Medicine, Mahatma Gandhi Medical College and Research Institute, Sri Balaji Vidyapeeth (Deemed to be University), Pondicherry, India.*
E-mail address: vignesshravee@gmail.com
doi: 10.66285/JAMS.2026.061
E-ISSN: 2539-6056

mortality and health-care costs.[1,2]

Prediabetes, defined by the American Diabetes Association (ADA) as impaired fasting glucose (IFG) 100-125 mg/dL, 2-hr oral glucose tolerance test (OGTT) 140-199 mg/dL, or HbA1c 5.7-6.4%, is common and clinically important because it signals elevated cardiometabolic risk.[3] Large meta-analyses show that prediabetes is associated with higher all-cause mortality and cardiovascular events compared with normoglycemia.[4] The risk of developing microvascular and macrovascular consequences, including nephropathy, retinopathy, peripheral neuropathy, cardiovascular events, and stroke emerge earlier in the transition from normal glucose to overt diabetes among these patients.[5]

IFG and impaired glucose tolerance (IGT) are two types of prediabetes that occur early in the spectrum of diabetes and have an estimated Relative Risk (RR) for ASCVD of 0.97 to 1.30 for IGT and 1.12 to 1.37 for IFG.[6-8] The risk associated with IFG and IGT is not well-established, despite a continuum of risk that extends into and below the pre-diabetic glucose range.

Subclinical atherosclerosis can be quantified noninvasively by carotid intima-media thickness (CIMT) using B-mode ultrasound, which correlates with traditional risk factors and future cardiovascular events. Contemporary reviews note standardized acquisition (segment proximal to the bifurcation, far-wall measurement, averaging multiple cardiac cycles) and support CIMT as an early marker of vascular injury.[9-11]

Age-stratified analyses also suggest CIMT improves prediction of cardiovascular death beyond traditional risk factors in adults <55 years.[12] Moreover, CIMT also corresponds strongly with ASCVD risk variables such blood pressure, lipid profiles, and smoking, as well as the severity of coronary artery disease (CAD) in both men and women.[13]

Inflammation is integral to dysglycemia-related atherogenesis. Systemic inflammation contributes to atherosclerosis, where high-sensitivity C-reactive protein (hs-CRP) is a practical inflammatory biomarker that tracks cardiometabolic risk.[14] Prospective data and meta-analyses in the cardiometabolic cohorts, link higher hs-CRP with incident type 2 diabetes mellitus (T2DM) and adverse cardiovascular outcomes across populations.[15-17] The hs-CRP levels also corresponded with the stage of insulin resistance and beta-cell dysfunction.[15] Findings from Indian cohorts further connect metabolic risk profiles with vascular changes relevant to CIMT.[18,19] In recent years, hs-CRP had emerged as a reliable test measure for predicting cardiovascular risk.

Insulin resistance (IR), commonly estimated by homeostasis model assessment (HOMA-IR), is mechanistically linked to endothelial dysfunction and atherosclerosis. Population data show positive associations between HOMA-IR and CIMT in non-diabetic adults after multivariable adjustment,[20] and contemporary work continues to relate cardiometabolic indices to carotid atherosclerosis burden.

The primary objective of this present study was to compare CIMT and hs-CRP levels between adults with prediabetes and normoglycemic controls, and the secondary objective were to evaluate IR using the HOMA-IR and to examine the associations of HOMA-IR with CIMT and hs-CRP within the study population.

Materials and methods

Study design and setting

A case-control study comparing adults with prediabetes with normoglycemic controls conducted in the Department of General Medicine, at a single tertiary care centre, Pondicherry, India. Methods and reporting followed STROBE recommendations.[21] The study was conducted over a period of 18 months, from January 2020 to June 2021. Ethical approval was obtained from Institutional Human Ethics Committee and informed consent obtained from the patients.

Study participants and definitions

Prediabetes was defined as per ADA criteria,[3] as FPG: 100-125 mg/dL; 2-hr PPG 140-199 mg/dL; or HbA_{1C} 5.7-6.4%. Controls were normoglycemic (FPG levels of 70-99 mg/dl, PPG levels <140 mg/dl, and $HbA1_c$ values <5.7%) by the same standards.

Patients with established history of diabetes (FPG>126mg/dl, 2-hr PPG>200mg/dl, or HbA1c>6.5%), and those with low FPG (<70 mg/dl) or currently on anti-diabetic therapy, individuals with were excluded. Additionally, participants diagnosed with other non-communicable diseases or undergoing treatment for hypertension, dyslipidaemia, or any condition that could influence CIMT were not included in the study.

Sample size and sampling technique

The sample size was estimated using CIMT as the primary variable, based on the findings of Chauhan *et al.*,[8] where the mean CIMT among prediabetic individuals was reported as 0.59±0.11 mm. The required sample size to estimate the mean CIMT with 95% confidence and an absolute precision (d) of 0.03mm was calculated using the formula; (It yields N~52. With a 1:1 case-control design, 52 prediabetes cases and 52 normoglycemic controls were included. A convenience sampling technique was employed to select participants.

Clinical and laboratory measurements

The participants' FPG, PPG, and HbA_{1c} levels were then determined after an overnight fast (8 hours), venous blood was collected in EDTA containers using an automated biochemical analyser. After two hours of 75 gm glucose consumption, blood samples were collected for PPG. The hs-CRP was measured using a validated high-sensitivity immunoassay and reported in mg/L.

HOMA-IR was used to estimate the IR as [fasting insulin (μIU/mL)*fasting glucose (mg/dL)]/405. CIMT was measured by B-mode ultrasound using a 7.5-10 MHz linear-array transducer with participant in the supine position and the head rotated contralaterally. CIMT was measured in the distal common carotid artery (CCA) on both sides, using a 10-mm (1-cm) plaque-free arterial segment located approximately 1cm proximal to the carotid bulb/bifurcation. Measurement was obtained from the far wall (leading edge of the lumen-intima interface to the leading edge of the media-adventitia interface), as recommended in standardised CIMT acquisition protocols. For each side, three end-diastolic measurements (at adjacent points within the selected 10-mm segment) were recorded and averaged to obtain a mean CIMT for that side. The overall mean CIMT was used for the analysis was calculated as the average of the right and left mean CIMT values. Protocol elements followed ASE consensus and recent reviews.[22]

Quality assurance

CIMT measurements were performed by a trained sonographer who was blinded to the participant's glycaemic status. All scans were obtained using a standardized acquisition protocol to minimize measurement variability. To assess the measurement repeatability, CIMT was re-measured in a randomly selected subset of participants on other periods by the same operator, blinded by the initial readings.

Outcomes

Primary outcome was the difference in mean CIMT and hs-CRP between prediabetes and controls, while the secondary outcomes show the association of HOMA-IR with CIMT and hs-CRP.

Statistical analysis

Analyses were performed using SPSS (ver_21.0). Categorical data (frequency (percentage)) and continuous variables (mean±SD and median (IQR)) based on the normality were provided. The Kolmogorov-Smirnov test was used to determine whether the data were normal. Non-parametric tests were performed if normality was refused. When the data sets weren't normally distributed, we used independent t-tests/ Mann-Whitney tests to compare quantitative variables between the two groups. The Chi-square test/Fisher's-Exact test was used to examine the relationship between qualitative variables. The correlation coefficient of Spearman rank was utilised to determine the relationship between various factors. The cause-and-effect link between several characteristics was discovered using univariate linear regression. The $p<0.05$ were considered statistically significant.

Results

Among the study population 104 participants, (52 prediabetes; 52 normoglycemic). Mean age was higher in prediabetes (49.77±7.94 years) than controls (45.10±10.70 years), mean difference 4.67 (95% CI 1.01-8.33; $p=0.013$). Sex distribution: 36 (69.2%)/16 (30.2%) (M/F) in prediabetes and 32 (61.5%)/20 (38.5%) (M/F) in controls. Baseline characteristics of the study participants were presented in Table 1. Glycemia and insulin measures between both groups were presented in Table 2.

Table 1. Baseline characteristics of the study participant.

Variables	Prediabetes** (N=52)	Normoglycemic** (N=52)	Mean differences (95% CI)	*p* value
BMI (kg/m^2)	25.87±5.02	24.88±2.69	0.98 (-0.58 to 2.5)	0.218
SBP (mmHg)	124.15±9.55	120.21±9.79	3.94 (0.179 to 7.705)	0.04*
DBP (mmHg)	81.15±5.4	77.56±7.86	3.59 (0.970 to 6.223)	0.08

Note: **Independent t-test at p<0.05 were statistically significant, **mean±SD, BMI: body mass index, SBP: systolic blood pressure, DBP: diastolic blood pressure, CI: confidence interval.*

Table 2. Glycemia and insulin measures among the study participants.

Variables	Prediabetes** (N=52)	Normoglycemia** (N=52)	Mean differences (95% CI)	*p* value
FBG (mg/dL)	107.69±7.31	92.32±8.26	15.32 (12.21-18.36)	<0.001*
PPBG (mg/dL)	157.92±23.71	114.87±21.42	43.05 (34.26-51.85)	<0.001*
HbA_{1C}	6.04±0.00	5.08±0.09	0.97 (0.837-1.163)	<0.001*
FI (mIU/mL)	13.92±4.27	5.21±2.46	8.71 (7.35-10.06)	<0.001*
HOMA-IR	3.70±1.10	1.21±0.58	2.38 (2.01-2.75)	<0.001*

Note: **Independent t-test at p<0.05 were statistically significant, **mean±SD, FBG: fasting blood glucose, PPBG: post-prandial blood glucose, HbA_{1C}: glycated hemoglobin, FI: fasting insulin, HOMA-IR: homeostasis model assessment insulin resistance, CI: confidence interval.*

Distribution of hs-CRP among the study participants were presented in Table 3. The mean hs-CRP (mg/L) among the prediabetes patients was 5.70±3.35 mg/L and the normoglycemic patients was 1.88±1.35 mg/L with the mean difference of 3.82 and were statistically significant (95% CI: 2.83-4.81; p<0.00). Similarly, the mean CIMT (mm) was 0.62±0.07 mm and 0.45±0.08 mm among the prediabetes and normoglycemic patients, respectively and were statistically significant with the mean difference of 0.16 (95% CI: 0.13- 0.21; p<0.001).

Table 3. hs-CRP levels among the study participants

hs-CRP (mg/L)	Prediabetes (N=52) N (%)	Normoglycemia (N=52) N (%)	Total	*p* value
<1	0 (0.0)	16 (30.8)	16 (15.3)	<0.001*
1-2.99	9 (17.3)	31 (59.6)	45 (43.2)	<0.001*
3-9.99	34 (65.4)	5 (9.6)	39 (37.5)	<0.001*
>10	9 (17.3)	0 (0.0)	9 (8.6)	<0.001*

Note: **Chi-square test, p<0.05 were statistically significant, hs-CRP: high sensitive C-reactive protein.*

Correlation was performed between the HOMA-IR with hs-CRP and CIMT showed positive correlation with both among pre-diabetes patients. For CIMT it showed stronger and positive correlation (r=0.631; p<0.001), while with hs-CRP it had moderate correlation (r=0.498; p<0.001) and both were statistically significant (Figure 1 and Figure 2).

Moreover, correlation between hs-CRP and CIMT in prediabetes study group were performed where the Pearson correlation coefficient was 0.786, indicating that the strong positive correlation between hs-CRP and CIMT among the prediabetes patient. The correlation was found to be statistically significant (p<0.001) (Figure 3).

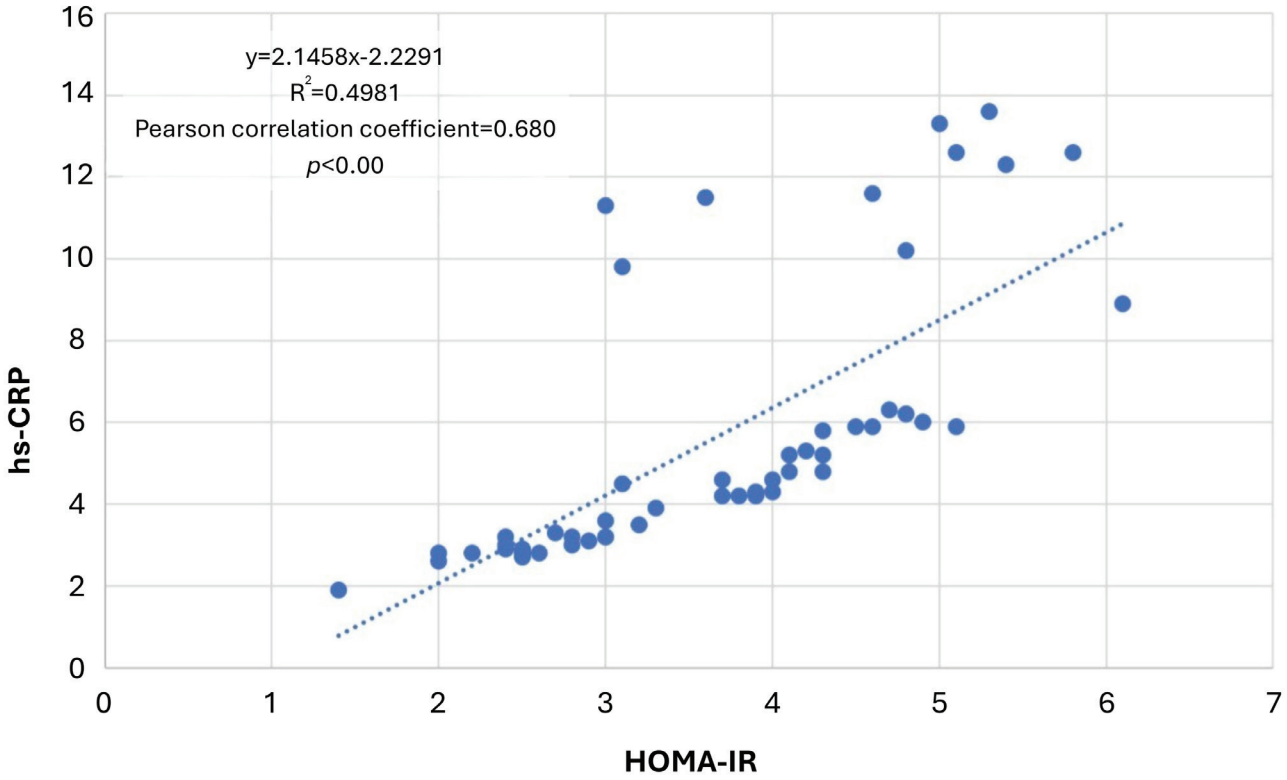


Figure 1. *Association between insulin resistance and CIMT in prediabetes. Scatter plot showing the relationship between HOMA-IR and mean carotid intima-media thickness (CIMT, mm) among adults with prediabetes (N=52). Each point represents one participant. Correlation coefficient (r)=0.631; p<0.001.*

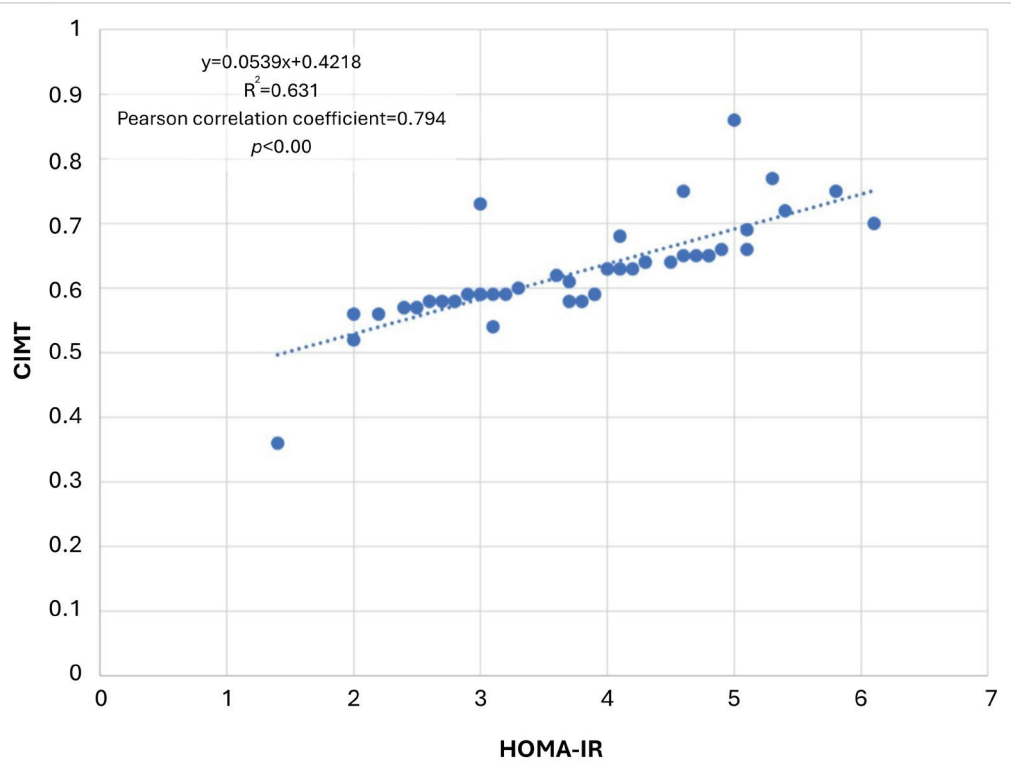


Figure 2. *Association between insulin resistance and hs-CRP in prediabetes. Scatter plot showing the relationship between HOMA-IR and high-sensitivity C-reactive protein (hs-CRP, mg/L) among adults with prediabetes (N=52). Each point represents one participant. Correlation coefficient (r)=0.498; p<0.001.*

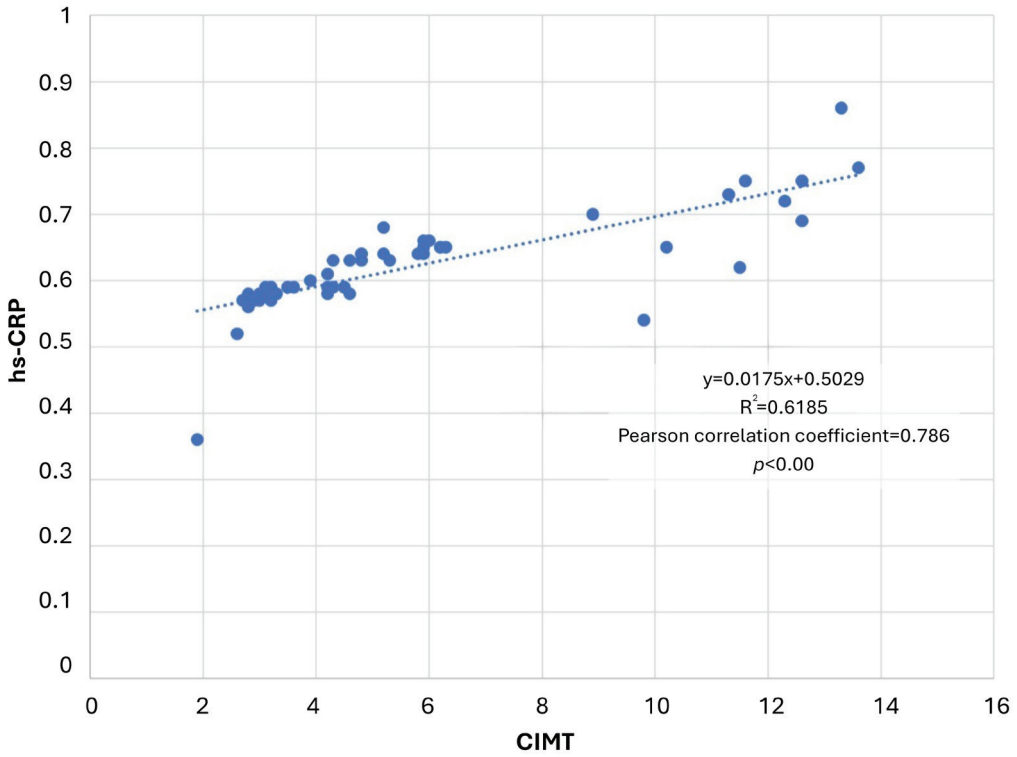


Figure 3. *Association between inflammation and CIMT in prediabetes. Scatter plot showing the relationship between hs-CRP (mg/L) and mean CIMT (mm) among adults with prediabetes (N=52). Each point represents one participant. Correlation coefficient (r)=0.786; p<0.001.*

Discussion

In this case-control study, adults with prediabetes showed higher subclinical atherosclerosis and inflammation than normoglycemic controls. IR correlated positively with both CIMT (r=0.631) and hs-CRP (r=0.498). Various studies have proven that cardiovascular secondary to atherosclerosis is more in diabetics when compared to non-diabetics. But not many studies have been done to prove the associated increase in risk of atherosclerotic cardiovascular disease in prediabetic stage.

An important contribution of this present work is that it provides contemporary data from an Indian tertiary-care setting on the combined assessment of subclinical atherosclerosis (CIMT) and systemic inflammation (hs-CRP) in adults with prediabetes compared with normoglycemic controls. While prior literatures,[4,6,7,9,10,12,13,15,17] established higher cardiovascular risk in overt diabetes, evidence specifically characterising early vascular change and inflammation in the prediabetic stage remains relatively limited, particularly across diverse Indian populations. By demonstrating higher CIMT and hs-CRP in prediabetes and their positive associations with HOMA-IR, our findings add population-relevant support to the concept that atherogenic remodelling and low-grade inflammation begin before clinically overt diabetes.

Prediabetes involves widespread insulin resistance, which disrupts nitric oxide signalling in blood vessels, increases oxidative stress, and speeds up changes in artery walls-all factors that lead to higher CIMT. Additionally, mild inflammation caused by cytokines from fat and the liver (such as IL-6) triggers the liver to produce CRP; elevated hs-CRP levels reflect this inflammatory state and are associated with vascular insulin resistance.[23] Contemporary reviews position CIMT as a validated marker of subclinical atherosclerosis and future ASCVD risk, consistent with our findings.[10,22,24,25] Lower inflammatory burden relates to glycaemic improvement, as cohorts show that lower hs-CRP is associated with regression from prediabetes to normoglycemia.[26] In Mahat et al it has shown that prediabetic subjects have abnormalities in atherogenic indices and CIMT, indicating a higher predilection for the development of cardiovascular disease. As a result, atherogenic indices, in addition to routine lipid parameters, can be used to better assess subclinical atherosclerosis in prediabetic subjects.[27] In Groop *et al*., it was shown that the underlying pathogenesis for IR in overt diabetes and prediabetes was microvascular complications that are secondary to diabetes.[28]

Our mean CIMT difference (0.62 vs 0.45 mm) aligns with multiple reports showing thicker carotid walls in prediabetes. In a 100-case/100-control study, Bhinder *et al*., reported CIMT 0.79±0.06 mm (prediabetes) vs 0.72±0.02 mm (controls) (p<0.05).[29] A larger cross-sectional series also found higher common and internal CIMT across impaired glucose states compared with normal metabolism.[30] A recent review summarizes that many-but not all-studies show greater CIMT in prediabetes, noting heterogeneity by measurement segment and population.[31] Earlier comparative data likewise showed higher CIMT in prediabetes vs controls.[32] Together, these support our observation of thicker CIMT in prediabetes.

We observed markedly higher hs-CRP in prediabetes (5.70±3.35 mg/L) than controls (1.88±1.35 mg/L). Population and clinic-based data similarly demonstrate higher hs-CRP with dysglycemia: Ghule *et al*., reported hs-CRP 2.17±0.72 mg/L in prediabetes vs 0.66±0.22 mg/L in controls (p<0.0001).[33] In an Indian cohort, median hs-CRP was higher in IFG/IGT (≈2.2-2.3 mg/L) vs NGT (1.64 mg/L).[34] More recently, NHANES analyses show the hs-CRP/HDL-C ratio rises across normal→prediabetes→diabetes and independently associates with prediabetes odds (OR per unit ≈1.13).[35] In Sabanayagam C *et al,* the study that was done in two Asian population has discovered the association between prediabetes and elevated hs-CRP levels.[36] In the study of Pfützner *et al*.'s, relation between hs-CRP and cardiovascular risk has shown that the risk exponentially rises in diabetics and also in prediabetes compared to non-diabetes.[37] These data corroborate our observation that inflammation is higher in prediabetes.

Chauhan et al. has studies between prediabetes and non-diabetes groups and measured CIMT and its correlation with HOMA-IR and hs-CRP.[8] It has proven to show positive correlation between all three parameters. Huang *et al*. has shown that when prediabetes was defined using IFG-ADA, IFG-WHO, IGT, and HbA_{1c}, studies show that it was associated with an increased risk of all-cause mortality, composite cardiovascular events, coronary heart disease, and stroke.[38]

Clinical implications

CIMT and hs-CRP, when interpreted alongside standard risk factors, can help flag higher ASCVD risk among adults with prediabetes. Elevated CIMT reflects early arterial remodelling, while higher hs-CRP signals active low-grade inflammation; together with HOMA-IR they identify a subgroup with greater subclinical burden and may prompt earlier lifestyle and cardiometabolic risk management (e.g., weight reduction, BP/lipid control). Although not diagnostic, CIMT progression relates to future events, supporting its role as an adjunct risk marker in preventive care.[39,40]

Limitations

First, the case-control, single-center design limits causal inference and generalizability. Second, modest sample size (N=104) and partial age imbalance may introduce residual confounding; multivariable adjustment beyond age/sex/BMI (e.g., smoking, lipids, medications) was not fully reported. Third, hs-CRP can be influenced by intercurrent inflammation; although major illnesses were excluded, single-timepoint

measurements may misclassify chronic inflammatory status. Finally, CIMT methodology (e.g., segment/side averaging, blinding, reproducibility metrics) should be standardized in future work to enhance comparability.

Conclusion

Adults with prediabetes demonstrated higher CIMT and hs-CRP compared with normoglycemic controls, and HOMA-IR showed positive associations with both CIMT and hs-CRP. These findings support the concept that subclinical atherosclerosis and low-grade inflammation may begin early in the dysglycemic continuum, before overt diabetes. Larger, well-designed prospective studies are needed to confirm these relationships after adjustment for cardiovascular risk factors and to determine whether CIMT and hs-CRP improve risk stratification and predict future cardiovascular events in prediabetes.

Ethical approval

Ethical clearance was obtained from the Institutional Human Ethics Committee (MGMCRI/Res/01/2019/45/IHEC/119).

Funding

This research received no external funding.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

R Prasanna Venkatesh: conceptualization, methodology, writing: review and edit; **Vignessh Raveekumara**: methodology, data analysis; **Kamani Bharath Teja**: data curation, data analysis, writing: original draft.

Acknowledgements

I sincerely acknowledge the staff nurses and laboratory technicians who helped me throughout the study.

References

[1] Magliano DJ, Boyko EJ, Committee IDA 10th Edition scientific. Global picture. In: IDF DIABETES ATLAS 10th Edition [Internet]. International Diabetes Federation; 2021 [cited 2025 Jan 27]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK581940/

[2] Yadi H. IDF Diabetes Atlas estimates for the global diabetes prevalence of adults older than 60 years. 2017 Nov 10; Available from: doi.org/10.26226/morressier.59d51843d462b80296ca3751.

[3] American Diabetes Association Professional Practice Committee. 2. Diagnosis and Classification of Diabetes: Standards of Care in Diabetes-2024. Diab Care. 2023; 47(Suppl_1): S20-42. doi: 10.2337/dc24-S002.

[4] Cai X, Zhang Y, Li M, Wu JH, Mai L, Li J, et al. Association between prediabetes and risk of all-cause mortality and cardiovascular disease: updated meta-analysis. BMJ. 2020;370: m2297. doi: 10.1136/bmj.m2297.

[5] Ohiagu FO, Chikezie PC, Chikezie CM. Pathophysiology of diabetes mellitus complications: Metabolic events and control. Biomed Res Ther. 2021; 8(3): 4243-57. doi:10.15419/bmrat.v8i3.663.

[6] Yip WCY, Sequeira IR, Plank LD, Poppitt SD. Prevalence of pre-diabetes across ethnicities: A review of impaired fasting glucose (IFG) and impaired glucose tolerance (IGT) for classification of dysglycaemia. Nutrients. 2017; 9(11): 1273. doi.org/10.3390/nu9111273.

[7] Neves JS, Newman C, Bostrom JA, Buysschaert M, Newman JD, Medina JL, et al. Management of dyslipidemia and atherosclerotic cardiovascular risk in prediabetes. Diabetes Res Clin Pract. 2022; 190: 109980. doi: 10.1016/j.diabres.2022.109980.

[8] Chauhan A, Ramesh P, Goyal P, Jilowa S, Singh A, Singhal A, et al. Carotid Intima Media Thickness and High Sensitivity C-Reactive Protein Levels in Prediabetes and its Correlation with Homeostatic Model Assessment for Insulin Resistance. J Clin Diagn Res. 2019; 13(9): 10-14. doi: 10.7860/JCDR/2019/42480.13157.

[9] Willeit P, Tschiderer L, Allara E, Reuber K, Seekircher L, Gao L, et al. Carotid intima-media thickness progression as surrogate marker for cardiovascular risk: Meta-analysis of 119 clinical trials involving 100,667 patients. Circulation. 2020;142(7): 621-42. doi: 10.1161/CIRCULATIONAHA.120.046361.

[10] Tschiderer L, Seekircher L, Izzo R, Mancusi C, Manzi MV, Baldassarre D, et al. Association of intima-media thickness measured at the common carotid artery with incident carotid plaque: Individual participant data meta-analysis of 20 prospective studies. J Am Heart Assoc. 2023; 12(12): e027657. doi: 10.1161/JAHA.122.027657.

[11] Randolph W, Dains JE. Ultrasound evaluation of carotid artery intima-media thickness: Effective early marker of carotid artery disease in adult head and neck cancer patients after neck radiation? J Adv Pract Oncol. 2022; 13(7): 683-94. doi: 10.6004/jadpro.2022.13.7.4

[12] Ge J, Jing F, Ji R, Tian A, Su X, Li W, et al. Age-related trends in the predictive value of carotid intima-media thickness for cardiovascular death: A prospective population-based cohort study. J Am Heart Assoc. 2023; 12(13): e029656. doi: 10.1161/JAHA.123.029656.

[13] Holland Z, Ntyintyane LM, Raal FJ, Gill GV. Carotid intima–media thickness is a predictor of coronary artery disease in South African black patients. Cardiovasc J Afr. 2009; 20(4): 237-9.

[14] Sun H, Yang J, Ma L, Wu Y. Association between

hs-CRP/HDL-C ratio and risk of prediabetes or diabetes: a cross-sectional study based on NHANES 2015-2023. BMC Endocr Disord. 2025; 25(1): 183. doi: 10.1186/s12902-025-02004-0.

[15] Basil B, Mohammed JA, Mba IN, Myke-Mbata BK, Akujieze JC. High-sensitivity C-reactive protein and cardiovascular disease risk assessment in a population of type 2 diabetes mellitus patients. BMC Cardiovasc Disord. 2025; 25(1): 498. doi: 10.1186/s12872-025-04975-3.

[16] Yang X, Tao S, Peng J, Zhao J, Li S, Wu N, et al. High-sensitivity C-reactive protein and risk of type 2 diabetes: A nationwide cohort study and updated meta-analysis. Diabetes Metab Res Rev. 2021; 37(8): e3446. doi: 10.1002/dmrr.3446.

[17] Liu Y, He W, Ji Y, Wang Q, Li X. A linear positive association between high-sensitivity C-reactive protein and the prevalence of cardiovascular disease among individuals with diabetes. BMC Cardiovasc Disord. 2024; 24(1): 411. doi: 10.1186/s12872-024-04091-8.

[18] Sajja V, Jeevarathnam D, James S, Rathinasamy J. A study on carotid artery intima-media thickness and metabolic risk factors in overweight and obese Indian children. Diabetol Int. 2019; 11(2): 142-9.

[19] Singh RP, Chauhan K, Tripathi A, Choudhary R. Association of adipokines and CIMT in metabolic syndrome in Western Uttar Pradesh population: a cross-sectional s tudy. Mædica. 2024;19(3): 494-501. doi: 10.26574/maedica.2024.19.3.494.

[20] Santos IS, Bittencourt MS, Goulart AC, Schmidt MI, Diniz M de FHS, Lotufo PA, et al. Insulin resistance is associated with carotid intima-media thickness in non-diabetic subjects. A cross-sectional analysis of the ELSA-Brasil cohort baseline. Atherosclerosis. 2017; 260: 34-40.

[21] Field N, Cohen T, Struelens MJ, Palm D, Cookson B, Glynn JR, et al. Strengthening the reporting of molecular epidemiology for infectious diseases (STROME-ID): an extension of the STROBE statement. Lancet Infect Dis. 2014; 14(4): 341-52. doi: 10.1016/S1473-3099(13)70324-4.

[22] Johri AM, Nambi V, Naqvi TZ, Feinstein SB, Kim ESH, Park MM, et al. Recommendations for the assessment of carotid arterial plaque by ultrasound for the characterization of atherosclerosis and evaluation of cardiovascular risk: From the American Society of Echocardiography. J Am Soc Echocardiogr. 2020; 33(8): 917-33. doi: 10.1016/j.echo.2020.04.021.

[23] Horton WB, Love KM, Gregory JM, Liu Z, Barrett EJ. Metabolic and vascular insulin resistance: partners in the pathogenesis of cardiovascular disease in diabetes. Am J Physiol-Heart Circ Physiol. 2025; 328(6): H1218-36. doi: 10.1152/ajpheart.00826.2024.

[24] Liu AB, Lin YX, Meng TT, Tian P, Chen JL, Zhang XH, et al. Associations of the cardiometabolic index with insulin resistance, prediabetes, and diabetes in U.S. adults: a cross-sectional study. BMC Endocr Disord. 2024; 24: 217. doi: 10.1186/s12902-024-01676-4.

[25] Mashaba RG, Phoswa W, Maimela E, Mokgalaboni K. Association of carotid intima–media thickness and dyslipidaemia in patients with type 2 diabetes: a protocol for systematic review and meta-analysis. BMJ Open. 2024; 14(1): e079209. doi:10.1136/bmjopen-2023-079209.

[26] Cheng ZJ, Wang YF, Jiang XY, Ren WY, Lei SF, Deng FY, et al. High sensitivity C-reactive protein and prediabetes progression and regression in middle-aged and older adults: A prospective cohort study. J Diabetes Investig. 2024; 15(1): 78-86. doi: 10.1111/jdi.14090.

[27] Mahat RK, Singh N, Rathore V, Gupta A, Shah RK. Relationship between atherogenic indices and carotid intima-media thickness in prediabetes: A cross-sectional study from Central India. Med Sci (Basel). 2018; 6(3): 55. doi: 10.3390/medsci 6030055.

[28] Groop PH, Forsblom C, Thomas MC. Mechanisms of disease: Pathway-selective insulin resistance and microvascular complications of diabetes. Nat Clin Pract Endocrinol Metab. 2005; 1(2): 100-10.

[29] Bhinder HPS, Kamble TK. The study of carotid intima-media thickness in prediabetes and its correlation with cardiovascular risk factors. J Datta Meghe Inst Med Sci Univ. 2018; 13(2): 79.

[30] Bulut A, Avci B. Carotid intima-media thickness values are significantly higher in patients with prediabetes compared to normal glucose metabolism. Medicine (Baltimore). 2019; 98(44): e17805. doi: 10.1097/MD.0000000000017805.

[31] Gateva A, Assyov Y, Karamfilova V, Kamenov Z. Common carotid artery intima media thickness (CIMT) in patients with prediabetes and newly diagnosed type 2 diabetes mellitus. J Diabetes Complications. 2024; 38(7): 108766.

[32] Parildar H, Gulmez O, Cigerli O, Dogruk Unal A, Erdal R, Guvener Demirag N. Carotid artery intima media thickness and hsCRP; Predictors for atherosclerosis in prediabetic patients? Pak J Med Sci. 2013; 29(2): 495-9. doi: 10.12669/pjms.292.3133.

[33] Ghule A, Kamble TK, Talwar D, Kumar S, Acharya S, Wanjari A, et al. Association of serum high sensitivity C-reactive protein with Pre-diabetes in rural population: A Two-Year Cross-Sectional Study. Cureus. 2025; 13(10): e19088. doi: 10.7759/cureus.19088.

[34] Jaiswal A, Tabassum R, Podder A, Ghosh S, Tandon N, Bharadwaj D. Elevated level of C-reactive protein is associated with risk of prediabetes in Indians. Atherosclerosis. 2012; 222(2): 495-501. doi: 10.1016/j.atherosclerosis.2012.02.034.

[35] Sun H, Yang J, Ma L, wu Y. Association between hs-CRP/HDL-C ratio and risk of prediabetes or

diabetes: a cross-sectional study based on NHANES 2015-2023. BMC Endocr Disord. 2025; 25(1): 183. doi: 10.1186/s12902-025-02004-0.

[36] Sabanayagam C, Shankar A, Lim SC, Lee J, Tai ES, Wong TY. Serum C-reactive protein level and prediabetes in two Asian populations. Diabetologia. 2011; 54(4): 767-75. doi: 10.1007/s00125-011-2052-5.

[37] Pfützner A, Forst T. High-sensitivity C-reactive protein as cardiovascular risk marker in patients with diabetes mellitus. Diabetes Technol Ther. 2006; 8(1): 28-36. doi: 10.1089/dia.2006.8.28.

[38] Huang Y, Cai X, Mai W, Li M, Hu Y. Association between prediabetes and risk of cardiovascular disease and all-cause mortality: systematic review and meta-analysis. BMJ. 2016; 355: i5953. doi: 10.1136/bmj.i5953.

[39] Liu AB, Lin YX, Meng TT, Tian P, Chen JL, Zhang XH, et al. Associations of the cardiometabolic index with insulin resistance, prediabetes, and diabetes in U.S. adults: a cross-sectional study. BMC Endocr Disord. 2024; 24: 217. doi: 10.1186/s12902-024-01676-4.

[40] Willeit P, Tschiderer L, Allara E, Reuber K, Seekircher L, Gao L, et al. Carotid intima-media thickness progression as surrogate marker for cardiovascular risk. Circulation. 2020; 142(7): 621-42. doi: 10.1161/CIRCULATIONAHA.120.046361.

Journal of Associated Medical Sciences 2026; 59(2): 307-315

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Dual-scale evaluation of pain after hypofractionated post-mastectomy radiotherapy

Pratibha Prashar[1*], Brish Bhanu Bhardwaj[2], Vandana Thakur[1], Poorva Vias[1], Ankita Pandey[3], Hardik Sharma[4], Kulbir Singh[5], Vivek Kumar[5]

[1]*Department of Radiotherapy and Oncology, Dr. Rajendra Prasad Government Medical College, Tanda, India.*
[2]*Department of Pathology, Dr. Rajendra Prasad Government Medical College, Tanda, India.*
[3]*Department of Radiotherapy and Oncology, Rohilkhand Medical College Hospital, Bareilly, India.*
[4]*Department of Radiotherapy and Oncology, All India Institute of Medical Sciences, Bilaspur,India.*
[5]*Department of Radiotherapy and Oncology, Dr. Radha Krishana Medical College, Hamirpur, India.*

ARTICLE INFO

Article history:
Received 17 December 2025
Accepted as revised 5 March 2026
Available online 18 March 2025

Keywords:
Neuropathic pain, somatic pain, post-mastectomy radiotherapy pain

ABSTRACT

Background: Post-mastectomy radiotherapy (PMRT) pain encompassing somatic shoulder and chest wall pain (SCWP) and neuropathic pain (NP) is a clinically significant late toxicity following hypofractionated radiotherapy (HFRT) for breast cancer. Limited data exist regarding its incidence, temporal evolution, and determinants across different HFRT regimens.

Objectives: To evaluate the incidence and predictors of SCWP and NP in patients receiving chest wall and regional nodal HFRT, across three HFRT schedules.

Materials and methods: Breast cancer patients were treated with one of three HFRT regimens: 26 Gy in 5 fractions, 34 Gy in 10 fractions, or 40 Gy in 15 fractions. PMRT pain was evaluated using the Common Terminology Criteria for Adverse Events, version 5.0 (CTCAE v5.0) for SCWP, and NP using the Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) scale. The temporal patterns of both pains were analyzed. Dosimetric parameters of the chest wall (CW), brachial plexus (BP), axillary levels I-III, and supraclavicular fossa (SCF), anatomical parameter (CW thickness), and clinical parameters (age, BMI) were correlated with pain outcomes.

Results: The mean duration of SCWP was 6 months (range, 3-36 months), which was most frequently observed with HFRT-C (40%), followed by HFRT-B (22.9%) and HFRT-A (23.3%) (p=0.09). Patients with Grade≥2 SCWP had significantly thinner CW (p=0.04) and higher CW D_{max} EQD2 values (p=0.04). NP occurred with a mean onset of 12 months (range, 3-24 months) and was most common in HFRT-C (60%), versus 40% in HFRT-A and 28.6% in HFRT-B (p=0.001).On multivariate analysis, Level III axillary dissection (0.049) and higher BP D_{max} EQD2(p=0.001), axillary level III D_{mean} EQD2 (p=0.049) were found to be significant predictors of NP.

Conclusion: HFRT regimen selection and radiation dose distribution significantly influence the incidence and severity of SCWP and NP. Patients with thinner chest walls, higher maximum doses to critical structures, and those undergoing Level III axillary dissection are at greater risk of developing these complications. These findings underscore the importance of refined dose-constraint strategies during HFRT planning to mitigate pain-related morbidity and improve long-term quality of life.

***Corresponding contributor.**
Author's Address: *Department of Radiotherapy and Oncology, Dr. Rajendra Prasad Government Medical College, Tanda, India.*
E-mail address: pratibhaprashar43@gmail.com
doi: 10.66285/JAMS.2026.062
E-ISSN: 2539-6056

Introduction

Post-mastectomy radiotherapy (PMRT) pain involving the chest wall (CW), axilla, and upper limb remains one of the most pervasive and functionally debilitating late effects in breast cancer survivors. Following mastectomy and adjuvant radiotherapy (RT), up to 25-60% of patients develop chronic pain that may persist for years, profoundly compromising quality of life and upper-limb function. This persistent PMRT pain constitutes a multifactorial late-toxicity syndrome, originating from the overlap of radiation-induced tissue injury, postoperative musculoskeletal impairment, and neuropathic sensory dysfunction.[1] Both somatic shoulder chest-wall pain (SCWP) and neuropathic pain (NP) frequently coexist in this population, underscoring the heterogeneous nature of PMRT-related pain.[2,3] Radiation-induced chest-wall pain (RICWP) is classically defined as pain originating from irradiated thoracic structures (ribs, intercostal musculature, fascia, subcutaneous tissues), typically presenting months to years after treatment completion.[4,5]

Concurrently, surgical factors, such as sacrifice of the intercostobrachial nerve (ICBN), and the extent of axillary dissection (AD), contribute to neuropathic manifestations, while postoperative musculoskeletal impairments, such as stiffness, reduced strength, and limited range of motion, add functional disability to the clinical picture. RT further exacerbates tissue injury through progressive fibrosis, microvascular compromise, chronic inflammation, and oxidative stress, resulting in reduced tissue compliance, myofascial pain, and impaired shoulder mobility. Dose-dependent injury to the brachial plexus (BP) can manifest as NP, paraesthesia, or upper-limb dysfunction, particularly when BP doses exceed contemporary tolerance thresholds.[6-8]

SCWP may also arise from radiation effects on CW, particularly when a larger irradiated volume is involved. The dose delivered to this region is influenced by factors such as wall thickness, thoracic anatomy, and overall patient morphology. These variables can increase tissue exposure, create dose heterogeneity, and elevate the risk of fibrosis, soft-tissue contracture, and osseous toxicity. Studies consistently demonstrate that increased CW dose directly heightens the risk of chronic pain following hypofractionated radiotherapy (HFRT).[9]

Host-related factors, including younger age, elevated body mass index (BMI), diabetes, and pre-existing pain syndromes, further increase susceptibility to both NP and SCWP. Despite growing recognition of these risk determinants, the contribution of detailed axillary and BP dosimetry, as well as modern CW dose metrics, remains insufficiently characterized for RICWP.[10]

A retrospective cohort study of post-mastectomy patients treated with contemporary HFRT was studied to bridge these knowledge gaps. The primary objective was to observe the incidence of SCWP and NP across the three HFRT regimens and to assess their correlation with patient factors, surgical details, regional nodal doses, and CW anatomy. The overlap of somatic and neuropathic symptoms in PMRT-related pain complicates accurate assessment, with limited evidence differentiating these phenotypes using validated dual-assessment tools in HFRT settings. To address this, a dual evaluation strategy was used: common terminology criteria of adverse events (CTCAE) version 5.0 was used to assess somatic and functional impairments, while LANSS (Leeds Assessment of Neuropathic Symptoms and Signs) scoring identified and quantified neuropathic pain. This combined approach allows precise characterization of both symptom types, enabling a better understanding of the multifactorial nature of persistent PMRT-related pain.

Materials and methods

This retrospective observational cohort study was conducted in the Department of Radiation Oncology, Dr. RPGMC Tanda, from October 2021 to August 2024. The study delineated the incidence and temporal patterns of SCWP and NP. The relationship between regional radiation dose parameters, anatomical, surgical, and clinical factors, and the development of CW pain syndromes in breast cancer patients treated with post-mastectomy radiotherapy (PMRT)was also evaluated.

Study population

Eligible participants were women with histologically confirmed breast carcinoma who underwent modified radical mastectomy (MRM) with AD followed by adjuvant PMRT to CW with regional nodes, availability of complete dosimetric data and planning CT scans, minimum 12-month clinical follow-up, and pain assessments beginning 3 months post RT. Patients were excluded if they had prior ipsilateral breast/thoracic irradiation or developed locoregional recurrence before pain evaluation. Patients received one of the following HFRT regimens: HFRT-A (26 Gy/5 fractions), HFRT-B (34 Gy/10 fractions), or HFRT-C (40 Gy/15 fractions). All patients were treated with 3-dimensional conformal radiotherapy (3D-CRT) with 6 Megavoltage (MV) photons. Treatment planning was performed using the Monaco® treatment planning system (Elekta AB, Stockholm, Sweden). Dose calculation was carried out using MONTE CARLO algorithm by a medical physicist.

Dosimetric, anatomical, and surgical parameters

Dosimetric parameters were derived from each patient's finalized, clinically approved RT plan. BP was contoured in accordance with the Radiation Therapy Oncology Group (RTOG) Brachial Plexus Atlas, and both the maximum dose (BP D_{max}) and mean dose (BP D_{mean}) were recorded (Figure 1A-E). Mean and maximum doses (D_{mean} and D_{max}) of axillary levels I–III and CW were extracted from structures delineated according

to the RTOG Breast Cancer Contouring Atlas (Figure 2A and Figure 2B). All contouring were performed by a single radiation oncologist to ensure uniformity and were independently reviewed by a senior consultant to maintain accuracy and consistency.

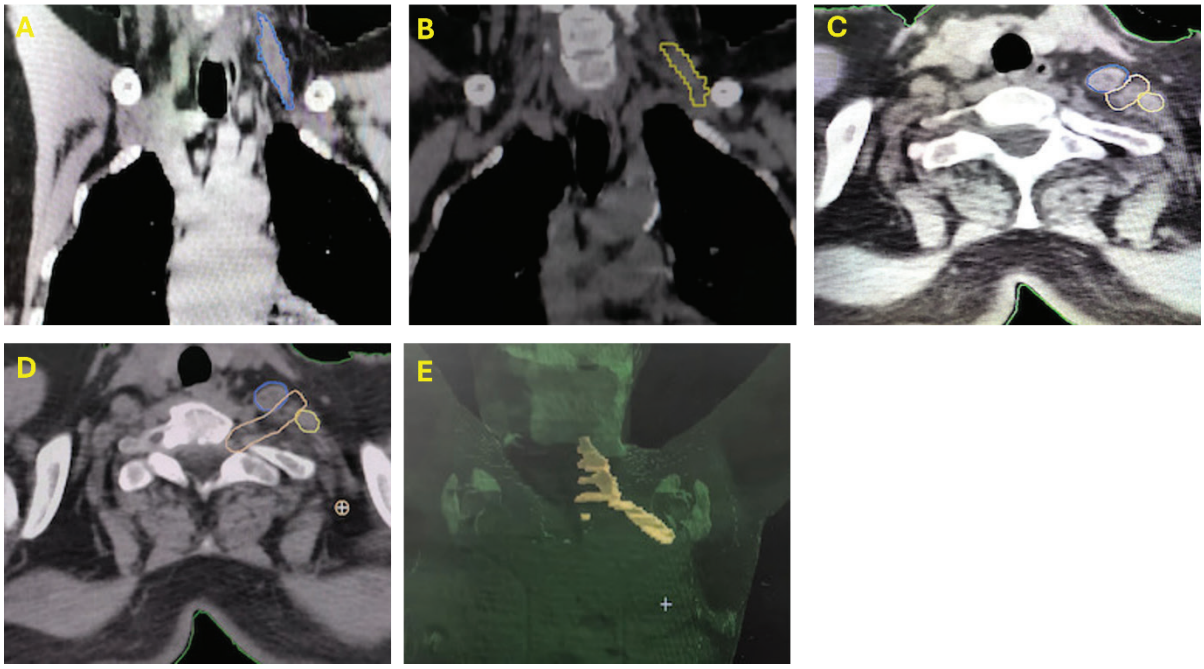


Figure 1. *Brachial plexus contouring per RTOG atlas. A-B: anterior scalene muscle (ASM) in blue color and the middle scalene muscle (MSM) in yellow color were first identified on a coronal view to establish clear anatomical landmarks, C: nerve roots C5-T1 (in dark yellow color) were delineated from each neural foramen into the interscalene space between ASM and MSM, D: when the foramen was not visible, only the scalene fat space was contoured. Below T1, contours followed the plexus posterior to the subclavian-axillary neurovascular bundle to 1-2 slices below the clavicle, E: 3-dimensional view of the contoured BP.*

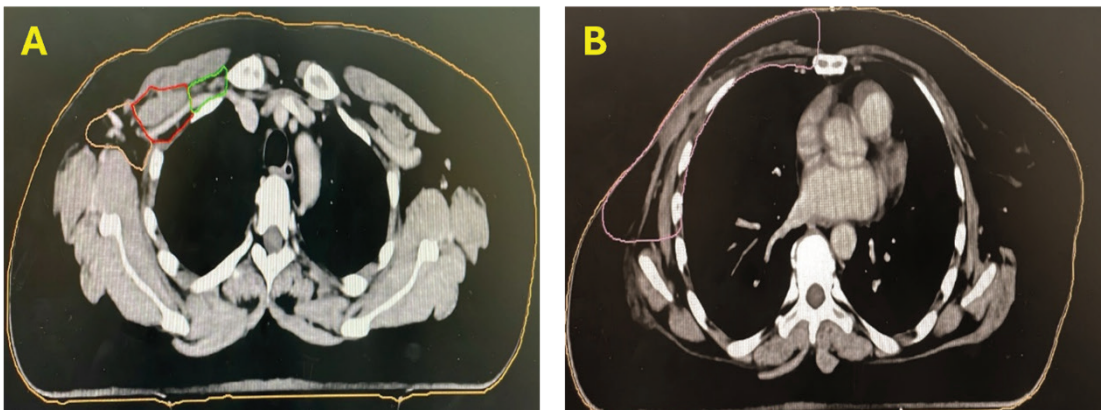


Figure 2. *Contouring of axilla. A: contouring of level I (golden color), level II (red color), and level III (green color), B: contouring of the left-sided chest wall (pink color) according to Breast Cancer RTOG contouring atlas.*

CW thickness was measured from the planning CT dataset at the 2nd and 5th intercostal spaces, including the narrowest and maximum thickness points, mid-axillary line, and mid-clavicular line. To minimize regional variability and obtain a representative value, the average of all recorded measurements was calculated and used for subsequent analysis.

Equivalent dose in 2Gy fractions (EQD2) was calculated for relevant dosimetric parameters using the linear-quadratic model with an α/β ratio of 3Gy for late responding tissues. The maximum dose and mean dose after EQD2 correction were denoted as D_{max} EQD2 and D_{mean} EQD2, respectively, and were used for further assessment.

Surgical factors were derived from intra-operative reports, with the extent of AD categorized according to the highest level removed (up to level II or up to level III) to account for the potential impact of surgical

manipulation on regional neural and soft-tissue integrity. Patient-related clinical characteristics, including age and BMI, were also extracted from medical records. These dosimetric, anatomical, surgical, and clinical parameters were analyzed as potential determinants of NP and SCWP. BP axillary dose parameters and extent of AD were primarily correlated with NP, whereas CW, axillary dose parameters, and CW thickness were evaluated in relation to SCWP.

Outcome assessment

Pain outcomes were assessed using standardized and validated instruments. Somatic shoulder and chest wall pain (SCWP) was evaluated according to CTCAE, version 5.0, and graded on a scale from 0 to 4, with grades ≥2 considered clinically significant. NP was assessed using the LANSS scale, with scores ≥12 indicating clinically significant NP. All pain assessments were conducted by trained personnel who were blinded to the study objectives and protocol to minimize assessment bias. The primary endpoint was the incidence of SCWP and NP. Clinical evaluations were conducted every three months during the first year and every six months thereafter, beginning three months after PMRT, and continued until SCWP reached Grade ≥2 or the LANSS score was ≥12. The correlations between these pains and patient factors (age, BMI), anatomical factors (CW thickness), extent of AD, dosimetric parameters (doses to axillary levels I-III, BP, supraclavicular fossa, and CW) were also analyzed.

Statistical analysis

Descriptive statistics were used to summarize patient, tumor, and treatment characteristics. Categorical variables were compared using the Chi-square or Fisher's exact test. Continuous variables were analyzed with One-way ANOVA, with post-hoc testing where appropriate. Time-to-event outcomes were estimated using the Kaplan-Meier (KM) method, and survival curves were compared using the log-rank test. Factors associated with pain outcomes were evaluated using Cox proportional hazards regression analysis to estimate hazard ratios with 95% confidence intervals. Variables demonstrating $p<0.10$ in univariable testing were included in multivariable models. A two-sided $p<0.05$ was considered statistically significant.

Results

A total of 90 patients met the eligibility criteria and were included in the final analysis. The median age at diagnosis was 54 years (range: 37-80) with a mean BMI of 22.59±2.3 kg/m^2. The median duration of follow-up was 24 months (range,3-36 months). All patients had previously undergone MRM with AD, followed by HFRT to the CW, supraclavicular fossa (SCF), and axillary levels I–III.

Most patients received HFRT B (38.8%), followed by HFRT-A (33.3%) and HFRT-C (27.7%). A 3D-printed tissue equivalent bolus of thickness 3-5 mm was used in patients with CW thickness <10mm (N=12/90), as per institutional protocol to ensure adequate coverage of planning target volume (PTV), given the potential for underdosage in thin CW patients. Bolus was used every other day in HFRT-B and HFRT-C, and daily in HFRT-A. Baseline demographic, anatomical, surgical, and treatment-related characteristics are summarized in Table 1.

Table 1. Patient, tumor, and surgical characteristics.

Characteristics	N
Age (years)	
<50	47
≥50years	43
Stage	
IIB	22
IIIA	33
IIIB	10
IIIC	25
Axillary dissection	
Level I and II	51
Level I, II, and III	39
Chemotherapy	
Yes	65
No	25
Hormonal therapy	
Tamoxifen	35
Letrozole	45
None	10
Radiotherapy	
34 Gy in 10 fractions	35
26 Gy in 5 fractions	30
40 Gy in 15 fractions	25
Histology	
Invasive ductal carcinoma	50
Infiltrating ductal carcinoma	36
Mucinous carcinoma	4
Laterality	
Right	51
Left	39
BMI (kg/m^2)	
<18	35
≥18	55

Somatic shoulder and chest wall pain (SCWP) outcomes

The overall incidence of Grade ≥2 SCWP was 30% (N=27). The incidence varied across HFRT regimens, with the highest frequency observed in HFRT-C (48%), followed by HFRT-A (23.3%), and HFRT-B (22.8%). However, the difference did not reach statistical significance ($p=0.09$).

KM analysis demonstrated a temporal pattern of SCWP occurrence across the three HFRT regimens, with

no statistically significant difference (log rank p=0.098) (Figure 3A). Median onset of SCWP was reached only in the HFRT C at approximately 18 months, while it was not reached in the HFRT B or HFRT C. The mean duration of SCWP was 6 months (range, 3-24 months).

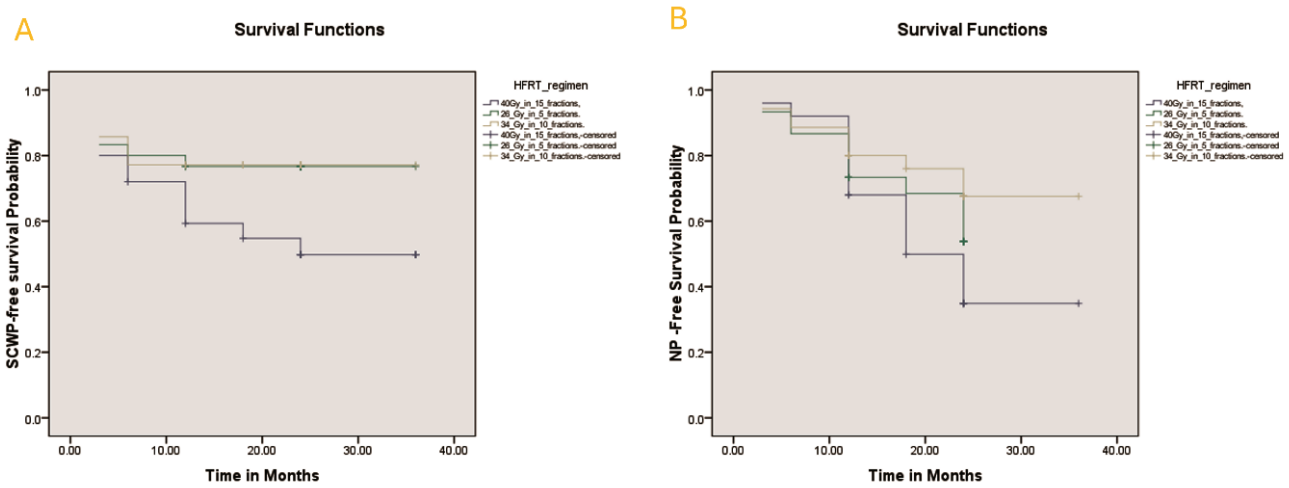


Figure 3. *KM curves. A: SCWP-free survival, B: NP-free survival according to the HFRT regimen. The plot compares Survival outcomes for patients treated with different HFRT regimens:40 Gy in 15 fractions (blue), 26 Gy in 5 fractions (green), and 34 Gy in 10 fractions (golden). ⁺censored observations.*

Patients developing Grade ≥2 SCWP had significantly thinner CW thickness and higher CW D_{max} EQD2.On Cox regression multivariate analysis, CW thickness (HR = 0.895; p = 0.04) and CW D_{max} EQD2(HR 1.45;0.04) were significantly associated with SCWP. No significant correlations were identified between SCWP and BP doses, BMI, age, level of AD, or dosimetry to axillary nodal levels (Table 2).

Table 2. Multivariate Cox regression analysis showing independent predictors of NP following PMRT.

Variable	SCWP (N=27)	NO SCWP (N=63)	HR (95% CI)	*p* value
CW thickness	18.3±8.9 mm	30.15±6.7 mm	0.895 (0.709-0.956	0.04
Mean CW D_{max}EQD2	55.3±1.63 Gy	48.23±1.89	1.456 (1.23-1.567)	0.04
Mean CW D_{mean} EQD2	44.98±2.3 Gy	43.4±1.65	1.12 (0.89-1.23)	0.5
AD level	Level III	Level II	1.01 (0.66-1.4)	0.79
Axillary level II D_{mean} EQD2	49.8±1.45	48.9±1.76	1.02 (0.89-1.2)	0.1
Axillary level III D_{mean} EQD2	51.2±1.89 Gy	50.5±1.3 Gy	1.015 (0.904-1.14)	0.089

Note: HR: hazard ratios, CI: confidence interval, EQD2: equivalent dose in 2 Gy fractions, D_{max}: maximum dose, D_{mean}: mean dose, AD: axillary dissection, SCWP: somatic chest wall and shoulder pain, CW: chest wall.

Neuropathic pain (NP) outcomes

The overall incidence of NP was 41.1% (N=37) with a significantly higher incidence observed in the HFRT C (60%) compared with HFRT A (40%) and HFRT B (28.57%) (p=0.001). KM analysis for time to NP onset demonstrated significant differences among the three HFRT (log rank p=0.01) (Figure 3B). Median onset of NP occurred exclusively in the HFRT C at 18 months. In contrast, HFRT A and HFRT B did not reach a median onset within the follow-up period. The mean duration of onset of NP was 12 months (range, 3-24 months).

NP was significantly associated with higher LANSS scores, AD up to level III, higher BP D_{max}EQD2, and higher axillary level III D_{mean} EQD2. Supraclavicular, chest wall doses, as well as clinical factors including age and BMI, were not significantly associated with NP (Table 3).

Table 3. Multivariate Cox regression analysis showing independent predictors of NP following PMRT.

Variable	NP (N=37)	No NP (N=53)	HR (95% CI)	*p* value
LANSS score	16.37±4.74	7.44±3.76	1.1 (1.08-1.23)	0.002
BP D_{max}EQD	46.85±2.4 Gy	43.26±2.14 Gy	1.08 (1.02-1.15)	0.001
BP D_{mean}EQD2	21.3±4.35 Gy	20.15±3.25 Gy	1.06 (0.75-1.56)	0.2
AD level	Level III	Level II	1.5 (1.3-1.67)	0.049
Axillary level II D_{mean} EQD2	50.5±1.23 Gy	45.03±2.33	1.03 (0.9-1.08)	0.05
Axillary level III D_{mean} EQD2	54.1±1.34 Gy	45.05±2.06	1.262 (1.109-1.3)	0.03
SCF D_{mean} EQD2	51.4Gy±1.5 Gy	47.3±4.8	1.1 (0.7-1.2)	0.056

Note: HR: hazard ratios, CI: confidence interval, EQD2: equivalent dose in 2 Gy fractions, D_{max}: maximum dose, D_{mean}: mean dose, BP: brachial

Discussion

PMRT-related pain remains an under-recognized but clinically significant toxicity that adversely affects long-term quality of life in breast cancer survivors.[13] The present study demonstrates that PMRT pain is not a homogeneous entity; rather, it represents a composite syndrome comprising both somatic and neuropathic elements.

The pattern of PMRT pain observed aligns closely with previously documented toxicity profiles following HFRT, where 10-30% of patients may develop moderate chest wall discomfort within the first year of treatment.[14,15] SCWP occurred early after RT, with most cases appearing within the first 3-6 months and gradually decreasing over time. This suggests that SCWP is likely related to acute or subacute RT-induced inflammation and tissue injury, which tends to improve as healing occurs. In contrast, NP developed later, most commonly at 12 months or beyond, and persisted longer. The delayed onset of NP indicates progressive radiation-related nerve damage or fibrosis, which evolves rather than immediately after treatment.[16-18] A pronounced inverse relationship between CW thickness and SCWP severity echoes prior analyses demonstrating that thinner CW is associated with higher effective dose deposition in superficial nociceptive structures (ribs and intercostal nerves). Dosimetric evidence indicates that patients with a “thin” CW are at particularly high risk for RICWP and rib-periosteum toxicity. In phantom and clinical-planning studies, thin CW geometry (≤15 mm) has been shown to alter dose deposition significantly. Furthermore, emerging evidence from advanced dosimetric and computational modelling demonstrates that the use of partial bolus techniques, combined with high-precision dose-calculation algorithms such as Monte Carlo, substantially improves dose accuracy in patients with thin CW and mitigates excessive dose deposition to the skin and underlying ribs, thereby potentially reducing toxicity in this high-risk subgroup.[19] This dosimetric effect has been reported in severe planning studies, demonstrating that thin CW anatomy contributes to elevated D_{max} values, higher rib doses, and increased heterogeneity within high dose regions. Clinically, it may manifest as a greater incidence of rib tenderness, costochondral pain, and CW discomfort.

SCWP was correlated strongly with higher CW D_{max} increased toxicity was observed in the 50-55 Gy equivalent dose range. The observed association between elevated CW D_{max} EQD2 and clinically significant SCWP is consistent with established dose-response relationships, which demonstrate a heightened risk of CW tenderness, fibrosis, and rib pain in areas exposed to high-dose CW hotspots. [20] In contrast, BP dosimetry showed no association with SCWP, consistent with the distinct anatomic pathways governing somatic CW sensation.

NP was most frequently observed in patients treated with 40 Gy in 15 fractions, followed by those treated with 26 Gy in 5 fractions, reflecting the well-established sensitivity of neural structures to higher biologically effective doses (BED). This aligns with the modelling study by Lee et al., which demonstrated that BP tolerance is strongly dependent on BED; a BP D_{max} of approximately 26 Gy delivered in 3-4 fractions (≈EQD2 ≈45-50 Gy for α/β=3 Gy) already produced a ~10% risk of plexopathy, with the probability rising steeply at higher BED levels-consistent with the dose-dependent neuropathic effects observed in current study. This is further supported by evidence showing that BP injury rates vary from less than 1 % to over 70 % as fraction sizes increase to 2.2-4.58 Gy (total doses 43.5-60 Gy), highlighting the remarkable sensitivity of the plexus to dose-per-fraction escalation.[21-23] Although BP D_{max} EQD2 values in the present study remained within conventionally accepted tolerance limits, the observed association with neuropathic pain suggests the presence of subclinical neural injury and/or cumulative multimodal effects rather than overt radiation-induced plexopathy. Neural tissue exhibits a spectrum of radiation responses, ranging from reversible functional disturbances to irreversible structural damage. At lower dose levels, radiation may induce microvascular endothelial injury, neuroinflammation, Schwann cell dysfunction, and altered axonal conduction, which can manifest clinically as NP without meeting radiographic

or electrophysiologic criteria for classical plexopathy. Importantly, the pattern of NP in our cohort suggests contributions from not only direct BP irradiation but also surgical and regional irradiation factors, particularly given that BP doses in this study remained well below established tolerance constraints. This interpretation is strongly reinforced by the findings of Lundstedt et al., who identified a clear dose-response relationship, with paraesthesia rates rising sharply only when the BP D_{max} exceeded approximately 55 Gy. Their results compellingly suggest that symptoms observed at lower dose levels are unlikely to be radiation-induced and are instead more plausibly explained by alternative mechanisms, most notably ICBN injury associated with extensive AD.[24-27] Apart from this, surgical manipulation can cause traction-related plexus stress and postoperative fibrosis, creating a vulnerable neural substrate. Subsequent regional nodal irradiation may then act as a biological sensitizer, amplifying inflammatory and fibrotic processes and lowering the effective tolerance of neural structures. This concept of reduced neural reserve following surgery provides a plausible explanation for the significant correlation between BP D_{max} EQD2 and LANSS scores observed in our cohort, despite doses remaining below traditional thresholds for overt plexopathy. In the present study, a borderline yet consistent positive association emerged between axillary nodal D_{mean} EQD2 and NP This pattern is biologically credible and strongly suggests a contributory role of non-BP neural structures within the axilla, highlighting a potential dose-sensitive mechanism beyond classical brachial plexopathy. Previous evidence has demonstrated that regional nodal irradiation, particularly when incorporating a posterior axillary boost (PAB), can precipitate late-onset neurotoxicity and brachial plexopathy. This supports the hypothesis that axillary irradiation may injure small sensory nerves, with effects potentially amplified by prior surgical trauma.[28,29] NP was more pronounced in patients undergoing level III AD, consistent with analysis, which showed that a greater number of removed lymph nodes markedly increases susceptibility to BP neuropathy, highlighting the synergistic impact of surgical disruption and regional nodal irradiation on neural injury risk.[30] Collectively, the evidence underscores that NP following PMRT is a multifactorial phenomenon driven by a combination of surgical factors, particularly extensive AD and the cumulative radiation dose delivered not only to BP but also to axillary nodal regions, underlining the importance of meticulous optimization of regional dose distribution in HFRT protocols.

While the current follow-up period captured most early and subacute toxicities, radiation-induced neuropathic complications are known to have a delayed onset, sometimes several years after treatment. Hence, extended follow-up would be necessary to fully characterize the long-term incidence of NP across HFRT regimens. Overall, our findings highlight the importance of optimized dose distribution, hotspot control, and adherence to organ at risk constraints during HFRT planning to minimize treatment-related morbidity while maintaining optimal therapeutic outcomes. Furthermore, with respect to SCWP, validation in larger, adequately powered cohorts is essential to derive statistically robust and clinically generalizable conclusions.

Limitations

As a retrospective, single-institution study, the findings may have limited generalizability to centers employing different radiotherapy planning techniques, dose calculation algorithms, contouring practices, and patient populations. Nevertheless, the homogeneity of treatment delivery and follow-up within a single institution allows for a consistent evaluation of dosimetric-toxicity relationships. This study did not evaluate individual rib dose parameters and ICBN involvement, which may further contribute to PMRT pain. The limited sample size and further subdivision into HFRT regimens may have reduced statistical power and restricted generalizability. Prospective studies with detailed nerve contouring and larger cohorts are needed

Conclusion

PMRT-related pain following HFRT is multifactorial, with somatic symptoms associated with higher CW dose and neuropathic symptoms influenced by surgical extent and regional noal irradiation. Given the larger equivalent doses delivered to late-responding tissues with hypofractionation, meticulous optimization of CW hotspots, regional nodal coverage, and BP dose during HFRT planning is recommended. A longer follow-up is warranted to better characterize the evolution of late pain outcomes and validate these dosimetric associations

Ethical approval

As this was a retrospective analysis of existing clinical data with no direct patient intervention, formal ethical committee approval was not obtained

Funding

None

Conflict of interest

None

CRediT authorship contribution statement

Pratibha Prashar: conceptualization, study design, data curation, formal analysis, writing: original draft; **Brish Bhanu Bhardwaj**: methodology, statistical analysis, writing: review and edit; **Vandana Thakur**: data collection, data curation; **Poorva Vias**: patient follow-up, data verification; **Hardik Sharma**: writing:

review and edit; **Kulbir Singh:** supervision, data curation; **Ankita Pandey**: writing: review and edit; **Vivek Kumar**: data curation.

Acknowledgements

The authors would like to thank the participants of the study. No specific external funding or technical assistance was received for this research.

References

[1] Andersen KG, Kehlet H. Persistent pain after breast cancer treatment: a critical review of risk factors and strategies for prevention. J Pain. 201; 12(7): 725-46. doi: 10.1016/j.jpain.2010.12.005.

[2] Smith HS, Wu SX. Persistent pain after breast cancer treatment. Ann Palliat Med. 2012; 1(3): 182-94. doi: 10.3978/j.issn.2224-5820.2012.10.13.

[3] Bennett MI, Rayment C, Hjermstad M, Aass N, Caraceni A, Kaasa S. Prevalence and aetiology of neuropathic pain in cancer patients: A systematic review. PAIN. 2012; 153(2): 359-65. doi: 10.1016/j.pain.2011.10.028.

[4] Voruganti IS, Donovan E, Walker-Dilks C, Swaminath A. Chest wall toxicity after stereotactic radiation in early lung cancer: a systematic review. Curr Oncol. 2020; 27(4): 179-89. doi: 10.3747/co.27.5959.

[5] Wang L, Guyatt GH, Kennedy SA, Romerosa B, Kwon HY, Kaushal A, et al. Predictors of persistent pain after breast cancer surgery: a systematic review and meta-analysis of observational studies. CMAJ. 2016; 188(14): E352-61. doi: 10.1503/cmaj.151276.

[6] Fijardo M, Kwan JYY, Bissey PA, Citrin DE, Yip KW, Liu FF. The clinical manifestations and molecular pathogenesis of radiation fibrosis. eBioMedicine. 2024; 103: 105089. doi: 10.1016/j.ebiom.2024.105089.

[7] Kim JS, Spiess AM. Surgical treatment of intercostal brachial nerve pain after mastectomy and axillary dissection. Plast Reconstr Surg Glob Open. 2021; 9(11): e3935. doi: 10.1097/GsOX.0000000000003935.

[8] Yan M, Kong W, Kerr A, Brundage M. The radiation dose tolerance of the brachial plexus: a systematic review and meta-analysis. Clin Transl Radiat Oncol. 2019; 18: 23-31. doi: 10.1016/j.ctro.2019.06.006.

[9] Andolino DL, Forquer JA, Henderson MA, Barriger RB, Shapiro RH, Brabham JG, et al. Maximum dose to chest wall greater than 60 Gy and volume of chest wall receiving greater than 30 Gy predict risk of chest wall toxicity following stereotactic body radiation therapy (SBRT). Int J Radiat Oncol Biol Phys. 2009; 75(3): S60. doi: 10.1016/j.ijrobp.2009.07.156.

[10] Calapai M, Esposito E, Puzzo L, Vecchio DA, Blandino R, Bova G, et al. Post-mastectomy pain: An updated overview on risk factors, predictors, and markers. Life. 2021; 11(10): 1026. doi: 10.3390/life11101026.

[11] Sun KY, Hall WH, Mathai M, Dublin AB, Gupta V, Purdy JA, et al. Validating the RTOG-endorsed brachial plexus contouring atlas: An evaluation of reproducibility among patients treated by intensity-modulated radiotherapy for head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2012; 82(3): 1060-4. doi: 10.1016/j.ijrobp.2010.10.035.

[12] Vargo JA, Beriwal S. RTOG chest wall contouring guidelines for post-mastectomy radiation therapy: is it evidence-based?. Int J Radiat Oncol Biol Phys. 2015; 93(2): 266. doi: 10.1016/j.ijrobp.2015.03.001.

[13] Kakati B, Nair N, Chatterjee A. Post mastectomy pain syndrome at an Indian tertiary cancer centre and its impact on quality of life. Indian J Cancer. 2023; 60(2): 275-81. doi: 10.4103/ijc.ijc_861_21.

[14] Yadav BS, Bansal A, Kuttikat PG, Das D, Gupta A, Dahiya D. Late-term effects of hypofractionated chest wall and regional nodal radiotherapy with two-dimensional technique in patients with breast cancer. Radiat Oncol J. 2020; 38(2): 109-18. doi: 10.3857/roj.2020.00129.

[15] Haviland JS, Bentzen SM, Bliss JM, Yarnold JR. Prolongation of overall treatment time as a cause of treatment failure in early breast cancer: An analysis of the UK START (standardisation of breast radiotherapy) trials of radiotherapy fractionation. Radiother Oncol. 2016; 121(3): 420-3. doi: 10.1016/j.radonc.2016.08.027.

[16] Delanian S, Lefaix JL, Pradat PF. Radiation-induced neuropathy in cancer survivors. Radiother Oncol. 2012; 105(3): 273-82. doi: 10.1016/j.radonc.2012.10.012.

[17] Kang KH, Okoye CC, Patel RB, Siva S, Biswas T, Ellis RJ, Yao M, Machtay M, Lo SS. Complications from stereotactic body radiotherapy for lung cancer. Cancers. 2015; 7(2): 98-1004. doi.org/ 10.3390/cancers7020820.

[18] Chang, P., Wu, S. & Emos, M.R. Identification, Evaluation, and Management of Post-breast Surgery Pain Syndrome. Curr Phys Med Rehabil Rep. *2024;* 12, 161-9. doi: 10.1007/s40141-024-00438-6.

[19] Lundstedt D, Lindberg A, Gustafsson M, Chakarova R, Karlsson P. Adjuvant radiation treatment of breast cancer after mastectomy: Advanced algorithms and partial bolus improve the dose calculation accuracy in the case of thin-chest-wall irradiation. Adv Radiat Oncol. 2023; 8(5): 101223. doi: 10.1016/j.adro.2023.101223.

[20] Kim D, Kim K, Kim JS, Kang S, Park JM, Shin KH. Near-maximum rib dose is the most relevant risk factor for ipsilateral spontaneous rib fracture: a dosimetric analysis of breast cancer patients after radiotherapy. Strahlenther Onkol. 2023; 199(1): 38-47. doi: 10.1007/s00066-022-01972-9.

[21] Kapitanova I, Biswas S, Divekar S, Kemmerer

EJ, Rostock RA, Forster KM, et al. Estimating the tolerance of brachial plexus to hypofractionated stereotactic body radiotherapy: a modelling-based approach from clinical experience. Radiat Oncol. 2021; 16(1): 98. doi: 10.1186/s13014-021-01822-5.

[22] Gałecki J, Hicer-Grzenkowicz J, Grudzień-Kowalska M, Michalska T, Załucki W. Radiation-induced brachial plexopathy and hypofractionated regimens in adjuvant irradiation of patients with breast cancer-a review. Acta Oncol. 2006; 45(3): 280-4. doi: 10.1080/02841860500371907.

[23] Olsen NK, Pfeiffer P, Johannsen L, Schrøder H, Rose C. Radiation-induced brachial plexopathy: neurological follow-up in 161 recurrence-free breast cancer patients. Int J Radiat Oncol Biol Phys. 1993; 26(1): 43-9. doi: 10.1016/0360-3016(93)90171-Q.

[24] Lundstedt D, Gustafsson M, Steineck G, Sundberg A, Wilderäng U, Holmberg E, et al. Radiation therapy to the plexus brachialis in breast cancer patients: Analysis of paresthesia in relation to dose and volume. Int J Radiat Oncol Biol Phys. 2015; 92(2): 277-83. doi: 10.1016/j.ijrobp.2015.01.016.

[25] Vecht CJ, Van de Brand HJ, Wajer OJ. Post-axillary dissection pain in breast cancer due to a lesion of the intercostobrachial nerve. Pain. 1989; 38(2): 171-6. doi: 10.1016/0304-3959(89)90235-2.

[26] Cui L, Fan P, Qiu C, Hong Y. Single institution analysis of incidence and risk factors for post-mastectomy pain syndrome. Sci. Rep. 2018; 8(1): 11494. doi: 10.1038/s41598-018-29946-x.

[27] Kaur N, Kumar R, Jain A, Saxena AK. Sensory changes and postmastectomy pain following preservation of intercostobrachial nerve in breast cancer surgery: a prospective randomized study. Indian J Surg Oncol. 2021; 12(1): 108-13. doi: 10.1007/s13193-020-01193-5.

[28] Goyal S, Menon D, Puzhakkal N, Makuny D. Brachial plexus doses in locoregional radiotherapy for breast cancer. TRO. 2019; 3. doi: 10.21037/tro.2019.08.03.

[29] Rudra S, Roy A, Brenneman R, Gabani P, Roach MC, Ochoa L, et al. Radiation-induced brachial plexopathy in patients with breast cancer treated with comprehensive adjuvant radiation Therapy. Adv Radiat Oncol. 2020; 6(1): 100602. doi: 10.1016/j.adro.2020.10.015.

[30] Wu SG, Huang SJ, Zhou J, Sun JY, Guo H, Li FY, et al. Dosimetric analysis of the brachial plexus among patients with breast cancer treated with post-mastectomy radiotherapy to the ipsilateral supraclavicular area: report of 3 cases of radiation-induced brachial plexus neuropathy. Radiat Oncol. 2014; 9: 292. doi: 10.1186/s13014-014-0292-5.

Journal of Associated Medical Sciences 2026; 59(2): 316-326

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Development of a YOLOv8-based deep learning framework for detection and classification of periapical radiolucencies in dental radiographs using the periapical index

Natdanai Hirata[1], Wannakamon Panyarak[2], Kittipit Klanliang[3], Arnon Charuakkra[2], Kittichai Wantanajittikul[1*]

[1]Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Division of Oral and Maxillofacial Radiology, Department of Oral Biology and Diagnostic Sciences, Faculty of Dentistry, Chiang Mai University, Chiang Mai Province, Thailand.
[3]Division of Endodontics, Department of Restorative Dentistry and Periodontology, Faculty of Dentistry, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 3 February 2025
Accepted as revised 18 March 2026
Available online 23 March 2026

Keywords:
Apical periodontitis, periapical radiograph, periapical index, deep learning, YOLO.

ABSTRACT

Background: Apical periodontitis (AP) is one of the most common oral diseases worldwide. Radiographs are widely used for diagnosing AP. However, their interpretation is subjective and time-consuming, contributing to increased clinical workload. Although deep learning has been introduced to assist diagnosis, most existing models are trained on limited sets of teeth, restricting their generalizability.

Objective: To investigate the potential of the YOLOv8 model for detecting and classifying periapical radiolucencies in periapical radiographs using the PAI criteria.

Materials and methods: A total of 1,191 periapical radiographs from both mandibular and maxillary jaws were collected and scored. Experts annotated 5,498 bounding boxes covering periapical root areas (PRAs). YOLOv8 models were developed to detect PRAs and classify PAI scores, which were grouped into three classes: PAI 1, PAI 2, and PAI 3-5. A custom post-processing method was implemented to improve inference by filtering overlapping detections. Filtered predictions from the test dataset were evaluated for both post-processing methods using precision, recall, F1-score, and mean average precision at an IoU threshold of 0.5 (mAP50). Confusion matrices were used to analyze misclassification patterns, while average IoU assessed overall detection performance.

Results: The best YOLOv8 model achieved a mean average precision at an intersection over union (IoU) threshold of 0.5 (mAP50) of 0.59 with a global average IoU of 0.62. A binary classification ablation model, grouping PAI 1-2 as “healthy” and PAI 3–5 as “diseased,” achieved an improved mAP50 of 0.66. The custom post-processing algorithm successfully eliminated all overlapping boxes and enhanced model performance.

Conclusions: Despite moderate performance, the YOLOv8-based models demonstrate potential as adjunct tools to assist dental practitioners in radiographic evaluation, particularly for less experienced clinicians. Further development is required to improve clinical applicability and robustness.

* **Corresponding contributor.**
Author’s Address: *Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: kittichai.wan@cmu.ac.th
doi: 10.66285/JAMS.2026.063
E-ISSN: 2539-6056

Introduction

Untreated dental health problems can lead to irreversible damage and negatively affect the quality of life of individuals.[1,2] Apical periodontitis (AP) is an infection of the root canals and surrounding tissues which progresses from caries.[3,4] A periapical radiograph (Pa) is an intraoral radiograph that provides two-dimensional high-resolution detail of tooth roots and surrounding structures. It is commonly used to assess the periapical health of one or several specific teeth. Evaluating or interpreting Pa usually takes less time compared to the other examinations. Additionally, it is non-invasive and widely available.[5] Nonetheless, clinical evaluation is essential for diagnosing periapical lesions, as relying solely on conventional radiography may lead to unnecessary surgical procedures.[6,7] Currently, there is increasing interest in developing radiograph-based diagnostic methods to streamline workflows and reduce dependence on human interpretation. Decades ago, the Periapical Index (PAI) scoring system was introduced as an early attempt to standardize radiographic assessment of periapical conditions.[8] It is one of the most widely used tools for assessing AP on radiographs. It utilizes a 5-point scale, ranging from normal (score 1) to severe AP (score 5). The system was developed based on correlations between radiographic features and histological findings, which are considered gold standard. While the PAI provided a valuable framework for consistent interpretation, it still heavily relied on expert visual assessment. Today, advances in artificial intelligence (AI) offer the potential to automate and enhance such radiographic evaluations,[9,10] building on the foundational goals established by tools like PAI.

Deep learning has been applied in dental imaging to overcome challenges in radiography. The models have shown strong performance in detection, classification, and segmentation, with potential to support dentists in diagnosis.[11-14] Sadr *et al*.[15] conducted a meta-analysis and reported high accuracy of deep learning models for detecting periapical lesions in dental radiographs despite some risk of bias across studies. Berne *et al*. used You Only Look Once (YOLO) version 3 to develop an automated lesion detection system on panoramic images with reliable clinical performance.[16] Moidu *et al*. developed a system that detects and classifies periapical lesions based on PAI scoring in mandibular Pa also using YOLO version 3.[17] The model achieved good performance in classifying all PAI scores. However, it struggled to differentiate the score with early-stage changes from the adjacent scores. Furthermore, their dataset consisted only of mandibular PAs, as they stated that excluding maxillary Pas reduced the complexity caused by the superimposition of maxillary anatomical structures.

YOLO version 8 (YOLOv8) was the latest and most stable version at the time of model development.[18] YOLOv8 provides improved detection accuracy and fast inference speed, making it suitable for real-time clinical applications.[19,20] In addition, its anchor-free detection architecture and enhanced feature extraction enable effective detection of small objects,[21] which is beneficial for identifying lesions in dental radiographs. Although newer versions have since been released, YOLOv8 remains a strong benchmark and has been successfully applied in dental imaging with promising results. Previously, we investigated the performance of deep learning models in classifying cropped tooth root images.[22] The models showed potential in distinguishing normal roots from diseased roots. However, to fully aid dental practitioners with the screening and diagnosing processes, automatic detection of the roots should be provided.

Therefore, the aim of this study was to develop an automatic detection and classification system for periapical root areas (PRAs) in Pa using YOLOv8, with the potential to assist dental practitioners in making diagnoses or evaluations based on radiographs alone, reducing reliance on clinical examinations or histological results, and ultimately enhancing the efficiency of the radiograph interpretation process.

Materials and methods

Data acquisition and inclusion criteria

Ethical approval for this retrospective study was obtained from the Institutional Ethical Review Board, Faculty of Dentistry, Chiang Mai University, Chiang Mai, Thailand. Digital Pas were obtained from the faculty's image database. Radiographic data were acquired from Heliodent Plus intraoral X-ray units (Dentsply Sirona, Charlotte, NC, USA) utilizing either Carestream 7600 or VistaScan Imaging Plates (DÜRR DENTAL, Bietigheim-Bissingen, Germany). The periapical radiographs were fully anonymized and subsequently viewed on a medical-grade display, cropped, and exported portable network graphic (PNG) files, with selected images having a minimum resolution of 800 by 1000 pixels. According to the criteria outlined by Bachani *et al*.,[23] radiographs were included if they demonstrated ideal imaging characteristics and contained at least one tooth showing evidence of pulp infection, a large filling, or a carious lesion approaching the pulp. Images with inadequate quality or artifacts that hindered assessment of the lamina dura were excluded. A total of 1,191 Pas were included from 649 patients (57.00% female, 43.00% male, aged 13-87). Both mandibular and maxillary Pas were used, with mandibular images comprising 76.00% of the total. An overview of the study workflow is presented in Figure 1.

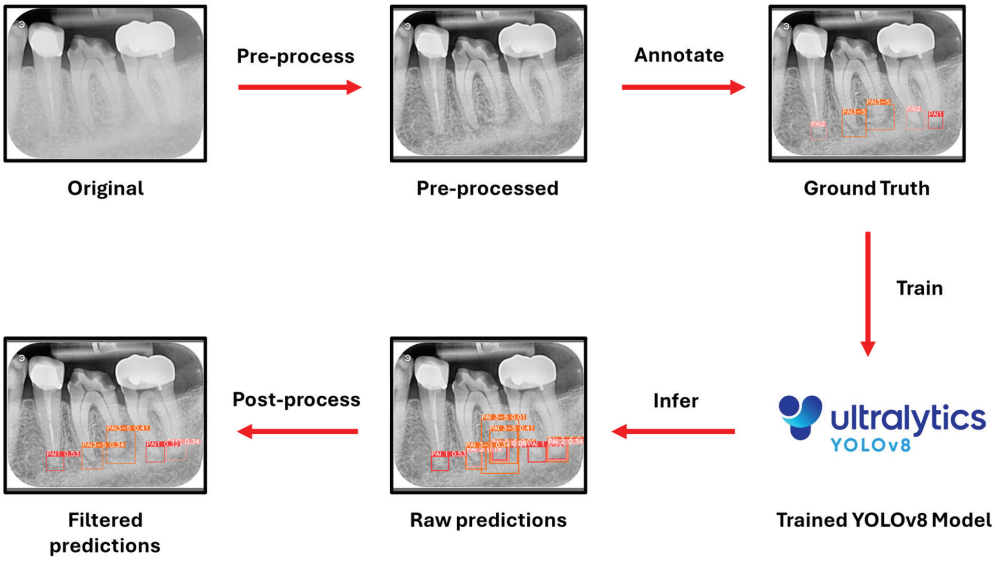


Figure 1. *A schematic illustration of the overall workflow used in this study.*

Image annotation and dataset splitting

The annotation process was conducted jointly by three experts-two oral and maxillofacial radiologists and one endodontist-with each having more than 10 years of clinical experience. The Computer Vision Annotation Tool (CVAT; available at www.cvat.ai) was employed to perform image labeling. A total of 5,498 bounding boxes were created, each covering a PRA. Every PRA was labeled with a corresponding class. PAI scores were divided into three classes for analysis: PAI 1, PAI 2, and PAI 3-5. Examples of PRAs from each class, along with the classification criteria, are shown in Figure 2. The test dataset consisted of 146 Pas, representing the most recent images collected from August through November 2024. The remaining images, 1045 Pas, taken between December 2011 and July 2024, were split into training and validation datasets using an 80:20 ratio (training=836 images, validation=209 images) prior to model training to ensure that no images were shared between the datasets. This ratio was selected as it is commonly used in machine learning studies to provide sufficient data for model training while maintaining a validation set for performance monitoring and hyperparameter tuning, while also ensuring that minority classes retain enough samples for meaningful evaluation. A stratified splitting strategy was used to divide the datasets while preserving the original class distribution to mitigate class imbalance. Class imbalance may affect model performance, as classes with more samples provide more training examples, while underrepresented classes may result in lower detection accuracy and recall. An overview of the class distribution across the training, validation, and test datasets is provided in Table 1, offering insights into the balance and representation of each class used in this study.

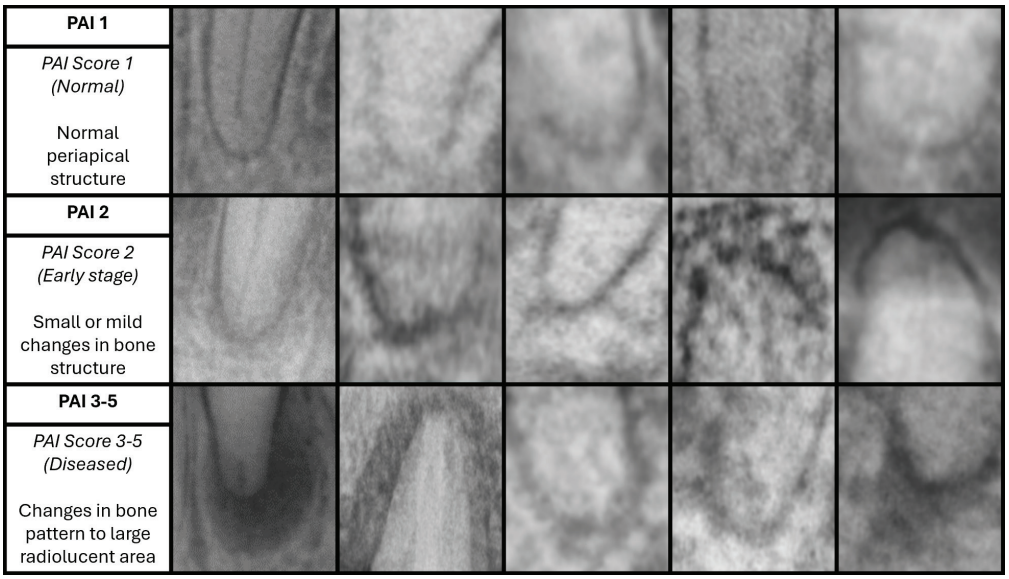


Figure 2. *The classification scheme used in this study, including the PAI scoring criteria for each class and representative example PRAs.*

Table 1. Distribution of dataset instances across classes.

Categories	Training set	Validation set	Test set	Total instances
PAI 1	2,740	716	452	3,908
PAI 2	456	108	84	648
PAI 3-5	622	161	159	942

Image processing and augmentation techniques

Contrast-limited adaptive histogram equalization (CLAHE) and bilateral filtering (BF) techniques were used on all images to normalize images acquired using two different scanning systems.[24-29] Prior to training, YOLOv8 also preprocesses input images by first resizing and padding them to a square of the designated size. The pixel values are then normalized to a range of 0-1, followed by the application of the image augmentation techniques. Image augmentation was applied to the training dataset to increase its variability. The augmentation methods included scaling, vertical flipping, horizontal flipping, and rotation, with respective parameters of 0.30, 0.50, 0.50, and 30°. Additionally, YOLOv8's mosaic augmentation was employed with a probability of 0.50, which combines four images into a single image to improve the model's ability to detect and classify infrequent objects. These techniques also help improve the model's exposure to variations of the minority class.[30]

Model development

Kaggle (available at www.kaggle.com) and Google Colab (available at www.colab.reseach.google.com) online coding platforms were used for model training and development with Python programming environments. We chose the YOLOv8 model architecture with no pre-trained weights. Hyperparameters were chosen using hyperparameter sweeping algorithms. The following training hyperparameters and options were used: a YOLOv8-m model, a square input image size of 1000 pixels, a batch size of 16, a dropout rate of 0.2, the AdamW optimizer with a learning rate of 0.001429, a momentum of 0.9, and a weight decay of 0.0005. The model was trained to provide a box for each PRA detected with its class and confidence score, which represents the model's estimated probability that an object is present within a predicted bounding box.

Post-processing algorithms

YOLOv8 uses Non-Maximum Suppression (NMS) to remove low-confidence and overlapping predictions based on confidence and intersection-over-union (IoU) thresholds. However, some overlapping predictions may remain. To address this, we developed a custom post-processing algorithm that analyzes bounding box coordinates to detect containment and prioritizes the more severe class when overlapping boxes have different labels. The conditions of the custom post-processing algorithm are illustrated in Figure 3.

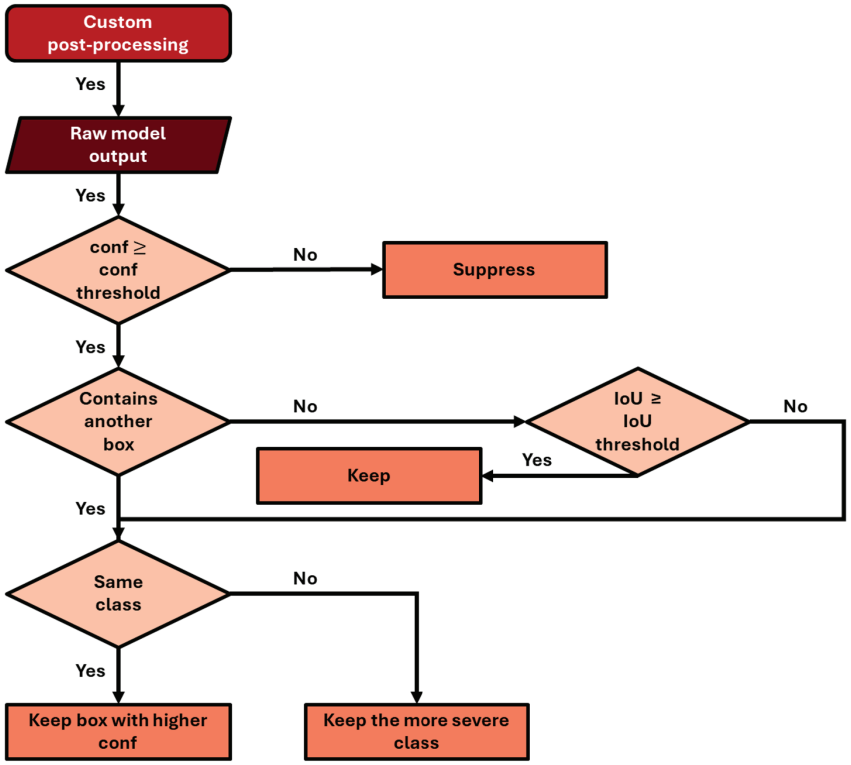


Figure 3. *A flow diagram of the custom post-processing algorithm. conf: confidence.*

Data analysis

Filtered predictions of the test dataset of both post-processing methods were evaluated and compared. Precision, recall, F1-score, and mean average precision at IoU threshold of 0.5 (mAP50) were used to describe and compare the models' performances. Additionally, generated confusion matrices were used to show detailed misclassification patterns, and average IoU was used to represent the models' overall ability in the detection task.

Results

The model achieved the best result at mAP50 of 0.56, recall of 0.57, precision of 0.58, and F1-score of 0.57. On test dataset, the application of both post-processing methods increased the model's mAP50 to a minimum of 0.57, with average IoU scores of 0.62 at 0.3 confidence and 0.59 at 0.5 confidence. The model incorporating the custom post-processing algorithm with a confidence threshold of 0.3 was selected as it demonstrated the best performance with mAP50 of 0.59, recall of 0.60, precision of 0.67, and F1-score of 0.63. The classification patterns are demonstrated in the confusion matrix generated by the model, as shown in Figure 4. In object detection, background class is added. From the confusion matrix, the background row represents missed detections, or when the model could not find the object. While the background column represents false detections, or when the model predicted where there is no object or the background.

Ablation study

An ablation study investigating the detection and classification of PRAs using YOLOv8 model on binary-class dataset was applied to reduce the complexity of the tasks and investigate the models' performances. The scores were dichotomized by grouping scores 1 and 2 as Healthy, and scores 3 to 5 as Diseased. Training parameters were also selected using hyperparameter sweeping process. The parameters were: a YOLOv8-m model, a square input image size of 800 pixels, a batch size of 32, a dropout rate of 0.2, the AdamW optimizer with a learning rate of 0.001, a momentum of 0.937, and a weight decay of 0.0005. The distribution of instances of each class in the dataset are provided in Table 2. As for the post-processing algorithm, the csutom algorithm was modified by removing only the special class condition parts.

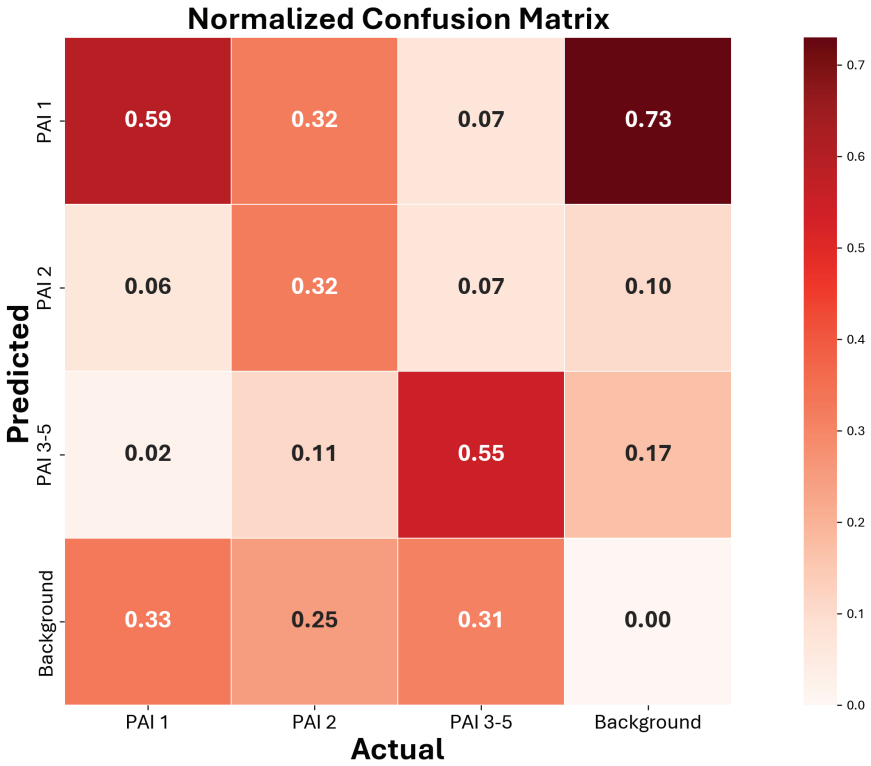


Figure 4. *Normalized confusion matrix of the model employing custom post-processing algorithms at a confidence threshold of 0.3.*

Table 2. Distribution of dataset instances across classes (binary classification).

Categories	Training set	Validation set	Test set	Total instances
Healthy	3247	773	536	4556
Diseased	622	161	159	942

The best-performing model trained on the binary class dataset achieved a mAP50 of 0.78, recall of 0.79, precision of 0.74, and F1 score of 0.76 on the validation dataset. On the test dataset, the raw model output yielded a mAP50 of 0.67. During inference, one overlapping case was observed at a confidence threshold of 0.3, while eight overlapping cases occurred at a threshold of 0.5. For the detection tasks, the model identified PRAs with an average IoU of 0.59 at both thresholds. The binary model employing the custom post-processing method with a confidence threshold of 0.5 was selected for further analysis with mAP50 of 0.66, recall of 0.64, precision of 0.76, and F1-score of 0.69. The corresponding confusion matrix is shown in Figure 4. Examples of common prediction errors after post-processing are shown in Figure 6-8.

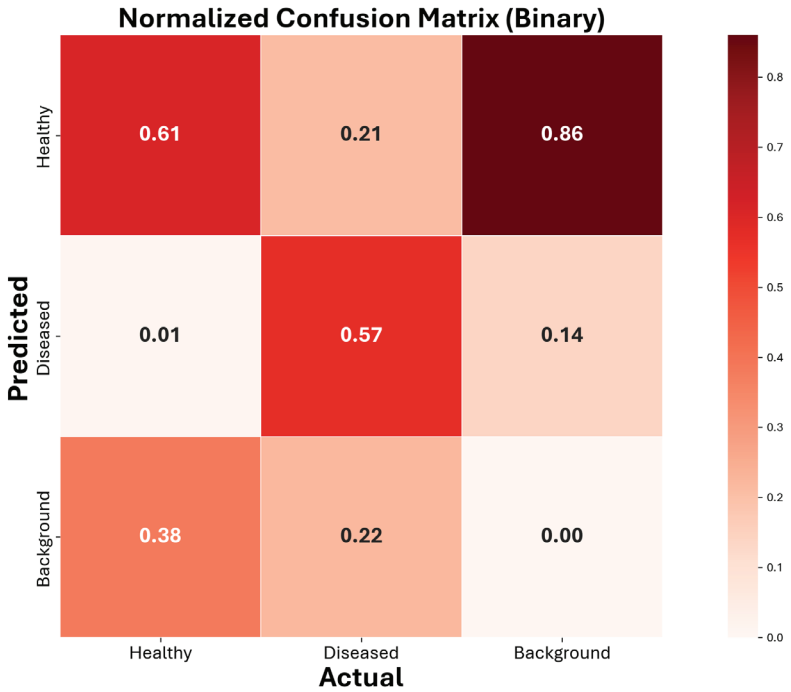


Figure 5. *Normalized confusion matrix of the binary model employing custom post-processing algorithms at a confidence threshold of 0.5.*

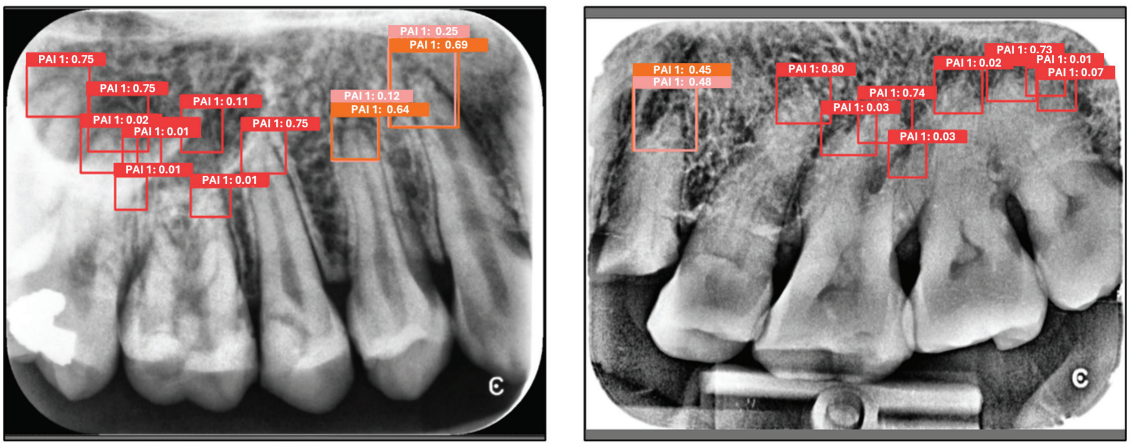


Figure 6. *Predictions on upper jaw Pas showing complete object coverage, but boxes with low confidence are suppressed.*

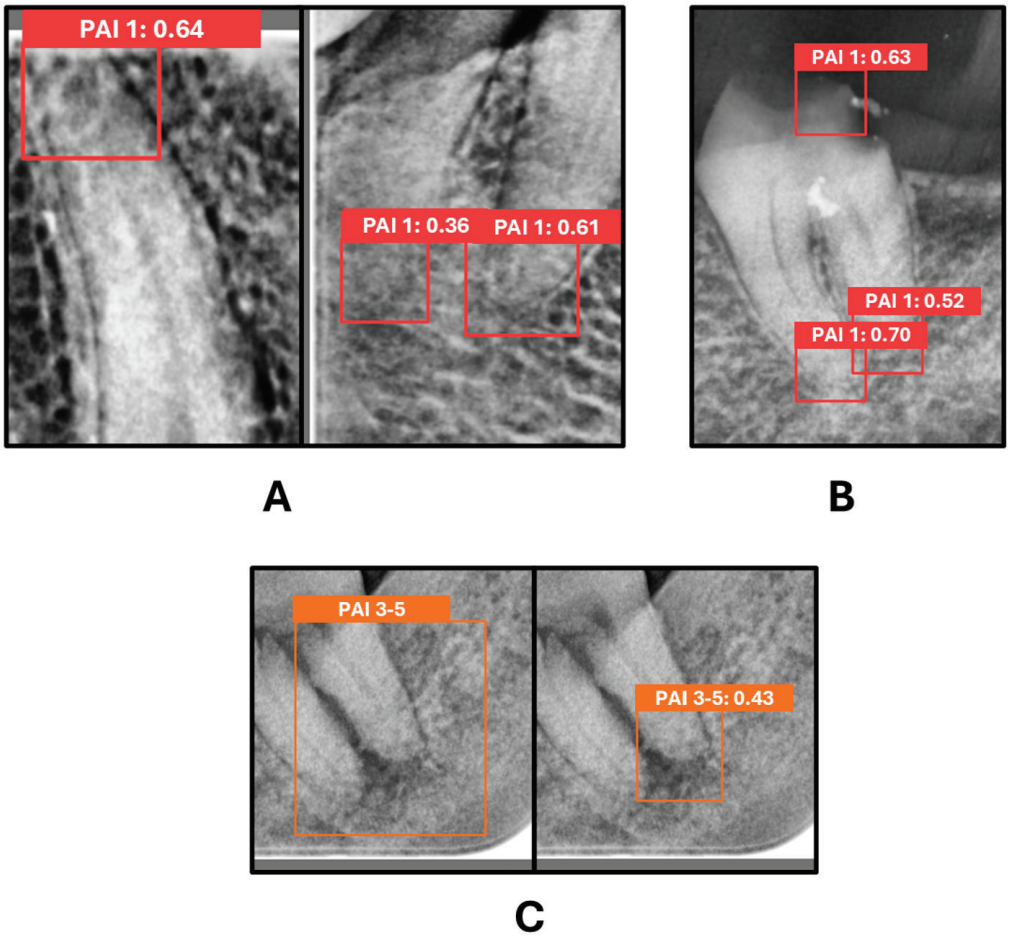


Figure 7. *False detection examples from the model: A: predicted boxes on partially visible roots (confidence 0.64 and 0.36), B: prediction on the tooth crown (confidence 0.63), C: smaller predicted box compared to left ground truth (IoU=0.26).*

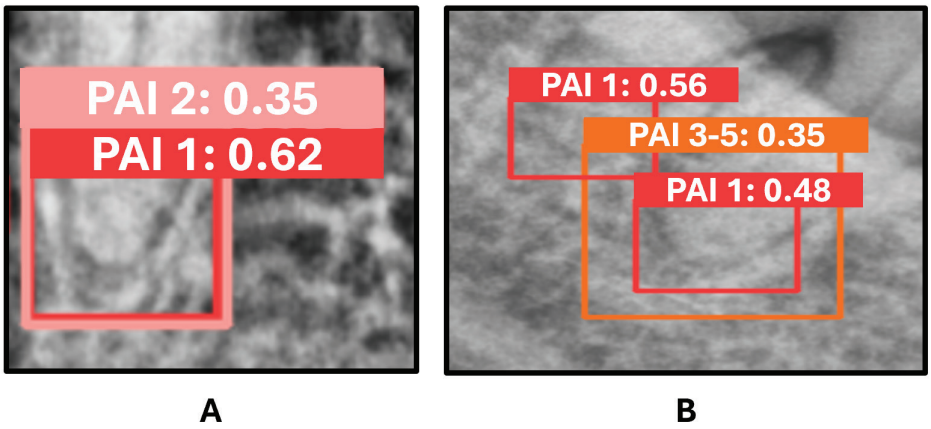


Figure 8. *Examples of bounding boxes not eliminated by NMS:*
A: near-complete overlap between PAI 1 and PAI 2 (IoU 0.99), B: PAI 1 entirely inside PAI 3-5 (IoU 0.40).

Discussion

Radiographs have long been a widely used method for screening and evaluating dental health conditions. PAI scoring system has been used to help assess periapical health in addition to radiograph features. However, interpreting and scoring dental radiographs manually are time-consuming and labor-intensive, placing a significant workload on dental practitioners, especially in high-volume clinical settings. Moreover, the procedures depend heavily on the expertise of trained dentists and radiologists. To address these limitations, deep learning has been introduced to streamline the evaluation process and support dental practitioners. Unlike previous studies that focused primarily on mandibular Pa, this study included images from both the maxillary and mandibular regions. This introduces greater anatomical variability and superimposition, reflecting a more realistic clinical scenario for automated PRA detection. The model achieved moderate accuracy in detecting and classifying PRAs with the highest mAP50 of 0.57.

As shown in Figure 4, the model had difficulty accurately classifying PAI 2, with a correct classification rate of only 0.32, which reduced the overall model accuracy. The moderate performance was influenced by several factors. Firstly, the model showed difficulty distinguishing between PAI 1 and PAI 2 due to subtle differences in radiographic features defined in the PAI scoring criteria, which led to frequent confusion between these classes. A similar pattern has also been reported in previous studies.[17,22,31] Secondly, a considerable number of raw predictions had low confidence scores and were removed during post-processing, resulting in increased missed detections. More than half of the raw predictions (56.00%) had confidence scores below 0.3. These errors were observed mostly in periapical radiographs taken from the upper jaw, as shown in the examples in Figure 6. Detection was more challenging in maxillary radiographs due to anatomical superimposition, such as the sinus floor and adjacent bones, which contributed to lower confidence predictions. In contrast, the mandibular region presents fewer sources of interference, with the mandibular canal being the primary structure that may complicate predictions. Finally, some background misclassifications occurred when normal roots were incorrectly detected as PAI 1 (71.00%), suggesting that the model occasionally confused normal roots with background regions. Figure 7 illustrates examples of these false detections that reduced model accuracy. Among them, size mismatch between the object and the predicted boxes-shown in Figure 7c-was the major contributor to background misclassification: the predicted box was counted as a false detection, while the unmatched object box was recorded as a missed detection.

In the binary classification study, the model predicted with higher confidence due to the lower task complexity. However, its correct classification rates remained moderate. Specifically, 21.00% of Diseased roots were misclassified as Healthy. Upon closer inspection, a single PAI 1 box was misclassified as PAI 3-5 (diseased), suggesting that PAI 2 exhibits characteristics like PAI 3-5, consistent with the PAI scoring system. Nonetheless, most errors in the binary model also stemmed from missed detections rather than misclassifications. Approximately 15.00% of the total predictions were suppressed by the post-processing algorithm at a confidence threshold of 0.3, with an additional 4.00% suppressed at 0.5. Thus, although the binary model generated fewer low-confidence predictions and achieved the highest recall, it still failed to detect some roots in the raw results and the missed detection rate remained high. While most of these predictions were ultimately suppressed, the three-class model was able to detect most complex PRAs. In comparison, the binary model showed stronger performance in detecting complex PRAs on mandibular periapical radiographs. Among false detections, 86% were assigned as healthy, though they comprised only 18.00% of all predictions. The model frequently produced high-confidence predictions for roots that were not fully visible, mostly around the edges. Like the previous model, size mismatch between the object and the predicted box remained the major contributor to background misclassifications. Overall, the model's performance was good, provided that early-stage lesion classification was not the focus.

The custom post-processing algorithm performed well in filtering fully contained boxes that the NMS method could not suppress. Compared with NMS, the model using the custom algorithm achieved lower recall because some overlapping cases were still left unsuppressed. However, this resulted in a slight improvement in precision, indicating that the algorithm and its special class conditions handled overlapping cases more effectively. Since only a small number of overlapping cases remained after confidence thresholding, the differences between NMS and custom post-processing models were minimal. At a confidence threshold of 0.5, no fully contained boxes were present after filtration, and both methods produced identical inference results. Some bias was introduced by forcing the model to prioritize the more severe class, which sometimes led to incorrect predictions. For example, in Figure 8a, the algorithm suppressed a contained PAI 1 box (with higher confidence) and retained the PAI 2 box, which was the correct score. In contrast, Figure 8b shows a case where the model was forced to keep a PAI 3-5 box and discard a PAI 1 box, resulting in an error. Despite this limitation, implementing class conditions increased the model's sensitivity to diseased roots, which is valuable for screening in clinical practice.[17] Nevertheless, having practitioners review the model's results can help prevent errors caused by such bias.

Moidu *et al.*[17] developed YOLOv3 models for PAI scoring in periapical radiographs. One was a five-class model covering all PAI scores, and the other was a binary model. Although a direct comparison is not appropriate due to differences in study settings, the results are discussed solely to highlight similarities in classification patterns. Their model reported PAI 1-5 accuracies of 90.90%, 30.00%, 60.00%, 71.00%, and 30.00%, respectively, while our model achieved 59.00% (PAI 1), 32.00% (PAI 2), and 55.00% (PAI 3-5). In binary classification, accuracies for healthy and diseased cases were 76.60% and 92.00% for their model, compared with 61.00% and 57.00% for ours. For both the multi-class and binary models, their results outperformed ours in all categories except PAI 2, indicating that classifying the early stage of the lesion was challenging for both studies. Their advantages were: 1) the data were acquired using a single scanning system, 2) a balanced dataset was used, and 3) most importantly, only mandibular periapical radiographs were included. Considering these advantages, their dataset was less complex and less biased, resulting in more robust models. They also noted that the model tended to overscore by one class. Based on this observation, we implemented a post-processing algorithm to correct predictions following this pattern, thereby enhancing model performance.

Limitations

Although the findings are encouraging, a few limitations should be acknowledged. Firstly, the dataset included a relatively low proportion of maxillary periapical radiographs, accounting for only 25%. To address the increased complexity caused by superimposed maxillary structures, a higher proportion of maxillary images would be necessary. A larger and more diverse dataset would also allow better class distribution.[32] Pa typically includes multiple teeth per image, which makes balancing the dataset practically challenging. In addition, external validation using datasets from other institutions was not performed in this study, and further evaluation on multi-center datasets is necessary to assess the generalizability of the proposed model. Second, our computing resources were limited. Hyperparameter sweeping and tuning were computationally expensive, which restricted the number of combinations we could explore. This reflects a trade-off between computational cost and model performance, as improving detection accuracy typically requires larger models, larger datasets, and more extensive hyperparameter optimization, all of which increase computational requirements. Nevertheless, training with larger models, a bigger dataset, and more hyperparameter combinations could potentially improve performance. Additionally, comparison with other architectures, including transformer-based models and approaches leveraging large-scale pre-training, would be valuable and could be explored in future studies. Finally, we used only CLAHE and BF for image preprocessing, as these methods are commonly employed for normalization in dental imaging. Although image preprocessing is often omitted for object detection models,[33,34] we applied these techniques to help manage dataset variability, since our comparatively smaller model-limited by computational resources-might not have been able to handle it effectively on its own.

While our models have not yet reached the performance of state-of-the-art methods, the limitations identified in this study offer valuable insights and highlight clear avenues for improvement. By addressing these challenges and implementing the strategies discussed, future models could achieve higher accuracy, improved computational efficiency, and greater adaptability to the complexities of dental imaging in clinical practice. Additionally, future work should include a comparison between the proposed model and the diagnostic performance of dentists to further evaluate its clinical applicability in detecting periapical lesions.

Conclusion

In this study, PRAs detection and classification models were developed using the YOLOv8 architecture. The multi-class model achieved moderate performance, with an mAP50 of 0.59, while the binary model reached an mAP50 of 0.66. Although not yet ready for clinical use, these models demonstrate potential as adjunct tools to support dental practitioners in radiographic evaluation, particularly for less experienced users. In a potential clinical workflow, such systems could automatically highlight suspicious regions in periapical radiographs and provide preliminary classification results, allowing clinicians to review the findings more efficiently while maintaining final diagnostic responsibility. However, further research, including larger datasets and improved model performance, is needed to develop a more robust and clinically applicable system.

Ethical approval

Ethical approval for this retrospective study was obtained from the Institutional Ethical Review Board, Faculty of Dentistry, Chiang Mai University, Chiang Mai, Thailand (Approval No. 26/2024) on July 5th, 2024. Due to the deidentification of data of the retrospective study, the informed consent was not required.

Funding

This research received no external funding.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Natdanai Hirata: conceptualization, methodology, data curation, formal analysis, software, writing: original

draft preparation, reviewing and editing; **Wannakamon Panyarak**: conceptualization, investigation, resources, writing: reviewing and editing; **Arnon Charuakkra**: investigation, resources, writing: reviewing and editing; **Kittipit Klanliang**: investigation, resources, writing: reviewing and editing; **Kittichai Wantanajittikul**: conceptualization, software, methodology, writing: original draft preparation, reviewing and editing.

Acknowledgements

The authors gratefully acknowledge the support of the master's degree program in Medical Radiation Sciences, Faculty of Associated Medical Sciences, Chiang Mai University, under the CMU Presidential Scholarship.

References

[1] Arias Z, Nizami MZI, Chen X, Xu B, Kuang C, Omori K, et al. Recent advances in apical periodontitis treatment: a narrative review. Bioengineering. 2023; 10(4):488. doi: 10.3390/bioengineering10040488.

[2] Tibúrcio-Machado CS, Michelon C, Zanatta FB, Gomes MS, Marin JA, Bier CA. The global prevalence of apical periodontitis: a systematic review and meta-analysis. Int Endod J. 2021; 54(5): 712-35. doi: 10.1111/iej.13467.

[3] Buonavoglia A, Zamparini F, Lanave G, Pellegrini F, Diakoudi G, Spinelli A, et al. Endodontic microbial communities in apical periodontitis. J Endod. 2023; 49(2): 178-89. doi: 10.1016/j.joen.2022.11.015

[4] Loesche WJ. Microbiology of Dental Decay and Periodontal Disease. In: Baron S, Editor. Medical Microbiology. 4th Edition. Galveston, University of Texas Medical Branch at Galveston, Chapter 99.

[5] Lo Giudice R, Nicita F, Puleio F, Alibrandi A, Cervino G, Lizio AS, et al. Accuracy of periapical radiography and CBCT in endodontic evaluation. Int J Dent. 2018; 2018:2514243. doi: 10.1155/2018/2514243.

[6] Tanomaru-Filho M, Jorge EG, Duarte MA, Gonçalves M, Guerreiro-Tanomaru JM. Comparative radiographic and histological analyses of periapical lesion development. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2009; 107(3): 442-7. doi: 10.1016/j.tripleo.2008.12.009.

[7] Karamifar K, Tondari A, Saghiri MA. Endodontic periapical lesion: an overview on the etiology, diagnosis and current treatment modalities. Eur Endod J. 2020; 5(2): 54-67. doi: 10.14744/eej.2020.42714.

[8] Ørstavik D, Kerekes K, Eriksen HM. The periapical index: a scoring system for radiographic assessment of apical periodontitis. Dent Traumatol. 1986; 2(1): 20-34. doi: 10.1111/j.1600-9657.1986.tb00119.x

[9] Rajmohamed RF, Chapala S, Shazahan MA, Wali P, Botchu R. Evaluating the accuracy and efficiency of AI-generated radiology reports based on positive findings-a qualitative assessment of AI in radiology. Acad Radiol. 2025; 35(12): 7035-40. doi: 10.1016/j.acra.2025.09.012.

[10] Kazimierczak N, Sultani N, Chwarścianek N, Krzykowski S, Janiszewska-Olszowska J, et al. Detection accuracy of an AI platform for dental treatment features on panoramic radiographs tooth- and patient-level analyses. Sci Rep, 2026; 16: 2436. doi: 10.1038/s41598-025-32226-0

[11] Huang C, Wang J, Wang S, Zhang Y. A review of deep learning in dentistry. Neurocomputing. 2023; 554: 126629. doi: 10.1016/j.neucom.2023.126629.

[12] Sohrabniya F, Hassanzadeh-Samani S, Ourang SA, Jafari B, Farzinnia G, Gorjinejad F, et al. Exploring a decade of deep learning in dentistry: a comprehensive mapping review. Clin Oral Invest. 2025; 29: 143. doi: 10.1007/s00784-025-06216-5.

[13] Khan HA, Haider MA, Ansari HA, Ishaq H, Kiyani A, Sohail K, et al. Automated feature detection in dental periapical radiographs by using deep learning. Oral Surg Oral Med Oral Pathol Oral Radiol. 2021; 131(6): 711-20. doi: 10.1016/j.oooo.2020.08.024

[14] Du M, Wu X, Ye Y, Fang S, Zhang H, Chen M. A combined approach for accurate and accelerated teeth detection on cone beam CT images. Diagnostics (Basel). 2022; 12(7): 1679. doi: 10.3390/diagnostics12071679.

[15] Sadr S, Mohammad-Rahimi H, Motamedian SR, Zahedrozegar S, Motie P, Vinayahalingam S, et al. Deep learning for detection of periapical radiolucent lesions: a systematic review and meta-analysis of diagnostic test accuracy. J Endod. 2023; 49(3): 248-61.e3. doi: 10.1016/j.joen.2022.12.007.

[16] Ver Berne J, Saadi SB, Politis C, Jacobs R. A deep learning approach for radiological detection and classification of radicular cysts and periapical granulomas. J Dent. 2023; 135: 104581. doi: 10.1016/j.jdent.2023.104581.

[17] Moidu NP, Sharma S, Chawla A, Kumar V, Logani A. Deep learning for categorization of endodontic lesion based on radiographic periapical index scoring system. Clin Oral Investig. 2021; 26(1): 651-8. doi: 10.1007/s00784-021-04043-y.

[18] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016: 779-88. doi: 10.1109/CVPR.2016.91.

[19] Sohan M, Sai Ram T, Rami Reddy ChV. A Review on YOLOv8 and its advancements. In: Jacob IJ, Piramuthu S, Falkowski-Gilski P, Editor. Data Intelligence and Cognitive Informatics. Singapore: Springer Nature; 2024: pp 529-45.

[20] Singh K, Kumar R, Singh S. Advanced YOLOv8 architecture for multi-class brain tumor detection. J Assoc Med Sci [internet]. 2026 Feb. 25 [cited 2026 Mar. 20];59(2):250-8. available from: https://he01.tci-thaijo.org/index.php/bulletinAMS/article/

view/281957.
[21] Yaseen M. What is YOLOv8: an in-depth exploration of the internal features of the next-generation object detector. ArXiv Preprint posted online August 28, 2024. doi: 10.48550/arXiv.2408.15857.
[22] Hirata N, Pudhieng P, Sena S, Torn-asa S, Panyarak W, Klanliang K, et al. Development of periapical index score classification system in periapical radiographs using deep learning. J Digit Imaging Inform Med. 2024; 38(5): 3125-33. doi: 10.1007/s10278-024-01360-y.
[23] Bachani L, Singh M, Anshul, Lingappa A. Ideal radiographs: an insight. IP Int J Maxillofac Imaging. 2020; 6(3): 56-64. doi: 10.18231/j.ijmi.2020.017.
[24] Choi JW, Han WJ, Kim EK. Image enhancement of digital periapical radiographs according to diagnostic tasks. Imaging Sci Dent. 2014; 44(1): 31-5. doi: 10.5624/isd.2014.44.1.31.
[25] Ari T, Sağlam H, Öksüzoğlu H, Kazan O, Bayrakdar İŞ, Duman SB, et al. Automatic feature segmentation in dental periapical radiographs. Diagnostics. 2022; 12(12): 3081. doi: 10.3390/diagnostics12123081.
[26] Chen IDS, Yang CM, Chen MJ, Chen MC, Weng RM, Yeh CH. Deep learning-based recognition of periodontitis and dental caries in dental X-ray images. Bioengineering (Basel). 2023; 10(8): 911. doi: 10.3390/bioengineering10080911.
[27] Çelik B, Savaştaer EF, Kaya HI, Çelik ME. The role of deep learning for periapical lesion detection on panoramic radiographs. Dentomaxillofacial Radiol. 2023; 52(8): 20230118. doi: 10.1259/dmfr.20230118.s.
[28] Taş İÇ. Application of panoramic dental x-Ray images denoising. Int J Innov Eng Appl. 2023; 7(1): 13-20. doi: 10.46460/ijiea.1134105.
[29] Hana FM, Maulida ID. Analysis of contrast limited adaptive histogram equalization (CLAHE) parameters on finger knuckle print identification. J Phys Conf Ser. 2021; 1764(1): 012049. doi: 10.1088/1742-6596/1764/1/012049.
[30] Ahmed S, Yoshiura N, Hassan S, Khan A. Towards better accuracy on imbalance image datasets using image augmentation and convolutional neural networks. Lecture Notes in Computer Science in International Conference on the Dynamics of Information Systems (LNCS-DIS 2024); 2025; 14661: 189-212. doi: 10.1007/978-3-031-81010-7_13.
[31] Maia Filho EM, Calisto AM, De Jesus Tavarez RR, De Castri Rizzi C, Bezerra Segato RA, Bezerra da Silva LA. Correlation between the periapical index and lesion volume in cone-beam computed tomography images. Iran Endod J. 2018; 13(2): 155-8. doi: 10.22037/iej.v13i2.15040.
[32] Mooijman P, Catal C, Tekinerdogan B, Lommen A, Blokland M. The effects of data balancing approaches: a case study. Appl Soft Comput. 2023; 132: 109853. doi: 10.1016/j.asoc.2022.109853.
[33] Kawaguchi K, Kaelbling LP, Bengio Y. Generalization in deep learning. ArXiv Preprint posted online October 16, 2017. doi: 10.48550/arXiv.1710.05468.
[34] Vakanski A, Xian M. Evaluation of complexity measures for deep learning generalization in medical image analysis. Proceedings of the 2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP); 2021: 1-6. doi: 10.1109/mlsp52302.2021.9596501.

Journal of Associated Medical Sciences 2026; 59(2): 327-338

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Brief review: Effects of textural properties of lamellar silica in drug delivery applications

Dyah Ellyawati Kusumaningtyas Maharani[1], Fajar Rakhman Wibowo[1], Rino Rakhmata Mukti[2*]

[1]Department of Chemistry, Faculty of Mathematics and Natural Sciences, Universitas Sebelas Maret, Surakarta, Indonesia.
[2]Research Center for Nanosciences and Nanotechnology, Institut Teknologi Bandung, Jalan Ganesha, Bandung, Indonesia.

ARTICLE INFO

Article history:
Received 7 February 2025
Accepted as revised 2 July 2025
Available online 25 March 2026

Keywords:
Dendritic, KCC-1, lamellar silica, textural properties, drug delivery

ABSTRACT

Background: Silica-based materials, especially those with lamellar morphology, usually have other names, such as dendritic, nanoflower, and wrinkle are often used in various applications due to their unique textural properties, such as radially centered pore structure, pore channels, and large pore volume.

Objectives: To determine the effect of lamellar silica texture properties, such as surface area, porosity, pore size, morphological shape, particle size, and surface charge value, as potential drug delivery material candidates.

Materials and methods: The texture properties of silica materials were reviewed in drug delivery applications, and four silica materials with different morphologies were reviewed for their texture properties and potential in drug delivery applications.

Results: The literature review revealed the effect of material texture properties on their potential in drug delivery applications. Drug loading in porous carriers can improve their activity, stability, and selectivity, and can reduce the drug dose to inhibit cell viability compared to unencapsulated drugs. Surface modification can also affect the zeta potential value and its performance in drug loading. In vitro drug release properties of the material and drug release rates showed different results related to their different texture properties. Particle size affects particle penetration into cells; the smaller the particle size, will cause deeper the penetration, but it has a tendency for particle aggregation.

Conclusion: Lamellar silica-based materials with various unique properties of their texture properties show great potential in drug delivery applications and are expected to be used in the future.

Introduction

Silica-based nanoparticles are often utilized in various applications, including MCM-41[1,2] and SBA15.[3,4] This is due to their tunable morphology, high surface area, easy modification, and biodegradable.[5-7] Although MCM-41 and SBA-15 have a high surface area, their tubular pore structure causes limited accessibility within the pores. In addition, when these silicas are loaded with guest molecules, there is limited loading of guest molecules and poor accessibility of active sites. Lastly, these materials are unstable at high temperatures and pressures due to their thin walls.[8] A potential candidate as a drug carrier material that has been widely developed is mesoporous silica nanoparticles (MSNs), which have several unique

* **Corresponding contributor.**
Author's Address: *Research Center for Nanosciences and Nanotechnology, Institut Teknologi Bandung, Jalan Ganesha, Bandung, Indonesia.*
E-mail address: rinomukti@itb.ac.id
doi: 10.66285/JAMS.2026.064
E-ISSN: 2539-6056

features. Recently, dendritic silica-based materials with center-radial pore structures, large pore channels, and highly accessible pore volumes were utilized as promising drug delivery systems.[9-13] The large pore nature of dendritic silica will bring the advantage of loading more drugs to be targeted.

The fibrous surface of dendritic silica allows it to be easily loaded with several ranges of organic groups, organometallic complex ionic liquids, polymers, peptides, enzymes, DNA, genes, drugs, and others.[14-16] The dendritic silica structure having radially distributed fibrous channels usually promotes increased functionalization capacity due to the presence of many accessible sites, which allows easy incorporation of different agents. This allows designing materials with modifications along with targeting systems.[17] The name of the dendritic itself has been a topic of discussion; some have mentioned it as fibrous, wrinkle, nanoflower, and most recently[18] found the fact that the morphological shape of this silica is lamellar morphology, so a new terminology is proposed in the form of biconcentric lamellar silica. In various types of terminology, the use of this material in drug delivery is limited, so it is necessary to optimize the formation of the material to maximize the performance of the material design, especially in drug delivery applications. The initial evaluation process of the potential material as a drug carrier can be done by evaluating its texture properties. By understanding how the texture of the material (such as pore size, surface area, particle shape, zeta potential value) affects the interaction with drugs, especially in biological tissues, the design of a more effective and safe drug delivery system can be carried out for further evaluation.

Textural properties

Some things that need to be considered in the formation of materials for drug delivery applications are the textural properties, including porosity and pore distribution,[19,20] surface area,[21] morphology, particle size,[22-24] zeta potential value[25] of the material. These properties will affect the drug loading, release profile, stability, and biocompatibility.[26] Several parameters that have a strong influence on the synthesis conditions, such as surfactants, silica precursors, solvents or cosolvents, pH of the media, reaction variations, or hydrothermal temperatures, will result in the formation of mesoporous silica nanoparticles with pore sizes, as well as the orientation of the formed pores, varied surface areas, and different morphologies.[27] Analysis of the textural properties of materials can be carried out by the absorption of gases such as N_2, He, Ar (adsorbate) at various relative pressures on solids (adsorbents), which provides information on textural properties including surface area, pore volume, and pore size.[28] Table 1 shows the analysis of the nature of dendritic silica texture with various morphological terminologies, using the specific method and results in their surface area and pore size.

Table 1. Analysis of the nature of dendritic silica texture with various morphological terminologies.

Material	Method	Surface area (m^2/gm)	Pore size (nm)	Type of isotherm and hysteresis	Reference
Dendritic fibrous silica	Open vessel conventional heating protocol	397-635	2-25	Type IV isotherm	31
KCC-1	Hydrothermal	435	4.56	Type IV isotherm with H3 hysteresis loop	32
Dendritic mesoporous silica	Aqueous-based dual-templating	784-809	3.2-9.8	Type IV isotherm	33
KCC-1	Microemulsion hydrothermal method	501.3	3.4	Type IV isotherm with H4 hysteresis loop	34
KCC-1	Microemulsion microwave-assisted	691	4.77	Type IV isotherm with H1 hysteresis loop	35
Discontinuous concentric lamellar	Low-pressure solvothermal	478-782	3.66-5.09	Type IV isotherm with an H3 hysteresis loop	36
Nanoflower silica	Light-assisted method	701	3.4	Type IV isotherm	37
KCC-1	Sol-gel method	442	2.76	Type IV with a type of hysteresis loop	38

The pore diameter can be adjusted from 2 to 30 nm depending on the surfactant used as a template and the synthesis conditions.[29] The formation of these pores and morphology will also affect the application of the silica material itself. The morphology of silica can also be affected by the formation of surfactant molecules into liquid micelles, which are often highly dependent on the hydrophobic-hydrophilic balance between the template, precursor material, hydrolysis reagent, and solvent.

In addition, the key to fibrous morphology as well as particle size is the control of the hydrolysis speed of TEOS as silica precursor by the hydrolysis reagent. However, with further increase in the amount of hydrolysis reagents, no change in fibrous morphology was observed except for an increase in particle size distribution, which may be due to the rapid (uncontrolled) hydrolysis of TEOS molecules by excess hydrolysis reagents.[30]

Porosity and pore size distribution

Porosity is defined as the ratio of total pore volume to particle or agglomerate volume.[39] In the IUPAC Classification, most silica materials have a mesoporous pore size (2-50 nm). Loading of drugs in porous carriers can improve their activity, stability, and selectivity, and can lower the dose of drugs to inhibit cell viability compared to unencapsulated drugs. This is one of the advantages of drug loading in pores.[40] In addition, drug loading inside a porous material shows an increase in the solubility of the drug because of spatial restriction in nano-sized pores. Crystalline drug molecules usually cannot form highly ordered crystals; instead, remaining in an amorphous or non-crystalline form, which is known to increase the rate of drug dissolution.[41] In addition to drug loading, drug release also depends on the pore size of the carrier. RSV drug release tests conducted by[39] showed materials with large pore size have more drug release percentage than carriers with smaller pores. Drug loaded in the pore can also be beneficial to achieve effective treatment concentrations quickly.[40] Type of pore shapes will also affect drug loading, worm-like pore ones are slightly had lower than hexagonal-porous ones, probably due to the smaller pore size and curved shape of worm-like pores, which makes the entry of drugs into the interior of the pore.[41]

In drug delivery applications, KCC-1, which has a fibrous morphology and a pore size of 4.77 nm, has increased diclofenac adsorption performance. Induction occurs at the beginning of 10 minutes, where the adsorption capacity and diclofenac removal efficiency increase with contact time. After 40 minutes, the diclofenac adsorption performance decreases drastically because the active sites of the composite for adsorption have been saturated with diclofenac.[34] In the work done,[42] the loading of CUR into the pore on the nanocarrier in the form of dendritic silica can enhance the cytotoxic activity because the nanocarrier maintains the optimum concentration of CUR molecules through sustained and controlled drug release and increases the solubility of CUR in PBS media. The formation of larger dendritic silica pore size to improve the performance of the material in drug loading can be done using auxiliary surfactants. One of the auxiliary surfactants used is ferrocene carboxylic acid (FCA) with different mass ratios of FCA to CTAB,[33] DLMSN can be used for selective loading of Dox, IR780, and Hb. The pore sizes of MSN, DLMSNFCA with FCA to CTAB mass ratio of 0.2, and DLMSNFCA with FCA to CTAB mass ratio of 0.3 were 2.6, 6.3, and 12.2 nm, respectively, which showed the highest selective loading content for Dox (34.4%), IR780 (5.7%), and Hb (43.3%).

Surface area

High surface area is one of the important things in drug delivery system applications. The drug loading capacity is positively correlated with the specific surface area; a large specific surface area will provide more absorption sites for the drug.[42] Drug loading studies were conducted on several commercial silica mesopores and showed that an increase in surface area (through smaller pore diameters) would not increase drug loading capacity, but rather a decrease, as the entire surface cannot be covered monomolecularly due to the spatial limitations of the narrow pores. Drug molecules cannot access the entire surface area, and the drug will precipitate outside the pores of conventional silica materials.[43] This is one of the advantages of the fibrous surface of lamellar silica that allows for the loading of various types of matrices as well as the ease of modification on its surface. The increase in pore diameter and surface area by the addition of organic functional groups further greatly enhances the loading and release of drug molecules and the kinetics.[40]

Surface functionalization of the material aims to introduce new functional groups in the framework of the carrier material to bind the drug without changing its structure, crystallinity, or morphology.[44] Functionalization on the surface will cause a decrease in the surface area of the material due to the closure of the pore with new functional groups.[45] Mesoporous amine-modified KCC-1 (KCC-1-NH_2) can be considered as a promising carrier for drug delivery systems to control the drug release rate influenced by molecular interactions, especially for highly soluble drugs, such as paracetamol.[46]

Morphology

As mentioned earlier, the name of this dendritic silica has been a topic of discussion, with some referring to it as dendritic fibrous, wrinkle, nanoflower, and most recently, the fact that the morphological shape of this silica is lamellar, so that a new terminology is proposed in the form of biconcentric lamellar silica.[18]

Here, we will explain some mentions of lamellar silica morphology, how it is formed, and its utilization in drug delivery applications.

Dendritic fibrous silica (KCC-1)

DFNS has a unique fibrous morphology (flower-like morphology consisting of thin sheets), unlike the porous structure of various conventional silica spheres. It has a high surface area due to its fibers (thin sheets) and, more importantly, it has improved accessibility to the internal surface and enhanced stability due to its good fiber/sheet thickness.[30] Some obtained DLMSNs had center-radial pore structures, unique open three-dimensional dendritic superstructures with large pore channels, and highly accessible internal surface areas.[32]

The formation mechanism of DFNS is an interaction between various processes, such as surfactant molecular assembly, micelle or microemulsion droplet formation, nucleation, and growth. In the five-step DFNS formation mechanism (Figure 1A), CTAB molecules self-assemble to form a flat phase and are further stabilized, where cosurfactant molecules are inserted between CTAB molecules. This is followed by the formation of Bicontinuous Microemulsion Droplets (BMDs) with water and oil channels. Some of these BMDs (depending on the alcohol chain length of the cosurfactant) are unstable, so their coalescence results in larger BMDs. Silica precursors already present in the oil channel are hydrolyzed at the water-oil interface. After supersaturation, nucleation and growth occur in the BMD channel.[47]

The mechanical stability of DFNS is due to the silica sheets being covalently and hence not mechanically disintegrating due to pressure or friction. The thermal stability is due to the better sheet (fiber) thickness of DFNS compared to the wall thickness of SBA-15 or MCM-41 so that it does not collapse even after heating up to 800 °C. The uniqueness of DFNS is its open dendritic fibrous morphology, which significantly improves its accessibility.[8]

The use of dendritic-based silica materials has been widely used for drug delivery applications.[48-54] In a study conducted,[48] the successful loading of the drug PTX in DMSNs was demonstrated by elemental carbon mapping. No appreciable carbon element was observed in the blank DMSN (Figure 1B, 1C), whereas the carbon 153 element appeared to be homogeneously distributed throughout the DMSN after PTX loading (Figure 1D, 1E).

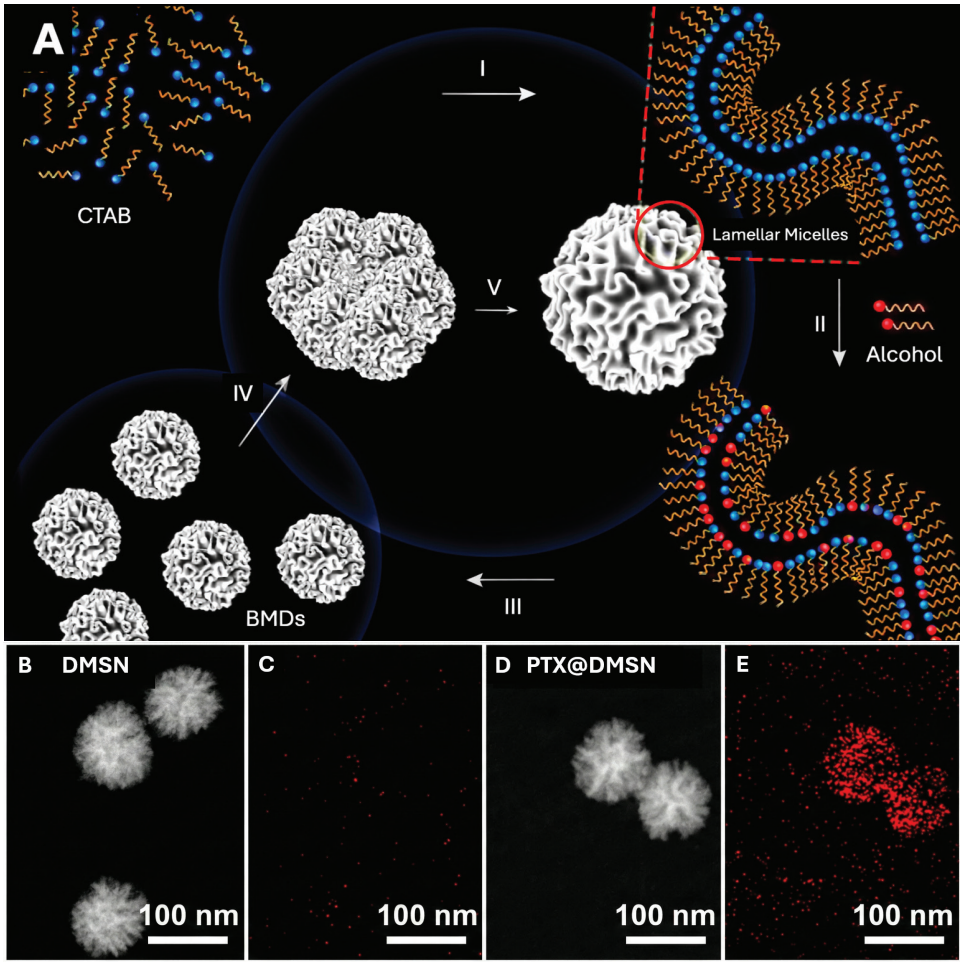


Figure 1. *Formation of DFNS and application of DMNS in drug delivery. A: formation mechanism of DFNS, B-C: carbon elemental mapping and the corresponding TEM images of DMSN, D-E: PTX@DMSN. (Adapted with permission from Deng C, et al.[48] Copyright 2021 Elsevier and Maity A, et al.[47] Copyright 2017 American Chemical Society).*

Wrinkle silica

During the synthesis of wrinkle silica nanoparticles, the silica fibers grow and spread radially from the center of the nanostructure, resulting in a cone-shaped pore with progressively larger parts moving outward. This feature helps to accommodate large macromolecules, such as proteins, from clogging the pore.[55] The formation of wrinkle silica is described in Figure 2 (Upper panel), which is formed from a bicontinuous microemulsion. It was confirmed that WSNs were generated in the bicontinuous microemulsion phase of the Winsor III system.[56] By using the phase behavior of the Winsor III system, which depends on the water-surfactant-oil mixing ratio, and by adding various cosolvents, the structure of silica nanoparticles can be precisely controlled from mesoporous shape to wrinkle shape, and the distance between wrinkles can be controlled.

Figure 2 (Lower panel) shows the effect of oil ratio as cosolvent in the formation of wrinkle silica. Based on the typical phase diagram of the Winsor III system, when additional oil is added gradually with a fixed amount of water and surfactant (sometimes including cosolvent) within a certain region of reaction conditions, the equilibrium state of the mixture will be reached.[56] WSM synthesis was carried out using various types of emulsion systems in the multiphase area as the structure-directing template can control the internal morphology and interparticle connecting structure (Figure 2 a-f).[57] The results show that mesoporous silica nanoparticles are interconnected and transformed into a wrinkled structure depending on the amount of oil in the emulsion system, which starts from direct micelles. In addition, it was confirmed that WSNs were interconnected and formed bulk gels interconnected with wrinkle structures as the amount of water in the emulsion system decreased, starting from the composition that resulted in the formation of spherical WSNs. The pores of the synthesized silica mesostructures with varying amounts of oil are interconnected and produce varying wrinkle structures. A few researchers have used this wrinkle-based silica material in drug delivery applications.[58-60] WMS presenting covalently coupled curcumin (WMS-Curc) demonstrates the flexibility of this methodology for the design and synthesis of new DDS, tailoring the surface functionalization of silica nanomaterials.[60]

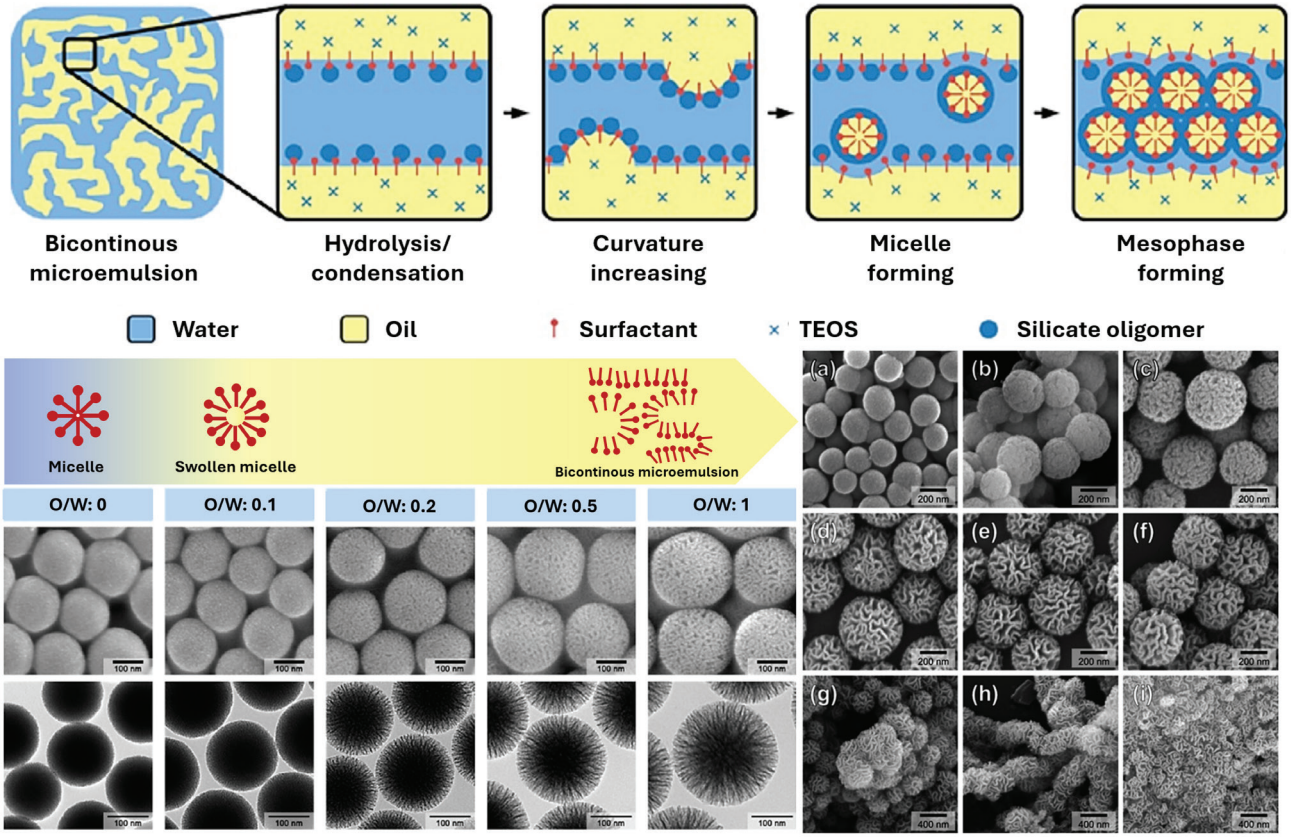


Figure 2. *Formation and morphology of MSN with radial wrinkle structure. Schematic illustration of the mesophase-forming mechanism. Upper panel: the schematic from very tight mesoporous to expanded mesoporous, mixed (or broken) mesoporous, and the wrinkled structure, Lower panel: SEM/TEM images of silica nanoparticles synthesized at different volume ratios of cyclohexane to 15 mL of aqueous solution of urea (0.3 gm) and CPB (0.5 gm) and isopropanol (0.46 mL). (Adapted with permission from Moon and Lee[56] Copyright 2012 American Chemical Society). A-F: SEM images of the silica materials synthesized from the emulsion system (Adapted with permission from Moon and Lee[57] Copyright 2014 American Chemical Society).*

Nanoflower

The term nanoflower arises due to the morphology of a hierarchical flower composed of a collection of petals oriented outward in three dimensions.[61] FESEM images of silica nanoflower nanomaterials[62] (Figure 3A) show a hierarchical flower morphology composed of a collection of thin petals intersecting each other. TEM images (Figure 3B) of the edges of the nanoflowers clearly show that the flowers are composed of elongated spines oriented outward in three dimensions. Based on the analysis, it is shown that the nano-flower petals develop radially from the core and assemble divergently, leading to the formation of nano-flowers with a dandelion-like morphology that makes it possible to obtain a large and easily accessible specific surface area along with fine-tuning of their chemical-physical properties. However, research on the synthesis of silica with well-defined morphology and controllable size still has a long way to go. Nanoflower silica is rarely reported.

Figure 3 shows the formation scheme of silica materials with nanoflower morphology. First, a water-dispersed-in-oil (W/O) emulsion is generated by stirring a mixture of water and cyclohexane in the presence of a surfactant, water droplets are dispersed in the emulsion and stabilized by an orderly array at the interface of water and oil, to form spherically radiating micelles. Secondly, negatively charged silicon oxide is generated through the hydrolysis of TEOS with urea and then arranged on the surface of emulsion droplets to form porous nanoflowers. Finally, spherical mesoporous nanoflowers can be obtained by removing the pattern using calcination. In the use of drug delivery systems, silica nanoflower materials show excellent properties for the controlled release of pyrene due to their unique structure, and can serve as ideal candidates for more potential biomedical applications, such as cellular imaging, biosensors, and targeted drug delivery.[64]

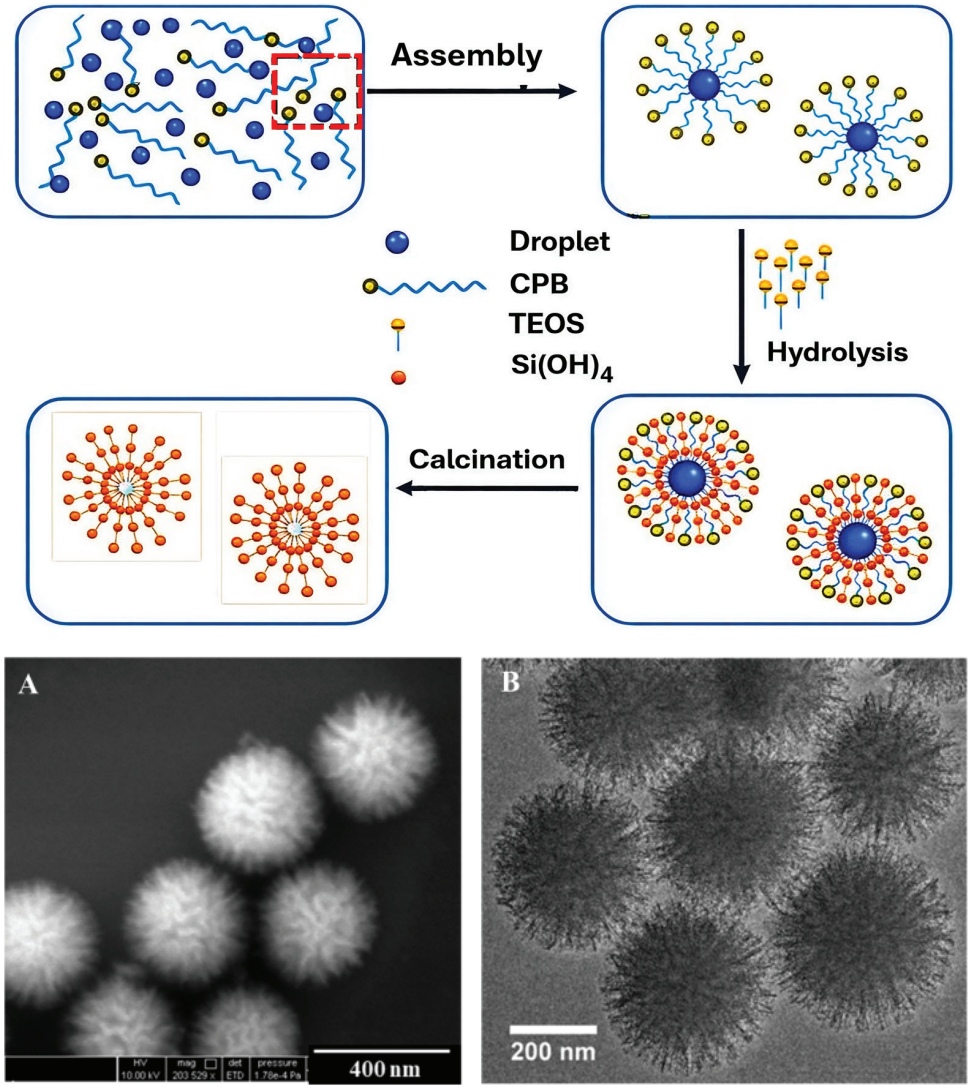


Figure 3. *Formation and morphology of nanoflower silica. Upper: schematic illustration of the possible formation mechanism of porous nanoflowers, Lower: micrograph image of silica-based chromogenic NCs, representing their flowerlike morphology (A: SEM, B: TEM). (Adapted with permission from Yang H, et al.[63] Copyright 2014 Elsevier and Kole K, et al.[37] Copyright 2020 American Chemical Society).*

Biconcentric lamellar silica

Lastly, a thorough study to revisit the morphological nomenclature of KCC-1 and its rational formation mechanism was reported.[18] The morphology of KCC-1 was interpreted after observing and analyzing SEM and TEM images, and shows a unique morphology called bicontinuous-concentric-lamellar (bcl) morphology, which is different from other silica materials (Figure 4). This morphology can be described as bicontinuous lamellar walls arranged vertically and concentrically to form spherical particles.[32] The wrinkles in bcl silica are lamellas, which are concentric and interconnected. This structure provides easily accessible active sites compared to conventional structures, and the internal particles of bcl silica consist of lamellas that are concentrically structured and evenly distributed in all directions.[65]

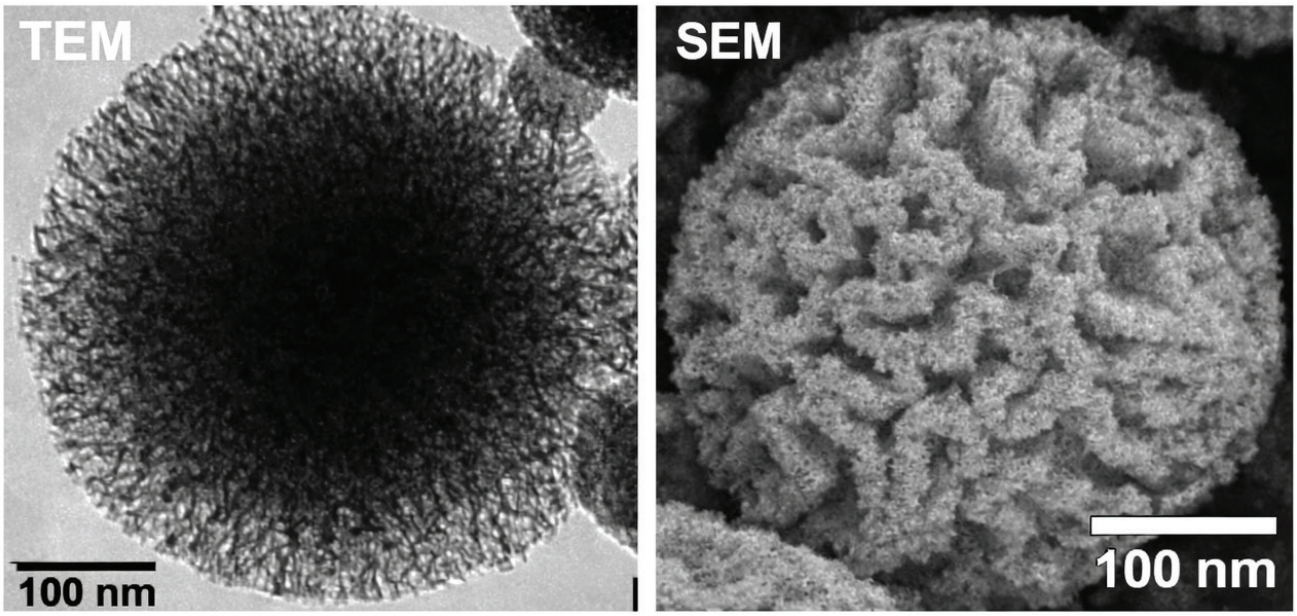


Figure 4. *SEM and TEM of KCC-1 showing bicontinuous concentric. (Adapted with permission from Febriyanti E, et al.[18] Copyright 2016 American Chemical Society).*

The reverse micelle plays an important role in the formation of bcl silica. The stages of bcl silica formation, which start from the hydrolysis process, are followed by the diffusion of hydrolysis products into the reverse micelle, before ending with the condensation and phase segregation process. The formation mechanism is proposed as a hydrolysis reaction occurring outside the inverted micelles in the vicinity of the interface. In contrast, the condensation reaction begins just after the diffusion of hydrolysis products into the inverted micelles, followed by a phase separation process. Particle growth during synthesis was observed, and showed that during the first 30 minutes of synthesis, bicontinuous lamellar features (bl) of bcl silica were formed. The particles then progressed to larger sizes with similar bl morphology and spherical shape. After the synthesis runs 1 to 2 hours, concentric features of bcl silica appear as a solid material. The change of lamellar orientation in bcl silica (i.e., parallel or perpendicular orientation) is a function of the density of the 232 lamellar phases.[18,65,66]

A schematic representation of the formation of the two morphologies (bl and bcl) is presented in Figure 5a and b. The bl morphology results from thin slabs of lamellar polysiloxane stacks evolving into perpendicular lamellar structures (Figure 5, part II, routes 4-a-bc-5). These structures are considered to have the bl morphology when we apply the topological transformation to the sphere. On the other hand, the bcl morphology is a product of thick slabs of lamellar polysiloxane stacks with complex orientations (Figure 5, part III, route 4-a-b'-c'-d'-6). Thin slabs of lamellar polysiloxane stacks will form perpendicular lamellar orientation and turn into bl morphology (reversible route 4-a-b-c-5), while thick slabs of lamellar polysiloxane stacks will turn into complex lamellar orientation and turn into bcl morphology (reversible route 4-a-b'-c'-d'-6). Based on the explanation of the advantages of the morphological shape formed, this material has the potential to be developed and utilized in drug delivery system applications. However, as far as this article is written, there has not been any utilization of this bcl morphology term in drug delivery system applications.

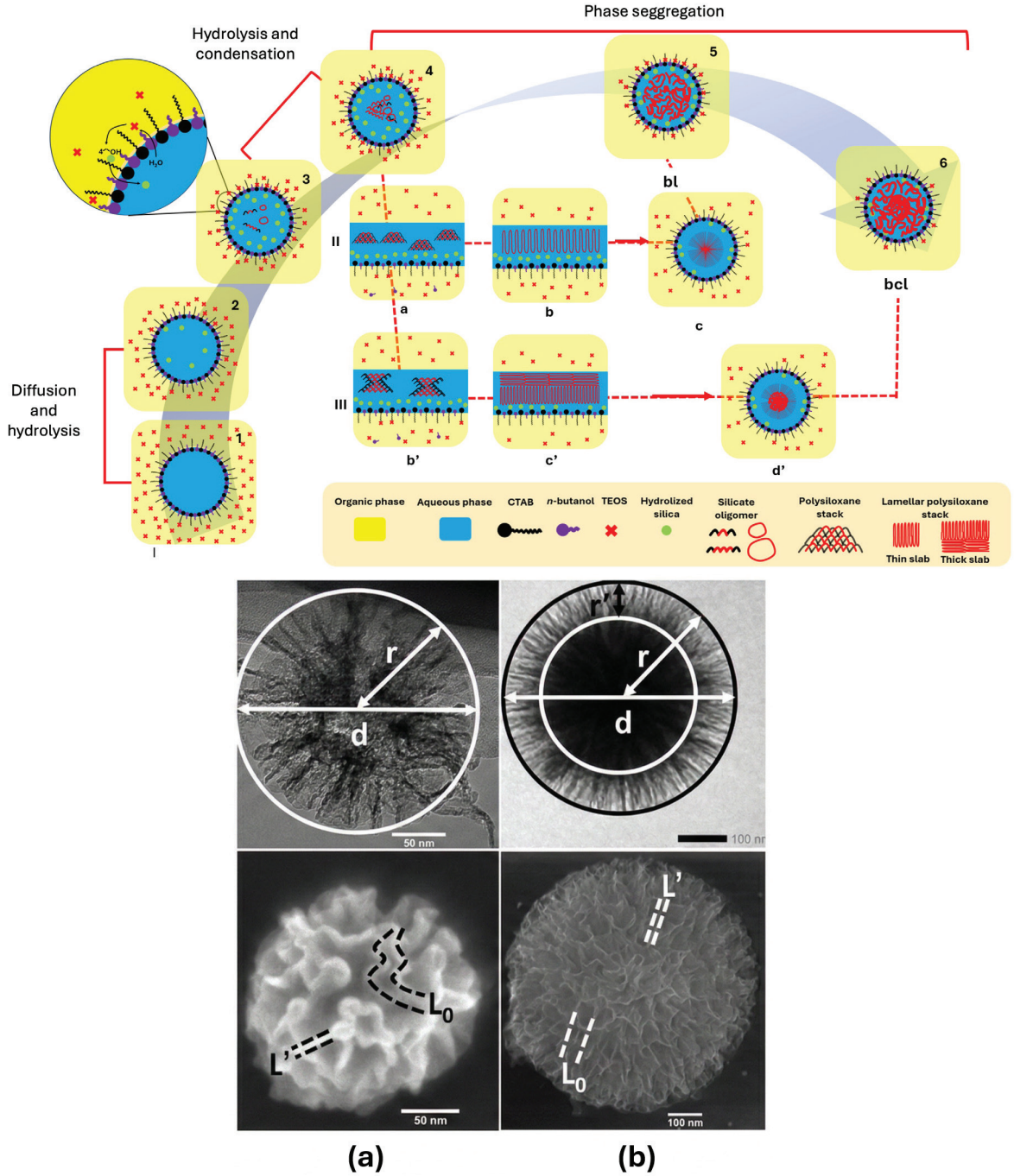


Figure 5. *Growth pathways of bcl silica and schematic description of the assigned morphological parameters of the lamella model on real bcl silica particles. Upper: growth pathway with I: irreversible, II-III: reversible (dashed red line) routes, representing the thermodynamic, Lower: schematic description of the assigned morphological parameters of the lamella model on real bcl silica particles (a: bicontinuous lamellar, bl, b: bicontinuous concentric lamellar, bcl) morphologies with perpendicular and complex orientations of lamellae. (Adapted with permission from Febriyanti E, et al.*[66] *Copyright 2022 American Chemical Society).*

Particle size

Hydrolysis and condensation of TEOS are two important reactions in the synthesis of silica-based materials. In the research conducted that controlled the particle size by Taguchi method-based optimization, the reaction was carried out concerning three factors: 1) the amount of TEOS, 2) the pH value, and 3) the reaction time, to control the size of the silica material.[67] The results show that the pH value greatly affects the mesostructural order and particle size. Longer reaction times under alkaline conditions have little effect on the structural order but lead to a decrease in the silica framework, resulting in smaller particle sizes. The more TEOS slightly increases the particle size of the synthesized silica. The basic nature of the catalyst is also considered when forming silica materials by varying the concentration of aqueous urea in the range of 0.08-2.0 M, and a certain amount of sodium

hydroxide or hydrochloric acid is added or used only to determine the effect of the catalyst component.[56] In the synthesis by the hydrothermal method,[31] the average particle size increased with increasing reaction time from 201 nm (for 2 hours) to 260 nm (for 6 hours) and then to 311 nm (for 9 hours), but did not increase with a further increase in reaction time (297 nm for 12 hours). The use of precursors also affects the size of the particles formed. When the amount of TEOS was 50 mL, DFNSs with a large particle size (715±106 nm) having very dense fibers were obtained. Increasing the amount of TEOS to 100 and 150 mL, there was a systematic decrease in the particle size to 260±11 nm and 118±16 nm, respectively. In drug delivery applications, in tests conducted,[24] smaller particles showed much deeper penetration. So far, there have been no further studies on the effect of particle size on lamellar silica in drug delivery applications.

Zeta potential

The zeta potential value indicates the surface charge value of a material. The higher the zeta potential value (negative or positive charge), the more stable material. In drug delivery applications, nanoparticles with a zeta potential greater than ±30 mV are considered strong cationic and strong anionic. In general, nanoparticles with higher zeta potential have stronger electrostatic repulsion, which can help stabilize them and prevent them from aggregating.[68] Surface modification can also affect the zeta potential value and its performance in drug loading.[42] Silica material has a negative zeta potential, and its nature will attract DOX molecules containing amino groups, which facilitates DOX loading. However, when the surface is modified with amino groups, the attractive effect will be weakened, resulting in a decrease in DOX loading. In addition, the functional groups modified on the surface have certain spatial site resistance, which will cause blockage of the pore channels, and this is also the reason for the decrease in DOX loading capacity.

Drug loading has also been reported[48] to affect the value of zeta potential, which can be inferred from the variation of particle zeta potential after hydrophobic PTX loading. Bare dendritic silica has a negatively charged surface, with a zeta potential of -24.8 mV in deionized water. The zeta potential of particles after 10 wt% and 50 wt% PTX loading increased to -20.2 mV and -3.77 mV, respectively. This is due to the shielding of negatively charged silicon hydroxyl groups by uncharged PTX molecules loaded on the particle surface. In contrast, the zeta potential of PTX-loaded dendritic silica through conventional impregnation loading only slightly increased to -23.5 mV, indicating that most of the silicon hydroxyl groups on DMSN were still exposed.

Conclusion

Lamellar silica-based materials with various unique properties of their texture properties show great potential in drug delivery applications and are expected to be used in the future.

Ethical approval

During the preparation of this work, the author used AI-assisted methods in order to increase the quality of some of the images in the manuscript without losing the originality of the data.

Funding

Penelitian Tesis Magister (No.1280.1/UN27.22/PT.01.03/2023).

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Dyah Ellyawati Kusumaningtyas Maharani: conceptualization, literature review, investigation, project administration, writing: original draft, review and edit; **Fajar Rakhman Wibowo**: review and edit, supervisor; **Rino Rakhmata Mukti**: supervisor, correspondence.

Acknowledgements

The authors would like to acknowledge the Directorate General of Higher Education, Ministry of Education, Culture, Research and Technology, Indonesia for funding this research through Penelitian Tesis Magister (No.1280.1/UN27.22/PT.01.03/2023).

References

[1] Esmaeili Y, Dabiri A, Mashayekhi F, Rahimmanesh I, Bidram E, Karbasi S, et al. Smart co-delivery of plasmid DNA and doxorubicin using MCM-chitosan-PEG polymerization functionalized with MUC-1 aptamer against breast cancer. Biomed Pharmacother. 2024; 173: 116465. doi: 10.1016/j.biopha.2024.116465.

[2] Septian Dwitya S, Lin K-S, Weng M-T, Vukile Mdlovu N, Yang M-T, Wu C-M. Synthesis and characterization of oleic acid-stabilized cobalt ferrite @MCM-41/nanocomposites for pH-responsive drug delivery. J Ind Eng Chem. 2024; doi: 10.1016/j.jiec.2024.08.036.

[3] Patel A, Mehta S. Functionalized SBA-15: engineering, detailed study on release and kinetics of alendronate as well as its anti-tumour properties towards osteosarcoma. RSC Pharm. 2024; 1(4): 797– 805. doi: 10.1039/D4PM00078A.

[4] Atiyah NA, Albayati TM, Atiya MA. Interaction behavior of curcumin encapsulated onto functionalized SBA-15 as an efficient carrier and release in drug delivery. J Mol Struct. 2022; 1260: 132879. doi: 10.1016/j.molstruc.2022.132879.

[5] Nguyen H-VT, Mai NXD, Thi Lai H, Nguyen LHT, Pham QT, Van Nguyen H, et al. Targeted, pH responsive quercetin delivery via hyaluronic acid-functionalized mesoporous nanoparticles

for cancer treatment. Colloids Surfaces A Physicochem Eng Asp. 2024; 135914. doi: 10.1016 /j.colsurfa. 2024. 135914.

[6] Lestari WA, Saputra OA, Díaz-García D, Wahyuningsih S, Gómez-Ruiz S, Wibowo FR. Design of experiments as a strategy for modulating the colloidal stability, physico-chemical properties and drugdelivery potential of small mesoporous silica nanoparticles. Adv Powder Technol. 2023; 34(10): 104191. doi: 10.1016/j.apt.2023.104191.

[7] Yang J, Liu B, Wang Q, Yan H, Li G, Wang X, et al. Carboxylated mesoporous silica nanoparticle-nucleic acid chimera conjugate-assisted delivery of siRNA and doxorubicin effectively treat drug-resistant bladder cancer. Biomed Pharmacother. 2024; 178: 117185. doi: 10.1016/j. biopha.2024.117185.

[8] Polshettiwar V. Dendritic Fibrous Nanosilica: Discovery, synthesis, formation mechanism, catalysis, and CO_2 capture-conversion. Acc Chem Res. 2022; 55(10): 1395-410. doi: 10.1021/acs. accounts.2c00031.

[9] Wang Y, Nor YA, Song H, Yang Y, Xu C, Yu M, et al. Small-sized and large-pore dendritic mesoporous silica nanoparticles enhance antimicrobial enzyme delivery. J Mater Chem B. 2016; 4(15): 2646-53. doi: 10.1039/C6TB00053C.

[10] Shen D, Yang J, Li X, Zhou L, Zhang R, Li W, et al. Biphase stratification approach to three-dimensional dendritic biodegradable mesoporous silica nanospheres. Nano Lett. 2014; 14(2): 923-32. doi: 10.1021/nl404316v.

[11] Yang X, Quan K, Wang J, Liu J, Liu B, Chen J, et al. Particle size and pore adjustment of dendritic mesoporous silica using different long alkyl-chain imidazolium ionic liquids as templates. Microporous Mesoporous Mater. 2022; 345: 112249. doi: 10.1016/j.micromeso.2022.112249.

[12] Sun Z, Cui G, Li H, Tian Y, Yan S. Multifunctional dendritic mesoporous silica nanospheres loaded with silver nanoparticles as a highly active and recyclable heterogeneous catalyst. Colloids Surfaces A Physicochem Eng Asp. 2016; 489: 142-53. doi: 10.1016/j.colsurfa.2015.10.052.

[13] Hao P, Peng B, Shan B-Q, Yang T-Q, Zhang K. Comprehensive understanding of the synthesis and formation mechanism of dendritic mesoporous silica nanospheres. Nanoscale Adv. 2020; 2(5): 1792- 810. doi: 10.1039/D0NA00219D.

[14] Maity A, Polshettiwar V. Dendritic fibrous nanosilica for catalysis, energy harvesting, carbon dioxide mitigation, drug delivery, and sensing. ChemSusChem. 2017; 10(20): 3866-913. doi: 10. 1002/cssc.201701076.

[15] Cocuzza C, Vincenzi C, Ottone C, Illanes A, Fino D, Cauda V, et al. Synthesis and characterization of mesoporous silicas with dendritic and spongy-like structures: Potential supports for human lactate dehydrogenase-based microreactors aimed at anticancer inhibitor screening. Microporous Mesoporous Mater. 2024; 376: 113182. doi: 10. 1016/j.micromeso.2024.113182.

[16] Li X, Yang S, Luan Y, Wang D, Du X. Polydopamine modification on dendritic porous silica surface for efficient adhesion of functional nanoparticles. Colloids Surfaces A Physicochem Eng Asp. 2024; 703: 135439. doi: 10.1016/j.colsurfa. 2024.135439.

[17] Ovejero-Paredes K, Díaz-García D, Mena-Palomo I, Marciello M, Lozano-Chamizo L, Morato YL, et al. Synthesis of a theranostic platform based on fibrous silica nanoparticles for the enhanced treatment of triple-negative breast cancer promoted by a combination of chemotherapeutic agents. Biomater Adv. 2022; 137: 212823. doi: 10.1016/j.bioadv.2022.212823.

[18] Febriyanti E, Suendo V, Mukti RR, Prasetyo A, Arifin AF, Akbar MA, et al. Further insight into the definite morphology and formation mechanism of mesoporous silica KCC-1. Langmuir. 2016; 32(23): 5802-11. doi: 10.1021/acs.langmuir.6b00675.

[19] Pérez-Moreno AM, Aranda CJ, Torres MJ, Mayorga C, Paris JL. Immunomodulatory potential of rapamycin-loaded mesoporous silica nanoparticles: pore size-dependent drug loading, release, and in vitro cellular responses. Drug Deliv Transl Res. 2024; 14(12): 3467-76. doi: 10.1007/s13346-0240 1575-0.

[20] Farzan M, Roth R, Schoelkopf J, Huwyler J, Puchkov M. The processes behind drug loading and release in porous drug delivery systems. Eur J Pharm Biopharm. 2023; 189: 133-51. doi: 10.1016/j.ejpb. 2023.05.019.

[21] Ly P-D, Ly K-N, Phan H-L, Nguyen HHT, Duong V-A, Nguyen H V. Recent advances in surface decoration of nanoparticles in drug delivery. Front Nanotechnol. 2024; 6. doi: 10.3389/fnano. 2024.1456939.

[22] Mok ZH. The effect of particle size on drug bioavailability in various parts of the body. Pharm Sci Adv. 2024; 2: 100031. doi: 10.1016/j.pscia.2023. 100031.

[23] Farahmandghavi F, Imani M, Hajiesmaeelian F. Silicone matrices loaded with levonorgestrel particles: Impact of the particle size on drug release. J Drug Deliv Sci Technol. 2019; 49: 132-42. doi: 10.1016/j.jddst.2018.10.029.

[24] Islam MA, Barua S, Barua D. A multiscale modeling study of particle size effects on the tissue penetration efficacy of drug-delivery nanoparticles. BMC Syst Biol. 2017; 11(1): 113. doi: 10.1186/ s12918-017-0491-4.

[25] Honary S, Zahir F. Effect of zeta potential on the properties of nano-drug delivery systems - A review (Part 1). Trop J Pharm Res. 2013; 12(2). doi: 10. 4314/tjpr.v12i2.19.

[26] Feng X, Jia P, Zhang D. Nanocarrier drug delivery system: promising platform for targeted depression therapy. Front Pharmacol. 2024; 15.

doi: 10.3389/fphar.2024.1435133.

[27] Naik SP, Elangovan SP, Okubo T, Sokolov I. Morphology control of mesoporous silica particles. J Phys Chem C. 2007; 111(30): 11168-73. doi: 10.1021/jp072184a.

[28] Atluri R, Jensen KA. Engineered nanomaterials: Their physicochemical characteristics and how to measure them. MODENA. 2017: pp3-23. doi: 10.1007/978-3-319-47754-1_1.

[29] Knežević NŽ, Durand J-O. Large pore mesoporous silica nanomaterials for application in delivery of biomolecules. Nanoscale. 2015; 7(6): 2199-209. doi: 10.1039/C4NR06114D.

[30] Polshettiwar V, Cha D, Zhang X, Basset JM. High-surface-area silica nanospheres (KCC-1) with a fibrous morphology. Angew Chemie Int Ed. 2010; 49(50): 9652-6. doi: 10.1002/anie.201003451.

[31] Maity A, Polshettiwar V. Scalable and sustainable synthesis of size-controlled monodisperse dendritic fibrous nanosilica quantified by E-factor. ACS Appl Nano Mater. 2018; 1(7): 3636-43. doi: 10.1021/acsanm.8b00761.

[32] Megantari V, Febriyanti E, Benu DP, Reza M, Steky F V., Yuliarto B, et al. Nanofluidic behavior of diatomic molecules in bicontinuous concentric lamellar (*bcl*) silica formed by polysiloxane sol-gel phase segregation as a reference in the mass transport through the open channel system. Polym Technol Mater. 2020; 59(12): 1359-69. doi: 10.1080/25740881.2020.1738478.

[33] Guo Z, Wu L, Wang Y, Zhu Y, Wan G, Li R, et al. Design of dendritic large-pore mesoporous silica nanoparticles with controlled structure and formation mechanism in dual-templating strategy. ACS Appl Mater Interfaces. 2020; 12(16): 18823-32. doi: 10.1021/acsami.0c00596.

[34] Yin S, Xu C, Yang H, Wu C, Wu M, Xu J, et al. Facilely preparing highly dispersed Ni-based CO2 methanation catalysts via employing the amino-functionalized KCC-1 support. Fuel. 2024; 365: 131162. doi: 10.1016/j.fuel.2024.131162.

[35] Lai LW, Teh LP, Timmiati SN, Kamarudin NHN, Setiabudi HD. A sustainable solution for diclofenac adsorption: Chitosan-modified fibrous silica KCC-1 adsorbent. J Environ Chem Eng. 2023; 11(6): 111295. doi: 10.1016/j.jece.2023.111295.

[36] Silmi N, Febriyanti E, Andriani A, Arsyad R, Steky FV, Mukti RR, et al. Textural control of bicontinuous concentric lamellar nanostructured mesoporous silica via low-pressure solvothermal method. Mater Chem Phys. 2021; 265: 124492. doi: 10.1016/j.matchemphys.2021.124492.

[37] Kole K, Halder A, Singh S, Samanta A, Das S, Kundu AK, et al. Chromogenic-functionalized silica nanoflower composites for the detection of carbon dioxide. ACS Appl Nano Mater. 2020; 3(5): 4321-8. doi: 10.1021/acsanm.0c00469.

[38] Palanichamy K, Umasankar S, Ganesh S, Sasirekha N. Highly coke resistant Ni–Co/KCC-1 catalysts for dry reforming of methane. Int J Hydrogen Energy. 2023; 48(31): 11727-45. doi: 10.1016/j.ijhydene.2022.12.076.

[39] Thommes M, Kaneko K, Neimark A V., Olivier JP, Rodriguez-Reinoso F, Rouquerol J, et al. Physisorption of gases, with special reference to the evaluation of surface area and pore size distribution (IUPAC Technical Report). Pure Appl Chem. 2015; 87(9–10): 1051-69. doi: 10.1515/pac-2014-1117.

[40] Latifi L, Sohrabnezhad S. Influence of pore size and surface area of mesoporous silica materilas (MCM41 and KIT-6) on the drug loading and release. J Sol-Gel Sci Technol. 2018; 87(3): 626-38. doi: 10.1007/s10971-018-4742-7.

[41] Jia L, Shen J, Li Z, Zhang D, Zhang Q, Liu G, et al. In vitro and in vivo evaluation of paclitaxel-loaded mesoporous silica nanoparticles with three pore sizes. Int J Pharm. 2013; 445(1-2): 12-9. doi: 10.1016/j.ijpharm.2013.01.058.

[42] Peng S, Huang B, Lin Y, Pei G, Zhang L. Effect of surface functionalization and pore structure type on the release performance of mesoporous silica nanoparticles. Microporous Mesoporous Mater. 2022; 336: 111862. doi: 10.1016/j.micromeso.2022.111862.

[43] Bavnhøj CG, Knopp MM, Madsen CM, Löbmann K. The role interplay between mesoporous silica pore volume and surface area and their effect on drug loading capacity. Int J Pharm X. 2019; 1: 100008. doi: 10.1016/j.ijpx.2019.100008.

[44] Mujtaba M, Negi A, King AWT, Zare M, Kuncova-Kallio J. Surface modifications of nanocellulose for drug delivery applications; a critical review. Curr Opin Biomed Eng. 2023; 28: 100475. doi: 10.1016/j.cobme.2023.100475.

[45] Wang J, Wang Y, Liu Q, Yang L, Zhu R, Yu C, et al. Rational design of multifunctional dendritic mesoporous silica nanoparticles to load curcumin and enhance efficacy for breast cancer therapy. ACS Appl Mater Interfaces. 2016; 8(40): 26511-23. doi: 10.1021/acsami.6b08400.

[46] Pishnamazi M, Hafizi H, Pishnamazi M, Marjani A, Shirazian S, Walker GM. Controlled release evaluation of paracetamol loaded amine functionalized mesoporous silica KCC1 compared to microcrystalline cellulose based tablets. Sci Rep. 2021; 11(1): 535. doi: 10.1038/s41598-020-79983-8.

[47] Maity A, Das A, Sen D, Mazumder S, Polshettiwar V. Unraveling the formation mechanism of dendritic fibrous nanosilica. Langmuir. 2017; 33(48): 13774-82. doi: 10.1021/acs.langmuir.7b02996.

[48] Deng C, Liu Y, Zhou F, Wu M, Zhang Q, Yi D, et al. Engineering of dendritic mesoporous silica nanoparticles for efficient delivery of water-insoluble paclitaxel in cancer therapy. J Colloid Interface Sci. 2021; 593: 424-33. doi: 10.1016/j.jcis.2021.02.098.

[49] Ma JZ, Sun HY, Zhang Y, Chen DW, Hu HY. Fabrication of epidermal growth factor imprinted and demethylcantharidin loaded dendritic mesoporous silica nanoparticle: An integrated drug vehicle for chemo-/antibody synergistic cancer therapy. J Drug Deliv Sci Technol. 2021; 62: 102387. doi:10.1016/j.jddst.2021.102387.

[50] Xu Z, Keller E, Beltrán AM, Zheng K, Boccaccini AR. Cerium doped dendritic mesoporous bioactive glass nanoparticles with bioactivity and drug delivery capability. J Non Cryst Solids. 2023; 620: 122578. doi: 10.1016/j.jnoncrysol.2023.122578.

[51] Pan Y, Zhou S, Li Y, Parshad B, Li W, Haag R. Novel dendritic polyglycerol-conjugated, mesoporous silica-based targeting nanocarriers for co-delivery of doxorubicin and tariquidar to overcome multidrug resistance in breast cancer stem cells. J Control Release. 2021; 330: 1106-17. doi: 10.1016/j.jconrel.2020.11.015.

[52] Wu L, Lv R, Wang H, Gao S, Wang Y, Zhang S, et al. Facile fabrication of glutathione-responsive and photothermal nanocarriers with dendritic mesoporous silica nanoparticles for the controlled drug delivery. J Nanoparticle Res. 2022; 24(7): 128. doi: 10.1007/s11051-022-05498-0.

[53] Ye M, Zhang W, Xu H, Xie P, Song L, Sun X, et al. Fe-doped biodegradable dendritic mesoporous silica nanoparticles for starvation therapy and photothermal-enhanced cascade catalysis in tumor therapy. J Colloid Interface Sci. 2025; 678: 378-92. doi: 10.1016/j.jcis.2024.08.172.

[54] Song S, Li X, Ji Y, Lv R, Wu L, Wang H, et al. GSH/pH dual-responsive and HA-targeting nano-carriers for effective drug delivery and controlled release. J Drug Deliv Sci Technol. 2021; 62: 102327. doi: 10.1016/j.jddst.2021.102327.

[55] Pota G, Andrés-Sanz D, Gallego M, Vitiello G, López-Gallego F, Costantini A, et al. Deciphering the immobilization of lipases on hydrophobic wrinkled silica nanoparticles. Int J Biol Macromol. 2024; 266: 131022. doi: 10.1016/j.ijbiomac.2024.131022.

[56] Moon D-S, Lee J-K. Tunable synthesis of hierarchical mesoporous silica nanoparticles with radial wrinkle structure. Langmuir. 2012; 28(33): 12341-7. doi: 10.1021/la302145j.

[57] Moon D-S, Lee J-K. Formation of wrinkled silica mesostructures based on the phase behavior of pseudoternary systems. Langmuir. 2014; 30(51): 15574-80. doi: 10.1021/la504207k.

[58] Lo TNH, Park I. Unprecedented superoleophobicity achieved with fluorinated wrinkle mesoporous silica. Surfaces and Interfaces. 2024; 46: 104090. doi: 10.1016/j.surfin.2024.104090.

[59] Flood-Garibay JA, Méndez-Rojas MA. Synthesis and characterization of magnetic wrinkled mesoporous silica nanocomposites containing Fe_3O_4 or $CoFe_2O_4$ nanoparticles for potential biomedical applications. Colloids Surfaces A Physicochem Eng Asp. 2021; 615: 126236. doi: 10.1016/j.colsurfa.2021.126236.

[60] Flood-Garibay JA, Méndez-Rojas MA. A simple method for the synthesis of carboxymethylated wrinkled mesoporous silica nanoparticles and preparation of a WMS-curcumin conjugate. J Sol-Gel Sci Technol. 2022; 102(1): 288-95. doi: 10.1007/s10971-022-05747-7.

[61] Kole K, Das S, Samanta A, Jana S. Parametric study and detailed kinetic understanding of CO_2 adsorption over high-surface-area flowery silica nanomaterials. Ind Eng Chem Res. 2020; 59(49): 21393-402. doi: 10.1021/acs.iecr.0c04531.

[62] Das S, Samanta A, Jana S. Light-driven synthesis of uniform dandelion-like mesoporous silica nanoflowers with tunable surface area for carbon dioxide uptake. Chem Eng J. 2019; 374: 1118-26. doi: 10.1016/j.cej.2019.05.114.

[63] Yang H, Liao S, Huang C, Du L, Chen P, Huang P, et al. Facile one-pot approach to the synthesis of spherical mesoporous silica nanoflowers with hierarchical pore structure. Appl Surf Sci. 2014; 314: 7-14. doi: 10.1016/j.apsusc.2014.06.128.

[64] Zhang H, Li Z, Xu P, Wu R, Jiao Z. A facile two step synthesis of novel chrysanthemum-like mesoporous silica nanoparticles for controlled pyrene release. Chem Commun. 2010; 46(36): 6783. doi: 10.1039/c0cc01673j.

[65] Silmi N, Arsyad R, Benu DP, Nugroho FG, Khasannah WL, Iqbal M, et al. A morphological study of bicontinuous concentric lamellar silica synthesized at atmospheric pressure and its application as an internal micro-reflector in dye-sensitized solar cells. Phys Chem Chem Phys. 2023; 25(35): 23792–807. doi: 10.1039/D3CP02876C.

[66] Febriyanti E, Silmi N, Suendo V, Mukti RR, Vivitasari PU, Adhika DR, et al. Thermodynamic Picture of Phase Segregation during the Formation of Bicontinuous Concentric Lamellar (*bcl*) Silica. Langmuir. 2022; 38(4): 1368–79. doi: 10.1021/acs.langmuir.1c02490.

[67] Chiang Y-D, Lian H-Y, Leo S-Y, Wang S-G, Yamauchi Y, Wu KC-W. Controlling particle size and structural properties of mesoporous silica nanoparticles using the Taguchi method. J Phys Chem C. 2011; 115(27): 13158–65. doi: 10.1021/jp201017e.

[68] Öztürk K, Kaplan M, Çalış S. Effects of nanoparticle size, shape, and zeta potential on drug delivery. Int J Pharm. 2024; 666: 124799. doi: 10.1016/j.ijpharm.2024.124799.

Journal of Associated Medical Sciences 2026; 59(2): 339-347

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Long-term effects of a structured training program on physical fitness in male football players aged 18-22 years

Peerawats Dejprom[1], Sothida Nantakool[2], Busaba Chuatrakoon[2*]

[1]Multidisciplinary and Interdisciplinary School, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Integrated Neuro-Musculoskeletal, Chronic Disease, and Aging Research Engagement Center (ICARE Center), Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 27 June 2025
Accepted as revised 26 March 2026
Available online 30 March 2026

Keywords:
Structured training, football players, aerobic capacity, muscle strength, anaerobic power, VO_2max, detraining period

ABSTRACT

Background: Football performance relies on multiple physical capacities, including aerobic endurance, anaerobic power, muscle strength, speed, agility, and technical skills. Comprehensive training programs that integrate various conditioning elements may lead to multidimensional improvements in player performance. However, limited evidence exists regarding the long-term effects of such integrated programs, particularly whether physical performance gains can be sustained after training cessation.

Objectives: This study aims to evaluate the long-term effects of a one-year structured training program on physical fitness components in male university football players, including assessment of sustained outcomes after a six-month detraining period.

Materials and methods: A retrospective-prospective cohort study was conducted involving 15 male university-level competitive football players aged 18-22 years. Participants underwent a structured training regimen that incorporated aerobic, anaerobic, strength, speed, and football-specific drills. Physical fitness parameters were measured before the training period, immediately after one year of training, and again at six months post-training. Statistical analysis involved repeated-measures ANOVA and Bonferroni post hoc testing.

Results: The program significantly improved VO_2max, leg strength, sprint speed, and anaerobic power ($p<0.05$). However, these gains were not sustained at the six-month follow-up. Other measures, such as handgrip strength, flexibility, and vertical jump, did not show significant changes.

Conclusion: Structured multimodal training is effective in improving essential fitness attributes for football performance. However, without continued stimulus, performance gains diminish over time, suggesting the need for long-term maintenance strategies.

* **Corresponding contributor.**
Author's Address: *Integrated Neuro-Musculoskeletal, Chronic Disease, and Aging Research Engagement Center (ICARE Center), Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: busaba.c@cmu.ac.th
doi: 10.66285/JAMS.2025.065
E-ISSN: 2539-6056

Introduction

Football is a high-demand intermittent sport characterized by intermittent high-intensity bouts interspersed with lower-intensity recovery reriods.[1] During a standard 90-minute match, players perform more than 1,000 activity changes, including sprinting, jumping, tackling, and rapid changes in direction.[2,3] These actions impose substantial physiological and neuromuscular demands, requiring the integration of multiple components of physical fitness. Key components include aerobic endurance, anaerobic

capacity, muscular strength, speed, agility, flexibility, and technical skills, all of which contribute to overall match performance.[4] Aerobic capacity, commonly represented by maximal oxygen uptake (VO_2max), plays a fundamental role in sustaining performance and enhancing recovery between high-intensity efforts.[5,6] In contrast, anaerobic capacity supports short-duration, high-intensity actions that are critical for decisive moments during match play.[6] Muscular strength, particularly in the lower limbs, has also been shown to be closely associated with sprint performance, jump ability, and overall physical performance in football players.[7,8]

Previous studies have demonstrated that specific training modalities, including aerobic conditioning, resistance training, and small-sided games (SSGs), can effectively improve individual components of physical fitness in football players.[4,9,10] Aerobic training has been shown to enhance VO_2max and endurance capacity, while resistance and plyometric training improve muscular strength, power, and sprint performance.[2,7] SSGs have gained popularity as an effective training method that simultaneously develops technical skills and physical conditioning under game-like conditions.[4,9,10] However, football performance is inherently multifactorial, requiring the interaction of several physiological and neuromuscular systems. Therefore, integrated training programs that combine multiple training modalities may provide more comprehensive benefits compared to isolated training approaches.

Despite the growing body of literature, most existing studies have focused on short-term interventions or individual training components, with limited investigation into the long-term effects of structured, multimodal training programs. Furthermore, there is limited evidence on the sustainability of training-induced adaptations following detraining, particularly in university-level football players.[4,10] Understanding whether improvements in physical fitness can be maintained after training cessation is important for designing effective long-term training strategies and preventing performance decline during off-season periods.

Therefore, this study aimed to investigate the long-term effects of a comprehensive structured training program on multiple components of physical fitness in university-level football players and to examine the sustainability of these adaptations following a six-month detraining period.

Materials and methods

Study design

This study employed a retrospective–prospective cohort design to evaluate the long-term effects of a structured training program on physical fitness in university-level football players. Baseline assessments were conducted before the training program. The intervention period lasted for one year (August 2023 to July 2024), followed by a six-month follow-up period (January 2025) to examine detraining effects. During the follow-up phase, participants were not provided with a structured training program and were instructed to continue their usual daily activities without supervised training, allowing the follow-up assessment to reflect a natural detraining condition (Figure 1).

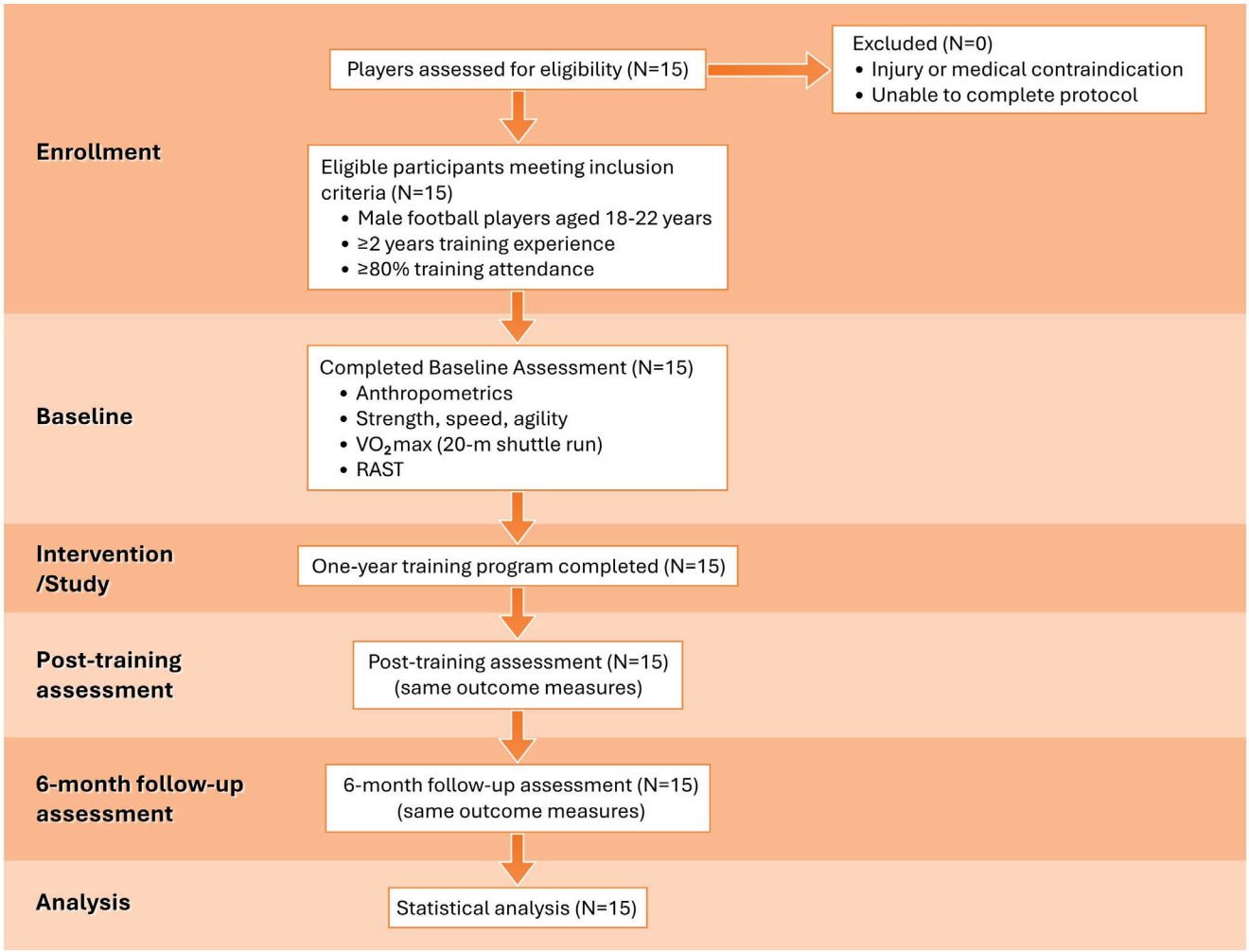


Figure 1. *Flow diagram of the study.*

Participants

Fifteen male university-level football players aged 18-22 years were recruited from the Chiang Mai university football team. All players who met the inclusion criteria were included; therefore, the sample represented the entire available team roster rather than being determined through an a priori sample size calculation. Participants had at least two years of continuous training experience and regularly participated in structured team training sessions approximately 3-5 days per week.

Participants were eligible for inclusion if they were male football players aged 18-22 years, actively registered with the university team, had at least two years of continuous training experience, and attended at least 80% of scheduled training sessions. Participants were excluded if they had any current musculoskeletal injury, cardiovascular or other medical conditions contraindicating high intensity exercise, or were unable to complete the full training protocol. Withdrawal criteria included absence from more than 20% of training sessions or injury during the study period.

All eligible participants completed the full training program and met the adherence criteria. This study was approved by the institutional ethics committee (AMSEC-67EX-040), and all participants provided written informed consent prior to participation.

Training protocol

Participants engaged in a structured training program lasting one year, developed under the supervision of a qualified sports scientist. The program was designed to enhance multiple components of physical fitness relevant to football performance, including aerobic capacity, sprint performance, muscular strength, flexibility, and sport-specific skills.

Aerobic training consisted of continuous running performed at moderate-to-high intensity (approximately 75-85% of maximum heart rate), conducted three times per week. Sprint training involved repeated 100-meter sprints alternated with jogging recovery, with progressive increases in training volume from 10 to 15 sets per session. Additional speed development was targeted using short-distance sprint repetitions (30 meters), performed in 5 to 8 sets per session, three times weekly.

Resistance training focused on lower-limb strength, using functional exercises such as squats and front lunges with a 4 kg medicine ball. Training was performed three times per week with 3-5 sets per

session. Flexibility training was incorporated three times per week and included dynamic stretching during the warm-up and static stretching during the cool-down. Dynamic stretching targeted major lower limb muscle groups using controlled movements, while static stretching was performed in standardized positions and held for 15-30 seconds, repeated 2-3 times per muscle group. In addition, participants performed football-specific drills, including ball possession, passing, attacking, and defending exercises, as well as SSGs (e.g., 3v3 and 5v5), for approximately 30-45 minutes per session.

Training progression was implemented according to the principle of progressive overload, with gradual increases in training volume, repetition number, and exercise complexity throughout the intervention period. Aerobic duration and sprint repetitions were progressively increased, while resistance training volume was adjusted across the season. Football-specific drills evolved from basic technical drills to more complex tactical scenarios. All training sessions were supervised by a qualified sports scientist and coaching staff. Training attendance was systematically recorded to monitor adherence, and participants were required to attend at least 80% of sessions to be included in the analysis (Table 1).[4,10,11]

Table 1. Structured one-year training program.

Component	Frequency	Intensity	Duration	Description	Progression
Flexibility training	3 sessions/week	Low intensity	10-15 minutes	Dynamic (pre-training) and static (post-training) stretching	Maintained throughout season
Aerobic training	3 sessions/week	75–85% HRmax	30- 45 minutes	Continuous running	Gradual increase in duration and intensity over season
Interval sprint training	3–4 sessions/week	High intensity	10-15 sets	100-m repeated sprints alternated with jogging	Sets progressed from 10 to 15 per session
Resistance training	3 sessions/week	Moderate load	3-5 sets/session	Squats and front lunges using 4-kg medicine ball	Gradual increase in training volume
Speed training	3 sessions/week	High intensity	5-8 sets	30-m sprint repetitions	Increased repetitions over time
Football-specific skill training	Daily	Sport-specific	30-45 minutes	Ball possession, passing, attacking, defending (1v1, 2v2), small-sided games (3v3, 5v5)	Increased tactical complexity and game intensity

Outcome measures

Physical fitness assessments were conducted at three time points: baseline (prior to training), post-training (immediately after completion of the one-year training program), and follow-up (six months after the training period). To ensure consistency and minimize fatigue effects, all assessments were conducted in a standardized sequence for all participants and at all time points.

Anthropometric measurements were obtained using a bioelectrical impedance analyzer (InBody, Biospace Co., Ltd., Seoul, Korea) and included body weight, body fat percentage, and body mass index (BMI). Flexibility was assessed using the sit-and-reach test, with the best of two trials recorded.[12] Explosive power was evaluated using a countermovement vertical jump test, and the highest value from two trials was used for analysis.[13] Muscle strength was assessed using a handgrip dynamometer (JAMAR, USA) and lower-body strength using a leg press test normalized to body weight.[14] Speed and agility were assessed using a 40-m sprint test and a shuttle run test, respectively, with the fastest performance from two trials recorded.[15,16]

Aerobic capacity was estimated using the 20-m multistage fitness test,[17] from which maximal oxygen uptake (VO_2max) was calculated based on the final stage achieved using standardized reference equations: VO_2max = (age factor × max oxygen uptake × 1,000) / body weight (kg)

Anaerobic performance was assessed using the Running-Based Anaerobic Sprint Test (RAST). Participants performed six maximal 35-m sprints separated by 10-s recovery intervals, and sprint time for each repetition was recorded.[18] Power output for each sprint was calculated as:

$$\text{Power} = (\text{body mass} \times (\text{distance})^2) / (\text{time})^3$$

From the six sprint trials, peak power, minimum power, mean power, and fatigue index (FI) were

calculated. The fatigue index was determined using the following equation:

FI = (peak power – minimum power) / total sprint time

All measurements were conducted by trained evaluators under the supervision of a qualified sports scientist, and the same assessor performed all measurements across all time points to minimize inter-rater variability. Participants were familiarized with all testing procedures prior to data collection, and standardized instructions and rest intervals were applied across testing sessions.

Training attendance was recorded by the coaching staff for each session to monitor exercise adherence. Participants were required to attend at least 80% of the scheduled training sessions to be included in the analysis. All participants met this criterion and completed the full training program.

Statistical analysis

All statistical analyses were performed using SPSS software (version 21.0, IBM Corporation, Chicago, IL). Descriptive statistics, including means and standard deviations, were calculated for all continuous variables. The normality of data distribution was assessed using the Shapiro–Wilk test due to the relatively small sample size. Since the data were approximately normally distributed, repeated-measures analysis of variance was conducted to examine changes in physical fitness outcomes over time. Where significant main effects were identified, Bonferroni-adjusted pairwise comparisons were performed to determine differences between time points (baseline, post-training, and follow-up). The assumption of sphericity was assessed using Mauchly's test, and the Greenhouse–Geisser correction was applied when necessary. Statistical significance was set at $p<0.05$.

Results

Participants' demographic and physical characteristics, including age, body composition, and baseline physical fitness measures, were collected before the training intervention to establish a reference point for evaluating program effects. These baseline data are summarized in Table 2.

Table 2. Baseline characteristics of male university football players.

Variables	Mean±SD
Age (year)	19.73±0.96
Weight (kg)	73.52±11.84
Height (cm)	174.41±6.27
Position (N)	15
Goalkeeper	1
Center back	3
Right back	2
Left back	2
Defensive midfielder	2
Attacking midfielder	3
Left winger	1
Right winger	1

Note: cm: centimeter, kg: kilogram.

The physical fitness profiles of participants, including body composition, flexibility, muscular strength, explosive power, agility, sprint speed, aerobic capacity, and anaerobic performance, were assessed at three time points. Mean and standard deviation values at baseline, after one year of training, and at the six-month follow-up are summarized to provide a clear overview of participants' physical status throughout the study (Table 3).

Table 3. Mean and standard deviation of physical fitness test scores among male university football players.

Variables	Baseline (N=15)	Post-training (N=15)	Follow-up (N=15)
Body mass index (BMI) (kg/m^2)	22.48±2.06	23.02±2.28	23.56±2.98
Body fat (%)	17.09±4.69	17.86±4.26	17.31±4.06
Sit and reach (cm)	18.67±6.96	17.00±6.20	17.27±6.22
Grip strength (kg/body weight)	0.61±0.10	0.60±0.10	0.61±0.10
Leg strength (kg/body weight)	2.21±0.60	2.50±0.46	2.26±0.39
Vertical jump (cm)	50.27±5.89	50.00±5.68	51.07±6.70
Shuttle run test (second)	10.12±0.50	9.97±0.56	10.50±0.86
Sprint 40-m test (second)	6.05±0.32	5.77±0.49	6.49±0.62
Multistage fitness test (VO_2max) (mL/min/kg)	43.68±5.82	46.39±5.70	41.62±9.16
RAST test (watt/kg)	38.25±6.51	28.52±7.72	42.91±7.97

Note: BMI: body mass index, cm: centimeter, m^2: square meter, VO_2max: maximal oxygen consumption.

Improvements in leg strength, sprint performance, VO_2max, and anaerobic power were observed following the training program, with statistically significant differences between baseline and post-training assessments. At the six-month follow-up, most outcomes that improved post-training were no longer significantly different from baseline, indicating partial loss of training- induced adaptations. However, 40-m sprint performance remained significantly improved compared with baseline at follow-up, despite a significant decline from post-training. No significant changes over time were observed in flexibility, handgrip strength, vertical jump, body fat percentage, or BMI (Table 4).

Table 4. Comparison of physical fitness measures at baseline (a), post-training (b), and follow-up (c) in male university football players (N=15).

Variables	b-a	*p* value	c-a	*p* value	c-b	*p* value	F	df	*p* (ANOVA)
Body mass index (BMI) (kg/m^2)	0.535	0.080	1.079	0.072	0.545	0.360	5.191	(2, 28)	0.012*
Body fat (%)	0.773	0.901	0.220	0.892	-0.553	1.000	0.497	(2,28)	0.613
Sit and reach (cm)	-1.667	0.101	-1.400	0.572	0.267	1.000	2.551	(1.39, 19.43)	0.118
Grip strength (kg/body weight)	-0.005	1.000	-0.003	1.000	0.002	1.000	0.031	(1.17, 16.37)	0.895
Leg strength (kg/body weight)	0.290	0.005*	0.050	1.000	-0.240	0.096	2.926	(1.38, 19.37)	0.093
Vertical jump (cm)	-0.267	1.000	0.800	0.937	1.067	0.203	2.692	(2, 28)	0.085
Shuttle run test (second)	-0.155	0.663	0.381	0.140	0.536	0.097	3.994	(2, 28)	0.030*
Sprint 40-m test (second)	-0.279	0.023*	0.443	0.018*	0.723	0.005*	10.412	(2, 28)	<0.001*
Multistage fitness test (VO_2max) (ml/kg/min)	2.713	0.013*	-2.060	0.658	-4.773	0.005*	2.769	(2, 28)	0.080
RAST test (watt/kg)	-9.733	0.005*	4.660	0.163	14.393	0.002*	6.647	(2, 28)	0.004

Note: BMI: body mass index, cm: centimeter, m^2: square meter, VO_2max: maximal oxygen consumption. Values are presented as mean differences between time points; a: baseline, b: post-training, c: follow-up. Statistical analysis was performed using repeated-measures ANOVA with Bonferroni-adjusted pairwise comparisons. Sphericity was assessed using Mauchly's test, and Greenhouse–Geisser correction was applied when assumptions were violated.,*statistical significance was set at $p<0.05$.

Discussion

The present study investigated the long-term effects of a structured training program on multiple components of physical fitness and examined the sustainability of these adaptations following a detraining period. The findings demonstrated that the one-year training program significantly improved aerobic capacity, lower-body strength, sprint performance, and anaerobic power. However, these improvements were not maintained after six months of detraining, indicating that continued training stimulus is required to sustain performance adaptations in university-level football players. These findings highlight the importance of continuous training and structured maintenance strategies to preserve physical performance in football players.

The significant increase in VO_2max observed post-training aligns with previous research emphasizing the importance of aerobic endurance for football performance. Aerobic capacity supports prolonged activity and facilitates recovery during intermittent high-intensity efforts typical of match play. The training protocol in this study included moderate-to-high-intensity continuous running and interval sprints, which likely contributed to the observed improvement. This finding is supported by Helgerud et al.,[11] who found that high-intensity aerobic training significantly enhances VO_2max and overall match-related endurance. Recent studies have also reported improvements in aerobic capacity and repeated sprint performance following structured training programs in football players.[19] Sprint performance and anaerobic power also improved significantly. These elements are essential for explosive actions such as accelerations, tackles, and transitions during matches. The inclusion of short-distance sprint drills, plyometric movements, and anaerobic running intervals likely contributed to these improvements.[20,21] RAST scores and 40-meter sprint times both reflected enhanced anaerobic output, supporting the view that repeated high-intensity bouts can lead to neuromuscular adaptations that benefit speed and power.[18] Lower body strength, as measured by leg press normalized to body weight, showed significant improvement after training. Resistance exercises such as squats and lunges were integrated into the weekly routine and performed consistently throughout the intervention period. Enhanced strength may also have contributed to better sprint performance and anaerobic output, indicating interdependence among physical fitness domains in football-specific activities.[20] The progressive overload applied throughout the one-year training program may also have contributed to these improvements. Gradual increases in training volume, sprint repetitions, and resistance training workload likely promoted physiological and neuromuscular adaptations that enhanced aerobic capacity, sprint performance, anaerobic power, and lower-body strength.

Despite these improvements, the six-month follow-up assessments indicated regression in all variables that had initially shown significant gains. Although values did not fall below baseline, they no longer differed significantly from pre-training levels. This suggests that the observed benefits were not sustained without continued training. The reversal of gains can be attributed to detraining, a phenomenon characterized by the partial or complete loss of training-induced adaptations when the training stimulus is reduced or withdrawn.[22-24] Mujika *et al*.[23] have reported that aerobic and strength performance can decline rapidly during periods of training cessation.

No significant changes were observed in flexibility, handgrip strength, vertical jump height, body fat percentage, or BMI across the study period. These results suggest that either the training program did not sufficiently target these components or that these variables are less sensitive to the type and volume of stimulus provided. For example, static flexibility and grip strength may require more specific or isolated interventions to elicit meaningful improvements.[25,26] Flexibility improvements may be achieved through structured stretching programs performed regularly throughout the training week. Dynamic stretching during warm-up and static stretching during cool-down sessions, performed for 10-15 minutes per session at least three times per week, may help improve joint mobility and muscle flexibility. In addition, grip strength and upper-body strength may benefit from specific resistance exercises such as pull-ups, farmer's carries, medicine ball throws, and handgrip training performed two to three times per week with progressive overload. These targeted interventions may complement football-specific training and help improve physical components that were not significantly affected in the present study.

The findings underscore the importance of consistent and ongoing conditioning for sustaining physical performance in football. For university athletes, training frequency often declines during off-season periods or during academic exam schedules, increasing the risk of detraining. Therefore, implementing maintenance programs, such as one to two sessions per week focusing on aerobic intervals and strength maintenance, may contribute to maintaining performance adaptations. Periodic assessment and individualized progression should also be considered as part of a long-term athlete development strategy. In addition, appropriate field-based assessments may help coaches and researchers monitor physical fitness variables in football players. For example, the 20-m multistage fitness test is commonly used to estimate aerobic capacity and requires minimal equipment, making it suitable for team training environments. Anaerobic performance can be effectively assessed using the Running-Based Anaerobic Sprint Test (RAST), which provides indices of peak power and fatigue. Sprint tests over short distances (e.g., 30-40 m) are practical for assessing running speed, while shuttle run tests provide information on agility and change-of-direction

ability. Muscular strength can be assessed using handgrip dynamometry or lower-body resistance tests, and flexibility may be evaluated using the sit-and-reach test. These field-based tests offer practical and accessible tools for monitoring performance adaptations in football training programs.

Limitations

This study contributes to the understanding of how structured training programs influence physical performance in university-level football players. While short-term improvements were observed following the training program, the findings highlight the importance of sustained training efforts to maintain these adaptations. However, several limitations should be acknowledged. First, the relatively small sample size (N=15) may reduce statistical power, as smaller samples limit the ability to detect small to moderate effects and increase the risk of Type II error. A post-hoc power analysis indicated that statistical power was acceptable for detecting moderate to large effects; however, the study may have been underpowered to detect smaller effects. Therefore, the findings should be interpreted with caution, particularly for variables that did not show statistically significant changes. Future studies with larger sample sizes would allow for more robust conclusions. Second, the study did not control several external factors, including dietary intake, sleep quality, academic demands, and unmonitored physical activity outside the training program. These factors are known to influence physical performance, recovery, and training adaptations in athletes. However, the present study was conducted under real-world collegiate training conditions, where strict control of these lifestyle variables was not feasible. Future studies may consider monitoring or controlling these factors to further improve internal validity. This study demonstrated acceptable internal validity using standardized training protocols, consistent outcome measurements, and repeated assessments across all time points. However, external validity may be limited due to the relatively small sample size and the inclusion of participants from a single university team. Therefore, caution should be exercised when generalizing the findings to other populations, particularly professional or elite football players. In addition, participants were recruited from a single university football team and represented university-level amateur athletes with limited formal competition exposure (approximately one official tournament per year). Therefore, the findings should be interpreted primarily within the context of collegiate amateur football players, and caution should be exercised when generalizing the results to professional or elite football populations. Future research should investigate the minimal effective training dose required to maintain performance adaptations and explore strategies for optimizing long-term training sustainability in football players.

Conclusion

A structured training program combining multiple exercise modalities over one year significantly improves VO_2max, strength, sprint speed, and anaerobic power in male football players. However, the cessation of training leads to a decline in these adaptations. Incorporating maintenance phases and consistent long-term planning is crucial to sustaining athletic performance.

Ethical approval

The study was performed in compliance with relevant laws and institutional guidelines and that the appropriate institutional committees have approved of them, along with an approval number AMSEC-67EX-040 (date of approval: 26 July 2024), and that participants have signed a written consent form.

Data availability statement

The data presented in this study are available on request from the corresponding authors.

Funding

This research received no external funding.

Conflict of interest

The authors declared no conflicts of interest in this work.

CRediT authorship contribution statement

Busaba Chuatrakoon: conceptualization, methodology, investigation, data curation, writing original draft, review and editing, supervision; **Sothida Nantakool:** conceptualization, methodology, review and editing; **Peerawats Dejprom:** data collection, data curation, writing original draft, review and editing.

Acknowledgements

The authors would like to thank all the participants who took part in this study. Thanks also to the Multidisciplinary and Interdisciplinary School, Chiang Mai University and the Faculty of Associated Medical Sciences, Chiang Mai University for supporting this research project.

References

[1] Chen Y, Lo TY, Xu F, Chang JH. Biomechanical characteristics for identifying the cutting direction of professional soccer players. Appl Sci. 2021; 11(16): 7193. doi:10.3390/app11167193.

[2] França C, Gouveia É, Caldeira R, Marques A, Martins J, Lopes H, et al. Speed and agility predictors among adolescent male football players. Int J Environ Res Public Health. 2022; 19(5): 2856. doi:10.3390/ijerph19052856.

[3] Stolen T, Chamari K, Castagna C, Wisløff U. Physiology of soccer. Sports Med. 2005; 35(6): 501-36. doi:10.2165/00007256-200535060-00004.

[4] Zaharia G, Badau D, Tudor V, Costache R, Geambasu A, Damian M, et al. The effects of 8 aerobic endurance training weeks of 4vs.4+GK small-sided games versus traditional training on physical fitness and skills among U18 football players. Appl Sci. 2023; 13(13): 7963. doi:10.3390/app13137963.

[5] Hoff J, Helgerud J. Endurance and strength training for soccer players: physiological considerations. Sports Med. 2004; 34(3): 165-80. doi: 10.2165/00007256-200434030-00003.

[6] Chmura P, Chmura J, Chodor W, Drozdzowski A, Rokita A, Konefał M. The effects of high-intensity interval training at the anaerobic and psychomotor fatigue thresholds on physiological parameters in young soccer players: a prospective study. Front Physiol. 2023; 14: 1221121. doi:10.3389/fphys.2023.1221121.

[7] Díaz-Hidalgo S, Ranchal-Sanchez A, Jurado-Castro JM. Improvements in jump height, speed, and quality of life through an 8-week strength program in male adolescent soccer players. Sports. 2024; 12(3): 67. doi:10.3390/sports12030067.

[8] Wisløff U, Castagna C, Helgerud J, Jones R, Hoff J. Strong correlation of maximal squat strength with sprint performance and vertical jump height in elite soccer players. Br J Sports Med. 2004; 38(3): 285-8. doi: 10.1136/bjsm.2002.002071.

[9] Owen AL, Wong DP, Paul D, Dellal A. Effects of a periodized small-sided game training intervention on physical performance in elite professional soccer. J Strength Cond Res. 2012; 26(10): 2748-54. doi: 10.1519/JSC.0b013e318242d2d1.

[10] Sousa H, Gouveia ÉR, Marques A, Sarmento H, Pestana M, Quintal T, et al. The influence of small-sided football games with numerical variability in external training load. Sustainability. 2022; 14(2): 1000. doi:10.3390/su14021000.

[11] Helgerud J, Engen LC, Wisløff U, Hoff J. Aerobic endurance training improves soccer performance. Med Sci Sports Exerc. 2001; 33(11): 1925-31. doi: 10.1097/00005768-200111000-00019.

[12] Jackson AW, Baker AA. The relationship of the sit and reach test to criterion measures of hamstring and back flexibility in young females. Res Q Exerc Sport. 1986; 57(3): 183-6. doi: 10.1080/02701367.1986.10605395.

[13] Markovic G, Dizdar D, Jukic I, Cardinale M. Reliability and factorial validity of squat and countermovement jump tests. J Strength Cond Res. 2004; 18(3): 551-5. doi: 10.1519/1533-4287(2004)18<551:RAFVOS>2.0.CO;2.

[14] Heyward VH, Gibson AL. Advanced fitness assessment and exercise prescription. 7th Ed. Champaign, IL: Human Kinetics; 2014.

[15] Little T, Williams AG. Specificity of acceleration, maximum speed, and agility in professional soccer players. J Strength Cond Res. 2005; 19(1): 76-8. doi: 10.1519/14253.1.

[16] Miller MG, Herniman JJ, Ricard MD, Cheatham CC, Michael TJ. The effects of a 6-week plyometric training program on agility. J Sports Sci Med. 2006; 5(3): 459-65.

[17] Léger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. J Sports Sci. 1988; 6(2): 93-101. doi: 10.1080/02640418808729800.

[18] Zagatto AM, Beck WR, Gobatto CA. Validity of the running anaerobic sprint test for assessing anaerobic power and predicting short-distance performances. J Strength Cond Res. 2009; 23(6): 1820-7. doi: 10.1519/JSC.0b013e3181b3df32.

[19] Slimani M, Nikolaidis PT. Anthropometric and physiological characteristics of male soccer players according to their competitive level, playing position and age group: a systematic review. J Sports Med Phys Fitness. 2019; 59(1): 141-63. doi: 10.23736/S0022-4707.17.07950-6.

[20] Suchomel TJ, Nimphius S, Stone MH. The importance of muscular strength in athletic performance. Sports Med. 2016; 46(10): 1419-49. doi: 10.1007/s40279-016-0486-0.

[21] Buchheit M, Mendez-Villanueva A, Delhomel G, Brughelli M, Ahmaidi S. Improving repeated sprint ability in young elite soccer players: repeated shuttle sprints vs. explosive strength training. J Strength Cond Res. 2010; 24(10): 2715-22. doi: 10.1519/JSC.0b013e3181bf0223.

[22] Silva JR, Nassis GP, Rebelo A. Strength training in soccer with a specific focus on highly trained players. Sports Med Open. 2015; 1(1): 17. doi:10.1186/s40798-015-0012-8.

[23] Mujika I, Halson S, Burke LM, Balagué G, Farrow D. An integrated, multifactorial approach to periodization for optimal performance in individual and team sports. Int J Sports Physiol Perform. 2018; 13(5): 538-61. doi: 10.1123/ijspp.2018-0093.

[24] Grgic J, Schoenfeld BJ, Davies TB, Lazinica B, Krieger JW, Pedisic Z. Effect of resistance training frequency on gains in muscular strength: a systematic review and meta-analysis. Sports Med. 2018; 48(5): 1207-20. doi: 10.1007/s40279-018-0872-x.

[25] Faigenbaum AD, Kraemer WJ, Blimkie CJ, Jeffreys I, Micheli LJ, Nitka M, et al. Youth resistance training: updated position statement paper from the National Strength and Conditioning Association. J Strength Cond Res. 2009; 23(5 Suppl): S60-S79. doi:10.1519/JSC.0b013e31819df407.

[26] Loturco I, Pereira LA, Kobal R, Zanetti V, Kitamura K, Abad CCC, et al. Transference of strength and power training to sports performance: a meta-analytical review. Sports Med. 2019; 49(2): 215-36. doi: 10.2478/hukin-2014-0054. eCollection 2014 Jun 28.

Journal of Associated Medical Sciences 2026; 59(2): 348-363

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Development and preliminary evaluation of a dynamic cognitive training program for healthy older adults

Warunya Sutthikampa, Suchitporn Lersilp, Autchariya Punyakaew, Supawadee Putthinoi*

Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 8 August 2025
Accepted as revised 30 March 2026
Available online 3 April 2026

Keywords:
Cognitive training, development, healthy ageing, older people, program

ABSTRACT

Background: Thailand is rapidly transitioning into a super-aged society, increasing the prevalence of age-related cognitive decline. Preserving cognitive function through evidence-based interventions is essential for active aging and independence.

Objectives: This study aimed to design, develop, and evaluate the content validity and feasibility of a multimodal dynamic cognitive training program through a systematic try-out process with healthy older adults.

Materials and methods: The study employed a three-phase research and development (R&D) framework: 1) program design based on neuroplasticity and the dynamic interactional model, 2) content validation by five experts, and 3) evaluating the program through a try-out with the target group (N=5, aged 64-73) over six weeks (12 sessions). Evaluation metrics included feasibility, acceptability, and preliminary cognitive outcomes using the Montreal Cognitive Assessment (MoCA), Digit Span Test, and subtests of the Dynamic Loewenstein Occupational Therapy Cognitive Assessment-Geriatric Version (DLOTCA-G).

Results: The cognitive training program was developed for improving attention, memory, visuospatial skills, and executive function. The cognitive training program, grounded in theories of information processing, neuroplasticity, cognitive retraining, and dynamic interaction, used a customized, multimodal framework. Over six weeks with biweekly 60-minute sessions, it combined cognitive exercises across three difficulty levels with physical activity. The program achieved an overall the Index of Item-Objective Congruence (IOC) of 0.98. During the Phase 3 try-out, participants demonstrated high engagement and 100% session attendance. Evaluation of cognitive outcomes showed positive trends in global cognition (100% of participants) and visuospatial function (100%).

Conclusion: The dynamic cognitive training program is valid and feasible. The try-out phase provided critical insights for program refinement, supporting its application in community-dwelling older adults to mitigate cognitive decline.

Introduction

The global demographic shift toward aging societies presents significant challenges, particularly regarding age-related cognitive decline.[1] Thailand has already transitioned into an aging society and is projected to become a super-aged society in the near future, with older adults constituting a rapidly increasing proportion of the population.[2] Age-related changes in brain structure and function may affect

* **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: supawadee.p@cmu.ac.th
doi: 10.66285/JAMS.2026.066
E-ISSN: 2539-6056

attention, memory, executive function, and decision-making, potentially progressing to cognitive impairment if not appropriately addressed.[3,4]

Memory impairment is frequently identified as one of the earliest and most significant cognitive concerns among older adults. Because memory processes are fundamental to higher-order cognitive functioning, deficits in this domain may subsequently contribute to impairments in attention, visuospatial abilities, and executive functioning. Therefore, the development of integrated training programs targeting these interrelated cognitive domains is essential for mitigating progressive cognitive decline and supporting functional independence in later life. Maintaining cognitive function is also recognized as a central component of healthy aging frameworks.[5]

Accumulating evidence indicates that cognitive functioning in later life remains modifiable. The theory of neuroplasticity suggests that sustained cognitive engagement and enriched interaction with the environment can promote neural adaptation and cognitive optimization throughout the lifespan.[6,7] Accordingly, cognitive training has been widely implemented as a preventive strategy to promote cognitive health among healthy older adults and those at risk of cognitive decline.[8] The development of the present program was further informed by theoretical constructs of information processing, consistent with emerging neuroscientific evidence suggesting that cognitive task performance depends on the efficient transmission of information across distributed brain networks and optimal neural pathways.[9] Recent evidence also indicates that interventions incorporating cognitively stimulating activities embedded within meaningful or everyday contexts may enhance engagement and support adaptive cognitive functioning in aging populations.[10]

Building on these perspectives, integrating physical movement with cognitive tasks has been shown to enhance cerebral blood flow and neural activation.[11] Additionally, game-based and multimodal interventions have demonstrated benefits in strengthening cognitive functioning and potentially reducing the risk of dementia.[12] Despite these promising findings, many existing cognitive training programs rely on standardized or fixed protocols that provide limited flexibility to accommodate individual differences in cognitive profiles, learning pace, and real-time performance changes. Such structured formats may restrict practical applicability, particularly for older adults whose cognitive abilities may fluctuate and require task adaptation, assistive strategies, or graded challenges to maintain optimal engagement.

Occupational therapists play a critical role in addressing age-related cognitive challenges by applying activity analysis and client-centered intervention principles to support cognitive functioning and occupational performance.[13,14] Activity analysis enables therapists to systematically examine task demands and environmental contexts, allowing activities to be selected and structured in ways that are meaningful and appropriate to individuals' roles, habits, values, and cognitive capacities.[15] Within occupational therapy practice, cognitive-oriented interventions encompass a broad range of approaches, including training embedded within everyday activities, combined physical-cognitive interventions targeting functional outcomes, and structured activity-based programs designed to support memory, attention, and performance in daily living tasks.[16] Although previous studies have reported promising outcomes from cognitive-oriented interventions, many publications provide limited descriptions of how task demands are progressively adjusted, how difficulty levels are tailored to individual performance, or how therapists actively mediate and adapt activities during intervention delivery. The absence of such detailed procedural information may reduce clinical reproducibility and hinder practical implementation, particularly in community-based settings where individualized adaptation and real-time clinical decision-making are essential components of occupational therapy practice.

Recent studies highlight increasing interest in the development of cognitive training interventions for older adults while also pointing to areas requiring further refinement. For example, Cruz *et al*. developed an online cognitive training program grounded in occupational therapy principles and reported positive user engagement and improvements in selected aspects of occupational performance and cognitive functioning among healthy older adults.[17] In the Thai context, the TEAM-V program has been implemented to address multiple cognitive domains, including executive function, attention, memory, and visuospatial abilities, with follow-up studies reporting variable outcomes across domains and time points.[18,19] Although some cognitive and functional benefits have been observed, the extent and consistency of effects across specific cognitive processes remain mixed, particularly for higher-order cognitive functions.

Despite this growing body of literature, several gaps remain. Many existing cognitive training programs rely on standardized task delivery and fixed session structures, with limited reporting on how task demands are dynamically adjusted in response to individual performance during intervention delivery.[20-22] Although multimodal elements are often incorporated, descriptions of real-time difficulty grading and therapist-mediated adaptation are frequently insufficient.[20,21] Furthermore, few studies explicitly describe intervention frameworks grounded in occupational therapy–based dynamic assessment principles or provide systematic accounts of graded mediation processes, which are essential for clinical transparency, reproducibility, and translation into community-based practice.[13,23]

In the present study, dynamic cognitive training is conceptualized as an occupational therapy–based intervention approach in which task selection, task demands, and therapist-mediated support are flexibly adjusted within each training session in response to individual performance and observed cognitive strategies. Consistent with occupational therapy practice principles, this approach emphasizes activity analysis, graded task demands, and clinical reasoning to support cognitive functioning within meaningful activity contexts. While the overall program structure, including the number, frequency, and duration of sessions, follows established cognitive training frameworks, adaptation occurs primarily at the activity level through real-time modification and graded mediation during task engagement.

Accordingly, this study aimed to develop and preliminarily evaluate a dynamic cognitive training program for healthy older adults using a structured yet flexible, therapist-mediated, and occupation-based framework. By integrating occupational therapy principles, dynamic assessment concepts, and culturally relevant activity design, the study provides methodological guidance for developing adaptable cognitive training programs suitable for community-dwelling older adults.

Materials and methods

Study design

This study employed a Research and Development (R&D) framework comprising three sequential phases: 1) program design and development, 2) expert content validation, and 3) program evaluation through a try-out with the target group. This framework was selected to ensure the systematic development, theoretical grounding, and practical applicability of the dynamic cognitive training program. An overview of the study workflow is presented in Figure 1. Data collection was conducted between October 2024 and May 2025.

Phase 1: Program design and development

The dynamic cognitive training program for healthy older adults was developed as an individualized, therapist-mediated intervention grounded in occupational therapy principles, activity analysis, and neuroplasticity-informed cognitive training frameworks. Phase 1 represented the conceptual and theoretical design stage of the study, focusing on establishing the program structure, core components, and operational mechanisms rather than evaluating intervention outcomes.

As illustrated in Figure 1, the development process was guided by theoretical constructs of information processing, neuroplasticity, and dynamic interaction. These constructs were applied as guiding design principles to inform task selection, activity sequencing, and intervention delivery rather than as direct outcome measures. References to neuroplasticity within the flowchart denote the hypothesized mechanism underlying the program design (specifically, the role of repeated cognitive engagement, adaptive challenge, and enriched task interaction in supporting cognitive function) rather than empirically verified neural-level changes. Because the present study did not include neurophysiological or neuroimaging assessments, no direct conclusions regarding neural plasticity were drawn.

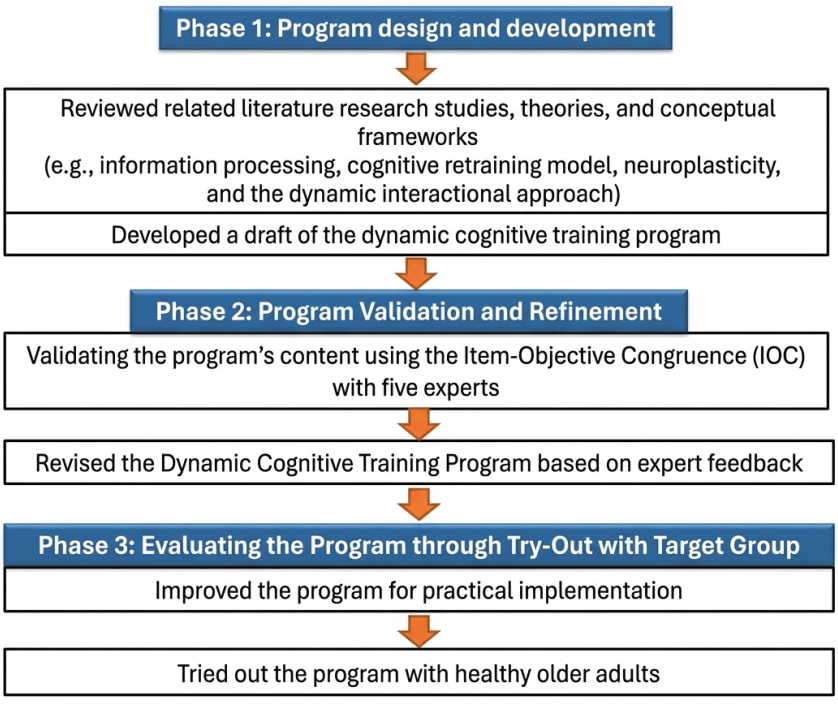


Figure 1. *Research and Development (R&D) framework*

Decisions regarding overall program duration, session length, and training frequency were informed by a structured synthesis of evidence from previous cognitive training and multimodal intervention studies involving older adults.[7,8,13,16-19] This synthesis emphasized three methodological principles: 1) training dosage sufficiency, whereby repeated exposure across multiple sessions supports cognitive engagement and learning processes, 2) feasibility and adherence, prioritizing intervention schedules that are tolerable and sustainable for older adults, and 3) consistency with commonly reported cognitive training protocols in the literature.

Each training session followed a standardized yet flexible structure consisting of two main components. The first component involved a movement-based activity intended to promote arousal and readiness for cognitive engagement. The second component comprised two core cognitive training activities targeting multiple cognitive domains, including attention, memory, visuospatial ability, and executive function. In addition, brief home-based practice was assigned between sessions. These homework tasks were derived from the cognitive activities practiced during the training sessions and were intended to reinforce strategy use and task familiarity. Because the program was delivered twice weekly, participants were encouraged to perform home-based practice on non-training days to maintain continuous engagement throughout the seven-day training cycle. Homework performance was briefly reviewed at the beginning of the subsequent session to support continuity and learning consolidation.

The defining feature of the program's dynamic nature lay in the adaptive design of the cognitive training activities. All cognitive tasks were developed with adjustable difficulty levels, allowing real-time modification based on individual cognitive performance, strategy use, and responsiveness during training. When a participant was unable to complete a task accurately or independently, graded mediation strategies were applied using three predefined levels of support. This individualized adjustment aligns with personalization principles shown to improve engagement in health-related interventions.[24] Level 1 involved general mediation (e.g., repetition of instructions or neutral prompts). Level 2 consisted of targeted feedback, including performance-related cues or strategy suggestions. Level 3 involved task gradation, whereby task complexity, stimulus load, or rule demands were modified to maintain an optimal challenge level while preserving the therapeutic goal.

This graduated mediation framework aligns conceptually with mediation principles embedded within the Dynamic Loewenstein Occupational Therapy Cognitive Assessment-Geriatric Version (DLOTCA-G), which emphasizes examiner-mediated adaptation to reveal cognitive potential rather than static performance alone.[23] Decisions regarding mediation level and task adjustment were guided by continuous observation of task accuracy, error patterns, response time, and the participant's ability to apply cognitive strategies.

All activities were additionally designed with explicit consideration of cultural relevance and contextual appropriateness for Thai older adults. Task content, visual materials, symbols, and examples were derived from familiar daily-life contexts to enhance ecological validity and meaningful engagement, consistent with occupational therapy principles emphasizing person-environment-occupation fit.

Phase 2: Program validation and refinement

The preliminary version of the dynamic cognitive training program was subjected to content validity assessment by a panel of five experts with recognized expertise in geriatric care. The panel comprised one geriatrician, one neurologist or family medicine specialist with a background in neuroscience, two occupational therapists with extensive experience in older adult rehabilitation, and one professor specializing in geriatric occupational therapy.

Content validity was systematically evaluated using the Index of Item-Objective Congruence (IOC). Experts independently assessed the program across six predefined domains: 1) alignment with the four targeted cognitive domains, 2) appropriateness of the session structure, 3) suitability of training frequency and interval, 4) logical sequencing of activities, 5) adequacy of mediation and task adaptation strategies, and 6) appropriateness of homework assignments.

Homework activities were included as a supplementary component of the intervention to encourage participants to maintain cognitive and physical engagement between training sessions. To ensure consistency in training exposure across participants, the homework activities were standardized according to the program protocol. Participants were assigned the same type of homework activity following each training session and were encouraged to complete the activity on days when they did not attend the scheduled training sessions. Each homework task required approximately 10-15 minutes and consisted of simplified versions of the cognitive or movement-based exercises introduced during the supervised sessions.

To monitor adherence, participants were asked to report completion of homework activities during the following training session, and the therapist briefly reviewed the tasks with participants when necessary. However, homework participation was intended as supportive practice rather than a mandatory component of the intervention. Therefore, the overall training intensity was primarily determined by the structured in-session activities delivered uniformly across participants. An IOC≥0.5 was considered acceptable.[25] Qualitative feedback provided by the experts was subsequently incorporated to refine and optimize the program prior to pilot implementation.

Phase 3: Evaluating the program through try-out with the target group.

This phase of the study focused on the development and preliminary evaluation of the intervention program. The program was a pilot-tested with five older adults to assess feasibility, clarity of procedures, and usability of the training activities prior to conducting a larger pilot study.

Participants

Participants included in this study were involved exclusively in Phase 3, which served as a preliminary evaluation aimed at determining the feasibility, acceptability, and practical implementation of the dynamic cognitive training program. In this study, "healthy older adults" were operationally defined as community-dwelling individuals aged 60-75 years who demonstrated intact global cognitive functioning and emotional status, and who were able to perform activities of daily living independently. Common age-related chronic conditions were not considered exclusion criteria, if participants did not have neurological or psychiatric disorders that could affect cognitive performance. Cognitive health was confirmed using the Thai version of the Montreal Cognitive Assessment (MoCA), with scores of 25 or higher indicating the absence of mild cognitive impairment. Emotional well-being was screened using the Thai Geriatric Depression Scale (TGDS-15), with scores below 5 indicating no clinically significant depressive symptoms. These criteria were applied to ensure that participants represented cognitively healthy older adults appropriate for preliminary feasibility testing of the dynamic cognitive training program. This phase was not intended to assess the efficacy of the intervention. Consequently, no formal sample size calculation based on statistical power was performed. Instead, a sample size of five participants was specifically designed as a formative usability testing and pre-pilot refinement process rather than a definitive quantitative feasibility study. This approach aligns with research and development (R&D) frameworks where small-scale try-outs are utilized to identify practical implementation challenges and refine intervention procedures prior to conducting larger scale trials.[26] To minimize the influence of confounding variables, participants were screened for stable medication and the absence of neurological or psychiatric conditions that could affect cognitive performance. Throughout the 6-week intervention, participants were formally instructed to maintain their regular daily routines, including their usual levels of physical exercise and social activities, and were requested to refrain from enrolling in any other structured cognitive or physical training programs.

Phase 3 was conducted to evaluate the feasibility of program implementation and to explore preliminary cognitive outcomes associated with the dynamic cognitive training program. A purposive sampling approach was adopted to ensure that participants possessed the characteristics necessary for feasibility testing. Five community-dwelling, cognitively healthy older adults aged 60-75 years were recruited based on predefined eligibility criteria.

Instruments

To ensure a comprehensive and rigorous evaluation during the Phase 3 try-out, a complementary battery of cognitive assessments was employed:

MoCA

The MoCA was utilized as the gold-standard clinical screening instrument to verify inclusion criteria and provide a measure of global cognition. It ensured that all participants were cognitively healthy (excluding MCI) before the intervention. The Thai version of the MoCA was used as a standardized cognitive screening instrument to verify participant eligibility and to provide an index of global cognitive functioning. The MoCA was administered to ensure that all participants were cognitively healthy and did not meet criteria for mild cognitive impairment prior to participation in the intervention. The Thai MoCA has been culturally adapted and validated for use in Thai older adult populations and has demonstrated adequate psychometric properties for discriminating amnestic mild cognitive impairment from Alzheimer's disease and cognitively healthy controls. Specifically, the instrument has shown good diagnostic accuracy, with reported areas under the receiver operating characteristic curve of 0.813 for amnestic mild cognitive impairment and 0.938 for Alzheimer's disease, as well as acceptable sensitivity and specificity using cutoff scores appropriate for Thai clinical settings. In addition, evidence of strong test-retest and inter-rater reliability support the reliability and applicability of the Thai MoCA as a cognitive screening tool in this population. [27]

TGDS-15:

The Thai version of the 15-item Geriatric Depression Scale (TGDS-15) was administered as a standardized screening instrument to identify depressive symptoms and to ensure participant eligibility for inclusion in the study. The TGDS-15 has been evaluated in Thai older adult populations and has demonstrated acceptable psychometric properties as a screening tool for major depressive disorder, including good sensitivity and specificity at a cutoff score of ≥5.[28] These findings support its appropriateness for detecting clinically relevant depressive symptoms in geriatric settings in Thailand.

In Phase 3, depression screening was conducted prior to baseline cognitive assessment. Participants who scored ≥5 on the TGDS-15, indicating the presence of clinically significant

depressive symptoms, were excluded from the study. Only participants with TGDS-15 scores below the cutoff were invited to proceed to baseline cognitive assessments and participate in the intervention. The TGDS-15 was used exclusively for eligibility screening and was not included as an outcome measure in the analysis. This procedure was implemented to reduce potential confounding effects of depressive symptoms on cognitive performance and responsiveness to the training program.

DLOTCA-G

The DLOTCA-G was administered by trained occupational therapists following the original administration guidelines. The assessment is composed of multiple subtests designed to evaluate different aspects of cognitive functioning, and individual subtests may be administered selectively when a focused evaluation of cognitive processes is required. In this study, three specific subtests-Everyday Life Objects (memory), Visuo-construction (visuospatial function), and Thinking Operations (executive function) were used to obtain in-depth cognitive data. These subtests were selected because they include a mediation component that directly aligns with the dynamic intervention framework used in this study. [23,29]

As the DLOTCA-G is a performance-based cognitive assessment involving therapist-guided tasks, the instructions were verbally explained in Thai to ensure participant comprehension, while the scoring procedures adhered to the standardized criteria described in the original manual. The subtests were interpreted according to the original scoring framework of instrument. [29]

Although attention was one of the targeted cognitive domains in the intervention program, it was assessed separately using the Digit Span Test to allow a focused evaluation of attentional capacity and working memory, which are not independently isolated within the selected DLOTCA-G subtests.

Digit Span Test

The Digit Span Test was administered to assess attentional capacity and verbal working memory, addressing aspects of attention that are not independently captured by the selected DLOTCA-G subtests. The test consists of two components: Digit Span Forward, which primarily reflects attention span and immediate auditory attention, and Digit Span Backward, which additionally requires working memory and mental manipulation. In cognitively healthy adults, performance typically approximates an average span of 7±2 digits, reflecting normal attentional capacity. The Digit Span Test has shown acceptable psychometric performance in older adult populations. Multiple studies, including de Paula et al.[30], report moderate to high reliability, with internal consistency coefficients typically ranging from 0.70 to 0.89 and test–retest reliability estimates of approximately 0.80 to 0.90 in older adults.

In the present study, both Digit Span Forward and Digit Span Backward were administered to evaluate attention and verbal working memory. The Digit Span Test was used as an outcome measure to describe individual-level changes in attentional performance following participation in the dynamic cognitive training program. All outcome measures were administered at two time points using the same assessment instruments: prior to the commencement of the intervention (pre-test) and immediately after completion of the 12-session dynamic cognitive training program (post-test). The Thai version of the MoCA, Digit Span Test (forward and backward), and selected subtests of the DLOTCA-G were used consistently at both pre-test and post-test to evaluate changes in global cognition, attention, memory, visuospatial ability, and executive function. The TGDS-15 was administered only at baseline as a screening tool to determine eligibility and was not included as an outcome measure.

Data analysis

Five older adults completed both baseline and post-intervention cognitive assessments across the targeted cognitive domains. At baseline, cognitive performance varied among participants. Given the exploratory and feasibility-oriented design of Phase 3 and the small sample size (N=5), data analysis was conducted descriptively to explore preliminary trends. Due to the small sample size, no definitive conclusions regarding feasibility in absolute quantitative terms were drawn; instead, the findings represent a preliminary assessment of the program's practical application.

Cognitive assessment scores were summarized using raw scores and individual-level changes between pre- and post-intervention assessments to illustrate patterns of change and variability across participants. This descriptive approach was adopted to explore preliminary trends and to support evaluation of program feasibility, rather than to test hypotheses or determine intervention efficacy. Accordingly, the findings from Phase 3 are interpreted cautiously and are intended to inform program refinement and the design of future studies with adequate sample sizes and statistical power.

Raw scores from the MoCA, Digit Span Test, and selected subtests of the DLOTCA-G were summarized and compared between pre- and post-intervention assessments. Although the intervention targeted four specific cognitive domains, the MoCA was included

to provide an overall indicator of global cognitive functioning and to contextualize domain-specific changes. Due to the standardized and proprietary nature of the assessment instruments, results are presented as raw scores with corresponding maximum possible scores to facilitate interpretation, rather than item-level reporting. The TGDS-15 was used exclusively as a screening instrument during participant recruitment to ensure eligibility for inclusion. TGDS-15 scores were not included in the outcome analysis and were not reported as part of the study results.

Results

Phase 1: Program design and development

Phase 1 resulted in the development of a structured dynamic cognitive training program for healthy older adults based on an integrated synthesis of occupational therapy theories, cognitive training principles, and evidence from prior studies.[7,8,13,16-19] The finalized program framework reflects a multimodal and adaptive approach that combines physical movement, multi-domain cognitive stimulation, and therapist-mediated task adjustment.

Based on the synthesized evidence and design principles described in Phase 1, the program was structured as a six-week intervention consisting of 12 sessions, each lasting 60 minutes and delivered twice weekly. Each session included a preparatory movement-based activity, followed by two core cognitive training activities designed to engage multiple cognitive domains simultaneously, and a structured home-based activity intended to reinforce learning.

The program incorporates a multimodal cognitive training framework that integrates physical movement with cognitive exercises to support neural activation and overall brain health.[31-33] Inspired by the "Brain Exercise" program described by Li *et al.*,[34] the intervention includes a preparatory phase consisting of movement-based warm-up activities, such as grid-based walking and hand coordination tasks,[20,35] designed to promote spatial awareness and neurophysiological readiness prior to cognitively demanding activities. Each session engaged participants in at least two cognitive domains simultaneously, consistent with principles of dual-task and integrated cognitive stimulation.

The training program included a variety of task types, including game-based challenges, skill-oriented tasks, simulated real-life scenarios, and analytical reading exercises. These activities were selected based on evidence suggesting their potential to enhance neural connectivity and promote cognitive resilience in aging populations. All cognitive tasks were designed with adjustable difficulty levels and implemented using a predefined graded mediation framework to ensure individualized adaptation during training.

The finalized structure of the dynamic cognitive training program (including session composition, core components, targeted cognitive domains, and underlying theoretical principles) is summarized in Table 1.

Table 1. Description of a dynamic cognitive training program for healthy older adults.

Program	Activities	Description	Time (min)	Cognitive domain			
				A	W	VP	EF
1	Physical activity 1	Song-based gross motor activity enhancing attention, working memory	10	✓	✓		
	Cognitive activity 1-1	Which medicine, when? Remember pill details (color, amount, time) for each fictional character	25	✓	✓		
	Cognitive activity 1-2	Find the match mission: memorize a face and find the matching image among distractors	25	✓	✓		
2	Physical activity 2	Swing and sway: full-body rhythmic movement to upbeat music improving attention, memory, and visuo-construction	10	✓	✓		
	Cognitive activity 2-1	Whose number? Match phone numbers to characters after brief memorization	25	✓	✓		
	Cognitive activity 2-2	Block matching: recreate a block pattern based on a visual model	25	✓	✓	✓	

Table 1. Description of a dynamic cognitive training program for healthy older adults (continue).

Program	Activities	Description	Time (min)	Cognitive domain			
				A	W	VP	EF
3	Physical activity 3	Joy all day long: midline-crossing dance moves improving bilateral coordination and cognitive engagement.	10	✓	✓		
	Cognitive activity 3-1	Wise storyteller: sequence pictures and narrate a logical story.	25	✓	✓		✓
	Cognitive activity 3-2	Mystery compass: solve spatial puzzles using compass directions.	25	✓	✓	✓	
4	Physical activity 4	Rain dance: rhythmic movements crossing the midline to fun music.	10	✓	✓		
	Cognitive activity 4-1	Memory test: observe images and answer detailed questions.	25	✓	✓		
	Cognitive activity 4-2	Count it right: count objects in complex scenes accurately.	25	✓	✓		
5	Physical activity 5	Touch your body: follow verbal instructions to touch various body parts in rhythm.	10	✓	✓		
	Cognitive activity 5-1	Memory detective: spot and describe differences between similar people.	25	✓	✓		
	Cognitive activity 5-2	Hidden numbers: identify the missing number from a scrambled set.	25	✓	✓	✓	
6	Physical activity 6	Tour around: mimic vehicle movements (e.g., train, airplane) with rhythm.	10	✓	✓		
	Cognitive activity 6-1	Pack it up: plan and select items for a trip based on scenario conditions.	25	✓	✓	✓	✓
	Cognitive activity 6-2	Explorer's journey: use maps and clues to identify item locations.	25	✓	✓	✓	✓
7	Physical activity 7	Coordinated hand-body movements to upbeat songs.	10	✓	✓		
	Cognitive activity 7-1	Read and answer: read text and answer comprehension questions.	25	✓	✓	✓	✓
	Cognitive activity 7-2	Treasure hunt: solve clues using maps and directions to find treasure.	25	✓	✓	✓	✓

Table 1. Description of a dynamic cognitive training program for healthy older adults (continue).

Program	Activities	Description	Time (min)	Cognitive domain			
				A	W	VP	EF
8	Physical activity 8	9-Box Grid Walk: Navigate 3x3 Floor Grid with Verbal Cues	15	✓	✓	✓	✓
	Cognitive activity 8-1	Top chef: choose ingredients from a catalog to follow recipe rules	20	✓	✓	✓	✓
	Cognitive activity 8-2	Fair share: divide animals evenly under rule constraints	25	✓	✓	✓	✓
9	Physical activity 9	9-box grid walk with music 1: diagonal stepping to rhythmic instructions	15	✓	✓	✓	✓
	Cognitive activity 9-1	Where am i? Interpret maps to determine correct location	20	✓	✓	✓	✓
	Cognitive activity 9-2	Overlay puzzle: identify hidden objects in layered images	25	✓	✓	✓	✓
10	Physical activity 10	9-box grid walk with music 2: cross-body footwork to choreographed music	15	✓	✓	✓	✓
	Cognitive activity 10-1	Market budgeting: calculate purchases to stay within a spending limit	20	✓	✓	✓	✓
	Cognitive activity 10-2	Math riddle: solve layered math problems from story-based questions	25	✓	✓	✓	✓
11	Physical activity 11	9-box grid walk with music 3: dynamic foot patterns in triangle formation	15	✓	✓	✓	✓
	Cognitive activity 11-1	Global awareness: analyze global landmarks through visual cues	20	✓	✓	✓	✓
	Cognitive activity 11-2	Plan a trip: schedule and budget for domestic travel scenarios	20	✓	✓	✓	✓
12	Physical activity 12	9-box grid walk with music 4: diamond pattern steps to upbeat rhythm	25	✓	✓	✓	✓
	Cognitive activity 12-1	Garden planner: distribute garden plots and trees per specific conditions	15	✓	✓	✓	✓
	Cognitive activity 12-2	Pack before travel: select suitable items within weight limits and restrictions	20	✓	✓	✓	✓

Note: *A: attention, W: working memory, VP: visuospatial function, EF: executive function.*

The mediation procedures were implemented using a dynamic cueing hierarchy conceptually adapted from the DLOTCA-G. In this program, mediation was applied as a responsive and performance-based process in which therapist support was dynamically adjusted according to participants' observed performance during task engagement.

Across all cognitive activities, therapist mediation followed a three-level cueing hierarchy progressing from general prompts to increasingly specific guidance. This hierarchical structure enabled participants to attempt independent problem-solving while allowing therapists to provide graduated support when difficulties emerged. To preserve the intended cognitive demands of each activity, no more than three cues were typically provided within a single task before considering modification of task parameters. At Level 1 (general prompt), therapists provided neutral prompts or repeated task instructions to support comprehension and encourage participants to review available information. At Level 2 (guided prompt), therapists provided strategy-oriented cues directing attention toward relevant task features or suggesting alternative cognitive strategies. At Level 3 (specific prompt), therapists delivered more explicit cues by identifying inconsistencies or directing participants to re-examine task elements.

An example of this dynamic mediation process can be observed in Cognitive Activity 1-1, "Which medicine, when?" (Table 1). In this activity, participants were required to remember and match medication details (such as pill color, dosage, and administration time) for several fictional characters. The task required participants to encode and recall multiple pieces of information, thereby engaging attention and working memory processes.

During task performance, therapist mediation was applied dynamically based on participant responses. When a participant showed uncertainty or produced an incorrect response, the therapist initially provided a Level 1 prompt, encouraging independent review of the task information (e.g., "Try reviewing the pill information again and check the color, amount, and time carefully."). If the participant continued to have trouble, the therapist provided a Level 2 prompt, directing attention toward relevant task features or suggesting cognitive strategies (e.g., comparing medication schedules or verifying whether pill characteristics matched the correct character). If errors persisted, a Level 3 prompt was provided, in which the therapist offered more explicit cues by highlighting potential inconsistencies and encouraging the participant to re-examine specific task elements.

Through this graduated mediation process, therapist support was continuously adjusted according to participant performance, allowing individualized guidance while preserving the cognitive challenge of the activity and maintaining the core principles of the dynamic cognitive training framework.

Phase 2: Content validation of the training program

The objective of this phase was to assess the relevance, clarity, and appropriateness of the physical and cognitive training activities included in the intervention program. Content validation was performed using the IOC method to evaluate the extent to which each activity aligned with the intended training objectives, as judged by expert consensus. A panel of five experts was assembled for the validation process, including one geriatrician, one neurologist, two occupational therapists with experience in geriatric cognitive rehabilitation, and one academic specialist in occupational therapy and aging. Each expert independently evaluated the training activities across all 12 sessions of the program.

The IOC scores were calculated for each physical and cognitive training activity to evaluate content validity. Physical training activities demonstrated high levels of expert agreement, with individual IOC values ranging from 0.88 to 1.00 and an overall mean IOC score of 0.98, indicating strong alignment with the training objectives. Cognitive training activities showed uniformly high agreement, with all activities receiving IOC scores of 1.00 and an overall cognitive activities IOC score of 1.00.

The overall IOC score for the entire program was 0.98, reflecting excellent content validity. These findings indicate that both the physical and cognitive components of the training program are conceptually appropriate and well aligned with the intended objectives for healthy older adults. A summary of the IOC scores for each activity and program component is presented in Table 2. In addition to the quantitative IOC ratings, the experts provided qualitative feedback regarding the clarity, relevance, and feasibility of the program components. Attention was given to items with IOC values below 1.0, which were mainly related to the clarity of activity instructions and the description of certain program components. Based on these comments, several aspects of the program were refined. For example, experts noted that the instructions for some activities were not sufficiently clear for older adults, which contributed to slightly lower IOC ratings. Based on these comments, the wording of the instructions was revised and simplified to improve clarity and usability. Similarly, some experts indicated that the difficulty level of certain activities might vary depending on participants' cognitive abilities. These comments led to the refinement of activity descriptions and the inclusion of clearer guidelines for graded mediation and task adjustment within the program. These modifications improved the clarity, usability, and participant-friendliness of the program, ensuring that the final version was appropriate for implementation in the subsequent effectiveness study.

Table 2. Item-objective congruence (IOC) scores of physical and cognitive training activities.

Program	Physical activities	Score	Cognitive training activities	Score
1	Physical activities 1	1.00	Activity: 1-1 Activity: 1-2	0.88 1.00
2	Physical activities 2	1.00	Activity: 2-1 Activity: 2-2	0.88 1.00
3	Physical activities 3	1.00	Activity: 3-1 Activity: 3-2	0.92 1.00
4	Physical activities 4	1.00	Activity: 4-1 Activity: 4-2	0.96 1.00
5	Physical activities 5	1.00	Activity: 5-1 Activity: 5-2	0.96 0.96
6	Physical activities 6	1.00	Activity: 6-1 Activity: 6-2	0.96 0.96
7	Physical activities 7	1.00	Activity: 7-1 Activity: 7-2	0.96 0.96
8	Physical activities 8	1.00	Activity: 8-1 Activity: 8-2	0.96 0.96
9	Physical activities 9	1.00	Activity: 9-1 Activity: 9-2	0.96 0.96
10	Physical activities 10	1.00	Activity: 10-1 Activity: 10-2	0.96 0.96
11	Physical activities 11	1.00	Activity: 11-1 Activity: 11-2	0.96 0.96
12	Physical activities 12	1.00	Activity: 12-1 Activity: 12-2	0.96 0.96
Overall physical activities IOC score		**1.00**	**Overall cognitive training activities IOC score**	**0.95**
Overall program IOC score: 0.98				

Note: *IOC scores were calculated based on expert ratings of item relevance to program objectives. An IOC≥0.50 indicates acceptable content validity.*

Phase 3: Evaluating the program through try-out with the target group

Following the confirmation of content validity, the dynamic cognitive training program was pilot-tested with a group of five cognitively healthy older adults (aged 64-73 years) to assess its feasibility, acceptability, and preliminary usability in a community-based setting. The primary objectives of this phase were to determine whether the program could be delivered as intended, identify any barriers to participation or implementation, and gather initial feedback from participants regarding their experiences. Participants completed pre- and post-intervention cognitive assessments covering five domains: global cognition, attention, memory, visuospatial function, and executive function.

At baseline, cognitive performance varied among participants. To ensure personalized engagement, intervention activities were selected according to everyone's cognitive profile, age, and educational background. The program included three levels of task difficulty, allowing for tailored progression. Adjustments to activity complexity and session pacing were made dynamically in response to participant feedback. These modifications were systematically documented using a real-time mediation system to ensure consistency and fidelity. Post-intervention results indicated improvements across all measured cognitive domains. Notably, all participants (100%) demonstrated increased MoCA scores, reflecting enhanced global cognitive functioning. Attention, as measured by the Digit Span test, improved in 60% of participants. Evaluation using the DLOTCA-G revealed gains in three domains: memory, visuospatial function, and executive function. Specifically, 80% of participants showed improvements in both memory and executive function, while visuospatial function improved in all participants (100%). These preliminary findings suggest that the individualized activity selection and responsive mediation strategies embedded within the program contributed meaningfully to the observed cognitive gains. Minor adjustments to session timing and pacing were made based on participant feedback

to optimize the flow between physical and cognitive components. These feasibility findings support the practical implementation of the program in larger-scale trials.

Five older adults (aged 64-73 years) completed both baseline and post-intervention assessments across five cognitive domains. At baseline, cognitive performance varied among participants. Following the intervention, improvements were observed across all assessed cognitive domains. All participants demonstrated increased MoCA scores (100%), indicating enhanced global cognitive functioning. Improvements in attention, as measured by the Digit Span Test, were observed in three participants (60%). The DLOTCA-G subtests revealed improvements in memory and executive function in four participants (80%), while visuospatial function improved in all participants (100%). These findings suggest that individualized activity selection and real-time, responsive mediation may have contributed to the observed cognitive improvements. Detailed results are presented in Table 3.

Table 3. Demographic characteristics and individual pre- and post-intervention cognitive scores of participants in Phase 3.

Variables		Participant 1	Participant 2	Participant 3	Participant 4	Participant 5
Gender		Female	Female	Female	Male	Male
Age (year)		64	67	70	65	73
Education		Secondary M3 (lower secondary education)	Secondary M6 (upper secondary education)	Secondary M6 (upper secondary education)	Secondary M6 (upper secondary education)	Secondary M3 (lower secondary education)
Underlying diseases		DM, HT	DM	DM, DLP	HT	HT
Occupation		Retired farmer	Housewife	Community volunteer	Retired merchant	Retired merchant
MoCA Total score=30	Pre-Test	27	26	28	25	27
	Post-Test	28	27	29	27	28
Attention (Digit Span Test) Total score=16	Pre-Test	5	5	4	5	4
	Post-Test	6	5	5	6	5
Memory (DLOTCA-G) Total score=4	Pre-Test	3	2	3	4	3
	Post-Test	4	3	4	4	3
Visuospatial function (DLOTCA-G) Total score=5	Pre-Test	3	3	3	4	3
	Post-Test	4	4	4	5	4
Executive function (DLOTCA-G) Total score=5	Pre-Test	3	3	3	4	3
	Post-Test	4	4	4	4	4

Note: *DM: diabetes mellitus, HT: hypertension, DLP: hyperlipidemia. Scores are presented as raw scores. Maximum possible scores for each instrument/subtest are provided to support interpretation.*

Discussion

This study describes the systematic development, content validation, and preliminary feasibility evaluation of a multimodal dynamic cognitive training program designed for healthy older adults. Grounded in occupational therapy theory, neuroscience, and activity analysis, the program integrates physical activity and cognitive stimulation within a structured yet adaptable framework. The high Index of Item-Objective Congruence (IOC) scores obtained during content validation (overall IOC=0.98) indicate strong agreement among experts that the program components were relevant, clearly defined, and appropriate for the intended training objectives and target population.

Findings from the try-out phase further demonstrate that the program was feasible and well tolerated in a community-based setting. Participants showed high engagement throughout the six-week intervention and completed all training sessions. Although the primary purpose of this phase was to examine feasibility rather than intervention effectiveness, descriptive trends suggested improvements across several cognitive domains, including global cognition, attention, memory, visuospatial ability, and executive function. Importantly, IOC findings represent evidence of program content validity and should not be interpreted as evidence of intervention effectiveness.

Given the exploratory nature of the try-out phase and the small sample size, the observed changes should be interpreted as preliminary descriptive trends rather than definitive evidence of intervention efficacy. Although participants were instructed to maintain their usual daily routines and avoid participation in other structured cognitive training programs during the intervention period, unmeasured lifestyle factors cannot be entirely ruled out. Future studies using controlled designs and larger sample sizes will therefore be necessary to determine the effectiveness of the program.

All participants demonstrated increases in MoCA scores following the intervention, suggesting improvements in global cognitive functioning. Improvements in Digit Span performance were observed in three participants, indicating potential gains in attentional capacity and verbal working memory. Attention is widely recognized as a fundamental cognitive process supporting information processing and daily functional performance, and it is often among the earliest domains to decline during aging.

These findings align with previous studies reporting cognitive benefits from cognitively demanding training activities. For example, Shatil[36] reported improvements in attention and other cognitive domains following computerized cognitive training, both alone and in combination with physical activity. Although the assessment methods used in that study differed from the neuropsychological measures employed in the present research, the findings support the broader premise that cognitively stimulating activities may contribute to improvements in attentional functioning in healthy older adults.

Improvements observed in memory, visuospatial function, and executive functioning using the DLOTCA-G subtests are also consistent with outcomes reported in other multimodal or adaptive cognitive training studies.[23,29] For example, the Training Project Amsterdam Seniors and Stroke (TAPASS) trial demonstrated improvements in executive functioning and processing speed following task-switching training among healthy older adults.[37] Similarly, integrated cognitive-physical training paradigms, including Training Under Cognitive Kinematics (TUCK) and Useful Field of View (UFOV) interventions, have shown concurrent cognitive and motor benefits, highlighting the potential advantages of combining physical movement with cognitive challenge.[38,39] The unique features of this study, such as individualized activity selection, difficulty-level adjustment, and real-time mediation, contributed to enhanced participant engagement, and improved intervention efficacy. Although the intervention was designed based on neuroplasticity-informed principles (such as graded challenge, task engagement, and therapist-mediated adaptation) neural plasticity was not directly measured. Therefore, references to neuroplasticity in this study describe the intended theoretical mechanism underlying the program design, not an empirically verified neural outcome. These tailored, responsive strategies echo findings from virtual-reality and adaptive programs that emphasize ecological validity and participant-centered design.[40,41] These studies are cited here to provide conceptual context for combining cognitive challenge with movement-based tasks, rather than to imply direct comparability of outcomes. In contrast to fixed or predominantly computerized protocols, the present program emphasizes therapist-mediated adaptation and occupation-based activity selection informed by ongoing performance.[22] Although the sample size was small, the study design and positive outcomes justify further investigation. Additionally, examining adherence, cost-effectiveness, and scalability across diverse settings will be essential for broader implementation.

Overall, the findings from this preliminary try-out suggest that the dynamic cognitive training program is feasible and acceptable for healthy older adults and may be associated with favorable descriptive trends across multiple cognitive domains. The individualized tailoring of activities, combined with real-time mediation and adaptive progression, may have contributed to participant engagement and observed changes in performance. All five participants demonstrated increased MoCA scores, indicating enhanced global cognitive performance. Improvements observed in the Digit Span Test among three participants indicate enhancements in attention, which is essential for daily

functioning and tends to be one of the first cognitive abilities to decline with age.[42] Furthermore, while previous reviews have discussed the potential cost-effectiveness and scalability of computerized cognitive training programs, [21] the present study did not evaluate economic outcomes. Cost-related considerations are therefore discussed only as broader contextual implications for future program development and implementation. LOTCA-G results further support these findings, with most participants showing improvements in memory and executive function, and all participants demonstrating enhanced visuospatial ability.

The individualized tailoring of activities, combined with real-time mediation and adaptive progression, may have contributed to participant engagement and observed changes in performance. These features are consistent with emerging evidence emphasizing the importance of personalized, ecologically relevant, and adaptive intervention designs in supporting cognitive health in aging populations.

Limitations

Despite the promising findings, several limitations of this study should be acknowledged. First, the small sample size (N=5) in the pilot phase limits the generalizability of the results and precluded the use of inferential statistical analyses. Although descriptive trends and qualitative feedback were positive, the absence of a control group restricts the ability to attribute observed cognitive changes solely to the intervention. Second, the short duration of follow-up prevents conclusions regarding the long-term sustainability of the observed cognitive changes. Third, while the intervention was personalized and flexibly delivered, it was implemented in a controlled, small-group setting. As a result, the feasibility of implementing the program in larger community or clinical contexts remains uncertain and may involve logistical, adherence, and resource-related challenges. Finally, the study focused exclusively on cognitively healthy older adults, which limits the applicability of the findings to other populations, such as individuals at risk for cognitive decline.

Future study

Future research should build on these preliminary findings by conducting larger-scale randomized controlled trials with adequate statistical power and appropriate control conditions to more rigorously evaluate intervention efficacy. Longitudinal study designs with extended follow-up periods are needed to assess the durability of cognitive changes and to determine whether booster sessions are necessary to maintain benefits. In addition, future studies should examine the scalability and real-world implementation of the program across diverse settings, including community centers, rehabilitation clinics, and home-based formats. Exploring the applicability of this dynamic cognitive training program in populations at risk for cognitive decline, such as older adults with mild cognitive impairment (MCI), may also help clarify its potential preventive and translational value.

Conclusion

This study demonstrates the feasibility and acceptability of a dynamic, therapist-mediated cognitive training program for healthy older adults. The program's individualized and adaptive design, grounded in occupational therapy principles, supports its potential relevance for promoting cognitive engagement in community-dwelling older adults. Although conclusions regarding efficacy cannot be drawn from this preliminary try-out, the findings provide a foundation for future research employing larger samples and controlled designs to further examine effectiveness, scalability, and functional impact in real-world settings.

Ethical approval

Ethical approval for this study was obtained from the Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University (Approval No. AMSEC-67EX-034). All participants provided written informed consent prior to participation.

Funding

Financial support for this study was provided through the Teaching Assistant and Research Assistant (TA/RA) Scholarships from Chiang Mai University, Thailand.

Conflict of interest

The authors declare no conflicts of interest concerning the research, authorship, or publication of this article. All authors approved the final manuscript.

CRediT authorship contribution statement

Supawadee Putthinoi: conceptualization and study design, writing: original draft, preparation and submission of the manuscript, corresponding author; **Warunya Sutthikampa**: data collection and analysis, development and design of training activities within the program; **Autchariya Punyakaew:** intellectual contribution to the development and design of training activities, data analysis: **Suchitporn Lersilp**: co-advisor, supervision of research methodology.

References

[1] Paitoonpong S. Promotion of active aging and quality of life in old age and preparation for a complete aged society in Thailand. TDRI Q Rev. 2023; 38(3): 3-14.

[2] Foundation of Thai Gerontology Research and Development Institute. Situation of the Thai Older Persons 2021. Nakhon Pathom: Institute for Population and Social Research, Mahidol

University; 2021 (in Thai).

[3] Samanez-Larkin GR, Knutson B. Decision making in the ageing brain: changes in affective and motivational circuits. Nat Rev Neurosci. 2015; 16(5): 278-89. doi: 10.1038/nrn3917.

[4] Prince JB, Davis HL, Tan J, Muller-Townsend K, Markovic S, Lewis DMG, et al. Cognitive and neuroscientific perspectives of healthy ageing. Neurosci Biobehav Rev. 2024; 161: 105649. doi: 10.1016/j.neubiorev.2024.105649.

[5] Cocchi C, Zazzara MB, Levati E, Calvani R, Onder G. How to promote healthy aging across the life cycle. Eur J Intern Med. 2025; 135: 5-13. doi: 10.1016/j.ejim.2025.03.003.

[6] Kumar J, Patel T, Sugandh F, Dev J, Kumar U, Adeeb M, et al. Innovative approaches and therapies to enhance neuroplasticity and promote recovery in patients with neurological disorders: a narrative review. Cureus. 2023; 15(7): e41914. doi: 10.7759/cureus.41914.

[7] Nguyen L, Murphy K, Andrews G. Cognitive and neural plasticity in old age: a systematic review of evidence from executive functions cognitive training. Ageing Res Rev. 2019; 53: 100912. doi: 10.1016/j.arr.2019.100912.

[8] Malmberg Gavelin H, Dong C, Minkov R, Bahar-Fuchs A, Ellis KA, Lautenschlager NT, et al. Combined physical and cognitive training for older adults with and without cognitive impairment: a systematic review and network meta-analysis of randomized controlled trials. Ageing Res Rev. 2021; 66: 101232. doi: 10.1016/j.arr.2020.101232.

[9] Wang Z, Yang Y, Huang Z, Zhao W, Su K, Zhu H, et al. Exploring the transmission of cognitive task information through optimal brain pathways. PLoS Comput Biol. 2025; 21(3): e1012870. doi: 10.1371/journal.pcbi.1012870.

[10] Ryan S, Brady O. Cognitive stimulation and activities of daily living for individuals with mild-to-moderate dementia: a scoping review. Br J Occup Ther. 2023; 86(8): 03080226231156517. doi: 10.1177/03080226231156517.

[11] van Balkom TD, van den Heuvel OA, Vriend C. The effects of cognitive training on brain network activity and connectivity in aging and neurodegenerative diseases: a systematic review. Neuropsychol Rev. 2020; 30(3): 267-86. doi: 10.1007/s11065-020-09440-w.

[12] Gui W, Cui X, Miao J, Zhu X, Li J. The effects of simultaneous aerobic exercise and video game training on executive functions and brain connectivity in older adults. Am J Geriatr Psychiatry. 2024; 32(10): 1244-58. doi: 10.1016/j.jagp.2024.04.009.

[13] American Occupational Therapy Association. Cognition, cognitive rehabilitation, and occupational performance. Am J Occup Ther. 2019; 73(Suppl 2): 7312410010p1-25.doi:10.5014/ajot.2019.73S201.

[14] Vasquez-Carrasco E, Jamett-Oliva P, Hernandez-Martinez J, Riquelme-Hernandez C, Villagran-Silva F, Branco BHM, et al. Effectiveness of occupational therapy interventions on activities of daily living, cognitive function, and physical function in middle-aged and older people with chronic stroke: a systematic review with meta-analysis. J Clin Med. 2025; 14(7): 2197. doi: 10.3390/jcm14072197.

[15] Thomas H. Occupational and activity analysis. 3rd ed. Abingdon: Routledge; 2022.

[16] Castellote-Caballero Y, Carcelen Fraile MC, Aibar-Almazan A, Afanador-Restrepo DF, Gonzalez-Martin AM. Effect of combined physical-cognitive training on the functional and cognitive capacity of older people with mild cognitive impairment: a randomized controlled trial. BMC Med. 2024; 22: 281. doi: 10.1186/s12916-024-03469-x.

[17] Cruz GP, Pereira LS, Pizzetti CF, Kiel LS, Canella TF, Raymundo TM. Development of an online cognitive training program for healthy older adults: an occupational therapy intervention. Cad Bras Ter Ocup. 2022; 30: e3124. doi: 10.1590/2526-8910.CTOAO239231242.

[18] Kuha O, Vidhyachak C, Phatcharapisitgul P, Kumniyom N, Mungklang P, Wisajun P, et al. Cognitive training for the elderly with mild neurocognitive disorder: utilizing TEAM-V program in urban and rural areas with a six-month follow-up. J Psychiatr Assoc Thailand [Internet]. 2023 [cited 2026 Jan 24]; 68(1): 102-14. Available from: https://he01.tci-thaijo.org/index.php/JPAT/article/view/254346 (in Thai).

[19] Srisuwan P, Nakawiro D, Kuha O, Kengpanich S, Gesakomol K, Chansirikarnjana S. Efficacy of a group-based 8-week multicomponent cognitive training on cognition, mood and activities of daily living among healthy older adults: a two-year follow-up of a randomized controlled trial. J Prev Alzheimers Dis. 2025; 12(2): 100033. Available from: https://repository.li.mahidol.ac.th/handle/123456789/104268.

[20] Versi N, Murphy K, Robinson C, Franklin M. Simultaneous dual-task interventions that improve cognition in older adults: a scoping review of implementation-relevant details. J Aging Res. 2022; 2022: 6686910. doi: 10.1155/2022/6686910.

[21] Kazazi L, Shati M, Mortazavi SS, Nejati V, Foroughan M. The impact of computer-based cognitive training intervention on the quality of life among elderly people: a randomized clinical trial. Trials. 2021; 22(1): 51. doi: 10.1186/s13063-020-05008-4.

[22] Lampit A, Hallock H, Valenzuela M. Computerized cognitive training in cognitively healthy older adults: a systematic review and meta-analysis of effect modifiers. PLoS Med. 2014; 11(11): e1001756. doi: 10.1371/journal.pmed.1001756.

[23] Katz N, Averbuch S, Bar-Haim Erez A. Dynamic Lowenstein Occupational Therapy Cognitive Assessment-Geriatric Version (DLOTCA-G): assessing

change in cognitive performance. Am J Occup Ther. 2012; 66(3): 311-9. doi: 10.5014/ajot.2012.002485.
[24] Hornstein S, Zantvoort K, Lueken U, Funk B, Hilbert K. Personalization strategies in digital mental health interventions: a systematic review and conceptual framework for depressive symptoms. Front Digit Health. 2023; 5: 1170002. doi: 10.3389/fdgth.2023.1170002.
[25] Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ. 2018; 48(1): 1-24. doi: 10.1007/s11165-016-9602-2.
[26] Hertzog MA. Considerations in determining sample size for pilot studies. Res Nurs Health. 2008; 31(2): 180-91. doi: 10.1002/nur.20247.
[27] Hemrungrojn S, Tangwongchai S, Charoenboon T, Panasawat M, Supasitthumrong T, Chaipresertsud P, et al. Use of the Montreal Cognitive Assessment Thai Version to discriminate amnestic mild cognitive impairment from Alzheimer's disease and healthy controls: machine learning results. Dement Geriatr Cogn Disord. 2021; 50(2): 183-94. doi: 10.1159/000517822.
[28] Wongpakaran N, Wongpakaran T. The use of GDS-15 in detecting MDD: a comparison between residents in a Thai long-term care home and geriatric outpatients. J Clin Med Res. 2013; 5(2): 101-11. doi: 10.4021/jocmr1239w.
[29] Katz N, Erez AB, Livni L, Averbuch S. Dynamic Loewenstein Occupational Therapy Cognitive Assessment (DLOTCA) [Manual]. Pequannock, NJ: Maddak Inc; 2011.
[30] de Paula JJ, Malloy-Diniz LF, Romano-Silva MA. Reliability of working memory assessment in neurocognitive disorders: a study of the Digit Span and Corsi Block-Tapping tasks. Braz J Psychiatry. 2016;38(3):262-3. doi:10.1590/1516-4446-2015-1879.
[31] Giuli C, Papa R, Lattanzio F, Postacchini D. The effects of cognitive training for elderly: results from My Mind Project. Rejuvenation Res. 2016; 19(6): 485-94. doi: 10.1089/rej.2015.1791.
[32] Chen Y, Denny KG, Harvey D, Farias ST, Mungas D, DeCarli C, et al. Progression from normal cognition to mild cognitive impairment in a diverse clinic-based and community-based elderly cohort. Alzheimers Dement. 2017; 13(4): 399-405. doi: 10.1016/j.jalz.2016.07.151.
[33] Weicker J, Hudl N, Frisch S, Lepsien J, Mueller K, Villringer A, et al. WOME: theory-based working memory training - a placebo-controlled, double-blind evaluation in older adults. Front Aging Neurosci. 2018; 10: 247. doi: 10.3389/fnagi.2018.00247.
[34] Li G, Xia H, Teng G, Chen A. The neural correlates of physical exercise-induced general cognitive gains: a systematic review and meta-analysis of functional magnetic resonance imaging studies. Neurosci Biobehav Rev. 2025; 169: 106008. doi: 10.1016/j.neubiorev.2025.106008.
[35] Xue D, Li PWC, Yu DSF, Lin RSY. Combined exercise and cognitive interventions for adults with mild cognitive impairment and dementia: a systematic review and network meta-analysis. Int J Nurs Stud. 2023; 147: 104592. doi: 10.1016/j.ijnurstu.2023.104592.
[36] Shatil E. Does combined cognitive training and physical activity training enhance cognitive abilities more than either alone? A four-condition randomized controlled trial among healthy older adults. Front Aging Neurosci. 2013; 5: 8. doi: 10.3389/fnagi.2013.00008.
[37] Buitenweg JIV, van de Ven RM, Prinssen S, Murre JMJ, Ridderinkhof KR. Cognitive flexibility training: a large-scale multimodal adaptive active-control intervention study in healthy older adults. Front Hum Neurosci. 2017; 11: 529. doi: 10.3389/fnhum.2017.00529.
[38] Theill N, Schumacher V, Adelsberger R, Martin M, Jancke L. Effects of simultaneously performed cognitive and physical training in older adults. BMC Neurosci. 2013; 14: 103. doi: 10.1186/1471-2202-14-103.
[39] Hudak EM, Bell KL, Hall C, Grismore LD, LaVere J, Edwards JD. Dynamic useful field of view training to enhance older adults' cognitive and motor function: a pilot study. J Cogn Enhanc. 2021; 5(4): 411-9. doi: 10.1007/s41465-021-00219-2.
[40] Man DWK, Chung JCC, Lee GY, Chow SMK. Evaluation of a virtual reality-based memory training programme for Hong Kong Chinese older adults with questionable dementia: a pilot study. Int J Geriatr Psychiatry. 2012; 27(5): 513-20. doi:10.1002/gps.2746.
[41] Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, et al. Video game training enhances cognitive control in older adults. Nature. 2013; 501(7465): 97-101. doi: 10.1038/nature12486.
[42] Lavner Y, Rabinowitz I. Increasing stimulus duration improves attention and memory performance in elderly with cognitive impairment. SAGE Open Med. 2015; 3: 2050312115621566. doi: 10.1177/2050312115621566.

Journal of Associated Medical Sciences 2026; 59(2): 364-371

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Evaluating the integration of assistive devices in occupational therapy across North India

Shivani Bhardwaj*, Bilal Ahmed, Saba Irem, Saqueba Shahi, Nikhat Sultana

Department of Occupational Therapy, SAHSR, Jamia Hamdard University, India.

ARTICLE INFO

Article history:
Received 18 September 2025
Accepted as revised 27 March 2026
Available online 4 April 2026

Keywords:
Assistive devices, occupational therapy, rehabilitation, North India, barriers, accessibility

ABSTRACT

Background: Assistive devices (ADs) are central to occupational therapy (OT) for enhancing independence and participation, yet their use in North India remains underexplored.

Objectives: This study aimed to evaluate how occupational therapists in North India utilize assistive devices, their level of knowledge and training, and the barriers influencing ADs prescription. Understanding these patterns is essential, as improved ADs integration can strengthen rehabilitation outcomes, reduce long-term disability burden.

Materials and methods: A descriptive cross-sectional survey was conducted using a structured, self-administered questionnaire developed from existing literature and validated by three OT experts. The instrument captured demographic details, types and frequency of assistive devices (ADs) use, knowledge and training related to ADs, and perceived barriers. Participants were recruited through purposive and snowball sampling across hospitals, rehabilitation centers, community health facilities, and private clinics in North India. Data were collected online via Google Forms over a three-month period, and descriptive and chi-square analyses were performed.

Results: A total of 51 occupational therapists participated in the study. Mobility and self-care devices were the most frequently prescribed, with 88.2% of respondents reporting some level of ADs use. Major barriers included limited availability (37%), lack of training (25%), high cost (21%), and client reluctance (17%).

Conclusion: Although occupational therapists value ADs, systemic and cultural challenges limit their integration. Improved training, access to affordable devices, and contextually relevant strategies are needed to optimize ADs use in OT across North India.

Introduction

Assistive devices (ADs) are an essential component of modern rehabilitation practice. They encompass a wide range of tools, equipment, and technologies that support individuals with disabilities in performing activities of daily living, participating in education, maintaining employment, and engaging in social life.[1] The World Health Organization (WHO) estimates that more than 2.5 billion people globally require one or more assistive products, a figure projected to rise to 3.5 billion by 2050.[2] Yet, despite this need, only about 10% of people worldwide currently have access to Ads.[2,3] This inequity highlights a critical challenge in rehabilitation and public health.

* **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, SAHSR, Jamia Hamdard University, India*
E-mail address: Shivani@jamiahamdard.ac.in
doi: 10.66285/JAMS.2026.067
E-ISSN: 2539-6056

In occupational therapy (OT), ADs are central to practice because they directly address limitations in self-care, productivity, and leisure, the three core areas of occupational performance. OT practitioners prescribe, adapt, and train clients in the use of ADs such as mobility aids, self-care tools, communication devices, and environmental modifications.[4] A growing body of research demonstrates that ADs improve functional independence, reduce caregiver burden, and enhance quality of life for individuals with physical, sensory, or cognitive impairments.[5-7] Furthermore, ADs are cost-effective interventions: by enabling independence, they reduce long-term health expenditures and reliance on institutional care.[8]

Despite these documented benefits, challenges to ADs integration persist worldwide. In high-income countries, barriers often relate to device design, training needs, and user adherence.[9] In contrast, in low- and middle-income countries (LMICs), systemic issues such as high costs, limited production, weak supply chains, and sociocultural stigma create significant obstacles.[10-12] A scoping review of global ADs access reported that affordability and lack of trained professionals are the two most frequently cited barriers in LMICs.[13] Additionally, inappropriate devices that do not match users' cultural or environmental contexts often result in abandonment.[14]

The global discourse has increasingly recognized assistive technology as a human right. The 2022 WHO-UNICEF Global Report on Assistive Technology (GReAT) emphasized the need for universal access to ADs as part of primary health care.[2] This framework stresses integration of ADs services into rehabilitation systems, development of national assistive technology policies, and inclusion of assistive products in universal health coverage packages. Occupational therapists are pivotal to this agenda, as they are trained to assess functional needs, select appropriate devices, and facilitate their effective use.[15]

In recent years, research in diverse regions has provided insights into ADs utilization. For example, studies from Europe and North America indicate that ADs significantly reduce participation restrictions among older adults and those with chronic illnesses.[16,17] In Scandinavian countries, policies that guarantee ADs provision as part of welfare systems have led to high levels of user satisfaction.[18] Conversely, research in Africa, Latin America, and South Asia reveals major disparities in availability and use.[19-21] In these regions, access to ADs is frequently dependent on non-governmental organisations or charity program rather than systematic public provision.[22]

In the Indian context, the situation mirrors these LMIC challenges but with unique regional variations. India has an estimated 26.8 million people with disabilities, representing 2.2% of the population.[23] Many of these individuals would benefit from ADs, yet only a small fraction has consistent access.[24] Government schemes such as the Assistance to Disabled Persons for Purchase/Fitting of Aids and Appliances (ADIP) program have attempted to bridge this gap, but distribution is uneven and often limited to mobility aids.[25] Studies have shown that rural populations, women, and people from lower socioeconomic strata are disproportionately underserved.[26,27]

Occupational therapy has been practiced in India since the mid-20th century, with a growing workforce concentrated in urban centers.[28] Indian occupational therapists are increasingly engaged in ADs prescription, particularly for individuals with neurological, musculoskeletal, and developmental disabilities.[29] However, professional challenges remain, including limited exposure to assistive technology in undergraduate curricula, inadequate clinical training, and lack of interdisciplinary collaboration.[30] Research from Tamil Nadu and Maharashtra suggests that although therapists recognize the importance of ADs, their use is often restricted by limited availability and client reluctance.[31,32] Stigma associated with disability and assistive product use is a further cultural barrier that reduces acceptance.[33]

Within this broader national picture, North India presents specific challenges. The region comprises several populous states with diverse socioeconomic conditions, including rural, peri-urban, and metropolitan areas. Rehabilitation services are unevenly distributed, with well-resourced tertiary centers in metropolitan cities and severe shortages in rural districts.[34] Occupational therapy practice is largely concentrated in pediatric and physical dysfunction specialties, with limited infrastructure for community-based rehabilitation.[35] However, systematic evidence on ADs integration in OT practice in this region is scarce.

Research gaps are notable in three areas. First, there is limited data on the types of ADs most prescribed by occupational therapists in North India. Second, little is known about therapists' knowledge, training, and perceptions regarding ADs. Finally, few studies have examined the systemic and cultural barriers that restrict effective integration. Without such evidence, efforts to improve access and build capacity remain fragmented. Therefore, the present study was designed as a cross-sectional study to evaluate the integration of ADs into occupational therapy practice in North India.

Specifically, the study aimed to a) describe the frequency and types of ADs prescribed, b) assess occupational therapists' knowledge and training, c) identify barriers to ADs utilization, and d) examine associations between therapist demographics, practice characteristics, and ADs prescription. By addressing these objectives, this study contributes to the growing evidence base on assistive technology in LMICs and offers insights to guide policy, education, and clinical practice in the Indian context.

Materials and methods

This study employed a descriptive, analytical cross-sectional design, suitable for assessing practices,

perceptions, and barriers related to the use of ADs among occupational therapists. The study was conducted in North India, encompassing multiple states with varied socioeconomic contexts. Occupational therapists were recruited from hospitals, rehabilitation institutes, community health centers, and private clinics. Eligibility criteria included, Registered or practicing occupational therapists, Minimum of one year of professional experience, currently involved in clinical service provision.

Therapists working exclusively in administrative or academic roles without direct client care were excluded. A total of 51 occupational therapists participated, after accounting for exclusions due to incomplete responses.

Sampling and recruitment

Participants were recruited through purposive and snowball sampling. Invitations were distributed electronically via email, professional networks, and institutional contacts. To increase participation, follow-up reminders were sent after two weeks. The sampling strategy was chosen to ensure representation from both urban and semi-urban clinical contexts, given the uneven distribution of occupational therapy services in North India.

Sample size

To determine the required sample size for this study, we use the formula for a single proportion:

$$N = (Z^2 \times p \times (1 - p)) / E^2$$

where:

- N = required sample size
- Z = Z-score corresponding to the desired confidence level (1.96 for 95% confidence)
- p = Estimated proportion of occupational therapists aware of assistive technology Based on the previous study, 85% (0.85) of respondents have prescribed an assistive device.
- E = Margin of error (precision level, usually 10% or 0.10

Thus, the minimum required sample size for the study is 49 to ensure statistically reliable results with a 95% confidence level and a 10% margin of error.[36,37]

Data collection tool

Data were collected using a structured, self-administered questionnaire, developed following a review of existing literature on ADs use in rehabilitation. The tool included four domains: demographic details, types and frequency of ADs prescription, knowledge and training related to assistive technology, and perceived barriers. Its content validity was reviewed by three senior OT experts, yielding a scale-level CVI of 0.89. The questionnaire was pilot tested with five therapists to ensure clarity and usability, and internal consistency for key domains ranged from 0.74 to 0.82, indicating acceptable reliability.[35] Data were collected online via Google Forms over a three-month period (January-March 2024).

Procedure

The survey was administered in online formats. Online responses were collected through a secure Google Forms link. Informed consent was obtained electronically or in writing prior to participation.[38] On average, the questionnaire required 10-15 minutes to complete.

Statistical analysis

Data were coded and analyzed using IBM SPSS Statistics version 25.0. Descriptive statistics (frequency, percentages, means, and standard deviations) were computed to summaries demographic data, frequency of ADs use, and barriers reported. Inferential analyses included chi-square tests to examine associations between demographic/practice variables and ADs use. The significance level was set at $p<0.05$.

This approach allowed exploration of whether therapist experience, practice area, or workplace setting influenced ADs integration. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp. was used for statistical analysis. The frequency of categorical variables was reported using descriptive statistical analysis.

The total number of enrolled subjects in the study were 51. The descriptive analysis was done for each item. Analysis and interpretation were done in accordance with the objectives laid down for the study. The purpose of analysis is to reduce the data into an interpretable and meaningful form so that the result can be compared and significance can be identified.

Ethical considerations

The study adhered to the ethical principles outlined in the Declaration of Helsinki (2013 revision).[38] Participation was voluntary, and respondents were assured of confidentiality and anonymity. Data were stored securely and accessed only by the research team.

Results

Participant characteristics

A total of 51 occupational therapists participated. The majority were female (66.7%), with males comprising 33.3% (Table 1 and Figure 1). Most respondents had 1-5 years of experience (70.6%), followed by 6-10 years (15.7%), 11-20 years (11.8%), and >20 years (2.0%) (Table 2 and Figure 2). Regarding practice areas, pediatrics (60.78%) and physical dysfunction (35.3%) were the most represented specialties (Table 3). The practice areas included 31 participants working in pediatric settings, 18 in physical dysfunction (orthopedics/neurology), 1 in community-based rehabilitation, and 1 in the "other" category, who reported working as a home-based occupational therapist.

Table 1. Gender distribution of participants.

Gender	Number	Percent
Female	34	66.7
Male	17	33.3

Table 2. Years of professional experience.

Experience (years)	Number	Percent
1-5	36	70.6
6-10	8	15.7
11-20	6	11.7
>20	1	2.0

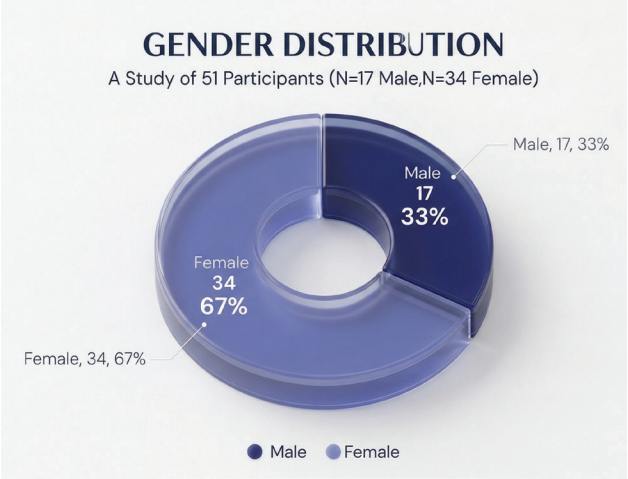


Figure 1. *Gender distribution of participants.*

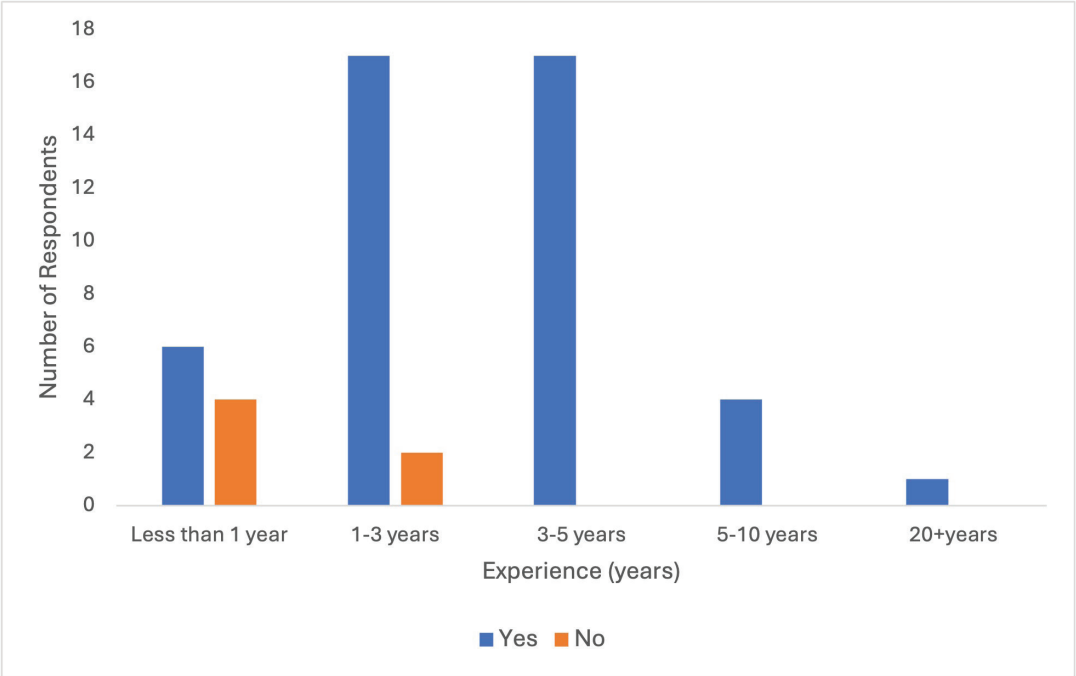


Figure 2. *Assistive devices prescription with years of professional experience.*

Table 3. Practice areas of participants.

Area of practice	Count	Percentage
Pediatrics	31	60.78
Physical dysfunction (orthopedics/neurology)	18	35.29
Others (please specify)	1	1.96
Community-based rehabilitation	1	1.96

Use of assistive devices

Out of 51 therapists, 88.2% reported prescribing assistive devices, while 11.8% had not. Frequency of prescription varied: "Sometimes" (47.1%) was most common, followed by "Rarely" (27.5%), "Often" (23.5%), and "Never" (2.0%) (Table 4 and Figure 3).

Table 4. Frequency of assistive device prescription.

Frequency	Count	Percent
Often	12	23.5
Sometimes	24	47.0
Rarely	14	27.5
Never	1	2.0

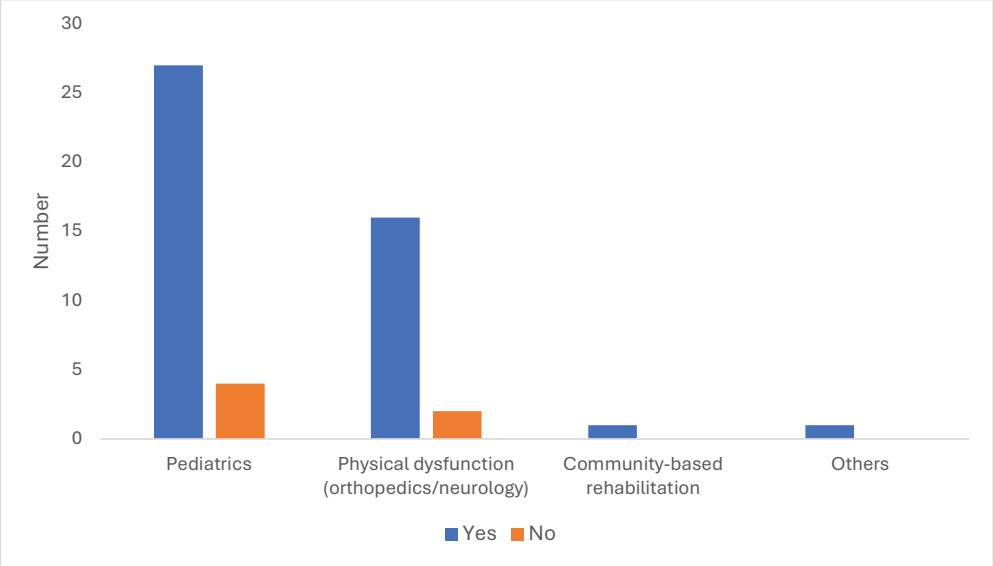


Figure 3. *Frequency of assistive device prescription by practice area.*

Barriers to integration

The most frequently reported barriers were limited availability (37%), lack of training (25%), high cost of devices (21%), and client reluctance due to stigma or cultural factors (17%) (Table 5 and Figure 4).

Table 5. Barriers to assistive device integration.

Barrier*	Percent
Limited availability	37
Lack of training	25
High cost	21
Client reluctance	17

Note: Multiple responses were chosen.

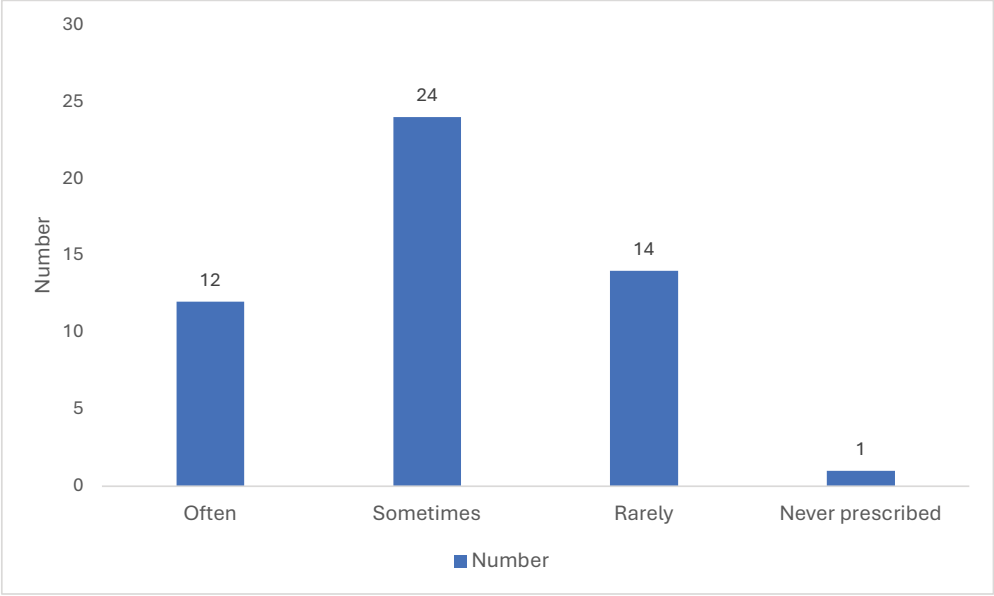


Figure 4. *Frequency of assistive devices prescription.*

Association with professional characteristics

Chi-square analysis revealed a significant relationship between years of experience and frequency of ADs prescription (χ^2=10.64, p=0.0309). Therapists with greater experience were more likely to prescribe ADs frequently. However, no significant relationship was found between practice area and ADs prescription (p=0.9576) (Table 6).

Table 6. Associations with professional characteristics.

Chi-square	*p* value	Degrees of Freedom	Significance
10.64	0.0309	4	Significant

Discussion

This study evaluated the integration of ADs in occupational therapy (OT) practice in North India. Findings indicate that while most occupational therapists prescribe ADs, usage is not yet routine, with most reporting that they prescribe "sometimes" rather than consistently. Mobility and self-care devices were the most prescribed, which aligns with international evidence highlighting mobility aids as the most widely used and accepted category of assistive products.[17,22]

A critical observation was the significant association between therapist experience and likelihood of prescribing ADs. This suggests that clinical exposure and accumulated practice knowledge positively influence integration, echoing earlier findings that professional confidence and competence are crucial for effective ADs use.[9] Conversely, "Although no significant association was found between practice area and the frequency of ADs prescription, this result should be interpreted cautiously. Other descriptive data from other studies show that many commonly prescribed ADs (e.g., heavyweight writing utensils, pencil grips, and fine-motor grippers) are predominantly pediatric-focused, likely reflecting the higher proportion of therapists practicing in pediatrics in India. However, global literature consistently identifies older adults as the major consumer group for ADs, suggesting that ADs needs vary substantially across specialties. Because our sample included few practitioners from geriatric rehabilitation, the study may not have fully captured specialty-specific prescribing trends.[34]

Reported barriers—limited availability, inadequate training, high device costs, and client reluctance-mirror challenges documented in other LMICs.[13,19] Limited device availability reflects weak supply chains and inconsistent governmental support, issues noted in both South Asia and sub-Saharan Africa.[11] The training gap underscores a broader concern regarding the limited inclusion of assistive technology in OT curricula in India, as reported by Arya and Pandey.[30] Cost and affordability remain central obstacles, with many families unable to purchase devices despite government schemes.[24] Finally, stigma and cultural reluctance remain significant deterrents to acceptance, consistent with qualitative evidence from India and other Asian contexts.[33]

Together, these findings reinforce the need for multipronged strategies to strengthen ADs integration. Education reforms should embed practical training on device assessment, selection, and customization. At a policy level, equitable distribution through national schemes such as ADIP must be improved, with broader coverage beyond mobility aids. Locally manufactured, cost-effective devices could reduce financial barriers. Finally, culturally sensitive awareness campaigns may improve acceptance among clients and caregivers.

Limitations

This study was cross-sectional, limiting causal inference. The sample size, although adequate for descriptive and chi-square analyses, was modest and region-specific, potentially limiting generalizability to all of India. Reliance on self-reported data may also have introduced reporting bias. Additionally, qualitative insights were limited, suggesting that future research should include interviews or focus groups to capture in-depth perspectives. One of the other limitations identified was that the questionnaire primarily focused on assistive device use in basic ADLs and did not comprehensively capture other occupational domains such as IADLs, work, education, and social participation, which may limit the understanding of the full scope of assistive device utilization in practice.

Conclusion

This cross-sectional study highlights that while occupational therapists in North India recognize the importance of ADs, their integration into routine practice remains inconsistent. Prescription is more frequent among experienced practitioners, suggesting that clinical exposure strongly influences utilization. However, systemic barriers-including limited availability, inadequate training, high device costs, and cultural reluctance-continue to restrict effective adoption. Addressing these challenges requires coordinated action through education, policy, and service delivery. Strengthening OT curricula with hands-on ADs training, expanding government schemes for equitable device provision, and promoting culturally sensitive awareness campaigns could substantially enhance ADs integration. Broader, multi-center research with larger samples is warranted to inform scalable interventions. Ultimately, improving access to and use of ADs has the potential to significantly enhance independence, participation, and quality of life for persons with disabilities in India.

Implications for rehabilitation

The findings of this study highlight the need for structured training programs that emphasize hands-on instruction in assistive devices (ADs) prescription, selection, and adaptation. Incorporating practical workshops, case-based learning, and supervised clinical exposure into undergraduate, postgraduate, and continuing professional development programs may enhance occupational therapists' competence and confidence in AD-related decision-making. Such training can directly improve the appropriateness and effectiveness of device prescription in diverse clinical settings.

In low-resource contexts, the promotion of affordable and locally produced assistive devices may help address financial barriers faced by clients. Collaboration between occupational therapists, local manufacturers, and rehabilitation engineers can support the development of cost-effective, culturally appropriate devices, thereby improving accessibility and long-term device use.

At the systems level, expansion and strengthening of national assistive device schemes are necessary to ensure equitable access across regions and practice settings. Policy initiatives that integrate occupational therapists into procurement, prescription, and follow-up processes may enhance service delivery and reduce disparities in device provision.

Finally, public awareness initiatives led by rehabilitation professionals can play a critical role in reducing stigma and improving client and family acceptance of assistive devices. Education through community outreach, counselling, and advocacy may promote positive attitudes toward ADs use, ultimately supporting participation in daily activities and social inclusion.

Ethical approval

Ethical clearances not required & Informed written consent taken from all subjects.

Funding

This study is not funded by any organisations and author.

Conflict of interest

'The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article'.

CRediT authorship contribution statement

All authors have equally contributed to this research.

Acknowledgements

The authors would like to express their sincere gratitude to all the occupational therapy practitioners who generously participated in this study and shared their valuable insights.

References

[1] Federici S, Meloni F, Borsci S. The abandonment of assistive technology in Italy: a survey of National Health Service users. Eur J Phys Rehabil Med. 2016; 52(4): 516-26. PMID: 26637347.

[2] World Health Organization. Global report on assistive technology. Geneva: WHO; 2022. Available from: https://www.who.int/publications/i/item/9789240068069.

[3] Matter R, Harniss M, Oderud T, Borg J, Eide AH. Assistive technology in resource-limited environments: a scoping review. Disabil Rehabil Assist Technol. 2017; 12(2): 105-14. doi: 10.3109/17483107.2016.1188170.

[4] American Occupational Therapy Association. Occupational therapy practice framework: domain and process, 4th Edition). Am J Occup Ther. 2020; 74(Suppl 2): 7412410010; 1-48. doi: 10.5014/ajot.2020.74S2001.

[5] Desmond D, Layton N, Bentley J. Assistive technology and people: a position paper from the Global Research, Innovation and Education on Assistive Technology (GREAT) Summit. Disabil Rehabil Assist Technol. 2018; 13(5): 437-44. doi: 10.1080/17483107.2018.1471169.

[6] Steel EJ, Layton N, Foster M, Bennett S, van der Zalm J, Moodie M, et al. The economic case for assistive technology: systematic review. Int J Environ Res Public Health. 2022; 19(6): 3671. doi: 10.3390/ijerph19063671.

[7] Mortenson WB, Demers L, Fuhrer MJ, Jutai JW, Lenker JA, Depa M. Effect of an assistive technology intervention on older adults' activity engagement: a multicenter randomized controlled trial. Arch Phys Med Rehabil. 2018; 99(11): 2160-6. doi: 10.1016/j.apmr.2018.03.010.

[8] Khasnabis C, Heinicke Motsch K, Achu K, Al Jubah K, Brodtkorb S, Chervin P, et al., editors. Assistive technology for children with disabilities: creating opportunities for education, inclusion and participation. Geneva: WHO; 2015.

[9] Phillips B, Zhao H. Predictors of assistive technology abandonment. Assist Technol. 1993; 5(1): 36-45. doi: 10.1080/10400435.1993.10132205.

[10] MacLachlan M, Banes D, Bell D, Borg J, Donnelly B, Fembek M, et al. Assistive technology policy: a position paper from the Global Research, Innovation and Education on Assistive Technology (GREAT) Summit. Disabil Rehabil Assist Technol. 2018; 13(5): 454-66. doi: 10.1080/17483107.2018.1468496.

[11] Holloway C, Austin V, Barbareschi G, Ramos Barajas F, Dyble M, Gallagher B, et al. Scoping research report on assistive technology: London: Global Disability Innovation Hub; 2018.

[12] Layton N, Borg J. Global perspectives on assistive technology: access, affordability, and sustainability. Technol Disabil. 2020; 32(1-2): 1-5. doi: 10.3233/TAD-200255.

[13] Smith EM, Gowran RJ, Mannan H, Donnelly B, Alvarez L, Bell D, et al. Barriers and facilitators to access and use of assistive technology: a systematic review. Int J Environ Res Public Health. 2018; 15(12): 2624. doi: 10.3390/ijerph15122624.

[14] Federici S, Borsci S. Providing assistive technology in Italy: the perceived adequacy of public services and the size of the phenomenon of abandonment. Disabil Rehabil Assist Technol. 2016; 11(1): 22-31. doi: 10.3109/17483107.2014.913710.

[15] Bauer GF, Jenny GJ. Development, implementation and dissemination of occupational health management (OHM): Putting salutogenesis into practice. In: Mittelmark MB, Sagy S, Eriksson M, editors. The Handbook of Salutogenesis. Cham: Springer; 2017: pp 211-24. doi: 10.1007/978-3-319-04600-6_21.

[16] Pettersson I, Appelros P, Ahlström G. Lifeworld perspectives utilizing assistive devices: individuals, lived experience after stroke. Scand J Occup Ther. 2020; 27(6): 446-56. doi: 10.1080/11038128.2019.1586780.

[17] Layton N, Steel E, Manjak NJ. An analysis of Australian survey data on assistive technology use in daily living. Disabil Rehabil Assist Technol. 2022; 17(8): 868-75. doi: 10.1080/17483107.2021.1886452.

[18] Borg J, Larsson S, Östergren PO. The right to assistive technology: for whom, for what, and by whom? Disabil Rehabil. 2011; 33(21-22): 1799-801. doi: 10.3109/09638288.2011.555999.

[19] Visagie S, Duffield S, Unger M, Scheffler E, Schraner I. Developing a sustainable assistive technology service delivery model for South Africa: outcome of a national consultative process. Disabil Rehabil Assist Technol. 2013; 8(5): 386-93. doi: 10.3109/17483107.2013.764516.

[20] Borg J, Larsson S, Östergren PO. Assistive technology provision: towards a national policy in Bangladesh. Disabil Rehabil. 2012; 34(10): 799-803. doi: 10.3109/0963 8288.2011.619621

[21] Oderud T. Surviving spinal cord injury in low-income countries. Afr J Disabil. 2014; 3(2): 80. doi: 10.4102/ajod.v3i2.80.

[22] Matter RA, Eide AH, Harniss M. Assistive technology in resource-limited settings: new evidence from Latin America. Disabil Rehabil Assist Technol. 2017; 12(3): 276-82. doi: 10.3109/17483107.2016.1188173.

[23] Government of India. Census of India 2011: Data on disability. Office of the Registrar General and Census Commissioner, India. Available from: https://censusindia.gov.in.

[24] Ghosh S, Banerjee R, Banerjee T. Assistive technology in India: unmet need, challenges, and opportunities. Indian J Public Health. 2021; 65(1): 95-9. doi: 10.4103/ijph.IJPH_611_20.

[25] Ministry of Social Justice and Empowerment. Assistance to Disabled Persons for Purchase/Fitting of Aids and Appliances (ADIP) Scheme. Government of India; 2020.

[26] Mukherjee S, Samanta R. Accessibility of assistive technologies for persons with disabilities in rural India. Disabil CBR Incl Dev. 2020; 31(1): 27-51. doi: 10.5463/dcid.v31i1.802.

[27] John A, Samuel R, George P. Exploring accessibility and challenges in the use of assistive technology among persons with disabilities in Kerala, India. Disabil Rehabil Assist Technol. 2022; 17(5): 531-8. doi: 10.1080/17483107.2020.1835231.

[28] Murthi KM. Evolution of occupational therapy practice in India: an overview of the historical foundation and current practice. Ann Int Occup Ther. 2019; 2(3): 141-8.

[29] Senjam SS, Mannan H. Assistive technology: The current perspective in India. Indian Journal of Ophthalmology. 2023; 71(5): 1804-9.

[30] Murthi KM. Evolution of occupational therapy practice in India: An overview of the historical foundation and current practice. Ann. Int. Occup. Ther. 2019: 2(3): 141-8. doi: 10.3928/24761222-20190314-01.

[31] Brien M, Krishna D, Borish M, Coutinho F, Bernardo A, Shah SR, et al. Enabling local provision of assistive products in rural South India: an organisational survey of needs, barriers, and facilitators. Disabil Rehabil Assis Technol. 2024; 19(8): 2904-14. doi: 10.1080/17483107.2024.2321601.

[32] Senjam SS, Manna S, Kishore J, Kumar A, Kumar R, Vashist P, et al. Assistive technology usage, unmet needs and barriers to access: a sub-population-based study in India. The Lancet Reg Health Southeast Asia. 2023; 15: 100213 doi: 10.1016/j.lansea.2023.100213.

[33] Singh P, Sharma R. Stigma and acceptance of assistive devices among persons with disabilities in India: a qualitative inquiry. Disabil Rehabil Assist Technol. 2022; 17(6): 623-31. doi: 10.1080/17483107.2020.1841309.

[34] Kumar SG, Roy G, Kar SS. Disability and rehabilitation services in India: Issues and challenges. J Family Med and Prim Care. 2012; 1(1): 69-73. doi: 10.4103/2249-4863.94458.

[35] Mani K. Utilization of assistive devices in occupational therapy practice in Tamil Nadu, India: a statewide survey. Indian J Occup Ther. 2020; 52(2): 43-9. doi: 10.4103/ijoth.ijoth_31_19.

[36] Bondoc S, et al. Assistive technology and occupational performance. Am J Occup Ther. 2016; 70(Suppl 2).

[37] Schaeffer RL, Mendenhall W, Ott L. Elementary survey sampling. 4th Ed. Belmont (CA): Duxbury Press; 1990.

[38] World Medical Association. World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. JAMA. 2013; 310(20): 2191-4. doi: 10.1001/jama.2013.281053.

Journal of Associated Medical Sciences 2026; 59(2): 372-381

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Role of vitamin C or vitamin E on platelet levels in dengue fever: A randomized controlled trial

Akshay Basireddy, Vignessh Raveekumaran*, Mohamed Hanifah

Department of General Medicine, Mahatma Gandhi Medical College and Research Institute, Sri Balaji Vidyapeeth (Deemed to be University), Pondicherry, India.

ARTICLE INFO

Article history:
Received 24 November 2025
Accepted as revised 7 April 2026
Available online 11 April 2026

Keywords:
Antioxidant, ascorbic acid, Dengue, thrombocytopenia, tocopherols, oxidative stress.

ABSTRACT

Background: Adjunctive antioxidant supplementation (vitamin C or E) has been proposed to improve hematologic recovery and hemoconcentration in dengue.

Objectives: To compare platelet recovery and packed cell volume (PCV) changes over 7 days among patients randomized to vitamin C, vitamin E, or control; and to explore differences in ICU admission.

Materials and methods: Randomized controlled trial with three arms (vitamin C (N=100, 500 mg three times daily (total 1.5 gm/day) and, vitamin E (N=100, 400 mg once daily) eight days with standard supportive care, control (N=100, received only standard supportive care), total participants 300). Primary endpoints were between-group differences in change for platelet count (PC) and PCV from day-1 to day-7, tested via Group×Time interaction (two time points) with baseline adjustment. ICU admission was compared across arms. Two-sided $p<0.05$ was statistically significant.

Results: By day-7, the mean PC were vitamin C 115,740±68,990/µL, vitamin E 119,760±70,214/µL, control 131,170±75,081/µL. PCV at day-7 was vitamin C 38.8±5.5%, vitamin E 40.8±7.7%, control 38.8±6.5%. The Group×Time interaction was not significant for PC ($p=0.563$) or PCV ($p=0.973$), indicating no between-group differences in change over 7 days. ICU admissions and the overall comparison were not significant ($p=0.105$).

Conclusion: In this three-arm randomized controlled trial, all groups demonstrated expected hematologic recovery by day-7. However, high-dose vitamin C or E did not show superior platelet recovery or hemoconcentration compared with standard care. ICU admission rates were also not significantly different across arms.

Introduction

Dengue (DENV) infection is a rapidly expanding mosquito-borne viral disease caused by four serotypes of the DENV virus, placing nearly half of the world's population at risk and resulting in an estimated 390 million infections annually.[1] In India, hyperendemic transmission and seasonal monsoon surges have driven case counts above 200,000 with hundreds of deaths annually.[2] Clinically, DENV presents with a triphasic course, include febrile, critical, and recovery phases during which thrombocytopenia arises from both suppression of bone-marrow megakaryopoiesis and immune-mediated platelet destruction via complement activation and cross-reactive antibodies.[3,4]

* **Corresponding contributor.**
Author's Address: *Department of General Medicine, Mahatma Gandhi Medical College and Research Institute, Sri Balaji Vidyapeeth (Deemed to be University), Pondicherry, India.*
E-mail address: vignesshravee@gmail.com
doi: 10.66285/JAMS.2026.068
E-ISSN: 2539-6056

Current guideline-based management remains primarily supportive, focusing on careful fluid management and monitoring for complciations.[5] Concurrently, dengue induces oxidative stress marked by elevated reactive oxygen species, lipid peroxidation, and protein carbonylation that correlated inversely with platelet counts and directly with disease severity.[6,7] Although endogenous enzymatic defenses such as superoxide dismutase and glutathione peroxidase mitigate basal oxidative insults, these systems are overwhelmed during acute infection.[8,9]

Exogenous antioxidants have thus become a focus of adjunctive dengue therapy. Vitamin C (ascorbic acid) scavenges aqueous-phase radicals and regenerates lipid-bound α-tocopherol, while vitamin E (α-tocopherol) quenches membrane peroxyl radicals and stabilizes endothelial barriers.[10–12] Clinical evidence for antioxidant supplementation in dengue remains limited and heterogeneous. As recent reviews of micronutrient/adjunctive interventions note that much of the available evidence for vitamin C and vitamin E in dengue from observational studies or small trials with variable methodology and outcomes, therefore, reported associations should be interpreted as hypothesis-generating rather than definitive evidence of efficacy of vitamin C and E in dengue.[5,13-15]

By comparison, the recommended dietary allowance (RDA) for vitamin C in healthy adults in 90m for men and 75 mg/day for women, and the tolerable upper intake level (UL) for adults is 2,000 mg/day.[16] Similarly, the adult RDA for vitamin E is 15 mg/day; high-dose supplementation can increase bleeding risk, and adult Uls varies.[16] In our trial, vitamin C was administer as 1.5 gm/day in divided doses and vitamin E as 400 mg/day, representing doses far above dietary requirements.

This present study seeks to address this critical evidence gap by directly comparing vitamin C and vitamin E in a randomized, blinded, controlled trial and to estimate their effects on platelet recovery and clinically relevant outcomes.

To compare the efficacy of high-dose vitamin C vs vitamin E on platelet recovery kinetics and clinical outcomes in adult dengue patients with thrombocytopenia. The primary objective was to assess and compare the between-group change in platelet count from baseline to day-7 among the vitamin C, vitamin E, and control arms. The secondary objectives were to compare PCV change over 7 days, incidence/severity of bleeding manifestations, requirement of platelet transfusion, and ICU admission across the three arms.

Materials and methods

Study design and setting

The study was a randomized, controlled, blinded, three-arm, parallel-group trial comparing vitamin C, vitamin E, and standard treatment alone, conducted in the Medicine Department of a tertiary care teaching hospital in Puducherry for the period of 18 months (October 2023 to April 2025) after Institutional Human Ethics Committee (IHEC) approval.

Study participants

Adult patients (18-60 years) of both genders. Dengue diagnosis required compatible clinical features pule laboratory evidence where the confirmed acute dengue was defined as positive NS1 antigen or NAAT/RT-PCR obtained within the first seven days of illness, while the presumptive recent dengue was defined as single-sample IgM positivity (MAC-ELISA) per public health guidance.[17] Also, paired serology was obtained when feasible and thrombocytopenia (platelets <150,000/mm^3) presenting within 72 hours of fever onset were enrolled. Baseline disease severity was categorized using WHO dengue classification (dengue with/without warning signs; severe dengue), only those without severe dengue at admission were eligible.

Exclusion criteria were: 1) major comorbidities other than diabetes and hypertension (e.g. chronic kidney disease (CKD), liver disorders or disease, immune thrombocytopenic purpura (ITP), sepsis), 2) patients presented with confirmed co-infections like scrub typus and malaria), 3) current use of antiplatelets/anticoagulants drugs or prior antioxidant supplementation (vitamin C and E), 4) pregnancy or lactation; and 5) severe dengue at presentation (per WHO criteria),[1] including shock, severe bleeding/active bleeding at admission, or severe organ involvement. Patients with isolated dengue 'warning signs' at baseline were not excluded; as warning signs were recorded and monitored during follow-up as part of the clinical assessment.

Eligible participants provided written informed consent and randomized 1:1:1 to receive either 400 mg α-tocopherol (vitamin E) once daily, 500 mg ascorbic acid (vitamin C) thrice daily, or control for eight days, in addition to standard supportive care (fluid management, antipyretics, platelet transfusion as needed). Daily clinical monitoring included vital signs, bleeding assessment, and warning signs of severe dengue, laboratory evaluation (complete blood count (CBC) with platelet counts and PCV; was performed at two time-points day-1 and day-7. Liver and renal function test (LFT, RFT)) were obtained as per routine clinical care when indicated (Figure 1).

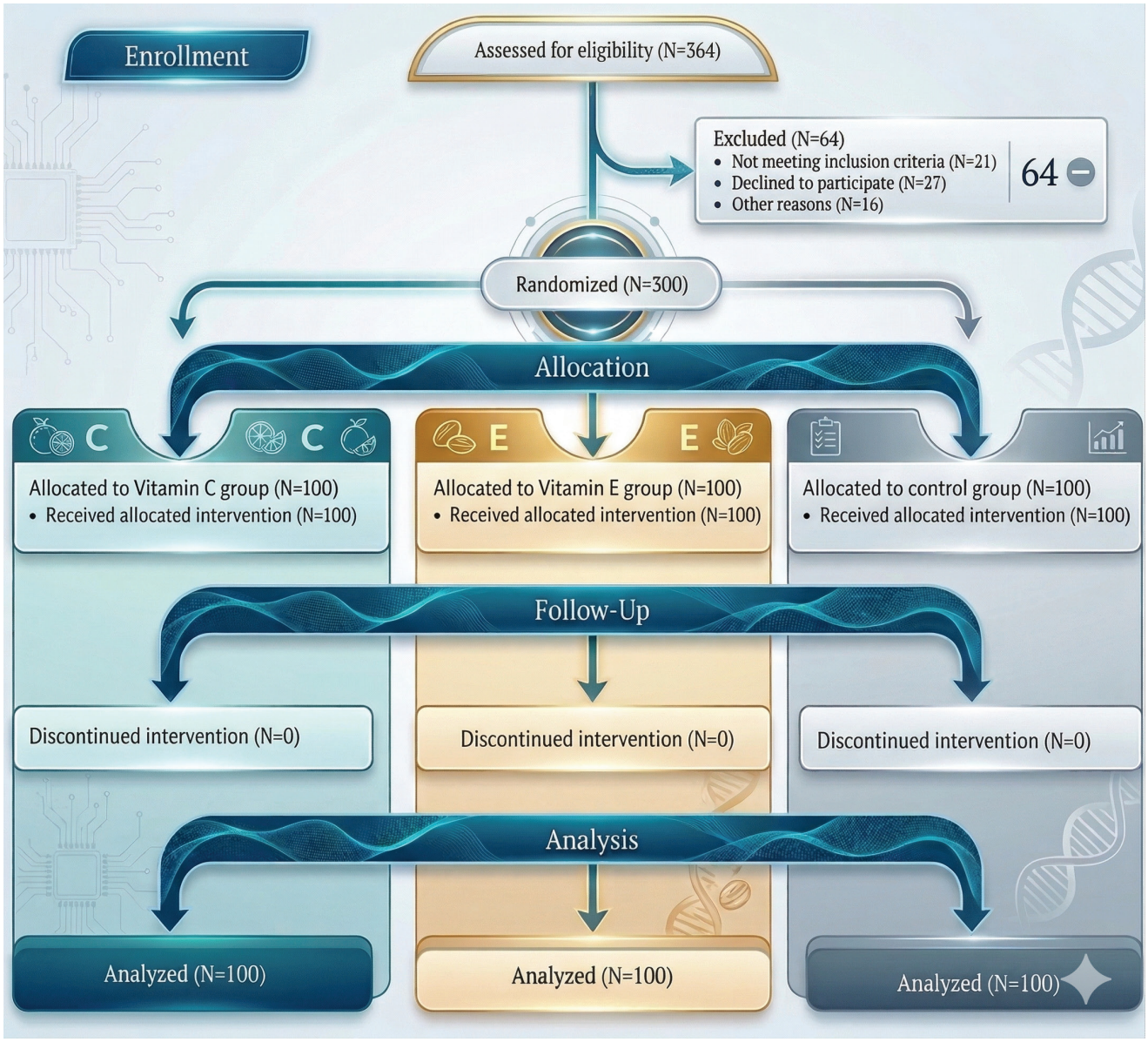


Figure 1. *CONSORT diagram*

Endpoints of the study

The primary endpoint was change in platelet count from baseline to day-7, and secondary endpoints included PCV variation, bleeding manifestations, platelet transfusion requirement, ICU admission, and length of hospital stay. Clinical improvement was defined as meeting all the following criteria at the time of discharge which includes,

1. Afebrile (body temperature <37.5 °C for at least 24 hours)
2. No bleeding manifestations
3. Hemodynamic stability without vasopressor support

Platelet count ≥100,000/mm^3[31]

Patients fulfilling these criteria were classified as 'improved'.

Sample size determination

From Chandra *et al*. study done on 2013 in Uttar Pradesh showed that the mean platelets count improved from 1.8±0.3x10^5/mm^3 at day-1 to 2.3±0.5 x10^5/mm^3 at day 8 with supplements of vitamin C or E.[13] Assuming that the between-group difference in platelet (Δplt) is 0.3x10^5/mm^3 clinically. The prespecified σΔ is 0.50x10^5/mm^3 which is conservative derived from the reported SDs (pooled with modest baseline-follow-up correlation and equal to the larger reported SD. Using a two-sided α'=0.025 (Bonferroni for two primary pairwise contrasts), and 80% power and the standard two-sample mean-difference formula:

$$n\,(per\ group) = 2\frac{\sigma_\Delta{}^2\left[Z_{1-\frac{\alpha'}{2}} + Z_{1-\beta}\right]^2}{\Delta^2}$$

Where $Z_{1-\alpha'/2}$ =2.24 and $Z_{1-\beta}$ = 0.84, giving N ~95

per group, which got rounded to highest figure of 100 per group a priori. The annual MRD (medical records department) count for dengue patients were ~1000 dengue admissions per annum considered for feasibility purpose only.

Randomization, blinding, and group allocation

An independent statistician/study pharmacist not involved in enrolment, treatment, or outcome assessment generated the randomization list using a computer-generated sequence with permuted blocks of size 2, allocation ratio 1:1:1 (vitamin C : vitamin E : control), with stratification by age and gender. Allocation was concealed using identical sequentially numbered, opaque, sealed envelopes (SNOSE) prepared by the pharmacist. Investigators and treating clinicians had no access to the sequence; assignment was revealed only after a participant was deemed eligible and enrolled. Because the study medications were administered/ dispensed by clinical staff, the duty medical officer (DMO) (not the investigator) and the nursing staff were necessarily aware of allocation. Participants were blinded to group assignment. Laboratory personnel and the statistician conducting the primary analyses were blinded to allocation with groups coded until database lock. Investigators responsible for outcome assessment and data interpretation remained blinded to group codes until the analysis plan was completed and the database was locked.

Group 1 (vitamin E): 400 mg of α-tocopherol (vitamin E as capsule. Evion, by Abbott composed of Tocopherol Acetate I.P. 400 mg equivalent to α-tocopherol 182 mg) once daily for eight days + standard of care dengue treatment as per WHO.[18]

Group 2 (vitamin C): 500 mg of ascorbic acid thrice daily for eight days (total daily dose 1.5 gm/day) (vitamin C as tab Celin, each uncoated chewable tablet contains, ascorbic acid i.p. 100 mg, sodium ascorbate I.P. 450 mg equivalent to ascorbic acid 400 mg, manufactured by Procter&Gamble Health Ltd.) to achieve high-dose daily intake + standard of care dengue treatment as per WHO.[18]

Group 3 (control): standard of care dengue treatment as per WHO.[18]

Data and sample collection procedure

Data were recorded on standardized case report forms and entered into a secure electronic database. Patients' dengue result and diagnosis were also noted. Baseline demographics, medical and physical examination including the early warning signs of severe dengue, such as bleeding manifestations, capillary leakage, and hydration status were documented on day-1. Baseline laboratory parameters PC, and PCV, were recorded for day-1 and repeated on day-7. Additional testing performed for all patients used for discretion of clinical management. Then ICU status of the patients and the outcome in the patient were also taken. Adherence to study medication was confirmed by daily pill counts.

Venous blood (4-5 mL) was drawn aseptically from the antecubital vein using a sterile syringe and needle, with patient identifiers recorded and samples held at room temperature. To minimize the aging artifacts, all assays were completed within one hour of collection at two time-points (day-1 and day-7). For platelet count, 3 mL of whole blood in K_2-EDTA tube, inverted gently 8-10 times, and analysed on a calibrated automated haematology (Sysmex XN-1000, Japan) analyser within four hours. PCV was measured via the microhematocrit method, two heparinized capillary tubes are filled to 75% capacity, sealed and centrifuged at 11,000xg for 5 minutes in the centrifuge (Beckman benchtop centrifuge, USA), and read to 0.1% accuracy; 10% of samples were duplicated, and weekly calibration and alignment checks ensured reproducibility.

Participants were actively monitored for adverse events (AEs) from the time of enrolment through the end of the study. Daily assessments included vital signs, symptoms on any bleeding or allergic manifestation, and laboratory safety and were all recorded in standardized case report forms.

Statistical analysis of data

Data were analysed in SPSS v24.0. Normality was assessed with Shapiro-Wilk and Q-Q plots. Continuous variables are presented as mean±SD (or median [IQR], as appropriate) and compared across arms (vitamin C, vitamin E, control) using one-way ANOVA with Bonferroni post-hoc tests. Categorical variables are counts (%) and compared to χ^2 or Fisher's exact tests. Longitudinal changes in platelet count (PC) and packed cell volume (PCV) from day-1 to day-7 were evaluated with paired t-tests within groups and repeated-measures ANOVA for between-group trends. For ICU admission, unadjusted risk ratios (RRs) with 95% CIs were computed from observed risks; adjusted RRs used a modified Poisson model (log link, robust variance) including prespecified prognostic covariates: sex, diagnostic modality (NS1 vs IgM vs IgG vs both), dehydration, bleeding manifestations, and vascular leak. Baseline comparability was summarized using standardized mean differences (SMDs) rather than significance testing; covariates showing SMD≥0.10 were retained for adjustment.

Sensitivity analyses were performed using inverse-probability-of-treatment weighting (IPTW) to reduce residual confounding across the three treatment arms. Propensity scores were estimated with one-vs-rest logistic models using the same covariates listed above; per-participant probabilities were normalized to sum to 1. Stabilized weights equalled the marginal probability of the received treatment divided by the participant-specific propensity; mild truncation at the 1st–99th percentiles limited extreme weights.

Weighted outcome models used logistic regression with control as reference. Two-sided α=0.05 analyses were intention-to-treat for primary efficacy endpoints.

Results

The mean age was 36.04±13.04 years that are ranged from 18 to 60 years, and among the patients in all three groups was 36.04±13.08 years, (p=1.000) respectively, while 53 patients (53%) and 47 patients (47%) were males and females in each group, (p=1.000) indicating its variability were identical in all groups. Dengue diagnostic characteristics and their symptoms were presented in Table 1**.** Diagnostic modality differed across arms. Therefore, adjusted for diagnostic modality in secondary models and performed analyses restricted to confirmed acute dengue. Moreover, standardized mean differences were calculated for all variables, where vitamin C vs control in diagnosis is 0.438, dehydration (yes) as 0.414, bleeding manifestation and vascular leak showed 0.307 and 0.206, while for vitamin E vs control, it was 0.718, 0.194, 0.16, 0.065 respectively. Thus, the baseline characteristics were summarized using standardized mean difference (SMD) rather than hypothesis tests. SMDs indicated imbalance between one or both supplement arms and control, and these variables were prespecified for covariate adjustment in outcome models.

Table 1. Dengue characteristics and outcome of the study based on vitamin supplementation (N=300).

Variables	Vitamin C (N=100)	Vitamin E (N=100)	Control (N=100)	*p* value
Diagnostic characteristics				
IgM antibody	34 (34.3)	51 (51.0)	17 (17.0)	<0.001[b]
NS1 antigen	63 (63.6)	48 (48.0)	83 (83.0)	
Both (IgM and NS1)	2 (2.0)	1 (1.0)	0 (0)	
Dehydration status				
No	47 (47.0)	36 (36.0)	27 (27.0)	0.315[b]
Yes	53 (53.0)	64 (64.0)	73 (73.0)	
Bleeding manifestations				
No	95 (95.0)	80 (80.0)	86 (86.0)	0.622[b]
Yes	5 (5.0)	20 (20.0)	14 (14.0)	
Vascular leakage				
No	79 (79.0)	67 (67.0)	70 (70.0)	0.546[b]
Yes	21 (21.0)	33 (33.0)	30 (30.0)	
ICU admission requirement				
Yes	22 (22.0)	31 (31.0)	34 (34.0)	0.105[b]
No	78 (78.0)	69 (69.0)	66 (66.0)	
Outcomes				
Improved	60 (60.0)	51 (51.0)	32 (32.0)	0.001
Deteriorated	6 (6.0)	19 (19.0)	29 (29.0)	
Static	34 (34.0)	30 (30.0)	39 (39.0)	
Platelet transfusion	0 (0)	0 (0)	(0)	-

Note: [a]ANOVA (continuous and parametric), [b]Chi-square test (nominal and non-parametric). The p<0.05 is statistically significant and indicated in boldface. Cells with 0 were considered as 0.5 for analysis. Values given in brackets were percentage. Age given in mean±SD. Ig: immunoglobulin, NS1: non-structured protein 1.

Platelet count and PCV of the study participants in all three groups changed significantly from day-1 to day-7 within each arm in Table 2. However, between-group comparisons of change over time were not statistically significant (Figure 2). The group*time interaction was non-significant for PC (p=0.563) and PCV (p=0.973), and pairwise contrasts vs control were also non-significant for PC (vitamin C vs control p=0.233; vitamin vs control p=0.456) and PCV (vitamin C vs control p=0.802; vitamin E vs control p=0.693). As effect estimates were imprecise, with wide 95% CI that included clinically meaningful benefit and harm, indicating low precision of the between-group comparisons (Table 3). ICU admission did not differ significantly across arms (p=0.105) (Table 1). Clinical status at discharge differed across groups (Table 1), but these differences should be interpreted cautiously given baseline imbalances and the exploratory nature of these analyses. No platelet transfusion was administered in any group during hospitalization.

Table 2. Platelet count and PCV on day-1 and day-7 among the study participants (N=300).

Variables	Day-1	Day-7	Mean difference	t, *p* value*
Platelet count (cells/mm^3)				
Vitamin C	71,970.00±56,210.36	115,740.00±68,990.20	43,770.00±75,657.50	5.79, <0.001
Vitamin E	81,515.00±74,413.98	119,760.00±70,214.26	38,245.00±98,718.00	3.87, <0.001
Control	10,2120.0±69,585.53	131,170.00±75,081.10	29,050.00±85,189.40	3.41; 0.001
Packed cell volume (%)				
Vitamin C	41.25±6.93	38.76±5.53	-2.50±4.60	-5.41, <0.001
Vitamin E	43.36±7.54	40.76±7.66	-2.60±7.40	-3.50, 0.001
Control	41.08±6.29	38.80±6.46	-2.30±4.60	-4.93, <0.001

Note: *Paired t-test (data were continuous and parametric). The p<0.05 is statistically significant and indicated in boldface. PCV: packed cell volume (%)

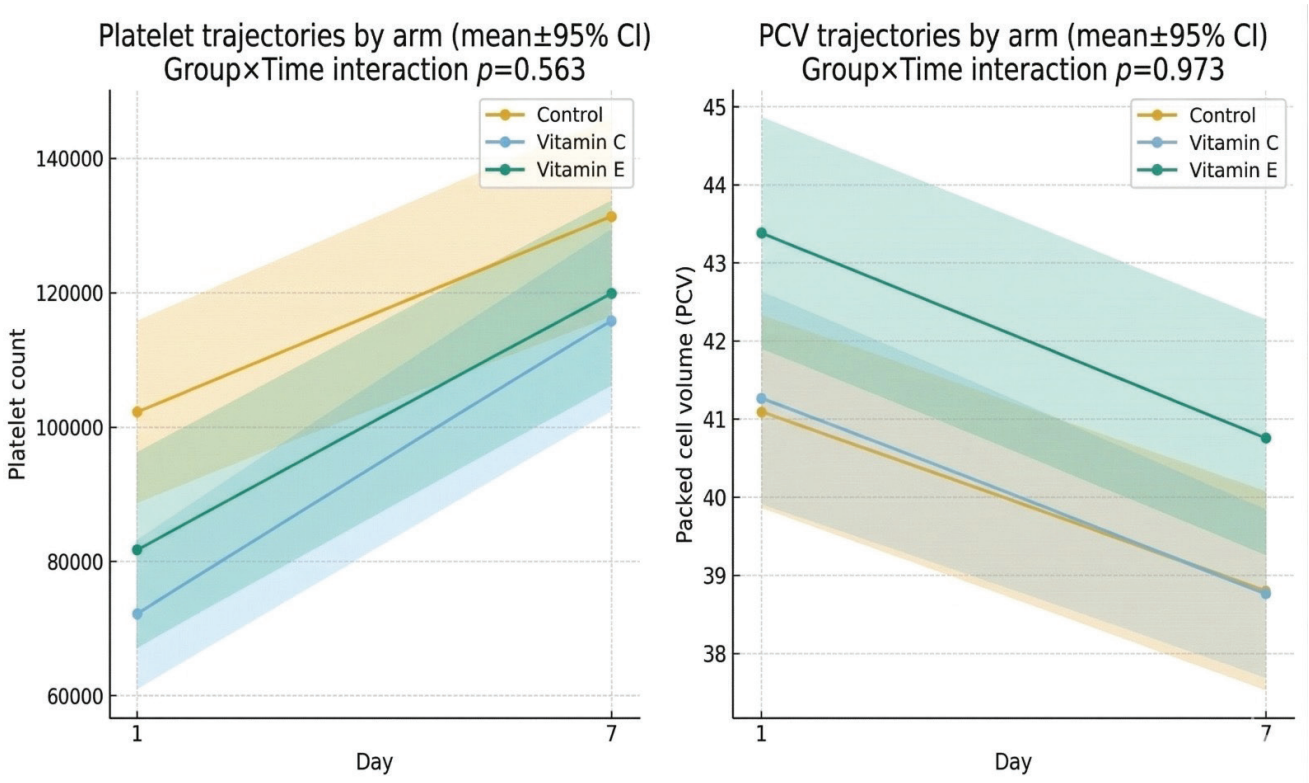


Figure 2. *Platelet count and PCV trajectories over 7 days across treatment arms. (platelet counts increased and PCV declined in all groups over time. Group*time interactions were non-significant, indicating no differential change over 7 days by treatment)*

Table 3. Difference change of platelet count and PCV.

Outcome (ΔDay-7 to Day-1)	**Comparison**	**Δ_estimate**	**CI_low**	**CI_high**	***p* value**
Platelet count	Vitamin C vs control	14720	-9,405.545	38,845.545	0.233
	Vitamin E vs control	9195	-14,930.545	33,320.545	0.456
PCV	Vitamin C vs control	-0.203	-1.786	1.38	0.802
	Vitamin E vs control	-0.316	-1.899	1.267	0.696

Compared to the control arm, vitamin C supplementation was associated with a 35% lower risk of ICU admission (RR (risk ratio) 0.65; 95% CI: 0.40-1.05), and vitamin E with a 9% lower risk (RR 0.91, 95% CI 0.59-1.41). However, neither reduction reached statistical significance, and the overall chi-square test did not detect a significant difference across groups (χ^2 4.50, *p*=0.105). Thus, while there is a trend towards fewer ICU admissions with vitamin C, the evidence is not strong enough to conclude a definitive effect.

The predictive model of the study for ICU admission estimated from a modified Poisson regression with log link and robust variance (sandwich). The model included group as vitamin C, vitamin E, control as reference, age, gender (male vs female), and diagnostic modality (NS1+IgM vs IgM/IgG). Results are reported as adjusted risk ratio (aRR) with 95% CI, where aRR>1 indicates higher ICU risk and <1 indicated lower ICU risk compared with the reference category, holding covariates constant. Estimates are based on complete-case analysis. The adjusted risk difference (RD) for vitamin C vs control was -0.024 and vitamin E vs control was 0.049. On the absolute scale the vitamin C arm's predicted ICU risk is ~2.4% lower than control, and the vitamin E arms' is ~4.9% higher, however, because the RR Cis cross 1, these differences are not statically significant. The only clear is male sex about 3=2.5 times higher adjusted risk. The diagnostic modality term trends higher but remains borderline (Figure 3).

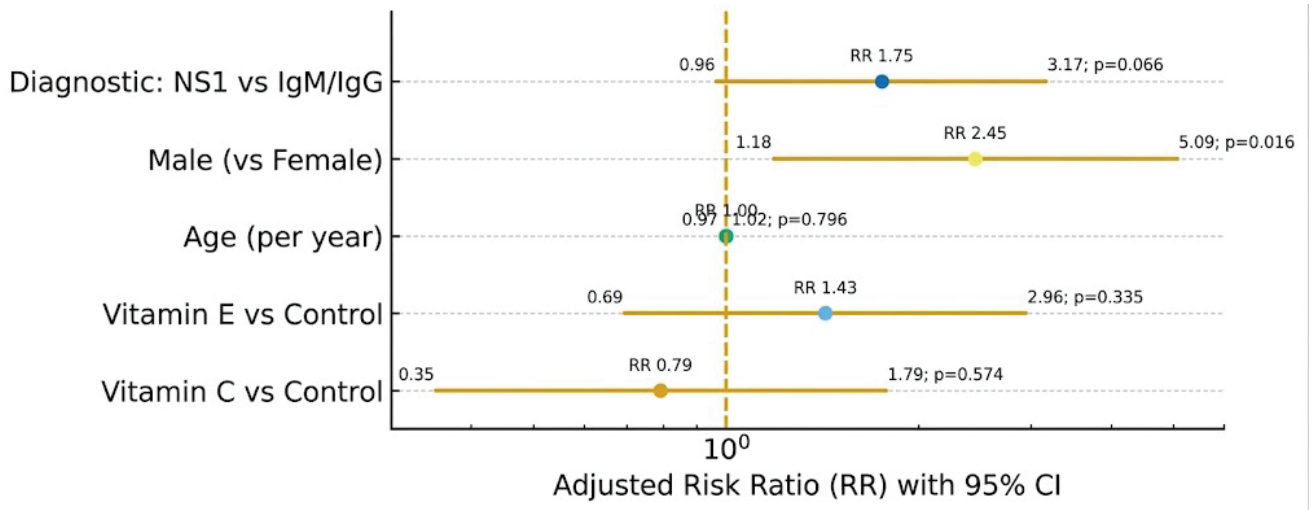


Figure 3. *Predictive model for ICU admission among the groups . (The forest plot shows the adjusted risk ratios for ICU admission according to key clinical and diagnostic predictors. Male sex was the only independent predictor associated with a significant higher risk of ICU admission, whereas vitamin C, vitamin E, age and NS-1 based diagnosis were not significantly associated with ICU admission).*

In IPTW-weighted models, ICU admission did not differ significantly from control for either vitamin arm; however, some discharge-status contrasts differed from control, and these findings should be interpreted cautiously given baseline imbalances, outcome-category subjectively, and imprecision (Table 4).

Table 4. IPTW-weighted models. Please use the full name of IPTW

Contrast	Comparison	OR	CI_low	CI_high	*p* value
ICU admission (Yes vs No)	Vitamin C vs control	2.492	0.971	6.395	0.058
ICU admission (Yes vs No)	Vitamin E vs control	1.978	0.755	5.187	0.165
Deteriorated vs improved	Vitamin C vs control	0.186	0.075	0.462	0
Deteriorated vs improved	Vitamin E vs control	0.408	0.187	0.887	0.024
Static vs improved	Vitamin C vs control	0.479	0.254	0.901	0.022
Static vs improved	Vitamin E vs control	0.612	0.326	1.15	0.127

In analysis, restricted to NS1 confirmed cases, effect estimates were directionally similar to primary analyses, and between-group comparisons remained non-significant. Models adjusted for diagnostic modality yielded comparable results. No adverse events, serious or otherwise, attributable to vitamin C or E supplementation were observed in this study arm during the study period.

Discussion

In our RCT of 300 adults with laboratory-confirmed dengue and thrombocytopenia randomized to vitamin C (500 mg/TDS/day), vitamin E (400 mg/OD/day), or Control. The mean age 36.04±13.08 years, 53% males, $p<1.000$, where all groups significant platelet rises by day-7 (vitamin C: +29,050±85,189 cells/mm^3, $p<0.001$; vitamin E: +38,245±98,718 cells/mm^3, $p<0.001$; Control: +43,770±75,658 cells/mm^3, $p<0.001$). However, the prespecified between-group comparisons of PC and PCV changes were non-significant. ICU admissions were lowest in the vitamin C arm (22% vs Control 34%, $p=0.105$), and overall clinical improvement by discharge was significantly higher with vitamin C (60%) compared to vitamin E (51%) and control (32%) ($p<0.001$).

Our overall results mirror prior evidence done by Soundravally *et al.*, and Ahmed *et al.* that high-dose vitamin C and E speed thrombopoiesis and correct dengue-related hematologic abnormalities.[19,20] Simmons *et al.* noted that hospitalized dengue cases in Southeast Asia increasingly comprise working-age adults, reflecting both evolving herd immunity patterns and occupational exposure risks,[4] similar to our study findings with male predominance. Also, by matching key prognostic factors such as age and sex at enrollment, minimized confounding and improved comparability of baseline prognostic factors across arms.[13,20]

Physiologically, platelet count in dengue typically reach a nadir around days 4-5, before recovering during convalescent phase.[3,21] In our study, by day-7 the vitamin C arm's mean rise of +43,770 cells/mm^3 ($p<0.001$) and the vitamin E arm's by +38,245 cells/mm^3 ($p<0.001$) compared with Control with a gain of +29,050 cells/mm^3 ($p<0.001$). PC increased from day-1 to day-7 within all three-arms, yet between-group differences in PC vs control were not statistically significant. Accordingly, our findings do not establish that vitamin C or E improves platelet recovery compared with standard care.[13,22,23] These results corroborate prior studies by Ahmed *et al.*, and Ramalingam *et al.* demonstrating that high-dose vitamin C and vitamin E accelerate platelet recovery in dengue in their studies.[19,22] Study done in Uttar Pradesh cohorts where oral ascorbate (1.5 gm/day) boosted PC from ~71,000 to >15,0000 cells/mm^3 by day-7, and tocopherol (400 mg/day) produced a 45% greater rise in platelets in day 5 than control.[13] Mechanistically, vitamin C mitigates oxidative stress in megakaryocytes preserving platelet-producing capacity while vitamin E shields circulating platelets from lipid peroxidation and apoptosis.[19,24] By shortening the duration and depth of thrombocytopenia, these antioxidants may reduce transfusion need and bleeding complications. These changes were significant within groups over time; we did not demonstrate statistically significant between-group differences in the magnitude of hematologic change

PCV a key indicator of hemoconcentration and plasma leakage in dengue,[4,19] decline modestly by day-7 in all arms of our study (-2.3 and -2.6% points in all groups, $p<0.001$), mirroring the WHO-2-5% critical-phase recovery threshold and suggesting that adjunctive antioxidant supplementation did not adversely affect the capillary integrity.[25] Preclinical data reveal that ascorbate's preserve endothelial tight-junction proteins (ZO-1, clauding-5) during dengue infection, further supports its potential role in vascular barrier stabilization.[18,26,27] Future studies should combine PCV monitoring with quantitative assessment of vascular leaks or biomarkers like syndecan-1 and angiopoietin-2 to delineate how antioxidant influences dengue-associated endothelial dysfunction.

ICU admissions were lower with vitamin C (22%) and vitamin E (31%) arms vs control (34%), were numerically lower in the vitamin C group but these differences did not reach statistically significant

(p<0.105), thus avoiding the inferring superiority of any supplement for escalation-of-care outcomes in this trial. Similarly, non-significant reductions reported by Ramalingam *et al*., which observed a 35% lower ICU risk with vitamin C (RR 0.65; 95% CI 0.40-1.05).[22] In multivariable analysis, neither antioxidant independently predicted ICU transfer, while larger platelet gains were paradoxically linked to higher ICU odds (OR 1.00, p<0.004), likely reflecting that patients with more severe baseline thrombocytopenia, who require ICU, demonstrate greater larger absolute recovery.

Patient-level outcomes favored antioxidant therapy where 60% of the vitamin C arm improved by discharge vs 51% and 32% in the vitamin E and control arms, respectively (p<0.001). This parallels with Jahnavi *et al*. findings that combined vitamin C and E shortened the leakage phase and reduced hemorrhagic complications in pediatric dengue.[28] Although our trial did not test combination therapy, the superior improvement rate in the vitamin C suggests that its aqueous-phase scavenging may be especially important in mitigating oxidative platelet injury.

Study strengths and limitations

This trials principal strength lies in its rigorous three-arm, randomized, blinded, Control-controlled design with robust allocation concealment, which not only minimizes selection and performance bias but also permits a direct head-to-head comparison of vitamin C vs vitamin E vs control and addressing a critical gap in adjunctive dengue therapy research. The powered sample size for haematological endpoint, and standardized, high-frequency monitoring of platelet count and PCV under real-world supportive-care with uniform dosing and adherence checks further enhanced internal validity. It is advisable to address the study's shortcomings. Being a single-centre study in a tertiary care hospital, the generalizability notably to paediatric population, diagnostic-stage heterogeneity, and insufficient power to detect rare but critical outcomes like major haemorrhages. Secondly, a subset of participants was enrolled based on single-sample IgM, which confers presumptive rather than confirmed acute dengue and may introduce non-differential disease misclassification, typically biasing treatment effects towards the null. Yet NS1-restricted and modality-adjusted analyses did not materially change the findings. Thirdly, we did not include a combine vitamin C+E arm. The study was intentionally designed to compare the monotherapy arms head-to-head against control to isolate individual supplement effects evaluating combination therapy would require an additional arm or factorial design with a larger sample size, which was not feasible within the recruitment period and available resources. Finally, outcomes were evaluated using two protocolized laboratory time-points as this short follow-up may not capture delayed clinical deterioration, late platelet nadir, or longer-term recovery patterns.

Conclusion

In this three-arm randomized trial, high-dose vitamin C (1.5 gm/day) and vitamin E did not demonstrate superior platelet recovery or PCV correction over 7 days compared with standard care. All groups showed expected within-group hematologic recovery. These findings do not support routine antioxidant supplementation for improving PC or PCV endpoints in dengue; larger trials are needed to evaluate any effects on clinical outcomes.

Ethics approval

All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Institutional Ethical Committee of MGMCRI (MGMCRI/Res/01/2023/122/IHEC/101).

Data availability

All data generated or analyzed during this study are included in this published article

Funding

This research did not receive specific funding from public, commercial, or non-profit agencies.

Conflicts of interest

The authors declared no competing interests.

CRediT authorship contribution statement

Vignessh Raveekumaran: conceived and designed the study, carried out all the experimental work, statistical analysis and interpreted, writing: reviewed and edited; **Akshay Basireddy**: carried out all the experimental work, statistical analysis and interpreted, writing: manuscript preparation; **Mohamed Hanifah**: conceived and designed the study, as the project administrator, writing: reviewed and edited the manuscript. All authors read and approved of the final manuscript**.**

Acknowledgements

We express our gratitude to the medical interns and nurses for their valuable assistance and support. I express my gratitude to the laboratory technicians who contributed to the study.

References

[1] Dengue and severe dengue. Accessed November 2, 2022. https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue.

[2] Home :: National Center for Vector Borne Diseases Control (NCVBDC). November 2, 2022. Accessed November 2, 2022. https://nvbdcp.gov.in/.

[3] Srichaikul T, Nimmannitya S. Haematology in dengue and dengue haemorrhagic fever. Best Pract Res Clin Haematol. 2000; 13(2): 261-76. doi: 10.1053/beha.2000.0073.
[4] Simmons CP, Farrar JJ, van Vinh Chau N, Wills B. Dengue. N Engl J Med. 2012; 366(15): 1423-32. doi: 10.1056/nejmra1110265.
[5] National Guidelines for Clinical Management of Dengue Fever 2023.pdf. Accessed March 4, 2026. https://ncvbdc.mohfw.gov.in/Doc/National%20Guidelines%20for%20Clinical%20Management%20of%20Dengue%20Fever%202023.pdf.
[6] Soundravally R, Sankar P, Bobby Z, Hoti SL. Oxidative stress in severe dengue viral infection: association of thrombocytopenia with lipid peroxidation. Platelets. 2008; 19(6): 447-54. doi: 10.1080/09537100802155284.
[7] Rajendiran S, Lakshamanappa HS, Zachariah B, Nambiar S. Desialylation of plasma proteins in severe dengue infection: Possible role of oxidative stress. Am J Trop Med Hyg. 2008; 79(3): 372-7. doi: 10.4269/ajtmh.2008.79.372.
[8] Gil L, Martínez G, Tápanes R, et al. Oxidative stress in adult dengue patients. Am J Trop Med Hyg. 2004; 71(5): 652-7. PMID: 15569800.
[9] Sies H, Stahl W. Vitamins E and C, beta-carotene, and other carotenoids as antioxidants. Am J Clin Nutr. 1995; 62(6 Suppl): 1315S-1321S. doi: 10.1093/ajcn/62.6.1315S.
[10] Padayatty SJ, Katz A, Wang Y, et al. Vitamin C as an antioxidant: Evaluation of its role in disease prevention. J Am Coll Nutr. 2003;22(1):18-35. doi: 10.1080/07315724.2003.10719272.
[11] Traber MG, Atkinson J. Vitamin E, antioxidant and nothing more. Free Radic Biol Med. 2007; 43(1): 4-15. doi:10.1016/j.freeradbiomed.2007.03.024.
[12] Traber MG. Current evidence of the role of vitamin E in prolonging a healthy life. Redox Exp Med. 2023; 2023(1). doi:10.1530/rem-23-0020.
[13] Chandra P, Sharma H, Gupta A, Rai Y. Role of antioxidant vitamin E and C on platelet levels in dengue fever. Int J Basic Appl Med Sci. 2013; 3(1): 287-91. ISSN: 2277-2103 (Online).
[14] Palanichamy Kala M, St. John AL, Rathore APS. Dengue: Update on clinically relevant therapeutic strategies and vaccines. Curr Treat Options Infect Dis. 2023; 15(2): 27-52. doi: 10.1007/s40506-023-00263-w.
[15] Langerman SD, Ververs M. Micronutrient supplementation and clinical outcomes in patients with dengue fever. Am J Trop Med Hyg. 2021; 104(1): 45-51. doi: 10.4269/ajtmh.20-0731.
[16] Office of Dietary Supplements - Vitamin C. Accessed March 4, 2026. https://ods.od.nih.gov/factsheets/VitaminC-HealthProfessional/.
[17] CDC. Guidelines for Classifying Dengue. Dengue. May 15, 2025. Accessed September 5, 2025. https://www.cdc.gov/dengue/hcp/clinical-signs/guidelines.html.
[18] Dengue. Accessed November 2, 2022. https://www.who.int/southeastasia/health-topics/dengue-and-severe-dengue.
[19] Ahmed S, Finkelstein JL, Stewart AM, et al. Micronutrients and dengue. Am J Trop Med Hyg. 2014; 91(5): 1049-56. doi: 10.4269/ajtmh.14-0142.
[20] Soundravally R, Sankar P, Hoti SL, Selvaraj N, Bobby Z, Sridhar MG. Oxidative stress induced changes in plasma protein can be a predictor of imminent severe dengue infection. Acta Trop. 2008; 106(3): 156-61. doi: 10.1016/j.actatropica.2008.03.001.
[21] Wills BA, Dung NM, Loan HT, et al. Comparison of three fluid solutions for resuscitation in dengue shock syndrome. N Engl J Med. 2005; 353(9): 877-89. doi: 10.1056/NEJMoa044057.
[22] Ramalingam K, Varghese C, Elias C, Mathew G, Balasubramanian A. A retrospective study on the effect of vitamin C in the management of dengue fever in three different states of India. Int J Res Pharm Sci. 2019; 10: 2670-3. doi: 10.26452/ijrps.v10i4.1525.
[23] Jahnavi T, Panchanathan S, Manoharan SD, Manokar CM. Vitamin C and vitamin E as treatment modalities in the management of dengue fever: A narrative review. J Med Sci. 2025; 11(1-4): 1-4. doi: 10.5005/jp-journals-10045-00310.
[24] Banerjee B. National Vector Borne Disease Control Programme-Dengue and chikungunya fever. DK Taneja's Health Policies Programme India. Published online 2017: 299-301. doi: 10.5005/jp/books/13071_18.
[25] World Health Organization. Dengue vaccine: WHO position paper, July 2016 – recommendations. Vaccine. 2017; 35(9): 1200-1. doi: 10.1016/j.vaccine.2016.10.070.
[26] Haider N, Hasan MN, Onyango J, Asaduzzaman M. Global landmark: 2023 marks the worst year for dengue cases with millions infected and thousands of deaths reported. IJID Reg. 2024; 13: 100459. doi: 10.1016/j.ijregi.2024.100459.
[27] Che P, Tang H, Li Q. The interaction between claudin-1 and dengue viral prM/M protein for its entry. Virology. 2013; 446(1-2): 1-2. doi: 10.1016/j.virol.2013.08.009.
[28] 4ff8af5cfb1162ad0a87195896c58320.pdf. Accessed June 12, 2025. https://www.afjbs.com/uploads/paper/4ff8af5cfb1162ad0a87195896c58320.pdf?utm_source=chatgpt.com.

Journal of Associated Medical Sciences 2026; 59(2): 382-392

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Radioprotective effect and antioxidant activities of Thai Gac fruit (*Momordica cochinchinensis* Spreng)

Anont Laorngnual[1], Sareeya Reungpatthanaphong[2], Paiboon Reungpatthanaphong[3*]

[1]Kanchanabhishek Institute of Medical and Public Health Technology, Faculty of Public Health and Allied Health Science, Praboromarajchanok Institute, Nonthaburi Province, Thailand.
[2]Cosmetic Science and Natural Products, College of Health and Wellness, Dhurakij Pundit University, Bangkok, Thailand.
[3]Department of Applied Radiation and Isotopes, Faculty of Science, Kasetsart University, Chatuchuck, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 26 August 2025
Accepted as revised 29 March 2026
Available online 11 April 2026

Keywords:
Momordica cochinchinensis Spreng, Thai Gac fruit, antioxidant, radioprotective, free radicals

ABSTRACT

Background: Radiation therapy or radiotherapy has been used to treat cancer which often causes significant adverse effects due to its cytotoxicity on both cancer and nearby intact cells. In addition, ionizing radiation generates free radicals which can damage cellular component leading to DNA damage and cell death. All cells and tissues have variable sensitivity and response to radiation injury. An antioxidant, a stable molecule, which can donate an electron to neutralize free radical. It delays or inhibits cellular damage mainly through its free radical scavenging property. Recently, there has been increasing interest in the prevention of disease mediated by free radicals. Gac fruit (*Momordica cochinchinensis* Spreng) is a major source of β-carotene and lycopene which can be found mainly in the red seed membrane. Both are well known antioxidant agents that can resist the negative effects of free radicals.

Objectives: The objective of this study was to determine the radioprotective effect, total phenolic content and antioxidant properties from various parts of Thai Gac fruit extracts.

Materials and methods: Each part of Thai Gac fruit including peel, pulp, aril, and seed were separately extracted using different solvents, 95% ethanol and ethyl acetate, for extraction. The total phenolic content and antioxidant activities were examined using DPPH radical scavenging assay and lipid peroxidation inhibition assay. Radioprotective effect of the extracts was evaluated by Dicentric Chromosome (DC) assay on TK6 lymphoblast cells after gamma irradiation.

Results: The ethanolic extraction of peel (EtP) showed the highest total phenolic content (140.43±0.63 mg GAE/gm), along with significant antioxidant activity which the values of SC_{50} were 3.72±0.12 mg/ml and 2.57±0.36 mg/ml for DPPH scavenging and lipid peroxidation inhibition, respectively. However, EtP and the ethyl acetate extract of pulp (EACN) exhibited higher LC_{50} values when compared to LC_{50} values of L-ascorbic acid (SC_{50}=0.25±0.04 mg/ml) and α-tocopherol (LC_{50}=0.85±0.14 mg/ml) as standard reagents. Cytotoxicity of Thai Gac fruit was studied on TK6 cells which revealed non-toxic at the concentrations up to 100 µg/ml and 25 µg/ml for EtP and EACN, respectively. Moreover, the DC assay showed that EtP exhibited radioprotective effect in a dose-dependent manner while EACN could reduce radiation-induced chromosomal abnormalities in TK6 lymphoblast cells.

* **Corresponding contributor.**
Author's Address: *Department of Applied Radiation and Isotopes, Faculty of Science, Kasetsart University, Chatuchuck, Bangkok, Thailand.*
E-mail address: fscipbr@ku.ac.th
doi: 10.66285/JAMS.2026.069
E-ISSN: 2539-6056

Conclusion: The ethanolic extract of peel (EtP) exhibited the highest level of antioxidant activities and radioprotective effect when compared to those of other solvent extractions, suggesting it might be used as a potential natural radioprotective agent. It was found that EtP could reduce the damage caused by radiation. Our findings suggested that the peel of Thai Gac fruit extracted by 95% ethanol exhibited a potent antioxidant activity and radioprotective effects in cell line.

Introduction

Radiotherapy is a fundamental approach in cancer treatment, used in around 60-80% of cases, due to its effectiveness in eradicating cancer cells.[1,2] Cancer is also known as salient "free radical" disease. The initiation and promotion of cancer have been associated with chromosomal defects and oncogene activation. Free radicals from those of endogenous reactions and initiated by ionizing radiation, may lead to tumor formation. However, ionizing radiation (IR) affects not only cancerous cells but also surrounding healthy tissues, resulting in various side effects. These adverse effects arise from both direct and indirect IR damage. Direct effects involve DNA strand breaks, while indirect effects are caused by reactive oxygen species (ROS) like hydroxyl radicals, which are produced from radiolysis of water within cells.[1] These ROS attack cellular structures, leading to inflammation, genomic instability, and long-term health risks such as fibrosis and secondary cancers.[2,3] The toxicity to normal tissues often limits the IR dose, highlighting the need for protective agents that can mitigate these negative impacts and optimize therapeutic outcomes.[4]

Amifostine, a widely studied synthetic radioprotector, has proven effective in minimizing radiation damage to healthy tissues. However, many other synthetic radioprotectors have been limited in use due to toxicity and short duration of effect.[4] This has led to a growing interest in natural products as safer, more sustainable alternatives. Natural compounds from medicinal plants have been reported to exhibit radioprotective effects through various mechanisms, including DNA repair enhancement, anti-inflammatory activity, upregulation of intracellular radioprotective enzymes, modulation of signaling and apoptotic pathways, immunomodulation, and free radical scavenging.[1,2,5,6] These agents, particularly antioxidants, neutralize reactive oxygen species (ROS) generated by radiation, reducing oxidative stress and protecting critical cellular components like DNA and lipids from damage.[7] All antioxidants act as radical scavenger, hydrogen donor, electron donor, peroxide decomposer, singlet oxygen quencher, enzyme inhibitor, synergist, and metal-chelating agents. Examples of natural compounds with radioprotective effects include polyphenols like quercetin and curcumin, as well as resveratrol and genistein, which have shown antioxidant and anti-inflammatory benefits in studies.[2,3] Additionally, the flavonoid apigenin, administered post-irradiation in an *in vivo* study, showed strong radioprotective effects by reducing oxidative DNA damage in bone marrow cells and suppressing inflammatory responses in bone marrow-derived macrophages.[8] These findings collectively highlight the potential of natural products as viable radioprotective agents.

Gac fruit (*Momordica cochinchinensis* Spreng.), native to Southeast Asia-particularly Thailand and Vietnam-is a dioecious vine traditionally valued for its medicinal properties and nutrient profile.[9,10] The aril and peel of Gac are rich in carotenoids, including β-carotene and lycopene, which are known for their antioxidant, anti-inflammatory, and anticancer properties.[9,11,12] Lycopene, in particular, has shown radioprotective effects by scavenging free radicals, enhancing cellular antioxidant defenses, and reducing radiation-induced DNA and lipid damage.[13] Additionally, several studies have identified phenolic compounds and flavonoids in Gac, further enhancing its antioxidant potential.[9] Considering the bioactive compounds and biological activity, Gac shows strong potential as a natural radioprotective agent. However, research on its specific radioprotective effects remains limited.

Therefore, this study aimed to address this gap by analyzing the total phenolic content and antioxidant activities of various Thai Gac extracts and assessing their radioprotective effects.

Materials and methods

Preparation and extraction

Thai Gac fruits (*M. cochinchinensis* Spreng.) were collected from Sai Noi district, Nonthaburi province, Thailand. The fruits were thoroughly washed and divided into 4 parts: peel, pulp, aril, and seeds (Figure 1). Each part was dried in a hot air oven at 60 °C and ground into a fine powder.[12] For extraction, 200 grams of each powdered sample was separately macerated in 1000 ml of ethyl acetate or 95% ethanol, in a ratio 1:5, at room temperature with occasional shaking. After 5 days, the extracts were filtered through cotton wool and then through filter paper using a Buchner funnel. The filtrates were pooled, and the solvent was removed using a rotary evaporator. The crude extracts were stored at 4 °C for future use, and the extraction yield was calculated based on the dry weight of each sample.

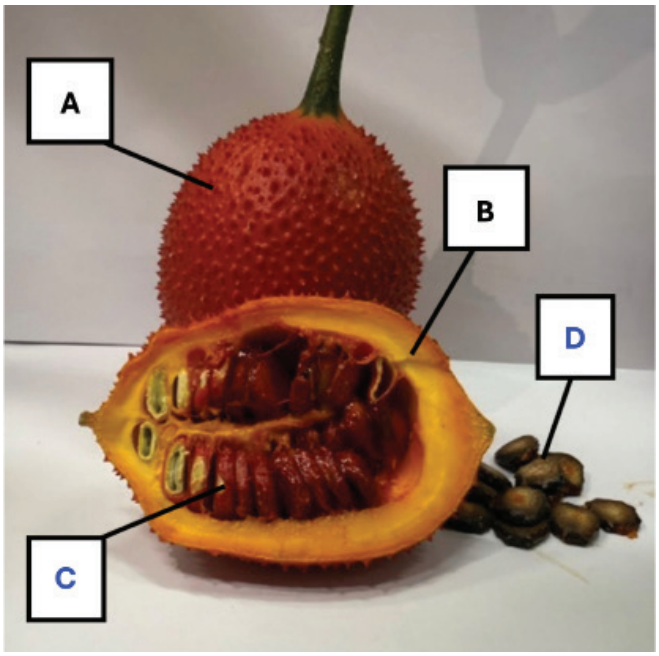


Figure 1. *Parts of Thai Gac fruit (Momordica cochinchinensis Spreng). A: peel, B: pulp, C: aril, D: seed*

Total phenolic content using Folin-Ciocalteu assay

The total phenolic content (TPC) was evaluated using the Folin-Ciocalteu assay.[14] A 50 µL aliquot of the extract (1 mg/mL) and 50 µL of Folin-Ciocalteu reagent (Sigma-Aldrich, USA) at 0.5 mg/mL in ethanol were added to the wells of a 96-well microplate. Then, 75 µL of 7.5% sodium carbonate solution (Loba Chemie, India) was added to each well. Distilled water was used as the control, and different concentrations of gallic acid (Sigma-Aldrich, USA) served as the standard. The reaction was allowed to proceed in the dark at 25 °C for 90 minutes. Absorbance was measured at 725 nm using a microplate reader. The total phenolic content was calculated based on a standard curve of gallic acid and reported as mg of gallic acid equivalents (GAE)/gm of extract. Each extract was tested in triplicate.

Antioxidant activity

DPPH free radical scavenging activity

The antioxidant activity of the Gac extracts was evaluated using the DPPH (2,2-diphenyl-1-picrylhydrazyl) free radical scavenging assay, following the method described by.[14] Each extract was prepared at varying concentrations and added (50 µL) into the wells of a 96-well microplate. Then, 50 µL of DPPH solution (Sigma-Aldrich, USA) at 0.5 mg/mL was added to each well. The reaction was allowed to proceed in the dark at 25 °C for 30 minutes. L-ascorbic acid (Sigma-Aldrich, USA) was used as the standard. Absorbance was measured at 515 nm using a microplate reader. The percentage of DPPH radical scavenging activity was calculated using the following formula (1):

$$\%Scavenging = [(A_0-A_1)/A_0]\times 100 \text{ — (1)}$$

Where:

A_0 represented the absorbance of the control,

A_1 represented the absorbance of the samples.

The percentage of DPPH radical scavenging activity was plotted against the extract concentrations, and the concentration required to achieve 50% antioxidant activity (SC_{50}) was calculated from the graph and reported. Experiments for each extract were performed in triplicate.

Lipid peroxidation inhibition

Lipid peroxidation inhibition was assessed using the Ferric-thiocyanate method as described elsewhere.[14] Aqueous solutions of each extract were prepared at varying concentrations (0.001-10 mg/mL). A 50 µL aliquot of each sample solution was mixed with 50 µL of linoleic acid dissolved in 50% DMSO (Sigma-Aldrich, USA) in a microplate. To initiate the reaction, 50 µL of 5 mM NH_4SCN and 50 µL of 2 mM $FeCl_2$ were added. The reaction mixture was incubated at 37 °C for 60 minutes in the dark. The α-tocopherol (Sigma-Aldrich, USA) was used as the standard. Absorbance was measured at 490 nm using a microplate reader, and the percentage of lipid peroxidation inhibition was calculated using the following formula.

$$\%lipid\ peroxidation\ inhibition = [(A_0-A_1)/A_0]\times 100 \text{ — (2)}$$

In this calculation, A_0 represented the absorbance of the control, and A_1 represented the absorbance of the sample. The graph showing the relationship between the percentage of lipid peroxidation inhibition and the extract concentration was plotted. The concentration required to achieve 50% inhibition of lipid peroxidation (LC_{50}) was then reported. Triplicate testing was conducted for each extract.

Cytotoxicity assay

Cell viability was evaluated to determine the effect of the extracts on cell survival using the MTT assay, measuring the number of live cells. This assay is based on the principle that living cells produce enzymes capable of reducing 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) into formazan which absorbs light at a wavelength of 570 nm, allowing quantification of live cells. In this experiment, 10,000 TK6 cells/well (50 µL) were cultured in 96-well plates with different concentrations of the extracts. The plates were incubated in a cell culture chamber at 37 °C with 5% CO_2 for 24 hours. Following incubation, MTT solution was added, and the microplates were incubated for an additional 3 hours. After that, 200 µL of DMSO was added, and absorbance was measured at 570 nm. The percentage of cell viability was then calculated in comparison to the control group. The IC_{20} values (the concentration that inhibits 20% of cell growth) were determined using nonlinear regression analysis (log(inhibitor) vs response). These concentrations were further confirmed to be non-toxic (p>0.05) using One-way ANOVA followed by Dunnett's post-hoc test compared to the untreated control group. Triplicate testing was conducted for each extract.

Radioprotective activity using the Dicentric assay

Radioprotective effect of the selected extracts was assessed based on the principle that radiation can cause DNA strand breaks, particularly double-strand breaks, leading to errors in the cell's repair process. As a result, dicentric chromosomes, having two centromeres indicating chromosomal abnormalities are formed. In this study, the Dicentric Assay was used to evaluate the radioprotective effects of the extracts. TK6 lymphocyte cells were treated with varying concentrations of the test extracts, as indicated by the cytotoxicity assay, and then exposed to gamma radiation (^{137}Cs) at a dose of 1 Gy with a dose rate of 3.13 Gy/min at the Nuclear Technology Research Center (NTRC). Amifostine served as the standard. A total of six experimental tubes were used for this assay as follows.[13,15]

Tube 1: TK6 lymphocyte cells control (no extract and radiation)

Tube 2: TK6 lymphocyte cells+1 Gy; (no extract)

Tube 3-5: TK6 lymphocyte cells+extract at 3 different concentration + 1 Gy

Tube 6: TK6 lymphocyte cells+Amifostine+1 Gy

Following irradiation, the cells were supplemented with RPMI-1640 culture medium with 10% Fetal bovine serum, then incubated at 37 °C with 5% carbon dioxide for 24 hours. After incubation, 100 µL of Colcemid solution was added to arrest cell division at metaphase by inhibiting spindle fiber formation, followed by an additional 24-hour incubation. Cells were then fixed and harvested by centrifugation, with the supernatant removed and replaced by 0.075 M potassium chloride (KCl) solution. The fixative of methanol and acetic acid (3:1) was gradually added with gentle mixing to prevent clumping. After three cycles of centrifugation and resuspension in fresh fixative, slides were prepared by placing the cell suspension onto slides and staining with 2% Giemsa in phosphate buffer (pH 6.8). After washing and drying, coverslips were mounted with a mounting medium. Dicentric chromosome counts were analyzed under a microscope, comparing irradiated and extract-treated samples with non-irradiated, extract-free tube as controls and Amifostine as the positive standard. Triplicate testing was conducted for each extract.

Statistical analysis

The results of total phenolic content and antioxidant activity were presented as the mean ± standard deviation (SD). For the cytotoxicity study, the IC_{20} values were determined using nonlinear regression analysis (log (inhibitor) vs. response). Statistical significance between the control and extract-treated groups was evaluated using one-way analysis of variance (ANOVA), followed by Dunnett's post-hoc test. A p<0.05 was considered statistically significant, while concentrations with p>0.05 compared to the control were selected as non-toxic doses for the radioprotective assay.

Results

Plant extraction

Extracts from different parts of the Gac fruit were obtained through maceration using ethyl acetate or 95% ethanol as solvents, at a material: solvent ratio of 1:5. After filtration, solvent removal, crude extracts were produced, with yields and extract characteristics summarized in Table 1.

Table 1. The percentage yield (%Yield) from different parts of Gac extracts.

Parts of Gac fruit	Solvent	Dried plant (gm)	Crude extract (gm)	Yield (%)
Peel	Ethyl acetate	200.21	0.86	0.43
	95% Ethanol	200.28	2.11	1.05
Pulp	Ethyl acetate	144.10	0.93	0.64
	95% Ethanol	154.63	5.14	3.32
Aril	Ethyl acetate	204.00	38.52	18.88
	95% Ethanol	204.87	26.78	13.07
Seed	Ethyl acetate	200.09	8.09	4.04
	95% Ethanol	201.12	11.09	5.51

The percentage yield (%Yield) from different parts of Gac fruit extracts was obtained, using ethyl acetate or 95% ethanol as solvents. The results indicated that the aril with ethyl acetate extract showed the highest %yield (18.88%), while the ethanolic extract showed 13.07%. Moreover, the ethyl acetate extract of the peel and pulp showed the lowest yield, at 0.43 and 0.64, respectively.

Total phenolic content using Folin-Ciocalteu assay

The total phenolic content (TPC) was analyzed using the Folin-Ciocalteu reagent method, based on the redox reaction of molybdotungstate ions. The total phenolic content was determined using a standard curve of gallic acid and expressed in milligrams of gallic acid equivalents per gram of extract (mg GAE/gm). The results showed in Table 2, the highest TPC was found in the ethanolic extract of peel (140.43±0.63 mg GAE/gm), followed by the ethyl acetate extract of peel (136.03±0.91 mg GAE/gm). The ethyl acetate extract of seed showed a substantial level (110.23±0.89 mg GAE/gm), while the lowest TPC was observed in the ethanolic extract of seed (6.18±0.24 mg GAE/gm).

Table 2. Total Phenolic Content (TPC) and antioxidant activity from different parts of Gac fruit's crude extracts

Parts of Gac fruit	Solvent	Total phenolic content	Antioxidant activity (mg/mL)	
		(mg GEA/gm)	DPPH assay (SC_{50})	Lipid peroxidation inhibition (LC_{50})
Peel	Ethyl acetate	136.03±0.91	10.15±1.55*	3.09±0.62*
	95% Ethanol	140.43±0.63	3.72±0.12*	2.57±0.36*
Pulp	Ethyl acetate	82.34±0.75*	3.79±0.14*	3.17±0.78*
	95% Ethanol	91.46±0.56*	6.10±0.10*	4.22±0.99*
Aril	Ethyl acetate	58.33±0.68*	10.21±0.85*	5.43±0.45*
	95% Ethanol	20.99±0.83*	5.76±0.15*	3.74±0.15*
Seed	Ethyl acetate	110.23±0.89*	36.70±1.36*	5.41±0.42*
	95% Ethanol	6.18±0.24*	13.88±0.94*	10.60±1.7*
L-ascorbic acid	-	ND	0.25±0.04*	ND
α-tocopherol	-	ND	ND	0.85±0.14*

Note: **significant differences at p<0.05, ND: not determined.*

Antioxidant activity

The study of antioxidant activity using the DPPH method, the ethanolic extract of peel's Gac fruit exhibited the highest free radical scavenging activity (SC_{50}=3.72±0.12 mg/mL), followed by the ethyl acetate extract of pulp (SC_{50}=3.79±0.14 mg/mL, respectively. As for the extraction of seed in ethyl acetate and 95% ethanol showed the lowest activity (SC_{50}=36.70±1.36 mg/mL and 13.88±0.94 mg/mL, respectively) when compared to other parts. All extracts were significantly less effective than the L-ascorbic acid standard (SC_{50}=0.25±0.04 mg/mL) at p<0.05. Similarly, in the lipid peroxidation inhibition assay, the ethanolic extract of peel's Gac fruit exhibited the highest activity (LC_{50}=2.57±0.36 mg/mL), followed by ethyl acetate extracts of peel and pulp (LC_{50}=3.09±0.62 mg/mL and 3.17±0.78 mg/mL, respectively). In addition, the ethanolic extract of seed showed the lowest lipid peroxidation inhibition (LC_{50}=10.60±1.71 mg/ml) when compared to other extracts. The activity of all extracts was significantly lower than α-tocopherol standard (LC_{50}=0.85±0.14 mg/mL) at p<0.05). The results of total Phenolic Content (TPC) and antioxidant activity from different parts of Gac fruit's crude extracts were shown in Table 2.

Cytotoxicity assay

Based on the results of total phenolic content and antioxidant activity using DPPH assay and Lipid peroxidation inhibition, the ethanolic extract of peel (EtP) and the ethyl acetate extract of pulp (EACN) from Gac fruit were identified as promising candidates for further study on radioprotective effects. Consequently, these two extracts were selected for additional evaluation of their protective effects against radiation-induced cellular damage. The investigation would begin with cytotoxicity study on TK6 cells, a lymphoblast cell line, at various extract concentrations to determine the optimal non-toxic concentration range for subsequent radioprotective studies.

The cytotoxicity assay showed that after a 24-hour incubation of TK6 cells with EtP extract at concentrations ranging from 6.25 to 200 µg/mL, control cells exhibited 100% viability (Figure 2). Cells treated with 0.1%DMSO as a solvent and EtP at concentrations between 6.25 and 100 µg/mL-maintained above 80% cell viability with values ranging from 80.32% to 86.61%. However, at 200 µg/mL, EtP reduced cell viability to 63.37%. Therefore, the non-toxic concentration range (IC_{20}) for EtP was identified as 6.25 to 100 µg/mL, which will be selected for radioprotective assay. Similarly, TK6 cells were treated with EACN extract (6.25 to 200 µg/mL), where control cells also exhibited 100% cell viability. Cells treated with DMSO and EACN concentrations between 6.25 and 25 µg/mL maintained over 80% cell viability, ranging from 81.70% to 90.11%. In contrast, EACN concentrations ranging from 50 to 200 µg/mL resulted in viabilities below 80%, with values ranging from 43.26% to 70.31%. Consequently, the non-toxic range (IC_{20}) for EACN was established at 6.25 to 25 µg/mL, which was selected for further radioprotective studies.

Radioprotective activity using the Dicentric assay

In the evaluation of radioprotective activity using the Dicentric assay, TK6 cells were incubated with extracts or Amifostine at non-toxic concentrations for 24 hours before exposing to 1 Gy of gamma radiation (^{137}Cs). Chromosomal abnormalities were assessed at the metaphase stage, with results recorded as the number of cells exhibiting dicentric chromosomes out of 50 analyzed metaphase cells. Untreated but irradiated cells served as the positive control, while untreated and unirradiated cells served as the negative control.

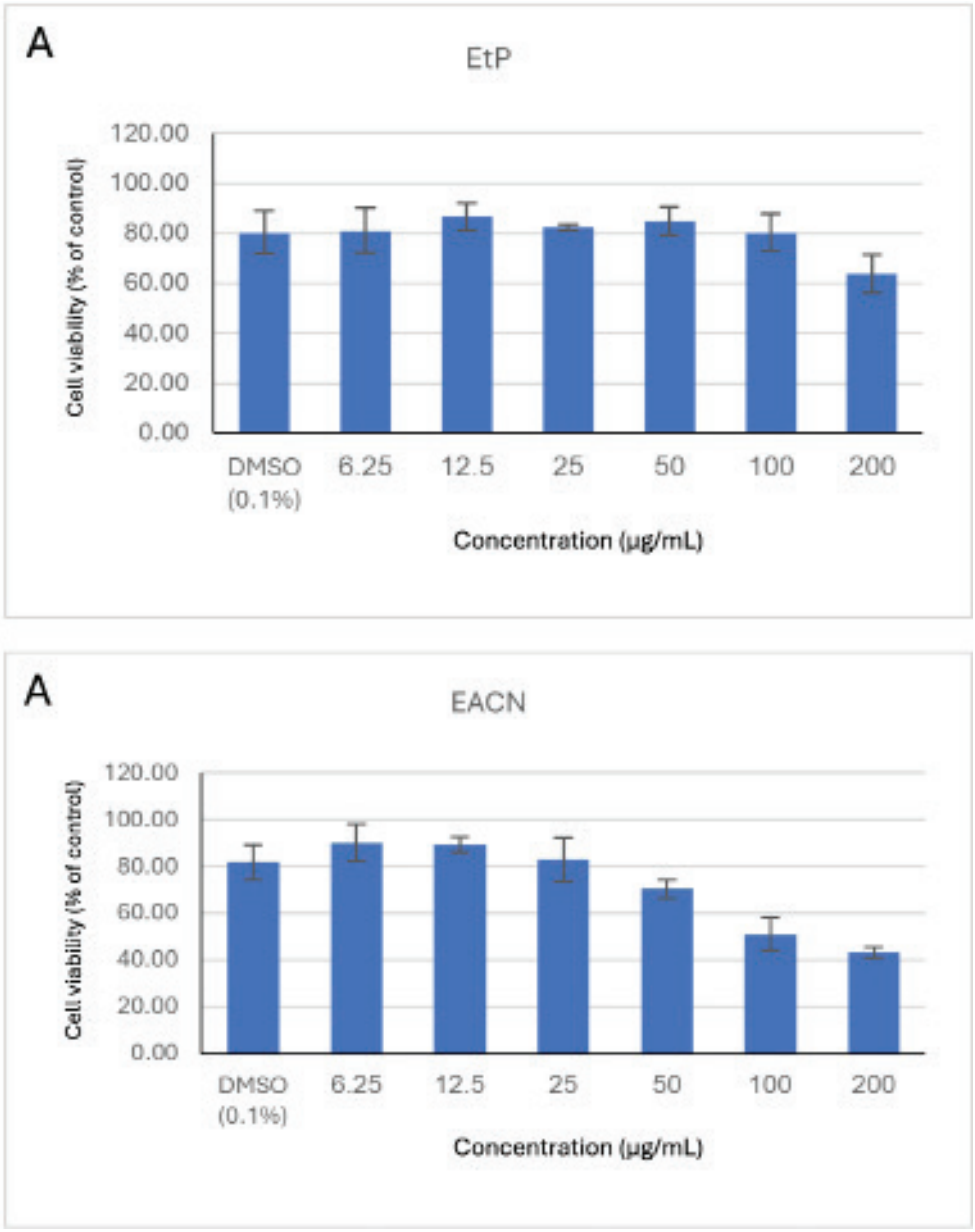


Figure 2. *Percentage viability of TK6 cells incubated with EtP (A) and EACN (A).*

Table 3 presents the percentage of TK6 cells displaying dicentric chromosomes (DC) across various treatments and concentrations. The ethanolic extract of the peel (EtP) demonstrated a promising radioprotective effect, with a dose-dependent decrease in DCs. The percentage of DCs was reduced to 31.3%, 50.7%, and 76.0% at EtP concentrations of 100, 50, and 25 µg/mL, respectively. The ethyl acetate extract of pulp (EACN) showed limited radioprotective effects, with high DC levels of 79.3%, 86.0%, and 96.0% at concentrations of 25, 12.5, and 6.25 µg/mL, respectively. As the standard, Amifostine demonstrated the strongest radioprotective effect, with DC levels ranging from 15.3% to 64.7% across different concentrations. The positive control (radiation treatment only) showed 96.7% DCs, while the negative control (no radiation treatment) showed no DCs, validating the experimental setup.

Discussion

Gac fruit (*Momordica cochinchinensis* Spreng.) is known for its abundance of bioactive compounds, including β-carotene, lycopene, and various phenolic compounds, which contribute to its antioxidant and health-promoting properties. Studies have shown that these bioactive compounds in Gac fruit possess antioxidant, anti-inflammatory, and anticancer activities.[9,11,12] Considering their protective effects against oxidative stress, these compounds have significant potential for radioprotective applications. To date, no studies have investigated the radioprotective activity of Gac fruit, making this research the first to explore this potential aspect.

In this study, the total phenolic content and antioxidant activity of Gac fruit demonstrated significant potential, particularly in the peel extract

with 95% ethanol, which exhibited the highest phenolic content. This aligns with findings from Abdulqader *et al*. which showed that aril and peel extracts of Gac possess high levels of carotenoids, phenolics, and flavonoids.[16] The total phenolic content in Gac fruit varies across different parts at different age, as reported by multiple studies.[16] Notably, phenolic compounds such as rutin, cinnamic acid, and 2-(3,4-dihydroxyphenyl)-7-hydroxy-5-benzene propanoic acid are particularly prevalent in the peel, while hydroxybenzoic acids (e.g., gallic acid, protocatechuic acid, p-hydroxybenzoic acid) and hydroxycinnamic acids (e.g., chlorogenic acid, caffeic acid, p-coumaric acid, ferulic acid, sinapic acid) are found throughout all parts of the fruit.[16,17] These phenolic compounds are potent antioxidants that act in a structure-dependent manner, with radical-scavenging and metal-chelating abilities being key mechanisms in oxidative stress protection. Multiple studies have demonstrated a strong correlation between total phenolic content and the antioxidant activity of plant extracts, supporting Gac's potential for health-promoting and therapeutic applications.[9,10,16,17]

Table 3. Percentage of cells with dicentric chromosomes (DC) following 1 Gy gamma radiation exposure (^{137Cs})

Samples	**Concentration** (µg/mL)	**Cells in metaphase**	**Dicentric cells (DC)** Mean±SD	**% Dicentric cells**
Ethanol extract of peel (EtP)	100	50	15.7±3.9	31.3
	50	50	25.3±1.7	50.7
	25	50	38.0±3.7	76.0
Ethyl acetate extract of pulp (EACN)	25	50	39.7±4.0	79.3
	12.5	50	43.0±3.7	86.0
	6.25	50	48.0±1.4	96.0
Amifostine	100	50	7.7±1.2	15.3
	50	50	11.7±2.6	23.3
	25	50	19.3±3.1	38.7
	12.5	50	24.0±4.3	48.0
	6.25	50	32.3±4.1	64.7
Positive control (radiation treatment only)		50	48.3±0.5	96.7
Negative control (no radiation treatment)		50	0	0

The antioxidant activity was confirmed by DPPH radical scavenging and lipid peroxidation inhibition assays, where ethanolic extract of peel (EtP) exhibited the strongest DPPH radical scavenging activity, followed by the ethyl acetate extract of pulp (EACN), similar to findings by Reungpatthanaphong *et al,*.[10,12] which showed that peel extracts from Thai Gac fruit displayed the highest overall antioxidant activities, particularly from FRAP assay. The antioxidant activity of EtP was positively correlated with its total phenolic content which was the highest among the extracts tested. This result is consistent with previous studies that report a strong correlation between the high phenolic content of Gac fruit and its potent antioxidant activity.[18] Additionally, the lipid peroxidation inhibition assay further supported these findings, with both the ethanolic (EtP) and ethyl acetate (EACP) extracts of peel, ranking as the most effective in inhibiting lipid peroxidation, a key mechanism for preventing oxidative degradation of cellular membranes and protecting against oxidative damage.[13] These extracts were found to contain the highest level of phenolic compounds which likely contribute to their protective effects against oxidative DNA damage.

Regarding the extraction yield, our results showed that while the aril exhibited the highest percentage yield (18.88%), it did not correlate with the highest antioxidant activity. This discrepancy suggests that a high yield in the aril is primarily composed of bulk lipids and carotenoids, whereas the ethanolic extract of the peel (EtP), despite its lower yield (1.05%), contains a more concentrated profile of polar phenolic compounds. These specific phenolics are more effective in DPPH radical scavenging and protecting against radiation-induced damage, indicating that the quality and concentration of bioactive phytochemicals are more critical than the total extraction yield.[20]

Previous studies have reported that phenolic compounds such as rutin and quercetin acted as a metal ion chelator inhibiting iron ion-dependent lipid peroxidation.[6,19,20] Interestingly, the ethyl acetate extract of pulp (EACN) also demonstrated notable lipid peroxidation inhibition despite its lower phenolic content. This may be attributed to the presence of

high levels of other bioactive compounds, particularly β-carotene and lycopene, which are known to enhance lipid protection.[16,17] While EtP demonstrated superior activity due to its high phenolic concentration, the notable effectiveness of EACN suggests a synergistic effect between its phenolic constituents and rich carotenoid content, such as β-carotene and lycopene. This explains why the aril, despite having the highest yield and carotenoid levels, did not exhibit the strongest antioxidant activity, as the specific combination and concentration of bioactive compounds in the peel and pulp may provide more potent protective mechanisms against oxidative stress. These consistent findings across studies emphasize the potential of Gac fruit, particularly its peel and pulp, as rich sources of antioxidants, which may be valuable for radioprotective applications, demonstrating the capacity to mitigate radiation-induced cellular damage via oxidative stress reduction.

Based on the total phenolic content analysis and antioxidant assays, the ethanolic extract of peel (EtP) and the ethyl acetate extract of pulp (EACN) were selected as candidates for evaluating radioprotective activity using the dicentric chromosome (DC) assay, a widely recognized method for assessing chromosomal damage and repair following radiation exposure.[16] The DC assay results indicated that EtP exhibited a dose-dependent reduction in dicentric chromosomes in a lymphoblast cell line (TK6 cells), showing greater effectiveness than EACN in reducing radiation-induced cellular damage. The safety profile of the extracts was confirmed using the TK6 human lymphoblast cell line, which is a standard model for genetic toxicology and radiation research due to its stable karyotype and p53 competency.[15] The IC_{20} values identified in this study (up to 100 µg/mL for EtP) demonstrate that the radioprotective effects occur at non-toxic concentrations. While this study focused on TK6 cells to maintain consistency with the Dicentric Chromosome Assay, further studies on a broader range of normal human cell lines would be beneficial to fully establish the clinical safety of Thai Gac fruit extracts. This finding aligned with Arora *et al.* who demonstrated that natural compounds with high phenolic content enhance DNA repair mechanisms and act as antioxidants, counteracting radiation-induced reactive oxygen species (ROS) damage.[6]

Beyond direct free radical scavenging, the radioprotective mechanisms of EtP may involve the upregulation of endogenous antioxidant enzymes, such as superoxide dismutase (SOD) and glutathione peroxidase (GPx), which reinforce the cellular defense system against oxidative stress. Furthermore, bioactive compounds in Gac peel, particularly phenolic acids and flavonoids, have been reported to enhance DNA repair pathways and modulate inflammatory signaling, such as inhibiting pro-inflammatory cytokines that contribute to secondary radiation injury.[13,21] Supporting studies by Srinivasan *et al.* and Cavusoglu and Yalcin demonstrated lycopene's ability to reduce radiation-induced chromosomal damage, such as dicentrics and micronuclei, in human lymphocyte by reducing lipid peroxidation and enhancing antioxidant enzyme activities like superoxide dismutase (SOD), catalase (CAT), and glutathione peroxidase (GPx).[13,21]

Limitations

This study is primarily limited by its focus on the TK6 human lymphoblast cell line, requiring further research on a broader range of normal human cell lines to fully establish clinical safety.

Conclusion

This study demonstrated the potent antioxidant and radioprotective activities of Thai Gac fruit (*Momordica cochinchinensis* Spreng.), particularly from the ethanolic extract of peel (EtP). Among the different parts tested, EtP exhibited the highest total phenolic content and showed superior performance in both DPPH radical scavenging and lipid peroxidation inhibition assays, which correlated well with its radioprotective efficacy. The results of Dicentric Chromosome (DC) assay revealed that EtP effectively reduced radiation-induced chromosomal damage, suggesting its potential role in radiation protection. Notably, the peel, often considered a byproduct, was identified as the most promising source of bioactive compounds, thus offering added economic and functional value.

Our findings suggest that Gac fruit peel extract had the potential to be developed as a natural radioprotective agent, providing an alternative to synthetic agents like amifostine, with the advantages of reduced side effects and broader clinical applicability. Comparison with the standard radioprotector, Amifostine, revealed that while Amifostine remains the more potent agent, EtP at its highest non-toxic concentration (100 µg/mL) showed a comparable trend in reducing dicentric chromosome formation. The clinical significance of EtP lies in its potential to overcome the limitations of Amifostine, which is often associated with systemic toxicity and adverse side effects such as hypotension. As a natural product derived from agricultural byproducts, EtP offers a safer, cost-effective, and more sustainable approach for radiation protection, particularly in scenarios involving repeated low-dose exposure or as a supportive dietary strategy during radiotherapy. Future studies focusing on the safety and mechanism of action of peel from Gac fruit as bioactive compounds are recommended to further explore its potential in clinical and therapeutic application.

Ethical approval

No study in human or animals was performed by any authors.

Funding

The study was performed without financial support.

Conflict of interest

The authors declare that there is no conflict of interest regarding in this research, including financial, personal, authorship, or other factors which could affect this research, and all results presented in this article.

CRediT authorship contribution statement

Anont Laorngnual and **Sareeya Reungpatthanaphong**: experimental design, experimental work, results analysis, investigation, writing: draft manuscript; **Paiboon Reungpatthanaphong**: experimental design, experimental work, results analysis, investigation, resources, writing: draft and final preparation.

Acknowledgements

The authors would like to express their sincere gratitude to the Department of Applied Radiation and Isotopes, Faculty of Sciences, Kasetsart University, for providing essential facilities and resources that significantly contributed to this research. We also extend our thanks to the Kanchanabhishek Institute of Medical and Public Health Technology, Faculty of Public Health and Allied Health Science, Praboromarajchanok Institute, for the invaluable support and collaboration. Special appreciation is given to Associate Professor Tewin Tencomnao from the Department of Clinical Chemistry, Faculty of Allied Health Sciences, Chulalongkorn University, for his expert guidance and insightful contributions throughout the study.

References

[1] Shivappa P, Bernhardt GV. Natural radioprotectors on current and future perspectives: A mini-review. J Pharm Bioallied Sci. 2022; 14(2): 57-71. doi: 10.4103/jpbs.jpbs_502_21.

[2] Mun G-I, Kim S, Choi E, Kim CS, Lee Y-S. Pharmacology of natural radioprotectors. Arch Pharm Res. 2018; 41(11): 1033-50. doi: 10.1007/s12272-018-1083-6.

[3] Yi J, Zhu J, Zhao C, Kang Q, Zhang X, Suo K, et al. Potential of natural products as radioprotectors and radiosensitizers: opportunities and challenges. Food & Function. 2021; 12(12): 5204-18. doi: 10.1039/d1fo00525a.

[4] Maurya DK, Devasagayam TP, Nair CKK. Some novel approaches for radioprotection and the beneficial effect of natural products. Indian J Exp Biol. 2006;44(2):93.

[5] Kumar S, Singh RK, Meena R. Emerging targets for radioprotection and radiosensitization in radiotherapy. Tumor Biology. 2016; 37(9): 11589-609. doi: 10.1007/s13277-016-5117-8.

[6] Arora R, Gupta D, Chawla R, Sagar R, Sharma A, Kumar R, et al. Radioprotection by plant products: present status and future prospects. Phytother Res. 2005;19(1):1-22. doi: 10.1002/ptr.1605.

[7] Shirazi A, Mihandoost E, Mahdavi SR, Mohseni M. Radio-protective role of antioxidant agents. Oncol Rev. 2012; 6(2): e16. doi: 10.4081/oncol.2012.e16.

[8] Rithidech KN, Tungjai M, Reungpatthanaphong P, Honikel L, Simon SR. Attenuation of oxidative damage and inflammatory responses by apigenin given to mice after irradiation. Mutat Res Genet Toxicol Envirol Mutagen. 2012; 749(1): 29-38. doi: 10.1016/j.mrgentox.2012.08.001.

[9] Chuyen HV, Nguyen MH, Roach PD, Golding JB, Parks SE. Gac fruit (Momordica cochinchinensis Spreng.): a rich source of bioactive compounds and its potential health benefits. Int J Food Sci Technol. 2015; 50(3): 567-77. doi: 10.1111/ijfs.12721.

[10] Reungpatthanaphong S, Kirdin T, Chinnawong M, Bamrungchai M, Reungpatthanaphong P. Antioxidant activity of accelerated solvent extraction from different fractions of Thai gac fruit (Momordica cochinchinensis Spreng). TJPS. 2018; 42(2018): 37-40. Available from: http://www.tjps.pharm.chula.ac.th/proceedings/backend/proceeding_file/9_PN_FP_Sareeya%20Reungpatthanaphong%20(37-40).pdf

[11] Abdulqader A, Ali F, Ismail A, Esa NM. Gac (Momordica cochinchinensis Spreng.) fruit and its potentiality and superiority in-health benefits. Journal of Contemporary Medical Sciences. 2018; 4(4). doi: 0.22317/jcms.v4i4.476

[12] Reungpatthanaphong S, Chawananorasest K, Kirdin T, Bamrungchai M, Reungpatthanaphong P. Development of Thai Gac Fruit Extraction as a Multifunctional Cosmeceutical Ingredient for Antioxidant, Melanogenesis and Collagen Stimulating Activities. Key Engineering Materials. 2019; 819: 104-10. doi: 10.4028/www.scientific.net/KEM.819.104.

[13] Srinivasan M, Devipriya N, Kalpana KB, Menon VP. Lycopene: An antioxidant and radioprotector against γ-radiation-induced cellular damages in cultured human lymphocytes. Toxicology. 2009; 262(1): 43-9. doi: 10.1016/j.tox.2009.05.004.

[14] Boonpisuttinant K, Taka T, Ruksiriwanich W, Chutoprapat R, Udompong S, Kansawang R, et al. Assessment of in vitro anti-skin aging activities of Phyllanthus indofischeri Bennet extracts for dermatological and aesthetic applications. Sci Rep. 2023; 13(1): 18661. doi: 10.1038/s41598-023-45434-3.

[15] Sudprasert W. Radiation dosimetry using biological methods. Bangkok: Extension and Training Office, Kasetsart University; 2018.

[16] Abdulqader A, Ali F, Ismail A, Esa NM. Antioxidant compounds and capacities of Gac (Momordica cochinchinensis Spreng) fruits. Asian Pac J Trop Biomed. 2019; 9(4): 158-67. doi: 10.4103/2221-1691.256729.

[17] Kubola J, Siriamornpun S. Phytochemicals and antioxidant activity of different fruit fractions (peel, pulp, aril and seed) of Thai gac (Momordica cochinchinensis Spreng). Food Chem. 2011;

127(3): 1138-45. doi: 10.1016/j.foodchem.2011.01.115.
[18] Tinrat S, Akkarachaneeyakorn S, Singhapol C. Evaluation of antioxidant and antimicrobial activities of Momordica Cochinchinensis Spreng (Gac fruit) ethanolic extract. Int J Pharm Sci Res. 2014; 5(8): 3163. doi: 10.13040/IJPSR.0975-8232. 5(8).3163-69.
[19] Afanas'ev IB, Dcrozhko AI, Brodskii AV, Kostyuk VA, Potapovitch AI. Chelating and free radical scavenging mechanisms of inhibitory action of rutin and quercetin in lipid peroxidation. Biochem Pharmacol. 1989; 38(11): 1763-9. doi: 10.1016/0006-2952(89)90410-3.
[20] Aherne SA, O'Brien NM. Mechanism of protection by the flavonoids, quercetin and rutin, against tert-butylhydroperoxide- and menadione-induced DNA single strand breaks in Caco-2 cells. Free Radic Biol Med. 2000; 29(6): 507-14. doi: 10.1016/S0891-5849(00)00360-9.
[21] Cavusoglu K, Yalcin E. Radioprotective effect of lycopene on chromosomal aberrations (CAs) induced by gamma radiation in human lymphocytes. J Environ Biol. 2009; 30(1): 113-7. PMID: 20112872. https://www.jeb.co.in/journal_issues/200901_jan09_spl/paper_18.pdf

Journal of Associated Medical Sciences 2026; 59(2): 393-420

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Obesity risk and body composition among Bangkok adolescents with autism spectrum disorder: Stratification by sensory and cognitive severity using data mining

Sarun Kunwittaya[1], Apilak Worachartcheewan[2], Chattip Silarangsri[3], Khanittha Punturee[4], Peeradech Thichanpiang[5], Kannika Permpoonputtana[1*]

[1]National Institute for Child and Family Development, Mahidol University, Nakhon Pathom Province, Thailand.
[2]Department of Community Medical Technology, Faculty of Medical Technology, Mahidol University, Nakhon Pathom Province, Thailand.
[3]Master of Science Program in Human Development, National Institute for Child and Family Development, Mahidol University, Nakhon Pathom Province, Thailand.
[4]Department of Medical Technology, Faculty of Associated Medical Sciences, Chang Mai University, Chiang Mai Province, Thailand.
[5]Occupational Therapy Division, Faculty of Physical Therapy, Mahidol University, Salaya, Nakhon Pathom Province, Thailand.

ARTICLE INFO

Article history:
Received 25 July 2025
Accepted as revised 26 March 2026
Available online 13 April 2026

Keywords:
Anthropometry, body composition, autism spectrum disorder (ASD), sensory/cognitive awareness, severity, association rule analysis

ABSTRACT

Background: Obesity is prevalent among children with autism spectrum disorder (ASD) and may worsen core symptoms. The unique challenges faced by this population may heighten their vulnerability to both general and ASD-specific obesity risk factors.

Objectives: This study examines body composition in adolescents with ASD, stratified by sensory and cognitive awareness severity and uses association rule analysis to identify key obesity-related factors for targeted intervention.

Materials and methods: This cross-sectional study was conducted at an inclusive school in Bangkok, Thailand, with 39 children aged 11-17 years diagnosed with ASD and their parents. ASD severity was assessed using the Thai Autism Treatment Evaluation Checklist (Thai-ATEC), and participants were categorized by sensory and cognitive awareness scores. Anthropometric data and body composition were measured using a bioelectrical impedance analyzer (InBody 770). BMI percentiles and body composition parameters such as body fat mass, skeletal muscle mass, and visceral fat area were recorded. Family and lifestyle data were also collected. Association rule analysis using a data mining technique was applied to explore variables linked to obesity.

Results: Different sensory/cognitive awareness severity levels (ATEC) had varying prevalences of obesity; the Mild group had the highest prevalence (86.67%), followed by the Moderate (66.67%) and Severe (58.33%) groups. The mild group consistently exhibited greater values for circumference, abdominal fat mass, body mass index, visceral fat area, waist-to-hip ratio, intracellular and extracellular water, skeletal muscle mass, and protein levels, indicating significant differences in anthropometric and body composition variables. Association rule analysis identified factors frequently linked to obesity, including exercise frequency, dietary behaviors, sensory/cognitive awareness severity, and digital technology use.

Conclusion: This study highlights obesity as a significant concern among adolescents with ASD with distinct body composition patterns across sensory/cognitive awareness severity levels. As the first comprehensive investigation of its kind in Thailand, it identifies key contributing factors-such as physical

* **Corresponding contributor.**
Author's Address: *National Institute for Child and Family Development, Mahidol University, Nakhon Pathom Province, Thailand.*
E-mail address: kannika.per@mahidol.ac.th
doi: 10.66285/JAMS.2026.070
E-ISSN: 2539-6056

inactivity, dietary habits, sensory processing differences, and digital media use-underscoring the need for personalized, lifestyle-based interventions in this population.

Introduction

Autism spectrum disorder (ASD) refers to a group of complex neurodevelopmental conditions characterized by impairments in social interaction and communication, along with repetitive behaviors and restricted interests. These symptoms often result in significant functional challenges in various environments, including the home, school, and community settings.[1] Diagnosis of ASD is typically possible by 18-24 months, as symptoms become distinguishable from typical developmental patterns and other developmental delays. Children with ASD often encounter difficulties in academic performance, behavior, and emotional regulation, necessitating a combination of pharmacological, educational, and psychosocial interventions.[2]

Educational and behavioral therapies remain the cornerstone of ASD treatment, as no pharmacological intervention currently addresses the core social deficits associated with the disorder. While antipsychotic medications such as risperidone can effectively reduce disruptive behaviors commonly observed in children with ASD, these treatments often have adverse effects, including significant weight gain that contributes to obesity.[3,4] Globally, the prevalence of ASD is approximately 1%, and its diagnosis has been increasing due to advancements in diagnostic criteria, improved screening tools, and more accurate identification of cases.[1,5] ASD is diagnosed about four times more frequently in males than in females.[6]

Obesity is an increasingly significant concern among children and adolescents with ASD, with evidence suggesting a positive correlation between obesity severity and the severity of ASD symptoms.[7,8] Contributing factors include selective eating habits, repetitive food-related behaviors, limited physical activity, sensory sensitivities, and the use of psychotropic medications.[9-12] Compared to neurotypical peers, children with ASD are at a disproportionately higher risk of obesity, though prevalence rates vary across studies.[13-15]

This condition presents serious health risks, such as increased likelihood of developing metabolic syndrome, type 2 diabetes, and cardiovascular disease.[13,16] Moreover, obesity can further aggravate ASD-related symptoms, compounding both behavioral and physical health challenges.[17,18] Sensory processing difficulties may result in food selectivity and avoidance of healthier food options, contributing to poor dietary quality.[19,20] Inadequate physical activity also contributes to the energy imbalance and weight gain commonly observed in this population.

Although the Body Mass Index (BMI) is a commonly used metric to evaluate overweight and obesity, it is limited in its ability to accurately reflect an individual's metabolic health status because it does not distinguish between lean and fat mass, especially in populations with abnormal growth and development patterns. This restriction may obscure differences in body composition that are clinically significant. On the other hand, more sophisticated methods like bioelectrical impedance analysis (BIA) enable comprehensive evaluations of skeletal muscle mass (SSM), body fat mass (BFM), and fat-free mass (FFM), providing a more accurate picture of physiological and nutritional health.[21]

According to recent epidemiological estimates, the prevalence of ASD in Thailand has increased sixfold in the last ten years, with an estimated 1 in 161 children under the age of five receiving a diagnosis.[22] There are still few studies conducted in Thai communities, despite an increase in research worldwide showing that children with ASD are more likely to be overweight or obese. Previous research has mostly focused on BMI as the only anthropometric measure, paying little attention to specific body composition measurements or how they relate to behavioral, cognitive, or sensory traits.[23]

The current study intends to investigate anthropo metric and body composition profiles among Thai adolescents with ASD, stratifying them according to sensory and cognitive severity levels, to fill in these knowledge gaps. Furthermore, this work makes use of association rule analysis, a machine learning method that is especially useful for identifying interdependent patterns and non-linear correlations between categorical data. In contrast to conventional statistical techniques, which frequently depend on linearity or normalcy assumptions and necessitate predetermined outcomes, association rule learning can capture frequent and significant co-occurrences in intricate, diverse datasets like those involving people with ASD. This method is especially useful in studies with small to medium sample sizes, where statistical power may be limited, as it evaluates associations based on support, confidence, and lift rather than group comparisons.[24,25] By identifying clinically relevant co-occurring risk factors, this technique offers insights that may be overlooked by conventional analyzes and supports the development of targeted obesity prevention strategies and interventions tailored for this high-risk population.

Materials and methods

Study population

This cross-sectional study, approved by the Institutional Review Board (MU-CIRB COA No. 2022/090.2608), was conducted at an inclusive school in Bangkok, Thailand, specializing in education for students with special needs, particularly those with ASD. Participants were selected using purposive sampling, with eligibility criteria requiring parents to be at least 18 years old, have children with ASD, and demonstrate

proficiency in the Thai language to facilitate effective communication. Each child had a clinical diagnosis of ASD confirmed by a child psychiatrist, based on the criteria specified in the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR),[26] or the 5th Edition (DSM-5).[6]

The study included 39 parent-child pairs, with children aged 11-17 years enrolled at the school during the research period from January to June 2023. All participants provided informed consent and voluntarily completed the questionnaire. Additionally, written consent was obtained from parents to allow their children's participation prior to the testing phase.

Measurement

Thai version of the Autism Treatment Evaluation Checklist (Thai-ATEC)

The severity of ASD symptoms in this study was assessed using the Thai version of the Thai-ATEC, a caregiver-administered instrument comprising four subscales: Speech/Language/Communication (14 items, 0-28), Sociability (20 items, 0-40), Sensory/Cognitive Awareness (18 items, 0-36), and Health/Physical/Behavior (25 items, 0-75). The first three subscales use a 0-2 scoring system, while the fourth uses a 0-3 scale, yielding a total score ranging from 0 to 179, with higher scores indicating greater symptom severity. All 39 participating parent-child pairs completed the Thai-ATEC in printed format, ensuring uniform administration. The tool has demonstrated robust psychometric properties: a total score ≤38 accurately classified mild ASD symptoms with 94% sensitivity and 61.9% specificity (AUC=0.90), while a score ≥68 effectively identified severe symptoms (94% sensitivity, 62.8% specificity, AUC=0.85). Additionally, the Thai-ATEC showed excellent inter-rater reliability (ICC=0.97), supporting its consistent use in caregiver assessments.[27]

Nutritional status

Height and weight measurements for each ASD participant were collected following WHO protocols.[28] Weight was recorded in kilograms to one decimal point, and height was measured in centimeters to one decimal point. BMI classification cut-offs in this study were based on WHO growth standards, with participants categorized into four groups using BMI-for-age percentiles: underweight (<5th percentile), normal weight (5th-84th percentile), overweight (85th-94th percentile), and obesity (≥95th percentile).

Family information of ASD

The family information of ASD such as child behavior (i.e., nutritional status, duration of exercise and duration of using digital technology) and parent's information (i.e., marital status, medical history, occupation and income) were evaluated in this study.

Body composition

In this study, body composition was assessed using bioelectrical impedance analysis (BIA) with the InBody 770 Body Composition Analyzer (InBody Co., Ltd., USA). BIA is a non-invasive technique that estimates body composition by passing a low-level electrical current through the body and measuring the resulting impedance. In line with the standard protocol for this device, the assessment was conducted in a standing position using hand and foot plate electrodes, which ensures consistent contact and segmental analysis across the limbs and trunk. Participants stood barefoot on the foot electrodes and lightly grasped the hand electrodes with arms slightly abducted to prevent contact with the torso.

To ensure accuracy, all participants were instructed to fast from food and beverages and refrain from physical activity for at least 12 hours prior to the measurement. The device provided multi-frequency impedance readings across five segments (four limbs and trunk), yielding comprehensive data including: body weight (kg), height (cm), abdominal circumference (cm), body fat mass (BFM, kg), percentage body fat (PBF, %), body mass index (BMI, kg/m^2), fat mass index (FMI, kg/m^2), visceral fat area (VFA, cm^2), waist-to-hip ratio (WHR), intracellular water (ICW, liters), extracellular water (ECW, liters), skeletal muscle mass (SMM, kg), fat-free mass (FFM, kg), fat-free mass index (FFMI), total body protein (kg), and obesity degree. These indicators were used to evaluate participants' physiological profiles and metabolic health characteristics relevant to neurodevelopmental risk in ASD.

Statistical analysis

Prior to analysis, the distribution of continuous variables was examined using the Shapiro-Wilk test to assess normality. The results indicated that the data were not normally distributed, and group sample sizes were unequal. Therefore, all continuous variables were summarized as mean±SD for descriptive purposes. Group differences in anthropometric and body composition parameters across the three sensory/cognitive awareness severity groups were examined using the Kruskal-Wallis's test. A $p<0.05$ was considered statistically significant.

To quantify the magnitude of group differences, effect sizes were calculated using epsilon-squared (ε^2) for each Kruskal-Wallis's test, computed as

$$\varepsilon^2 = \frac{H-k+1}{n-k}$$

where H denotes the Kruskal-Wallis's Chi-square statistic, k the number of groups, and n the total sample size. Effect sizes were interpreted as small ($\varepsilon^2 \approx 0.01$), medium ($\varepsilon^2 \approx 0.06$), and large ($\varepsilon^2 \geq 0.14$), reflecting negligible to substantial practical significance beyond statistical significance alone.[29] All statistical analyses were performed using IBM SPSS Statistics for Windows, version 30.0 (IBM Corp., Armonk, NY, USA).

Additionally, association rule analysis, a data

mining technique, was used to identify patterns of co-occurring variables associated with obesity in children with ASD. The analysis was performed using SPSS Clementine, version 11.1 (SPSS Inc., USA), employing the Apriori algorithm.[24] Only categorical variables were included in the model: sensory/cognitive severity level, obesity classification, dietary behavior, physical activity frequency and duration, digital technology use, and medication status. Minimum thresholds were set at 40% for support and 70% for confidence, to ensure robust and frequent associations.

Results

Participants characteristics

The demographic characteristics of adolescents with ASD (N=39), as presented in Table 1, indicate that the mean age was 12.28 years (SD=3.19), with the majority being male (84.62%) and a smaller proportion female (15.38%). Obesity was observed in 71.79% of the participants, with an average weight of 67.50 kg (SD=27.07) and an average height of 154.66 cm (SD=17.21). In terms of medication use, 48.72% of participants were treated with risperidone, 2.56% with Ritalin, while 48.72% were not on any medication. Parent-reported autism severity levels were distributed as 38.46% mild, 30.77% moderate, 30.77% severe. The prevalence of obesity across sensory/cognitive awareness severity levels (ATEC categories) showed notable differences, with 86.67% of participants in the mild group, 66.67% in the moderate group, and 58.33% in the severe group classified as obese. Exercise habits varied, with 43.58% not engaging in exercise, 28.21% exercising 1-2 days per week, and 28.21% exercising 3-7 days per week. The duration of exercise was less than 10 minutes per day for 41.02%, 10-20 minutes per day for 17.95%, 21-30 minutes per day for 17.95%, and over 30 minutes per day for 23.08%. Digital technology usage was reported as 43.59% using 1-3 hours per day, followed by 20.51% using less than 1 hour per day, 17.95% using 4-5 hours per day, 10.26% using more than 5 hours per day, and 7.7% not using digital technology at all.

Table 1. Demographic characteristics of adolescents with ASD (N=39).

Demographic variable	**Number** (%)
Age, Mean (SD)	12.28 (3.19)
Gender	
Male	33 (84.62)
Female	6 (15.38)
Obesity status	
Obese	28 (71.79)
Non-obese	11 (28.21)
Weight, Mean (SD)	67.50 (27.07)
Height, Mean (SD)	154.66 (17.21)
Medication used	
Risperidone	19 (48.72)
Ritalin	1 (2.56)
Not being treated with medicine	19 (48.72)
Autism severity levels (total ATEC)	
Mild	15 (38.46)
Moderate	12 (30.7)
Severe	12 (30.7)
% Obesity across sensory/cognitive awareness severity levels (ATEC Categories)	
Mild (N=15)	13 (86.67)
Moderate (N=12)	8 (66.67)
Severe (N=12)	7 (58.33)

Table 1. Demographic characteristics of adolescents with ASD (N=39)(continue).

Demographic variable	Number (%)
Exercise	
Not exercise	17 (43.58)
1-2 days/week	11 (28.21)
3-7 days/week	11 (28.21)
Duration of exercise	
<10 minute/day	16 (41.02)
10-20 minute/day	7 (17.95)
21-30 minute/day	7 (17.95)
>30 minute/day	9 (23.08)
Duration of using digital technology	
Not using	3 (7.69)
<1 hour/day	8 (20.51)
1-3 hours/day	17 (43.59)
4-5 hours/day	7 (17.95)
>5 hours/day	4 (10.26)

The anthropometric characteristics and body composition of adolescents with ASD

Participants with ASD were stratified into three groups based on the severity of sensory/cognitive awareness using the ATEC subscale: mild (N=15), moderate (N=12), and severe (N=12). Significant differences were observed across several anthropometric and body composition parameters (Table 2 and Figure 1). The Mild group consistently exhibited higher values than the other groups in abdominal circumference (p=0.010, ε^2=0.202), body fat mass (BFM; p=0.024, ε^2=0.154), body mass index (BMI; p=0.028, ε^2=0.144), visceral fat area (VFA; p=0.043, ε^2=0.119), and waist-to-hip ratio (WHR; p=0.027, ε^2=0.145), all reflecting large to medium-large effect sizes.

Additionally, the Mild group also demonstrated significantly elevated levels of intracellular water (ICW; p=0.003, ε^2=0.259), extracellular water (ECW; p=0.003, ε^2=0.260), skeletal muscle mass (SMM; p=0.003, ε^2=0.261), and protein content (p=0.003, ε^2=0.261), all with large effect sizes. While differences in fat mass index (FMI) approached significance (p=0.071, ε^2=0.091), suggesting a medium effect, no significant differences were found for percent body fat (PBF; p=0.300, ε^2=0.011), fat-free mass (FFM; p=0.235, ε^2=0.025), or obesity degree (p=0.152, ε^2=0.049), all of which had small or negligible effects. Notably, the Severe group consistently showed the lowest values across most variables.

Table 2. Anthropometric characteristics and body composition of ASD by sensory/cognitive awareness severity (ATEC Categories).

Anthropometric data and body composition	**Mean±SD**				**Kruskal-Wallis Chi-square**	***p* value**
	Total (N=39)	**Mild** (N=15)	**Moderate** (N=12)	**Severe** (N=12)		
Weight (kg)	67.50±27.07	84.07±24.91	64.82±26.46	49.47±17.47	1.691	0.429
Height (cm)	154.66±17.21	163.66±13.36	156.41±14.92	141.66±16.61	1.815	0.404
Circumference abdomen (cm)	86.71±19.94	98.01±18.49	84.54±18.47	74.75±16.20	9.283	0.010**
Body fat mass (BFM, kg)	25.72±14.62	33.56±14.85	23.55±14.04	18.08±10.41	7.458	0.024*
Percentage body fat (PBF, %)	36.20±10.79	38.98±10.92	33.75±9.79	35.18±11.67	2.407	0.300
Body mass index (BMI, kg/m^2)	27.20±7.18	31.07±7.01	25.53±7.20	24.05±5.44	7.173	0.028*
Fat mass index (FMI, kg/m^2)	10.41±5.11	12.62±5.37	9.15±4.74	8.92±4.52	5.284	0.071
Visceral fat area (VFA, cm^2)	122.73±64.98	155.25±60.39	109.50±61.54	95.32±60.87	6.290	0.043*
Waist-to-hip ratio (WHR)	0.87±0.08	0.91±0.07	0.86±0.06	0.82±0.09	7.242	0.027*
Intracellular water (ICW, L)	18.90±6.92	22.91±6.59	18.65±6.29	14.15±4.92	11.335	0.003**
Extracellular water (ECW, L)	11.76±4.17	14.22±3.89	11.59±3.82	8.88±3.00	11.346	0.003**
Skeletal muscle mass (SMM, kg)	22.64±9.03	27.88±8.57	22.30±8.22	16.44±6.41	11.395	0.003**
Fat-free mass (FFM, kg)	42.32±15.47	47.24±18.22	42.12±9.59	36.39±15.54	2.899	0.235
Protein (kg)	8.17±2.99	9.91±2.86	8.05±2.72	6.12±2.11	11.395	0.003**
Obesity degree	145.23±33.58	156.80±33.97	135.25±33.18	140.75±31.81	3.768	0.152

Note: **Significance tested at 0.05, **significance tested at 0.01.*

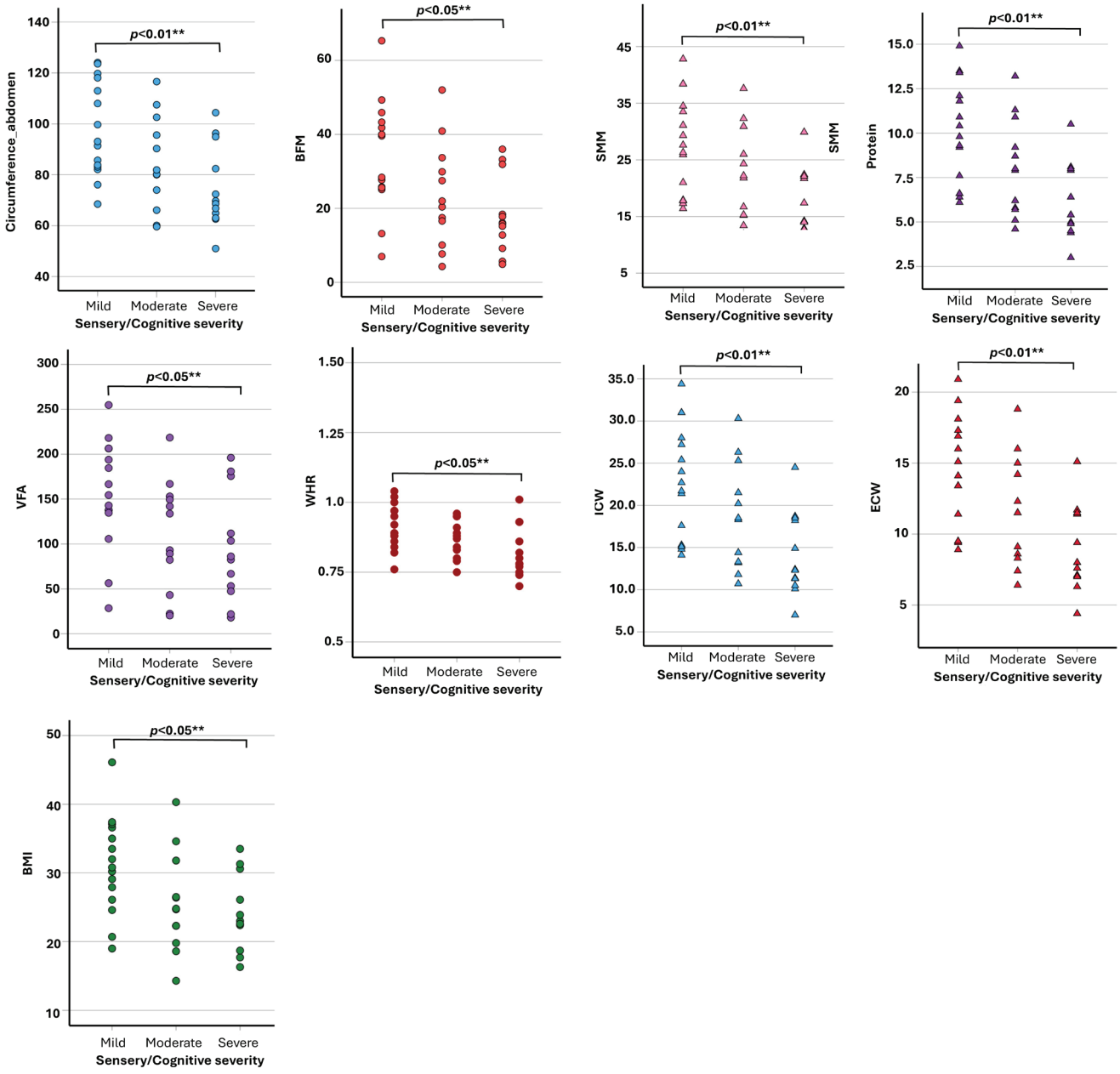


Figure 1. *Group differences in body composition parameters among adolescents with ASD by sensory/cognitive severity.*

Association rule analysis

A total of 278 association rules were generated using the Apriori algorithm, which is a well-established data mining technique for uncovering frequent patterns and conditional relationships in categorical datasets. The analysis included variables such as obesity classification, sensory/cognitive awareness level, physical activity, digital media use, and dietary behavior. To ensure robustness, rules were filtered using thresholds of minimum support ≥40% and confidence ≥70%. These criteria ensured that the resulting rules reflected both statistically significant and practically meaningful associations within the sample, where support values indicated that each rule was observed in at least 40% of participants and confidence values suggested a high probability (≥70%) that the consequent (obesity classification) occurred when the antecedent conditions were met.

Table 3 presents selected rules that exemplify the most dramatically meaningful patterns linking behavioral and physiological profiles to obesity. For instance, Rule #1 indicates that adolescents who exhibited high BMI, elevated visceral fat area, low physical activity, and poor dietary behavior were classified as obese with a confidence of 100%, a support of 47.1%, and a lift of 1.41. This suggests that this pattern occurred in nearly half the sample and was more predictive of obesity than would be expected by chance. Another illustrative rule showed that participants with mild sensory/cognitive awareness combined with high BMI and low physical activity were obese in 100% of cases (support: 42.4%), underscoring a deterministic-like relationship for this specific subgroup.

Across many of the high-confidence rules, mild sensory/cognitive awareness severity frequently appeared as a key antecedent factor, with confidence values ranging from 71.8% to 100%. This consistent appearance across rules implies a meaningful link between this sensory/cognitive profile and obesity risk in adolescents with ASD. The presence of several rules with 100% confidence suggests that certain combinations of behavioral and physiological traits almost invariably co-occurred with obesity in this population. Full details of the 278 association rules, including all support, confidence, and lift values, are provided in Supplementary data.

Table 3. Example of association rules identified for the ASD-obesity group.

Consequent	**Antecedent**	**Support** (%)	**Confidence** (%)
Obesity	BMI = Upper, Obesity degree = Upper, VFA = Upper, Consumption behavior = Unsuitable	58.54	100.00
Obesity	Duration using digital technology day = Average, BMI = Upper, Consumption behavior = Unsuitable	53.66	100.00
Obesity	Duration using digital technology day = Average, BMI = Upper, Obesity degree = Upper, Gender = Male	51.22	100.00
Obesity	BMI = Upper, Obesity degree = Upper, VFA = Upper, Gender = Male, Consumption behavior = Unsuitable	48.78	100.00
Obesity	Duration exercise day = Poor, Exercise week = Mild, BMI = Upper, Obesity degree = Upper	43.90	100.00
Obesity	Duration exercise day = Poor, Exercise week = Mild, Consumption behavior= Unsuitable	41.46	100.00
Obesity	ATEX Sensory cognitive awareness = Mild	41.46	76.47

Note: *Minimum antecedent support=40%, Minimum rule confidence=70%, Maximum number of antecedents=5*

Discussion

Overweight and obesity among children and adolescents with ASD continue to represent an important public health concern. The present study contributes novel evidence from Thailand by providing a detailed examination of anthropometric measures and body composition among adolescents with ASD, an age group and population that have received limited attention in prior local research. By stratifying participants according to sensory and cognitive awareness severity using the ATEC subscale, this study was able to examine not only differences in obesity prevalence but also the behavioral and contextual factors associated with excess weight, using association rule analysis to explore recurring patterns within the data.

The prevalence of obesity observed among Thai adolescents with ASD in this study was high and comparable to rates reported in other countries. Previous studies have shown that approximately one quarter of school-aged children with ASD in Japan meet criteria for obesity,[30] while research from Australia indicates that more than one third of youth with ASD are classified as overweight or obese.[17] Similar estimates have been reported in the United States, where combined rates of overweight and obesity among children with ASD range from 35% to 40%.[8,13,31,32] These consistent findings across regions suggest that obesity among individuals with ASD is a widespread issue rather than one confined to specific cultural or healthcare contexts, reinforcing the need for population-specific prevention strategies.

The significant change in obesity-related anthropometric and body composition characteristics among adolescents with ASD across varying degrees of sensory/cognitive awareness severity was one of the study's main findings. Interestingly, the group with mild ASD had the highest obesity prevalence and noticeably higher values for several variables, such as skeletal muscle mass, VFA, WHR, BMI, BFM, and circumference. In addition to being statistically significant, these differences also showed substantial effect sizes, indicating that the observed disparities were not merely the result of measurement variability in abdominal indicators but rather reflected meaningful differences in physical health status. According to this trend, the risk of obesity in ASD does not merely rise with the severity of symptoms; rather, it seems to be impacted by the distinct functional and behavioral characteristics of each severity level.

Adolescents with mild ASD often demonstrate relatively preserved cognitive and adaptive functioning, which may lead to greater autonomy in food selection and digital media use. However, persistent sensory sensitivities-particularly to taste, texture, and smell-can drive preferences toward familiar, energy-dense foods such as packaged snacks and sugar-sweetened beverages.[10,20,33-35] When this pattern of selective eating is coupled with reduced caregiver oversight, it may result in higher overall caloric intake, potentially explaining the large effect sizes observed in abdominal and visceral fat accumulation in this group.[10,36]

Conversely, adolescents with moderate to severe ASD frequently present with more rigid routines and pronounced sensory defensiveness, which manifest in behaviors such as food refusal, strict dietary limitations, and aversion to unfamiliar foods.[11,37] These feeding difficulties may result in lower total energy intake and less body fat accumulation, consistent with the lower mean values and small effect sizes observed in the severe group for many body composition metrics. Prior studies support that while food selectivity is common across the spectrum, the type of selectivity varies by severity-with mild ASD being more associated with selective preference for palatable, processed foods, and severe ASD involving global avoidance or very narrow food repertoires.[35] These distinctions may account for the inverse relationship between ASD severity and obesity indicators, and underscore the need for severity-specific, multi-dimensional interventions.

The behavioral and physiological characteristics linked to obesity in this group were further elucidated by the results of the association rule mining research. Several high-confidence rules showed that obesity was consistently associated with higher obesity levels, increased visceral fat area, elevated BMI, and unhealthy dietary behaviors. Other trends showed that among adolescents who were classified as obese, moderate daily screen use, male sex, and limited physical exercise were common. Crucially, the research demonstrated that adolescents with weak sensory and cognitive awareness were significantly more likely to be obese than those in other severity groups, confirming the idea that this subgroup represents a particularly vulnerable population.

Behavioral data collected in this study were consistent with these patterns. A significant number of participants reported prolonged daily screen use, short durations of exercise, and low levels of physical activity. Youth with ASD frequently report engaging in such sedentary habits, which are known to contribute to excessive weight gain.[17,38,39] Adolescents with mild ASD may be especially susceptible, as increased independence can facilitate longer screen time[40,41] while social communication difficulties and sensory sensitivities may reduce participation in organized sports or group-based physical activities.[42] Medication use also warrants consideration, as approximately half of the participants in this study were prescribed risperidone, a medication associated with increased appetite and weight gain.[43] The impact of such medications may be amplified among adolescents who have sufficient autonomy to access preferred foods independently.

From a clinical perspective, these association rules provide practical guidance for early screening even before overt obesity is clinically evident, adolescents

who exhibit poor food habits, minimal physical exercise, and mild sensory/cognitive awareness intensity may be deemed a high-risk subtype. Schools and families can emphasize structured movement programs and early lifestyle counseling by using these combined characteristics as a simplified risk profile. Early detection of at-risk adolescents may be enhanced by using behavioral indicators such as eating habits and screen use rather than relying solely on anthropometric measures.

Taken together, these findings suggest that obesity in adolescents with ASD arises from the interaction of multiple factors rather than a single determinant. Adolescents with mild ASD appear to be at particularly high risk due to a combination of sensory-driven food preferences, greater behavioral independence, sedentary lifestyles, and exposure to medications with metabolic side effects. The use of association rule mining provided additional support for this interpretation by identifying consistent co-occurring patterns across behavioral and physiological variables. These results underscore the importance of moving beyond symptom severity alone when assessing obesity risk in ASD and highlight the need for targeted, functionally informed prevention and intervention strategies.

Limitations

Several limitations should be acknowledged when interpreting the findings of this study. First, the results may not be as applicable to larger populations of adolescents with ASD in Thailand or other countries due to the limited sample size and the fact that it was taken from a single inclusive school in Bangkok. The results should be interpreted cautiously, even if the sample was sufficient for exploratory analyses and for using association rule mining, a method that finds common co-occurring patterns instead of relying solely on large-sample inferential statistics.

In addition, the study did not include a typically developing comparison group. As a result, it is not possible to determine whether the body composition patterns observed are specific to adolescents with ASD or reflect trends present in the broader adolescent population. To determine if sensory and cognitive severity adds specifically to obesity risk beyond typical adolescent patterns, future studies would benefit from including age- and sex-matched peers with typical development. It may also be possible to clarify how contextual factors, such as opportunities for physical activity, dietary environments, and school-based supports, influence obesity risk in this population by conducting comparative studies across various educational environments, such as inclusive schools and specialized institutions.

Conclusion

Obesity is becoming a growing health concern among adolescents with ASD, especially for those with milder symptoms who often have cognitive and sensory abilities that are reasonably intact. Nearly half (43.6%) of the participants reported that they did not regularly participate in physical activity. For those who exercised, the average session lasted no more than half an hour. Furthermore, 10.3% of participants reported using screen-based media for more than five hours per day.

These lifestyle behaviors-characterized by a persistently high amount of screen time and a lack of physical activity-have been identified in adolescents with ASD and have been linked in previous research to disturbances in energy control and a higher risk of weight gain. Since they frequently have more independence in their daily activities, youth with higher-functioning ASD may experience this problem more acutely. This independence can occasionally result in less engagement in organized physical activity and unsupervised access to electronic devices. Additionally, many adolescents with ASD experience social communication and sensory processing difficulties that can limit their participation in team sports or physical education classes. By reducing the diversity and context of possibilities for physical exercise, this group's propensity for familiar settings, routines, or food choices can exacerbate these health hazards.

At the policy level, these results imply that national obesity screening and prevention frameworks should specifically identify adolescents with ASD as a high-risk population, especially those with mild symptom intensity. Regular metabolic and body composition monitoring, as well as a systematic evaluation of dietary practices, physical activity, and screen time in this population, may be recommended by national guidelines. Early intervention may be made possible by integrating family-based behavioral counseling, modified physical activity programs, and nutrition education into public health and special education systems. Additionally, to fully address medication-related weight gain and lifestyle risk factors, cooperation between pediatric, psychiatric, and educational services is crucial. To better understand the long-term metabolic and health trajectories of adolescents with ASD as they enter adulthood and to inform evidence-based policy modification, longitudinal, multi-site research is necessary.

Ethical approval

The study was approved by the Human Research Ethics Committee of Mahidol University, Thailand (MU-CIRB COA No. 2022/090.2608).

Funding

This research received no external funding.

Conflicts of Interest

There are no conflicts of interest to report.

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

Data used for the analysis are available from the corresponding author upon request.

CRediT authorship contribution statement

Sarun Kunwittaya: conceptualization, data collection, statistical analysis, writing-original draft, **Apilak Worachartcheewan:** methodology, statistical analysis, data interpretation, **Chattip Silarangsri:** data collection, methodology, **Khanittha Punturee:** conceptualization, methodology conceptualization, **Peeradech Thichanpiang:** conceptualization, methodology, **Kannika Permpoonputtana:** conceptualization, methodology, investigation, statistical analysis, data interpretation, writing-review and editing.

Acknowledgements

Research participants, parents and teachers were gratefully acknowledged.

References

[1] Zeidan J, Fombonne E, Scorah J, Ibrahim A, Durkin MS, Saxena S, et al. Global prevalence of autism: A systematic review update. Autism Res. 2022; 15(5): 778-90. doi:10.1002/aur.2696.

[2] Okoye C, Obialo-Ibeawuchi CM, Obajeun OA, Sarwar S, Tawfik C, Waleed MS, et al. Early diagnosis of autism spectrum disorder: A review and analysis of the risks and benefits. Cureus. 2023; 15(8). doi:10.7759/cureus.43226.

[3] Mano-Sousa BJ, Pedrosa AM, Alves BC, Galduróz JCF, Belo VS, Chaves VE, et al. Effects of risperidone in autistic children and young adults: a systematic review and meta-analysis. Curr Neuropharmacol. 2021; 19(4): 538-52. doi:10.2174/1570159X18666200529151741.

[4] Scahill L, Jeon S, Boorin SJ, McDougle CJ, Aman MG, Dziura J, et al. Weight gain and metabolic consequences of risperidone in young children with autism spectrum disorder. J Am Acad Child Adolesc Psychiatry. 2016; 55(5): 415-23. doi:10.1016/j.jaac.2016.02.016.

[5] Durkin MS, Maenner MJ, Newschaffer CJ, Lee L, Cunniff CM, Daniels JL, et al. Advanced parental age and the risk of autism spectrum disorder. Am J Epidemiol. 2008; 168(11): 1268-76. doi:10.1093/aje/kwn250.

[6] American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. 5th Ed. 2013. doi:10.1176/appi.books.9780890425596.

[7] Ho HH, Eaves LC, Peabody D. Nutrient intake and obesity in children with autism. Focus Autism Dev Dis. 1997; 12(3): 187-92. doi:10.1177/108835769701200308.

[8] Egan AM, Dreyer ML, Odar CC, Beckwith M, Garrison CB. Obesity in young children with autism spectrum disorders: prevalence and associated factors. Child Obes. 2013; 9(2): 125-31. doi:10.1089/chi.2012.0028.

[9] Curtin C, Jojic M, Bandini LG. Obesity in children with autism spectrum disorder. Harv Rev Psychiatry. 2014; 22(2): 93-103. doi:10.1097/HRP.0000000000000031.

[10] Bandini LG, Anderson SE, Curtin C, Cermak S, Evans EW, Scampini R, et al. Food selectivity in children with autism spectrum disorders and typically developing children. J Pediatr. 2010; 157(2): 259-64. doi:10.1016/j.jpeds.2010.02.013.

[11] Ferguson BJ, Marler S, Altstein LL, et al. Psychophysiological associations with gastrointestinal symptomatology in autism spectrum disorder. Autism Res. 2017; 10(2): 276-88. doi:10.1002/aur.1646.

[12] Dhaliwal KK, Orsso CE, Richard C, Haqq AM, Zwaigenbaum L. Risk factors for unhealthy weight gain and obesity among children with autism spectrum disorder. Int J Mol Sci. 2019; 20(13): 3285. doi:10.3390/ijms20133285.

[13] Curtin C, Anderson SE, Must A, Bandini L. The prevalence of obesity in children with autism: a secondary data analysis using nationally representative data from the national survey of children's health. BMC Pediatr. 2010; 10: 1-5. doi:10.1186/1471-2431-10-11.

[14] de Vinck-Baroody O, Shui A, Macklin EA, Hyman SL, Leventhal JM, Weitzman C. Overweight and obesity in a sample of children with autism spectrum disorder. Acad Pediatr. 2015; 15(4): 396-404. doi: 10.1016/j.acap.2015.03.008.

[15] Zulkifli MN, Kadar M, Hamzaid NH. Weight status and associated risk factors of mealtime Behaviours among children with autism spectrum disorder. children. 2022; 9(7): 927. doi:10.3390/children9070927.

[16] Kittana M, Ahmadani A, Williams KE, Attlee A. Nutritional atatus and feeding behavior of children with autism spectrum disorder in the Middle East and North Africa region. A Systematic review Nutrients. 2023; 15(3): 711. doi:10.3390/nu15030711.

[17] Granich J, Lin A, Hunt A, Wray J, Dass A, Whitehouse AJ. Obesity and associated factors in youth with an autism spectrum disorder. Autism. 2016; 20(8): 916-26. doi:10.1177/1362361315616345.

[18] Kantawala B, Abu-Bakr A, Kasini B, Ndayambaje M, Soh SI, Nazir A, et al. Exploring the landscape of autism in Africa: challenges in diagnosis, support, and resources–a short communication. Ann Med Surg. 2023; 85(11): 5838-41. doi:10.1097/MS9.0000000000001376.

[19] Doreswamy S, Bashir A, Guarecuco JE, Lahori S, Baig A, Narra LR, et al. Effects of diet, nutrition, and exercise in children with autism and autism spectrum disorder: A literature review. Cureus.

2020; 12(12). doi: 10.7759/cureus.12222.
[20] Tomchek SD, Dunn W. Sensory processing in children with and without autism: a comparative study using the short sensory profile. Am J Occup Ther. 2007; 61(2): 190-200. doi:10.5014/ajot.61.2.190.
[21] Androutsos O, Zampelas A. Body composition in children: What does it tell us so far? Children. 2022; 9(8): 1199. doi:10.3390/children9081199.
[22] Asia-Pacific Development Center on Disability. Autism at a Glance in ASEAN Under the Autism Mapping Project in the ASEAN Region: Part II Country Profiles. Bangkok: Thailand: Asia-Pacific Development Center on Disability Foundation; 2020 [Available from: https://efaidnbmnnnibpcajpcglclefindmkaj/https://www.apcdfoundation.org/sites/default/files/2020-09/AW_Autism%20at%20a%20Glance_Part%20II.pdf.
[23] Khongpiboonkit R, Chaiyakit N, Wannapaschaiyong P, Sittanomai N. Prevalence and Factors Associated with Overweight and Obesity in Thai Children and Adolescents with Autism Spectrum Disorder: A Retrospective Longitudinal Study. J Autism Dev Disord. 2025. doi: 10.1007/s10803-025-06974-x
[24] Agrawal R, Imieliński T, Swami A. Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD international conference on Management of data: 207-216.
[25] Tan PN, Steinbach M, Kumar V. Introduction to data mining. Pearson Education India; 2016.
[26] American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV-TR®. 2000.
[27] Sunakarach K, Kessomboon P. Validity and reliability of the Thai version of the Autism Treatment Evaluation Checklist: A two-phase diagnostic accuracy study. F1000Research. 2018; 7: 538.
[28] WHO Expert Committee. Physical status : the use of and interpretation of anthropometry , report of a WHO expert committee. World Health Organ Tech Rep Ser. 1995: 312-44.
[29] Tomczak M, Tomczak E. The need to report effect size estimates revisited. Trends Sport Sci. 2014; 21(1): 19-25.
[30] Sasayama D, Kuge R, Toibana Y, Honda H. Trends in autism spectrum disorder diagnoses in Japan, 2009 to 2019. JAMA Network Open. 2021; 4(5): e219234-. doi:10.1001/jamanetworkopen.2021.9234.
[31] Broder-Fingert S, Brazauskas K, Lindgren K, Iannuzzi D, Van Cleave J. Prevalence of overweight and obesity in a large clinical sample of children with autism. Acad Pediatr. 2014; 14(4): 408-14. doi:10.1016/j.acap.2014.04.004.
[32] Zuckerman KE, Hill AP, Guion K, Voltolina L, Fombonne E. Overweight and obesity: prevalence and correlates in a large clinical sample of children with autism spectrum disorder. J Autism Dev Disord. 2014; 44: 1708-19. doi:10.1007/s10803-014-2050-9.
[33] Cermak SA, Curtin C, Bandini LG. Food selectivity and sensory sensitivity in children with autism spectrum disorders. J Am Diet Assoc. 2010; 110(2): 238-46. doi: 10.1016/j.jada.2009.10.032.
[34] Sharp WG, Berry RC, McCracken C, Nuhu NN, Marvel E, Saulnier CA, et al. Feeding problems and nutrient intake in children with autism spectrum disorders: a meta-analysis and comprehensive review of the literature. J Autism Dev Disord. 2013; 43: 2159-73. doi:10.1007/s10803-013-1771-5.
[35] Lockner DW, Crowe TK, Skipper BJ. Dietary intake and parents' perception of mealtime behaviors in preschool-age children with autism spectrum disorder and in typically developing children. J Am Diet Assoc. 2008; 108(8): 1360-3. doi:10.1016/j.jada.2008.05.003.
[36] Evans EW, Must A, Anderson SE, Curtin C, Scampini R, Maslin M, et al. Dietary patterns and body mass index in children with autism and typically developing children. Res Autism Spectr Disord 2012; 6(1): 399-405. doi:10.1016/j.rasd.2011.06.014.
[37] Hyman SL, Stewart PA, Schmidt B, Cain U, Lemcke N, Foley JT, et al. Nutrient intake from food in children with autism. Pediatrics. 2012; 130(Supplement_2): S145-S53. doi:10.1542/peds.2012-0900L.
[38] Corvey K, Menear KS, Preskitt J, Goldfarb S, Menachemi N. Obesity, physical activity and sedentary behaviors in children with an autism spectrum disorder. Matern Child Health J 2016; 20: 466-76. doi:10.1007/s10995-015-1844-5.
[39] McCoy SM, Jakicic JM, Gibbs BB. Comparison of obesity, physical activity, and sedentary behaviors between adolescents with autism spectrum disorders and without. J Autism Dev Disord 2016; 46: 2317-26. doi:10.1007/s10803-016-2762-0.
[40] Bandini LG, Curtin C, Phillips S, Anderson SE, Maslin M, Must A. Changes in food selectivity in children with autism spectrum disorder. J Autism Dev Disord 2017; 47: 439-46. doi: 10.1007/s10803-016-2963-6.
[41] Matheson BE, Douglas JM. Overweight and obesity in children with autism spectrum disorder (ASD): A critical review investigating the etiology, development, and maintenance of this relationship. Rev J Autism Dev Disord 2017; 4(2): 142-56. doi:10.1007/s40489-017-0103-7.
[42] O'Flaherty M, Hill J, Bourke M, Fortnum K, Thomas G, Gomersall SR, et al. Unpacking the relationship between autism, sport participation, and adolescent health: A causal mediation analysis. Disabil Health J. 2025; 3: 101832. doi: 10.1016/j.dhjo.2025.101832.
[43] Arkesteyn A, Van Damme T, Thoen A, Cornelissen

V, Healy S, Vancampfort D. Physical activity correlates in children and adolescents with autism spectrum disorder: a systematic review. Disabil Rehabil. 2022; 44(22): 6539-50. doi: 10.1080/09638288.2021.1970251.

Supplementary data

Table S1. Classification of body composition, behavior and family information for association rule analysis.

Parameter/Factors		Interpretation
Gender	Male	Male
	Female	Female
Age group	1= 7-12 years	School age children
	2= 13-17 years	Adolescent
Drug used	Loratadine	Loratadine
	Risperidone	Risperidone
	Ritalin	Ritalin
	Not Used	Not Used
ATEX speech language communication	Mild	Mild
	Moderate	Moderate
	Severe	Severe
ATEX sociability	Mild	Mild
	Moderate	Moderate
	Severe	Severe
ATEX health physical behavior	Mild	Mild
	Moderate	Moderate
	Severe	Severe
ATEX summary	Mild	Mild
	Moderate	Moderate
	Severe	Severe
Consumption behavior	1= Suitable	Not consume sweet, oily, and salty foods
	2= Unsuitable	Consume sweet, oily, and salty foods
Exercise/week	Not exercise	
	Mild	1-2 days/week
	Moderate	3-4 days/week
	Often	>4 days/week
Duration of exercise/day	Not exercise	
	Poor	< 20 mins/day
	Average	21-30 mins/day
	Good	>30 mins/day
Duration of using digital technology/day	Not used	
	Lower average	<1 hour/day
	Average	1-3 hours /day
	Above average	4-5 hours /day
	Too much	>5 hours /day
Duration of using smartphone/week	Not used	
	Rarely	1-2 times/week
	Sometime	3-4 times/week
	Often	5-6 times/week
	Nearly always	7-8 times/week
	Always	≥9 times/week

Table S1. Classification of body composition, behavior and family information for association rule analysis (continue).

Parameter/Factors		Interpretation
Body mass index group (BMI, kg/m^2)	*Male (11-17 years)*	
	Lower	<15.3
	Normal	15.3-27.0
	Upper	>27.0
	Female (11-17 years)	
	Lower	<16.6
	Normal	16.6-26.0
	Upper	>26.0
Body fat mass (BFM, kg)	*Male (11-17 years)*	
	Lower	<4.9
	Normal	4.9-21.5
	Upper	>21.5
	Female (11-17 years)	
	Lower	<9.3
	Normal	9.3-29.1
	Upper	>29.1
Percentage body fat (PBF, %)	*Male (11-17 years)*	
	Lower	<12.3
	Normal	12.3-27.1
	Upper	>27.1
	Female (11-17 years)	
	Lower	<22.4
	Normal	22.4-38.8
	Upper	>38.8
Fat mass index (FMI, kg/m^2)	*Male (11-17 years)*	
	Lower	<4.0
	Normal	4.0-9.6
	Upper	>9.6
	Female (11-17 years)	
	Lower	<5.5
	Normal	5.5-12.5
	Upper	>12.5
Fat free mass (FFM, kg)	*Male (11-17 years)*	
	Lower	<40.2
	Normal	40.2-62
	Upper	>62.0
	Female (11-17 years)	
	Lower	<34.3
	Normal	34.3-46.3
	Upper	>46.3

Table S1. Classification of body composition, behavior and family information for association rule analysis (continue).

Parameter/Factors		Interpretation
Fat free mass index (FFMI)	*Male (11-17 years)*	
	Lower	<18.1
	Normal	18.1-21.3
	Upper	>21.3
	Female (11-17 years)	
	Lower	<14.6
	Normal	14.6-17.6
	Upper	>17.6
Skeletal muscle mass (SMM, kg)	*Male (11-17 years)*	
	Lower	<25.6
	Normal	25.6-35.0
	Upper	>35.0
	Female (11-17 years)	
	Lower	<15.8
	Normal	15.8-23.2
	Upper	>23.2
Visceral fat area (VFA, cm^2)	*Male (11-17 years)*	
	Normal	<100
	Upper	>100
	Female (11-17 years)	
	Normal	<100
	Upper	>100
Circumference of abdomen (cm)	*Male (11-17 years)*	
	Lower	<79.9
	Normal	79.9-102.5
	Upper	>102.5
	Female (11-17 years)	
	Lower	<72.8
	Normal	72.8-109
	Upper	>109
Waist-hip ratio (WHR)	*Male (11-17 years)*	
	Lower	<0.90
	Normal	0.90-0.95
	Upper	>0.95
	Female (11-17 years))	
	Lower	<0.80
	Normal	0.80-0.86
	Upper	>0.86

Table S1. Classification of body composition, behavior and family information for association rule analysis (continue).

Parameter/Factors		Interpretation
Intracellular water (ICW, L	*Male (11-17 years)*	
	Lower	<18.8
	Normal	18.8-27.0
	Upper	>27.0
	Female (11-17 years)	
	Lower	<14.5
	Normal	14.5-18.3
	Upper	>18.3
Extracellular water (ECW, L)	*Male (11-17 years)*	
	Lower	<13.3
	Normal	13.3-19.3
	Upper	>19.3
	Female (11-17 years)	
	Lower	<10.9
	Normal	10.9-16.3
	Upper	>13.6
Obesity degree group	*Male (11-17 years)*	
	Lower	<90
	Normal	90-110
	Upper	>110
	Female (11-17 years)	
	Lower	<90
	Normal	90-110
	Upper	>110
Gender of caregiver	Male	Male
	Female	Female
Caregiver's relationship	Father	Father
	Mother	Mother
Education of caregiver	Junior high school	Junior high school
	Diploma	Diploma
	Bachelor's degree	Bachelor's degree
	Master's degree	Master's degree
Occupation of caregiver	Company employee	Company employee
	General employee	General employee
	Government officer	Government officer
	Merchant	Merchant
	Unemployed	Unemployed

Table S2. Analysis of association rules for obese individuals with ASD (Total 278 rules).

No.	Consequent	Antecedent	Support %	Confidence %
1	Obesity	BMI = Upper	70.73	100.00
2	Obesity	BMI = Upper, Obesity Degree = Upper	70.73	100.00
3	Obesity	BMI = Upper, Consumption Behavior = Unsuitable	65.85	100.00
4	Obesity	BMI = Upper, Obesity Degree = Upper, Consumption Behavior = Unsuitable	65.85	100.00
5	Obesity	BMI = Upper, VFA = Upper	63.41	100.00
6	Obesity	BMI = Upper, Obesity Degree = Upper, VFA = Upper	63.41	100.00
7	Obesity	BMI = Upper, Gender = Male	60.98	100.00
8	Obesity	BMI = Upper, Obesity Degree = Upper, Gender=Male	60.98	100.00
9	Obesity	Duration using digital technology Day = Average, BMI = Upper	58.54	100.00
10	Obesity	BMI = Upper, Disease caregiver = Normal	58.54	100.00
11	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper	58.54	100.00
12	Obesity	BMI = Upper, Obesity Degree = Upper, Disease caregiver = Normal	58.54	100.00
13	Obesity	BMI = Upper, VFA = Upper, Consumption Behavior = Unsuitable	58.54	100.00
14	Obesity	BMI = Upper, Obesity Degree = Upper, VFA = Upper, Consumption Behavior = Unsuitable	58.54	100.00
15	Obesity	BMI = Upper, Gender = Male, Consumption Behavior= Unsuitable	56.10	100.00
16	Obesity	BMI = Upper, Obesity Degree = Upper, Gender = Male, Consumption Behavior = Unsuitable	56.10	100.00
17	Obesity	Duration using digital technology Day = Average, BMI = Upper, VFA = Upper	53.66	100.00
18	Obesity	Duration using digital technology Day = Average, BMI = Upper, Consumption Behavior = Unsuitable	53.66	100.00
19	Obesity	BMI = Upper, VFA = Upper, Disease caregiver = Normal	53.66	100.00
20	Obesity	BMI = Upper, VFA = Upper, Gender = Male	53.66	100.00
21	Obesity	BMI = Upper, Disease caregiver = Normal, Consumption Behavior = Unsuitable	53.66	100.00
22	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, Visceral fat area = Upper	53.66	100.00
23	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, Consumption Behavior = Unsuitable	53.66	100.00
24	Obesity	BMI = Upper, Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal	53.66	100.00
25	Obesity	BMI = Upper, Obesity Degree = Upper, VFA = Upper, Gender = Male	53.66	100.00
26	Obesity	BMI = Upper, Obesity Degree = Upper, Disease caregiver = Normal, Consumption Behavior = Unsuitable	53.66	100.00
27	Obesity	Duration using digital technology Day = Average, BMI = Upper, Gender = Male	51.22	100.00
28	Obesity	BMI = Upper, Disease caregiver = Normal, Gender= Male	51.22	100.00

Table S2. Analysis of association rules for obese individuals with ASD (Total 278 rules)(continue).

No.	Consequent	Antecedent	Support %	Confidence %
29	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, Gender = Male	51.22	100.00
30	Obesity	BMI = Upper, Obesity Degree = Upper, Disease caregiver = Normal, Gender = Male	51.22	100.00
31	Obesity	FFMI = Normal, BMI = Upper	48.78	100.00
32	Obesity	FFMI = Normal, BMI = Upper, Obesity Degree = Upper	48.78	100.00
33	Obesity	Duration using digital technology Day = Average, BMI = Upper, Disease caregiver = Normal	48.78	100.00
34	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, Disease caregiver = Normal	48.78	100.00
35	Obesity	Duration using digital technology Day= Average, BMI = Upper, VFA = Upper, Consumption Behavior= Unsuitable	48.78	100.00
36	Obesity	BMI = Upper, VFA = Upper, Disease caregiver = Normal, Consumption Behavior= Unsuitable	48.78	100.00
37	Obesity	BMI = Upper, VFA = Upper, Gender = Male, Consumption Behavior = Unsuitable	48.78	100.00
38	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, Visceral fat area, Consumption Behavior = Unsuitable	48.78	100.00
39	Obesity	BMI = Upper, Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal, Consumption Behavior= Unsuitable	48.78	100.00
40	Obesity	BMI = Upper, Obesity Degree = Upper, VFA = Upper, Gender = Male, Consumption Behavior = Unsuitable	48.78	100.00
41	Obesity	FFMI = Normal, BMI = Upper, Consumption Behavior = Unsuitable	46.34	100.00
42	Obesity	FFMI = Normal, BMI = Upper, Obesity Degree = Upper, Consumption Behavior = Unsuitable	46.34	100.00
43	Obesity	Duration using digital technology Day = Average, BMI = Upper, Visceral fat area = Upper, Gender = Male	46.34	100.00
44	Obesity	Duration using digital technology Day = Average, BMI = Upper, Gender = Male, Consumption Behavior= Unsuitable	46.34	100.00
45	Obesity	BMI = Upper, VFA = Upper, Disease caregiver = Normal, Gender= Male	46.34	100.00
46	Obesity	BMI = Upper, Disease caregiver = Normal, Gender= Male, Consumption Behavior = Unsuitable	46.34	100.00
47	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, VFA = Upper, Gender= Male	46.34	100.00
48	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, Gender = Male, Consumption Behavior = Unsuitable	46.34	100.00
49	Obesity	BMI = Upper, Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal, Gender= Male	46.34	100.00
50	Obesity	BMI = Upper, Obesity Degree = Upper, Disease caregiver = Normal, Gender= Male, Consumption Behavior = Unsuitable	46.34	100.00
51	Obesity	Duration Exercise Day = Poor	43.90	100.00
52	Obesity	Duration Exercise Day = Poor, Exercise Week = Mild	43.90	100.00

Table S2. Analysis of association rules for obese individuals with ASD (Total 278 rules)(continue).

No.	Consequent	Antecedent	Support %	Confidence %
53	Obesity	Duration Exercise Day = Poor, BMI = Upper	43.90	100.00
54	Obesity	Duration Exercise Day = Poor, Obesity Degree = Upper	43.90	100.00
55	Obesity	Exercise Week = Mild, BMI = Upper	43.90	100.00
56	Obesity	Duration Exercise Day = Poor, Exercise Week = Mild, BMI = Upper	43.90	100.00
57	Obesity	Duration Exercise Day = Poor, Exercise Week = Mild, Obesity Degree = Upper	43.90	100.00
58	Obesity	Duration Exercise Day = Poor, BMI = Upper, Obesity Degree = Upper	43.90	100.00
59	Obesity	Exercise Week = Mild, BMI = Upper, Obesity Degree = Upper	43.90	100.00
60	Obesity	FFMI = Normal, BMI = Upper, VFA = Upper	43.90	100.00
61	Obesity	FFMI = Normal, BMI = Upper, Gender = Male	43.90	100.00
62	Obesity	Duration Exercise Day = Poor, Exercise Week = Mild, BMI = Upper, Obesity Degree = Upper	43.90	100.00
63	Obesity	FFMI = Normal, BMI = Upper, Obesity Degree = Upper, VFA = Upper	43.90	100.00
64	Obesity	FFMI = Normal, BMI = Obesity, Obesity Degree = Upper, Gender = Male	43.90	100.00
65	Obesity	Duration using digital technology Day = Average, BMI = Upper, VFA = Upper, Disease caregiver = Normal	43.90	100.00
66	Obesity	Duration using digital technology Day = Average, BMI = Upper, Disease caregiver = Normal, Gender = Male	43.90	100.00
67	Obesity	Duration using digital technology Day = Average, BMI = Upper, Disease caregiver = Normal, Consumption Behavior= Unsuitable	43.90	100.00
68	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal	43.90	100.00
69	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, Disease caregiver = Normal, Gender = Male	43.90	100.00
70	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, Disease caregiver = Normal, Consumption Behavior = Unsuitable	43.90	100.00
71	Obesity	Duration Exercise Day = Poor, Consumption Behavior = Unsuitable	41.46	100.00
72	Obesity	WHR = Lower, BMI = Upper	41.46	100.00
73	Obesity	Caregivers' relationship = Mother, BMI = Upper	41.46	100.00
74	Obesity	Duration Exercise Day = Poor, Exercise Week = Mild, Consumption Behavior= Unsuitable	41.46	100.00
75	Obesity	Duration Exercise Day = Poor, BMI = Upper, Consumption Behavior = Unsuitable	41.46	100.00
76	Obesity	Duration Exercise Day = Poor, Obesity Degree = Upper, Consumption Behavior = Unsuitable	41.46	100.00
77	Obesity	Exercise Week = Mild, BMI = Upper, Consumption Behavior = Unsuitable	41.46	100.00
78	Obesity	WHR = Lower, BMI = Upper, Obesity Degree = Upper	41.46	100.00
79	Obesity	FFMI = Normal, Duration using digital technology Day = Average, BMI = Upper	41.46	100.00

Table S2. Analysis of association rules for obese individuals with ASD (Total 278 rules)(continue).

No.	Consequent	Antecedent	Support %	Confidence %
80	Obesity	Caregivers' relationship= Mother, BMI = Upper, Obesity Degree = Upper	41.46	100.00
61	Obesity	FFMI = Normal, BMI = Upper, Gender = Male	43.90	100.00
62	Obesity	Duration Exercise Day = Poor, Exercise Week = Mild, BMI = Upper, Obesity Degree = Upper	43.90	100.00
63	Obesity	FFMI = Normal, BMI = Upper, Obesity Degree = Upper, VFA = Upper	43.90	100.00
64	Obesity	FFMI = Normal, BMI = Upper, Obesity Degree = Upper, Gender = Male	43.90	100.00
65	Obesity	Duration using digital technology Day = Average, BMI = Upper, VFA = Upper, Disease caregiver = Normal	43.90	100.00
66	Obesity	Duration using digital technology Day = Average, BMI = Upper, Disease caregiver = Normal, Gender = Male	43.90	100.00
67	Obesity	Duration using digital technology Day = Average, BMI = Upper, Disease caregiver = Normal, Consumption Behavior = Unsuitable	43.90	100.00
68	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal	43.90	100.00
69	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, Disease caregiver = Normal, Gender = Male	43.90	100.00
70	Obesity	Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper, Disease caregiver = Normal, Consumption Behavior = Unsuitable	43.90	100.00
71	Obesity	Duration Exercise Day = Poor, Consumption Behavior = Unsuitable	41.46	100.00
72	Obesity	WHR = Lower, BMI = Upper	41.46	100.00
73	Obesity	Caregivers' relationship = Mother, BMI = Upper	41.46	100.00
74	Obesity	Duration Exercise Day = Poor, Exercise Week = Mild, Consumption Behavior = Unsuitable	41.46	100.00
75	Obesity	Duration Exercise Day = Poor, BMI = Upper, Consumption Behavior = Unsuitable	41.46	100.00
76	Obesity	Duration Exercise Day = Poor, Obesity Degree = Upper, Consumption Behavior = Unsuitable	41.46	100.00
77	Obesity	Exercise Week = Mild, BMI = Upper, Consumption Behavior = Unsuitable	41.46	100.00
78	Obesity	WHR = Lower, BMI = Upper, Obesity Degree = Upper	41.46	100.00
79	Obesity	FFMI = Normal, Duration using digital technology Day = Average, BMI = Upper	41.46	100.00
80	Obesity	Caregivers' relationship = Mother, BMI = Upper, Obesity Degree = Upper	41.46	100.00
81	Obesity	Duration Exercise Day = Poor, Exercise Week = Mild, BMI = Upper, Consumption Behavior = Unsuitable	41.46	100.00
82	Obesity	Duration Exercise Day = Poor, Exercise Week = Mild, Obesity Degree = Upper, Consumption Behavior = Unsuitable	41.46	100.00
83	Obesity	Duration Exercise Day = Poor, BMI = Upper, Obesity Degree = Upper, Consumption Behavior = Unsuitable	41.46	100.00
84	Obesity	Exercise Week = Mild, BMI = Upper, Obesity Degree = Upper, Consumption Behavior = Unsuitable	41.46	100.00
85	Obesity	FFMI = Normal, Duration using digital technology Day = Average, BMI = Upper, Obesity Degree = Upper	41.46	100.00

Table S2. Analysis of association rules for obese individuals with ASD (Total 278 rules)(continue).

No.	Consequent	Antecedent	Support %	Confidence %
86	Obesity	FFMI = Normal, BMI = Upper, VFA = Upper, Consumption Behavior = Unsuitable	41.46	100.00
87	Obesity	FFMI = Normal, BMI = Upper, Gender = Male, Consumption Behavior = Unsuitable	41.46	100.00
88	Obesity	Duration Exercise Day = Poor, Exercise Week = Mild, BMI = Upper, Obesity Degree = Upper, Consumption Behavior = Unsuitable	41.46	100.00
89	Obesity	FFMI = Normal, BMI = Upper, Obesity Degree = Upper, VFA = Upper, Consumption Behavior = Unsuitable	41.46	100.00
90	Obesity	FFMI = Normal, BMI = Upper, Obesity Degree = Upper, Gender = Male, Consumption Behavior = Unsuitable	41.46	100.00
91	Obesity	Duration using digital technology Day = Average, BMI = Upper, VFA = Upper, Gender= Male, Consumption Behavior = Unsuitable	41.46	100.00
92	Obesity	BMI = Upper, VFA = Upper, Disease caregiver = Normal, Gender= Male, Consumption Behavior = Unsuitable	41.46	100.00
93	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper	60.98	96.00
94	Obesity	Duration using digital technology Day = Average, VFA = Upper	56.10	95.65
95	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, VFA = Upper	56.10	95.65
96	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, Consumption Behavior= Unsuitable	56.10	95.65
97	Obesity	Duration using digital technology Day = Average, Gender= Male	53.66	95.45
98	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, Gender = Male	53.66	95.45
99	Obesity	FFMI = Normal, Obesity Degree = Upper	51.22	95.24
100	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, Disease caregiver = Normal	51.22	95.24
101	Obesity	Duration using digital technology Day = Average, VFA = Upper, Consumption Behavior= Unsuitable	51.22	95.24
102	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, VFA = Upper, Consumption Behavior = Unsuitable	51.22	95.24
103	Obesity	FFMI = Normal, Obesity Degree = Upper, Consumption Behavior = Unsuitable	48.78	95.00
104	Obesity	Duration using digital technology Day = Average, VFA = Upper, Gender = Male	48.78	95.00
105	Obesity	Duration using digital technology Day = Average, Gender = Male, Consumption Behavior = Unsuitable	48.78	95.00
106	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, VFA = Upper, Gender = Male	48.78	95.00
107	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, Gender = Male, Consumption Behavior= Unsuitable	48.78	95.00
108	Obesity	Exercise Week = Mild, Obesity Degree = Upper	46.34	94.74
109	Obesity	FFMI = Normal, Obesity Degree = Upper, VFA = Upper	46.34	94.74
110	Obesity	FFMI = Normal, Obesity Degree = Upper, Gender = Male	46.34	94.74
111	Obesity	Duration using digital technology Day = Average, VFA = Upper, Disease caregiver = Normal	46.34	94.74

Table S2. Analysis of association rules for obese individuals with ASD (Total 278 rules)(continue).

No.	Consequent	Antecedent	Support %	Confidence %
112	Obesity	Duration using digital technology Day = Average, Disease caregiver = Normal, Gender= Male	46.34	94.74
113	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal	46.34	94.74
114	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, Disease caregiver = Normal, Gender= Male	46.34	94.74
115	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, Disease caregiver = Normal, Consumption Behavior = Unsuitable	46.34	94.74
116	Obesity	WHR = Lower, Obesity Degree = Upper	43.90	94.44
117	Obesity	Exercise Week = Mild, Obesity Degree = Upper, Consumption Behavior = Unsuitable	43.90	94.44
118	Obesity	FFMI = Normal, Duration using digital technology Day = Average, Obesity Degree = Upper	43.90	94.44
119	Obesity	FFMI = Normal, Obesity Degree = Upper, VFA = Upper, Consumption Behavior = Unsuitable	43.90	94.44
120	Obesity	FFMI = Normal, Obesity Degree = Upper, Gender = Male, Consumption Behavior = Unsuitable	43.90	94.44
121	Obesity	Duration using digital technology Day = Average, VFA = Upper, Gender = Male, Consumption Behavior = Unsuitable	43.90	94.44
122	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, VFA = Upper, Gender = Male, Consumption Behavior = Unsuitable	43.90	94.44
123	Obesity	BMI = Upper, Obesity Degree = Upper	41.46	94.12
124	Obesity	FMI = Normal, Obesity Degree = Upper	41.46	94.12
125	Obesity	BMI = Upper, Obesity Degree = Upper, VFA = Upper	41.46	94.12
126	Obesity	Exercise Week = Mild, Obesity Degree = Upper, VFA = Upper	41.46	94.12
127	Obesity	Exercise Week= Mild, Obesity Degree = Upper, Disease caregiver = Normal	41.46	94.12
128	Obesity	Exercise Week= Mild, Obesity Degree = Upper, Gender = Male	41.46	94.12
129	Obesity	FMI = Normal, Obesity Degree = Upper, Consumption Behavior = Unsuitable	41.46	94.12
130	Obesity	WHR = Lower, Obesity Degree = Upper, Consumption Behavior = Unsuitable	41.46	94.12
131	Obesity	FFMI = Normal, Duration using digital technology Day = Average, VFA = Upper	41.46	94.12
132	Obesity	FFMI = Normal, Obesity Degree = Upper, Disease caregiver = Normal	41.46	94.12
133	Obesity	FFMI = Normal, Duration using digital technology Day = Average, Obesity Degree = Upper, VFA = Upper	41.46	94.12
134	Obesity	FFMI = Normal, Duration using digital technology Day = Average, Obesity Degree = Upper, Consumption Behavior= Unsuitable	41.46	94.12
135	Obesity	FFMI = Normal, Obesity Degree = Upper, VFA = Upper, Gender = Male	41.46	94.12
136	Obesity	Duration using digital technology Day = Average, VFA = Upper, Disease caregiver = Normal, Gender = Male	41.46	94.12

Table S2. Analysis of association rules for obese individuals with ASD (Total 278 rules)(continue).

No.	Consequent	Antecedent	Support %	Confidence %
137	Obesity	Duration using digital technology Day = Average, VFA = Upper, Disease caregiver = Normal, Consumption Behavior = Unsuitable	41.46	94.12
138	Obesity	Duration using digital technology Day = Average, Disease caregiver = Normal, Gender = Male, Consumption Behavior = Unsuitable	41.46	94.12
139	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal, Gender = Male	41.46	94.12
140	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal, Consumption Behavior = Unsuitable	41.46	94.12
141	Obesity	Duration using digital technology Day = Average, Obesity Degree = Upper, Disease caregiver = Normal, Gender, Consumption Behavior = Unsuitable	41.46	94.12
142	Obesity	Obesity Degree = Upper	75.61	93.55
143	Obesity	Obesity Degree = Upper, Consumption Behavior = Unsuitable	70.73	93.10
144	Obesity	Obesity Degree = Upper, VFA = Upper	68.29	92.86
145	Obesity	Obesity Degree = Upper, Gender= Male	65.85	92.59
146	Obesity	Duration using digital technology Day= Average	63.41	92.31
147	Obesity	Obesity Degree = Upper, Disease caregiver = Normal	63.41	92.31
148	Obesity	Obesity Degree = Upper, VFA = Upper, Consumption Behavior = Unsuitable	63.41	92.31
149	Obesity	Obesity Degree = Upper, Gender = Male, Consumption Behavior = Unsuitable	60.98	92.00
150	Obesity	Duration using digital technology Day= Average, Consumption Behavior = Unsuitable	58.54	91.67
151	Obesity	Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal	58.54	91.67
152	Obesity	Obesity Degree = Upper, VFA = Upper, Gender = Male	58.54	91.67
153	Obesity	Obesity Degree = Upper, Disease caregiver = Normal, Consumption Behavior = Unsuitable	58.54	91.67
154	Obesity	Obesity Degree = Upper, Disease caregiver = Normal, Gender = Male	56.10	91.30
155	Obesity	Duration using digital technology Day = Average, Disease caregiver = Normal	53.66	90.91
156	Obesity	Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal, Consumption Behavior = Unsuitable	53.66	90.91
157	Obesity	Obesity Degree = Upper, VFA = Upper, Gender = Male, Consumption Behavior = Unsuitable	53.66	90.91
158	Obesity	Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal, Gender= Male	51.22	90.48
159	Obesity	Obesity Degree = Upper, Disease caregiver = Normal, Gender = Male, Consumption Behavior= Unsuitable	51.22	90.48
160	Obesity	Duration using digital technology Day = Average, Disease caregiver = Normal, Consumption Behavior = Unsuitable	48.78	90.00
161	Obesity	Exercise Week= Mild, Consumption Behavior = Unsuitable	46.34	89.47

Table S2. Analysis of association rules for obese individuals with ASD (Total 278 rules)(continue).

No.	Consequent	Antecedent	Support %	Confidence %
162	Obesity	FFMI = Normal, Duration using digital technology Day = Average	46.34	89.47
163	Obesity	Caregivers' relationship = Mother, obesity Degree = Upper	46.34	89.47
164	Obesity	Obesity Degree = Upper, VFA = Upper, Disease caregiver = Normal, Gender = Male, Consumption Behavior = Unsuitable	46.34	89.47
165	Obesity	Exercise Week= Mild, VFA = Upper	43.90	88.89
166	Obesity	FFMI = Normal, Duration using digital technology Day = Average, Consumption Behavior= Unsuitable	43.90	88.89
167	Obesity	Caregivers' relationship= Mother, Obesity Degree = Upper, Gender= Male	43.90	88.89
168	Obesity	Drug Used= Risperidone, Obesity Degree = Upper	41.46	88.24
169	Obesity	Drug Used= Risperidone, VFA = Upper	41.46	88.24
170	Obesity	Age = School age children, Obesity Degree = Upper	41.46	88.24
171	Obesity	Drug Used = Risperidone, Obesity Degree = Upper, VFA = Upper	41.46	88.24
172	Obesity	Exercise Week = Mild, VFA = Upper, Consumption Behavior = Unsuitable	41.46	88.24
173	Obesity	Exercise Week = Mild, Disease caregiver = Normal, Consumption Behavior = Unsuitable	41.46	88.24
174	Obesity	Exercise Week = Mild, Gender = Male, Consumption Behavior = Unsuitable	41.46	88.24
175	Obesity	Age = School age children, Obesity Degree = Upper, Consumption Behavior = Unsuitable	41.46	88.24
176	Obesity	Caregivers' relationship = Mother, Obesity Degree = Upper, Disease caregiver = Normal	41.46	88.24
177	Obesity	Caregivers' relationship = Mother, Obesity Degree = Upper, Consumption Behavior= Unsuitable	41.46	88.24
178	Obesity	Exercise Week = Mild	51.22	85.71
179	Obesity	FFMI = Normal, VFA = Upper	51.22	85.71
180	Obesity	FFMI = Normal, Gender = Male	51.22	85.71
181	Obesity	FFMI = Normal, VFA = Upper, Consumption Behavior = Unsuitable	48.78	85.00
182	Obesity	FFMI = Normal, Gender = Male, Consumption Behavior = Unsuitable	48.78	85.00
183	Obesity	BMI = Upper	46.34	84.21
184	Obesity	BMI = Upper, VFA = Upper	46.34	84.21
185	Obesity	Exercise Week = Mild, Disease caregiver = Normal	46.34	84.21
186	Obesity	Exercise Week = Mild, Gender = Male	46.34	84.21
187	Obesity	FFMI = Normal, VFA = Upper, Gender = Male	46.34	84.21
188	Obesity	FFMI = Normal	58.54	83.33
189	Obesity	VFA = Upper, Gender = Male, Consumption Behavior = Unsuitable	58.54	83.33
190	Obesity	ATEX Health Physical Behavior = Mild	43.90	83.33

Table S2. Analysis of association rules for obese individuals with ASD (Total 278 rules)(continue).

No.	Consequent	Antecedent	Support %	Confidence %
191	Obesity	FFMI = Normal, VFA = Upper, Disease caregiver = Normal	43.90	83.33
192	Obesity	FFMI = Normal, Disease caregiver = Normal, Gender = Male	43.90	83.33
193	Obesity	FFMI = Normal, VFA = Upper, Gender = Male, Consumption Behavior = Unsuitable	43.90	83.33
194	Obesity	VFA = Upper, Consumption Behavior= Unsuitable	70.73	82.76
195	Obesity	FFMI = Normal, Consumption Behavior= Unsuitable	56.10	82.61
196	Obesity	PBF = Normal	41.46	82.35
197	Obesity	BMI = Upper, Gender= Male	41.46	82.35
198	Obesity	SMM = Normal, FFMI = Normal	41.46	82.35
199	Obesity	FMI = Normal, VFA = Upper	41.46	82.35
200	Obesity	BMI = Upper, VFA = Upper, Gender = Male	41.46	82.35
201	Obesity	Exercise Week = Mild, Disease caregiver = Normal, Gender = Male	41.46	82.35
202	Obesity	FMI = Normal, VFA = Upper, Consumption Behavior = Unsuitable	41.46	82.35
203	Obesity	FFMI = Normal, VFA = Upper, Disease caregiver = Normal, Gender = Male	41.46	82.35
204	Obesity	FFMI = Normal, VFA = Upper, Disease caregiver = Normal, Consumption Behavior= Unsuitable	41.46	82.35
205	Obesity	FFMI = Normal, Disease caregiver = Normal, Gender = Male, Consumption Behavior = Unsuitable	41.46	82.35
206	Obesity	VFA = Upper, Gender= Male	65.85	81.48
207	Obesity	VFA = Upper	78.05	81.25
208	Obesity	VFA = Upper, Disease caregiver = Normal, Gender= Male, Consumption Behavior = Unsuitable	51.22	80.95
209	Obesity	VFA = Upper, Disease caregiver = Normal, Consumption Behavior = Unsuitable	60.98	80.00
210	Obesity	FFMI = Normal, Disease caregiver = Normal	48.78	80.00
211	Obesity	VFA = Upper, Disease caregiver = Normal, Gender= Male	58.54	79.17
212	Obesity	FFMI = Normal, Disease caregiver = Normal, Consumption Behavior= Unsuitable	46.34	78.95
213	Obesity	VFA = Upper, Disease caregiver = Normal	68.29	78.57
214	Obesity	ATEX Sociability = Mild, Consumption Behavior= Unsuitable	43.90	77.78
215	Obesity	SMM = Normal, Gender = Male	43.90	77.78
216	Obesity	WHR = Lower, Disease caregiver = Normal	43.90	77.78
217	Obesity	WHR = Lower	53.66	77.27
218	Obesity	ATEX Sensory Cognitive Awareness = Mild	41.46	76.47
219	Obesity	FFM = Lower, Consumption Behavior = Unsuitable	41.46	76.47
220	Obesity	ECW = Lower, Consumption Behavior = Unsuitable	41.46	76.47

Table S2. Analysis of association rules for obese individuals with ASD (Total 278 rules)(continue).

No.	Consequent	Antecedent	Support %	Confidence %
221	Obesity	ATEX Sociability = Mild, Disease caregiver = Normal	41.46	76.47
222	Obesity	ATEX Sociability = Mild, Gender= Male	41.46	76.47
223	Obesity	ICW = Lower, WHR = Lower	41.46	76.47
224	Obesity	SMM = Normal, VFA = Upper	41.46	76.47
225	Obesity	ECW = Normal, FFMI = Normal	41.46	76.47
226	Obesity	FMI = Normal, Gender = Male	41.46	76.47
227	Obesity	FFM = Lower, ECW = Lower, Consumption Behavior = Unsuitable	41.46	76.47
228	Obesity	ICW = Lower, WHR = Lower, Consumption Behavior = Unsuitable	41.46	76.47
229	Obesity	SMM = Normal, Gender = Male, Consumption Behavior = Unsuitable	41.46	76.47
230	Obesity	Drug Used = Risperidone, Disease caregiver = Normal, Consumption Behavior = Unsuitable	41.46	76.47
231	Obesity	FMI = Normal, Gender = Male, Consumption Behavior = Unsuitable	41.46	76.47
232	Obesity	WHR = Lower, Disease caregiver = Normal, Consumption Behavior= Unsuitable	41.46	76.47
233	Obesity	Age = School age children, VFA = Upper, Consumption Behavior = Unsuitable	41.46	76.47
234	Obesity	Caregivers' relationship = Mother, VFA = Upper, Gender = Male	41.46	76.47
235	Obesity	WHR = Lower, Consumption Behavior = Unsuitable	51.22	76.19
236	Obesity	ATEX Sociability = Mild	48.78	75.00
237	Obesity	SMM = Normal	48.78	75.00
238	Obesity	Gender = Male, Consumption Behavior = Unsuitable	75.61	74.19
239	Obesity	SMM = Normal, Consumption Behavior = Unsuitable	46.34	73.68
240	Obesity	Drug Used = Risperidone, Disease caregiver = Normal	46.34	73.68
241	Obesity	Drug Used = Risperidone, Consumption Behavior = Unsuitable	46.34	73.68
242	Obesity	WHR = Lower, BFM = Normal	46.34	73.68
243	Obesity	WHR = Lower, Gender = Male	46.34	73.68
244	Obesity	Caregivers' relationship = Mother, VFA = Upper	46.34	73.68
245	Obesity	WHR = Lower, BFM = Normal, Consumption Behavior = Unsuitable	46.34	73.68
246	Obesity	Caregivers' relationship = Mother, Gender = Male, Consumption Behavior = Unsuitable	46.34	73.68
247	Obesity	Disease caregiver = Normal, Gender = Male, Consumption Behavior = Unsuitable	63.41	73.08
248	Obesity	Consumption Behavior= Unsuitable	90.24	72.97
249	Obesity	FMI = Normal	53.66	72.73
250	Obesity	FMI = Normal, Consumption Behavior = Unsuitable	53.66	72.73

Table S2. Analysis of association rules for obese individuals with ASD (Total 278 rules)(continue).

No.	Consequent	Antecedent	Support %	Confidence %
251	Obesity	Caregivers' relationship = Mother, Gender= Male	53.66	72.73
252	Obesity	FFM = Lower	43.90	72.22
253	Obesity	ECW = Lower	43.90	72.22
254	Obesity	FFM = Lower, ECW = Lower	43.90	72.22
255	Obesity	Age = School age children, VFA = Upper	43.90	72.22
256	Obesity	WHR = Lower, Gender = Male, Consumption Behavior = Unsuitable	43.90	72.22
257	Obesity	Caregivers' relationship = Mother, VFA = Upper, Disease caregiver = Normal	43.90	72.22
258	Obesity	Gender = Male	85.37	71.43
259	Obesity	Drug Used = Risperidone	51.22	71.43
260	Obesity	Disease caregiver = Normal, Consumption Behavior = Unsuitable	75.61	70.97
261	Obesity	Gender Caregiver = Male	41.46	70.59
262	Obesity	FFM = Lower, Disease caregiver = Normal	41.46	70.59
263	Obesity	ECW = Lower, Disease caregiver = Normal	41.46	70.59
264	Obesity	Residence = Apartment, Consumption Behavior = Unsuitable	41.46	70.59
265	Obesity	FFM = Lower, ECW = Lower, Disease caregiver = Normal	41.46	70.59
266	Obesity	PBF = Upper, VFA = Upper, Gender = Male	41.46	70.59
267	Obesity	PBF = Upper, VFA = Upper, Consumption Behavior= Unsuitable	41.46	70.59
268	Obesity	Drug Used = Risperidone, Disease caregiver = Normal, Gender = Male	41.46	70.59
269	Obesity	Drug Used = Risperidone, Gender = Male, Consumption Behavior = Unsuitable	41.46	70.59
270	Obesity	WHR = Lower, BFM = Normal, Gender = Male	41.46	70.59
271	Obesity	Age = School age children, VFA = Upper, Disease caregiver = Normal	41.46	70.59
272	Obesity	BFM = Normal, Caregivers relationship = Mother, Consumption Behavior = Unsuitable	41.46	70.59
273	Obesity	Caregivers' relationship = Mother, VFA = Upper, Consumption Behavior = Unsuitable	41.46	70.59
274	Obesity	WHR = Lower, BFM = Normal, Gender = Male, Consumption Behavior = Unsuitable	41.46	70.59
275	Obesity	Caregivers' relationship = Mother, Disease caregiver = Normal, Gender = Male, Consumption Behavior = Unsuitable	41.46	70.59
276	Obesity	Disease caregiver = Normal, Gender= Male	73.17	70.00
277	Obesity	PBF = Upper, VFA = Upper	48.78	70.00
278	Obesity	Caregivers' relationship = Mother, Disease caregiver = Normal, Gender = Male	48.78	70.00

Note: *Minimum antecedent support=40%, Minimumrule confidence=70%, Maximun number of antecedents=5*

Journal of Associated Medical Sciences 2026; 59(2): 421-427

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Efficacy of the modified carbapenem inactivation method (mCIM) for detecting carbapenemase-producing *Pseudomonas aeruginosa* in a 1,000-bed tertiary hospital in Thailand

Jakkrit Khampakul[1,2], Noppadon Jumroon[1], Narong Chaihongsa[2], Pitak Santanirand[2*]

[1]Faculty of Allied Health Science, Naresuan University, Phitsanulok Province, Thailand.
[2]Department of Pathology, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 3 January 2026
Accepted as revised 16 April 2026
Available online 23 April 2026

Keywords:
Carbapenem-resistant Pseudomonas aeruginosa (CRPA), carbapenemase-producing P. aeruginosa (CPPA), modified carbapenem inactivation method (mCIM).

ABSTRACT

Background: Carbapenem-resistant *Pseudomonas aeruginosa* (CRPA) is a prominent healthcare-associated pathogen with worldwide implications for public health. Distinguishing between carbapenemase-producing and non-carbapenemase-producing strains over an extended hospitalization period is vital in assessing the risk of treatment failure in infected patients.

Objectives: To study the prevalence of carbapenemase genes and to evaluate the efficacy of the modified carbapenem inactivation method (mCIM) for detecting carbapenemase production among CRPA isolated from Thailand.

Materials and methods: A total of 164 CRPA isolates were phenotypically tested for carbapenemase production by the mCIM comparing with genotypic detection of carbapenemase genes using the direct flow chip kits and conventional PCR. In addition, carbapenemase genes were identified by sequencing.

Results: The carbapenemase genes, including bla_{IMP}, bla_{VIM}, bla_{NDM}, bla_{GES}, and $bla_{SIM,}$ were found in 83 (50.61%) isolates. The bla_{IMP} showed to be the most prevalent gene, followed by bla_{VIM} and bla_{NDM}, respectively. Of these isolates, 68/83 (82%) of carbapenemase gene-positive isolates were positive by the mCIM assay. The 15 mCIM-negative isolates, six (40%) carried bla_{VIM-2} gene, while five (33%) isolates contained bla_{GES-5}. Moreover, bla_{IMP} was detected in 3 isolates (20%) and one isolate (7%) of a combination between bla_{VIM} and bla_{IMP} genes. Furthermore, we noted that isolates carrying bla_{VIM-2} were the most common among the mCIM-negative carbapenemase-producing *P. aeruginosa* (CPPA), followed by $bla_{GES-5.}$

Conclusion: This study represents a critical report on using mCIM to detect CPPA. It highlights that relying solely on mCIM may not be sufficient for the comprehensive detection of CPPA. Further approaches are needed to enhance the accuracy of detection in this context.

Introduction

Carbapenem-resistant *Pseudomonas aeruginosa* (CRPA) is a significant causative agent of healthcare-associated infection worldwide.[1] Several pivotal mechanisms underlie carbapenem resistance. The primary mechanism involves drug efflux, which occurs due to the over-expression of the MexAB-OprM efflux pump,[2] the overproduction of AmpC beta-lactamase and the inactivation of the OprD outer membrane protein, and carbapenemase production. Although the

* **Corresponding contributor.**
Author's Address: *Department of Pathology, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.*
E-mail address: pitak.san@mahidol.ac.th
doi: 10.66285/JAMS.2026.071
E-ISSN: 2539-6056

latter mechanism is relatively less frequent among *P. aeruginosa* isolates, the emergence of carbapenemase production is showing an increasing trend, adding a new dimension to carbapenem resistance mechanisms.[1] This mechanism of carbapenem resistance assumes paramount importance as it significantly compromises the effectiveness of widely employed antipseudomonal agents, including ceftazidime, cefepime, piperacillin-tazobactam, along with the more recently introduced beta-lactam/beta-lactamase inhibitor combinations, such as ceftolozane-tazobactam, imipenem-relebactam, and ceftazidime-avibactam. Notably, the carbapenem resistance determinants carried by *P. aeruginosa* are frequently encoded on various mobile genetic elements, including plasmids, such as the *IncP* type, class I integrons housing the bla_{VIM} gene, and other mobile genetic elements, particularly insertion sequences with a common region (ISCRs). These genetic elements enhance the capacity to disseminate resistance across multiple species, underscoring the significance of this mechanism in the context of antimicrobial resistance.[2] Furthermore, these isolates often harbor supplementary resistance determinants that reduce the clinical effectiveness of fluoroquinolones and aminoglycosides.

Carbapenemase-producing *P. aeruginosa* (CPPA) frequently resists most of the therapeutic options, elevating the risk of treatment failure. Additionally, CPPA has been linked to nosocomial transmission, necessitating the implementation of infection prevention measures.[1] *Pseudomonas aeruginosa* isolates have been documented to harbor a diverse array of carbapenemases on a global scale. In Latin America, this spectrum encompasses carbapenemases such as *Klebsiella pneumoniae* carbapenemase (KPC), Guiana extended-spectrum (GES), Imipenemase (IMP), Verona integron-encoded metallo-β-lactamase (VIM), New Delhi metallo-beta-lactamase (NDM), and Sao Paolo metallo-beta-lactamase (SPM).[3] A notable percentage of CRPA isolates in various regions, ranging from 30 to 69%, possessed carbapenemases. In contrast, only 2% of CRPA isolates originating from the USA were found to carry carbapenemases, with the highest incidence recorded in South and Central America. Globally, bla_{KPC-2} emerged as the most prevalent carbapenemase, closely followed by bla_{VIM-2}, bla_{IMP-1}, bla_{NDM-1}, and bla_{GES-5}. However, it is worth noting that different regions exhibited distinctive predominant carbapenemases.[4]

In 2022, a study conducted in South Korea involving patients with urinary tract infections caused by *P. aeruginosa* demonstrated the presence of high-risk clones, specifically ST235 with bla_{IMP-6} and ST773 with bla_{NDM-1}. The study revealed a carbapenem resistance rate of approximately 46.7% among the isolated *P. aeruginosa* strains, exhibiting a multidrug-resistant (MDR) phenotype, and most CPPA isolates carried carbapenemase genes, such as bla_{IMP-6} or bla_{NDM-1}.[5] Another report in the same year using 807 CRPA isolates collected from 17 centers across 12 countries between 2019 and 2021, demonstrated approximately 33% exhibited phenotypic positivity for carbapenemases when tested using the modified Carbapenem Inactivation Method (mCIM). Among these phenotypically positive cases, 86% were genotypically positive, with VIM being the most prevalent carbapenemase, followed by GES. While carbapenemase producers were expected in Middle Eastern centers based on previous epidemiological data, a surprisingly high prevalence and diversity were also observed in European, South American, and African centers. In contrast, the CPPA in the United States was identified in between 3 to 30% of CRPA cases.[1] These contemporary global findings underscore the significance of carbapenemase testing in CRPA cases. The fundamental question arises regarding whether laboratories should perform phenotypic or genotypic testing for carbapenemase production, utilizing PCR to identify specific carbapenem-resistance genes in CRPA isolates. Such comprehensive testing is pivotal in assisting antimicrobial stewardship programs in making informed decisions regarding the selection of appropriate therapies for pseudomonal infections.[1]

Phenotypic methods have been documented for detecting carbapenemase producers within the *Enterobacteriaceae*, with some adaptable for application to non-fermentative bacteria. The methods endorsed by the Clinical and Laboratory Standards Institute (CLSI) for clinical microbiology laboratories to detect carbapenemase-producing *Enterobactarales* and *P. aeruginosa* encompass the Carba NP test and the modified Carbapenem Inactivation Method (mCIM).[5] The Carba NP test stands out for its simplicity and suitability for routine use in clinical laboratories. On the other hand, the mCIM has exhibited notable sensitivity and specificity in numerous recent studies.[6] However, each test has shown to have limitations against certain carbapenemase genes. For instance, the carbaNP test failed to detect $bla_{OXA-181}$ and $bla_{OXA-232}$, which are commonly found in Southeast Asian countries,[7] while the mCIM has been reported to have a limitation on particular IMP and GES genes.[8,9]

In addition, there is limited information regarding the distribution of carbapenemase genes among *P. aeruginosa* in Thailand; therefore, our study aimed to assess the effectiveness of the mCIM for detecting CPPA.

Materials and methods

Bacterial isolates

A total of 164 isolates of CRPA from non-duplicated patients admitted to Ramathibodi Hospital, a 1000-bed tertiary care hospital in Bangkok, Thailand, between 2019 and 2022, were used. The isolates exhibited the least intermediate or resistance to either imipenem, doripenem, or meropenem according to the clinical breakpoint recommended by the Clinical Laboratory Sciences Institute (CLSI).[10]

Modified carbapenem inactivation method (mCIM) testing

All isolates were tested for carbapenemase using the mCIM test per the CLSI guidelines. Briefly, a 10-µL loop full of CRPA was suspended in 2 mL of brain heart infusion broth (BHI). A 10 µg meropenem disc (Becton, Dickinson, USA) was added to the suspension and incubated at 37 °C for 4 hours. Following incubation, the disc was removed and placed on Muller-Hinton agar previously spread with *Escherichia coli* ATCC 25922 adjusted to a 0.5 McFarland standard. The plates were incubated at 37 °C for 24 hours. The inhibition zones were measured for interpretation according to CLSI guidelines.[10] *Klebsiella pneumoniae* ATCC BAA-1705 (bla_{KPC}) and PS249 *P. aeruginosa* (bla_{GES}, bla_{IMP}, bla_{VIM}, and bla_{NDM}) were used as the positive controls, while *P. aeruginosa* ATCC 27853 and *K. pneumoniae* ATCC BAA-1706 strains were negative controls.

Molecular detection of carbapenemase genes

For molecular testing, a direct flow kit (Master Diagnostica, Spain) was used to detect carbapenemase genes. The principle of the test kit is based on the targeted DNA amplification by multiplex PCR and the hybridization of the PCR products by specific probes coated on the membrane. Fifteen carbapenemase genes, including bla_{KPC}, bla_{NDM}, bla_{IPM}, bla_{IMI}, bla_{VIM}, bla_{GES}, bla_{GIM}, bla_{SME}, bla_{SIM}, $bla_{NMC/IMI}$, bla_{SPM}, $bla_{OXA-48-like}$, $bla_{OXA-23-like}$, $bla_{OXA-24-like}$, $bla_{OXA-51-like}$, and $bla_{OXA-58-like}$, were tested. Briefly, parts from a few colonies were suspended in 500 µL DNase/RNase-free double-distilled water. The suspension was vortexed until homogeneity was obtained. Thirty microliters of the suspension were added to the PCR tube containing the lyophilized master mix (Master Diagnostica, Spain). The PCR tubes were placed in a thermocycler, and PCR was run under the following conditions: one cycle of 10 minutes at 25 °C, then 1 cycle of 3 minutes at 95 °C, then 40 cycles of denaturation at 95 °C for 15 seconds, annealing at 55 °C for 45 seconds, and extension at 75°C for 1 minute. The PCR amplicons were further hybridized with specific probes. The amplicons were heated at 95 °C for 8-10 minutes and rapidly cooled in ice for at least 2 minutes. The hybridization process was automatically performed by hybriSpot in HS24 platform (Master Diagnostica, Spain). The hybridization result was analyzed by the Hybrispot software.

Conventional PCR and sequencing of bla_{VIM}, bla_{IMP}, and bla_{GES} genes

Regarding some limitation of the AMR flow kit on detecting certain types of bla_{VIM} and bla_{IPM}, the additional PCR determinations of these two genes were tested in all isolates. In addition, those isolates that revealed a negative result of mCIM with positive of bla_{GES} underwent PCR and sequencing. The primers of bla_{VIM}, bla_{IPM} and bla_{GES} is shown in Table 1.

Table 1. Primers of bla_{VIM}, bla_{IMP}, and bla_{GES}.[11,12]

Gene	*Sequence (5' to 3')*	*Size (bp)*
VIM-F	TGTCCGTGATGGTGATGAGT	456
VIM-R	GTGCTTCCGGGTAGTGTTGT	
IMP-F	GGAATAGAGTGGCTTAAYTCTC	382
IMP-R	GGTTTAAYAAAACAACCACC	
GES-F	TCACTCTGCATATGCGTCGG	692
GES-R	ACTTGACCGACAGAGGCAAC	

Thermal cycling was used under the following conditions: initial denaturation at 94 °C for 10 minutes, 36 cycles of denaturation at 94 °C for 30 seconds, annealing at 52 °C for 40 seconds, extension at 72 °C for 50 seconds, and final extension at 72 °C for 5 minutes. The PCR amplicons were sequenced by bi-directional DNA sequencing (Celemic, The Republic of Korea). The *beta-lactamase* genes were blasted with the database on the NCBI website at http://www.ncbi.nlm.nih.gov/blast.cgi

Results

Distribution of carbapenemase genes in CPPA strains

Of 164 isolates of CRPA, 83 (50.61%) isolates carried at least one *carbapenemase* gene. Five genes, including bla_{VIM}, $bla_{IMP\ (3/15/19\ group)}$, bla_{NDM}, bla_{GES} and $bla_{SIM,}$ were detected in 75 isolates by the direct flow kit. The conventional PCR detected another 8 isolates containing bla_{IPM} other than the 3/15/19 group. The bla_{IMP} and bla_{VIM} were the most common genes found

in 32.53% and 26.51% of the isolates, respectively. Among these, 67/83 (80.72%) contained a single *carbapenemase* gene, while approximately 20% carried more than one resistant gene. Most of these isolates (10/16) had a combination of bla_{NDM} and $bla_{IMP\ (3/15/19\ group)}$. Interestingly, one isolate contained at least four *carbapenemase* genes (Table 2).

Table 2. Distribution of *carbapenemase* genes of CPPA detected by AMR direct flow chip and conventional PCR (N=83).

***Carbapenemase* gene**	**Number of isolates (%)**
bla_{VIM}	22 (26.51)
$bla_{IMP\ (3/15/19\ group)}$	19 (22.89)
bla_{NDM}	12 (14.46)
bla_{NDM} + $bla_{IMP\ (3/15/19\ group)}$	10 (12.05)
$bla^{*}_{IMP\ (not\ 3/15/19\ group)}$	8 (9.64)
bla_{GES}	5 (6.02)
bla_{VIM} + $bla_{IMP\ (3/15/19\ group)}$	2 (2.41)
bla_{VIM} + bla_{NDM}	2 (2.41)
bla_{SIM}	1 (1.20)
bla_{VIM} + bla_{GES}	1 (1.20)
bla_{VIM} + bla_{GES} + bla_{NDM} + $bla_{IMP\ (3/15/19\ group)}$	1 (1.20)

Note: *$bla_{IMP\ (not\ 3/15/19\ group)}$ was detected by conventional PCR.

Detection of CPPA by mCIM

To establish the efficacy of mCIM for detecting CPPA, the results of mCIM were compared with molecular testing. Of the 164 isolates, sixty-eight revealed positive results with mCIM, and ninety-six were negative. In these 96 isolates of mCIM-negative, fifteen isolates provided positive results with the *carbapenemase* genes (Table 3). The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the mCIM were 81.93, 100, 100, and 84.38%, respectively. Among these isolates, 6 (40%), 5 (33.33%), and 3 (20%) carried bla_{VIM}, bla_{GES} and bla_{IMP}, respectively, and a single isolate containing both bla_{VIM} and bla_{IPM} (Table 4).

Table 3. The efficacy of mCIM for detecting carbapenemase-producing *P. aeruginosa* (CPPA) and non-carbapenemase-producing *P. aeruginosa* (non-CPPA) against *carbapenemase* genes detection.

	CPPA* (N=83)	**Non-CPPA**** (N=81)
mCIM positive	68	0
mCIM negative	15	81

Note: **carbapenemase* genes detected, ***carbapenemase* genes not detected.

Table 4. Efficacy of mCIM versus various *carbapenemase* genes.

***Carbapenemase* genes**	**Number of isolates** (N=83)	
	mCIM positive (N=68)	**mCIM negative** (N=15)
bla_{VIM}	16	6
$bla_{IMP\ (3/15/19\ group)}$	16	3
bla_{NDM}	12	0
bla_{NDM} + $bla_{IMP\ (3/15/19\ group)}$	10	0
bla_{GES}	0	5
$bla^{*}_{IMP\ (not\ 3/15/19\ group)}$	8	0
bla_{VIM} + $bla_{IMP\ (3/15/19\ group)}$	1	1
bla_{VIM} + bla_{NDM}	2	0
bla_{SIM}	1	0
bla_{VIM} + bla_{GES}	1	0
bla_{VIM} + bla_{GES} + bla_{NDM} + $bla_{IMP\ (3/15/19\ group)}$	1	0

Note: *$bla_{IMP\ (not\ 3/15/19\ group)}$ was detected by conventional PCR.

Of the fifteen mCIM-negative isolates, bla_{VIM-2} was demonstrated in six isolates, with one isolate containing bla_{VIM-2} and bla_{IPM-62}. Most of these isolates revealed a high level of resistance to all three carbapenems (Table 5). A single isolate had $bla_{VIM-63.}$ However, it is worth noting that not all bla_{VIM-2} and bla_{VIM-63} revealed negative results by mCIM. Approximately 40 and 75% of isolates carrying bla_{VIM-2} and bla_{VIM-63}, respectively, were detected by mCIM (data not shown). The only *imipenemase* gene (IMP) type related to a negative mCIM result was bla_{IMP-62}. All other bla_{IMP}-containing isolates were mCIM positive. Interestingly, all bla_{GES-5} isolates were unable to be detected by mCIM, despite their high-level resistance in 4/5 isolates.

Table 5. Distribution of *carbapenemase* genes in mCIM-negative carbapenemase-producing (CP) strains of *P. aeruginosa* (N=15).

***Carbapenemase* genes**	**Number of isolates**	**MIC of each isolate**		
		Imipenem	Meropenem	Doripenem
bla_{VIM-2}	5	≥32	≥32	≥32
		≥32	≥32	≥32
		≥32	≥32	16
		≥32	16	16
		4	8	1
bla_{VIM-63}	1	≥32	≥32	≥32
bla_{IMP-62}	3	≥32	≥32	≥32
		≥32	≥32	16
		16	4	4
bla_{IMP-62} + bla_{VIM-2}	1	8	16	4
bla_{GES-5}	5	≥32	≥32	≥32
		≥32	≥32	≥32
		≥32	≥32	≥32
		≥32	≥32	≥32
		8	2	1

Discussion

Carbapenem-resistant *Pseudomonas aeruginosa* (CRPA) has become a global health concern because of its association with not only limited therapeutic options and high mortality rates, but also plays a major role in healthcare-associated infection. Unlike the carbapenem-resistant *Enterobacterales* (CRE), for which most of the resistant mechanisms are based on the production of carbapenemase enzymes, the data of our study demonstrate that approximately equal percentages of carbapenemase and non-carbapenemase-producing distributed in the area. However, it is crucial to identify the isolates that contain the *carbapenemase* genes since the genes can transfer among *P. aeruginosa* or across species.[13]

In addition, *Enterobacterales,* such as *Escherichia coli* and *Klebsiella pneumoniae,* in which most of the *carbapenemase* genes belong to bla_{KPC}, bla_{NDM}, and bla_{OXA-48}, the most common carbapenem-resistant genes among *P. aeruginosa* found in this study were bla_{IMP}, bla_{VIM}, and bla_{NDM}, each of which was detected in approximately 23 to 38% of the isolates. Therefore, it reflected that nearly 90% of the CPPA carried metallo-beta-lactamase genes, leading to even smaller choices of antibiotics that effectively treat CPPA in the region.

The modified carbapenem inactivation method (mCIM) has been recognized for accurately detecting carbapenemases. It is recommended as a standard test by the CLSI for *Enterobacterales* and *P. aeruginosa*.[10] In carbapenemase-producing *Enterobacterales* (CPE), the test is excellent against the organisms isolated in Thailand and Southeast Asia compared with other standard tests, e.g., CarbaNP.[7] The efficacy of each test varied between different parts of the world and is strongly related to the distribution of bacterial strains in each area.[7] However, the efficacy of mCIM against *P. aeruginosa* strains isolated in the region has not been widely reported. The results from our study revealed that mCIM reached 100% specificity, but the sensitivity was less than 90%. The low sensitivity was due to the existence of specific genes, including bla_{VIM-2}, bla_{VIM-63}, bla_{IPM-62} and bla_{GES-5}. The number of these strains was over 15% of the CPPA. Therefore, the efficacy of mCIM for detecting carbapenemase in *P. aeruginosa* relied on the percentage of certain resistant genes in the region. Nevertheless, it remains unclear regarding the factors that cause false negatives by mCIM, despite the high level of carbapenem resistance in most of these isolates. This may be due to the low expression or low activity of the enzymes in certain strains. In 2020, Christian M. Gill et al. reported that approximately 50% of Guiana extended spectrum (GES) enzymes were detected as positive using the standard mCIM. These genes were identified as bla_{GES-1}, bla_{GES-19}, and bla_{GES-20}.[14] They hypothesized that the lower catalytic activity of GES might be a contributing factor. To address this, the investigators increased the bacterial inoculum, speculating that it could enhance the mCIM test's sensitivity. Subsequent retesting with a higher inoculum showed improved mCIM performance. A similar outcome occurred in the prolonged incubation of the meropenem disc for eight hours. Lisboa and colleagues encountered a similar challenge, where the mCIM could not detect 4 out of 16 *P. aeruginosa* isolates carrying the bla_{GES-5} gene.[8] In our study, although only 5 isolates were found containing bla_{GES-5}, all were negative by the mCIM.

Limitations

Although the data demonstrated the efficacy of the mCIM against CPPA isolates, the number of certain uncommon *carbapenemase* genes was limited. Therefore, the expansion of bacterial isolates, particularly those uncommon genes, would provide a clearer understanding.

Conclusion

In conclusion, three major findings were revealed from this study. Firstly, the resistance mechanism of carbapenem-resistant *P. aeruginosa* (CRPA) was approximately equal between carbapenemase-producing and non-carbapenemase-producing mechanisms. Secondly, the majority of carbapenemase enzymes were metallo-beta-lactamase, which were encoded by bla_{VIM}, $bla_{IPM,}$ and bla_{NDM} genes. Finally, the mCIM was shown to be an effective test for detecting CPPA. However, there were limitations in detecting certain genes; therefore, the information on the distribution of *carbapenemase* genes in the area would be necessary. Although this study was based on *P. aeruginosa* isolated mainly in Bangkok, Thailand, it provided solid evidence regarding the distribution of *carbapenemase* genes among *P. aeruginosa* as well as the efficacy of the mCIM for detecting CPPA. Nevertheless, further study using isolates from various regions in the country would demonstrate a clearer picture of the antibiotic resistance of this medically important organism.

Ethical approval

This study was conducted with the approval of the Institutional Review Board at Naresuan University, Thailand, under ethical approval number P1-0040/2566.

Informed consent statement

None

Funding

A partial scholarship from the Faculty of Allied Health Sciences, Naresuan University.

Conflicts of interest

The authors declare no conflicts of interest.

Data availability statement

The data presented in this study are available on request directly to the corresponding author. Due to patient privacy concerns, the data are not publicly available.

CRediT authorship contribution statement

Jakkrit Khampakul: conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing: original draft preparation; **Noppadon Jumroon**: conceptualization, methodology, investigation, writing: review and editing, supervision; **Narong Chaihongsa**: software; **Pitak Santanirand**: conceptualization, methodology, formal analysis, investigation, resources, writing: original draft preparation, review and editing, visualization, supervision, project administration, funding acquisition. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

We would like to thank Sudaluk Tienkeaw for bacterial strains collection and storage.

References

[1] Tenover FC, Nicolau DP, Gill CM. Carbapenemase-producing *Pseudomonas aeruginosa*-an emerging challenge. Emerg Microbes Infect. 2022; 11 (1): 811-4. doi: 10.1080/22221751.2022.2048972.

[2] Yoon E-J, Jeong SH. Mobile carbapenemase genes in *Pseudomonas aeruginosa*. Front Microbiol. 2021; 12: 614058. doi: 10.3389/fmicb.2021.614058.

[3] Escandón-Vargas K, Reyes S, Gutiérrez S, Villegas MV. The epidemiology of carbapenemases in Latin America and the Caribbean. Exp Rev Anti-infect Ther. 2017; 15 (3): 277-97. doi: 10.1080/14787210.2017.1268918.

[4] Reyes J, Komarow L, Chen L, Ge L, Hanson BM, Cober E, et al. Global epidemiology and clinical outcomes of carbapenem-resistant *Pseudomonas aeruginosa* and associated carbapenemases (POP): a prospective cohort study. Lancet Microbe. 2023; 4 (3): e159-e70. doi: 10.1016/S2666-5247(22)00307-2.

[5] Choi YJ, Kim YA, Kim J, Jeong SH, Shin JH, Shin KS, et al. Emergence of NDM-1–producing *Pseudomonas aeruginosa* sequence type 773 clone: shift of carbapenemase molecular epidemiology and spread of 16S rRNA methylase genes in Korea. Ann Lab Med. 2023; 43(2): 196-9. doi: 10.3343/alm.2023.43.2.196.

[6] Butler-Wu SM, Abbott AN. Is this the carbapenemase test we've been waiting for? A multicenter evaluation of the modified carbapenem inactivation method. J Clin Microbiol. 2017; 55 (8): 2309-12. doi: 10.1128/ JCM.00321-17.

[7] Laolerd W, Akeda Y, Preeyanon L, Ratthawongjirakul P, Santanirand P. Carbapenemase-producing carbapenem-resistant *Enterobacteriaceae* from Bangkok, Thailand, and their detection by the Carba NP and modified carbapenem inactivation method tests. Microb Drug Resist. 2018; 24 (7): 1006-11. doi: 10.1089/mdr.2017.0363.

[8] Lisboa LF, Turnbull L, Boyd DA, Mulvey MR, Dingle TC. Evaluation of a modified carbapenem inactivation method for detection of carbapenemases in *Pseudomonas aeruginosa*. J Clin Microbiol. 2018; 56 (1): e01234-17. doi: 10.1128/JCM.01234-17.

[9] Tsai Y-M, Wang S, Chiu H-C, Kao C-Y, Wen L-L. Combination of modified carbapenem inactivation method (mCIM) and EDTA-CIM (eCIM) for phenotypic detection of carbapenemase-producing *Enterobacteriaceae*. BMC Microbiol. 2020; 20 (1): 315. doi: 10.1186/s12866-020-01978-7.

[10] Clinical and Laboratory Standards Institute. Performance standards for antimicrobial susceptibility testing. 35th Ed. Wayne (PA): Clinical and Laboratory Standards Institute; 2025. CLSI supplement M100.

[11] Poirel L, Walsh TR, Cuvillier V, Nordmann P. Multiplex PCR for detection of acquired carbapenemase genes. Diag Microbiol Infect Dis. 2011; 70 (1): 119-23. doi: 10.1016/j.diagmicrobio.2010.12.002.

[12] Murugan N, Malathi J, Therese KL, Madhavan HN. Application of six multiplex PCR's among 200 clinical isolates of *Pseudomonas aeruginosa* for the detection of 20 drug resistance encoding genes. Kaohsiung J Med Sci. 2018; 34 (2): 79-88. doi:10.1016/j.kjms.2017.09.008.

[13] Yoon EJ, Jeong SH. Mobile carbapenemase genes in *Pseudomonas aeruginosa*. Front Microbiol. 2021; 12: 614058. doi: 10.3389/fmicb.2021.614058.

[14] Gill CM, Lasko MJ, Asempa TE, Nicolau DP. Evaluation of the EDTA-modified carbapenem inactivation method for detecting metallo-β-lactamase-producing *Pseudomonas aeruginosa*. J Clin Microbiol. 2020; 58 (6): e02015-19. doi: 10.1128/JCM.02015-19.

Journal of Associated Medical Sciences 2026; 59(2): 428-437

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Prevalence and characterization of β-thalassemia minor in a premarital screening at a hereditary blood diseases center in Babylon, Iraq

Shekair WH[1], Al-Maaroof ZW[2], Aljawazery SAA[3], Haddawi KH[4*], Al-Khyate SAM[5]

[1]Department of Pediatrician, College of Medicine Al-Mustaqbal University, Babylon, Hillah, Iraq.
[2]Department of Pathology, College of Medicine, University of Babylon, Babylon, Iraq.
[3]Department of Laboratory, Al Hilla Teaching General Hospital, Iraq.
[4]Department of Clinical Biochemistry, College of Medicine, University of Al-Qadisiyah, Iraq.
[5]Al Zahra'a Hospital / Wasit Health Directorate, Iraq.

ARTICLE INFO

Article history:
Received 3 November 2025
Accepted as revised 16 April 2026
Available online 24 April 2026

Keywords:
β-thalassemia, premarital screening, HbA2, iron deficiency anemia, genetic counseling.

ABSTRACT

Background: β-thalassemia is a significant hereditary hematological disorder that is highly prevalent in Iraq. Premarital screening is a critical preventive strategy for identifying couples at risk of having children with thalassemia major. This study evaluated the effectiveness of a premarital screening program in Babylon, Iraq.

Objectives: To determine the prevalence of β -thalassemia minor among couples with abnormal red cell indices and characterize the associated hematological parameters.

Materials and methods: A cross-sectional study was conducted from 2016 to 2019 on 960 individuals (480 couples) referred to the Hereditary Blood Diseases Center. Participants with a mean corpuscular volume (MCV) <80 fL and/or mean corpuscular hemoglobin (MCH) <27 pg underwent confirmatory testing via high-performance liquid chromatography (HPLC) for HbA2 quantification and serum ferritin analysis.

Results: The prevalence of β-thalassemia minor was 20.1% (193/960), with a significant sex disparity (males: 27.7%, females: 12.5%). Iron deficiency anemia (IDA) was identified in 46.6% of participants. The initial screening step using MCV/MCH cut-offs showed 100% sensitivity but 53.4% specificity, primarily because of IDA. Fifteen high-risk couples (1.7%) in which both partners were carriers were identified.

Conclusion: The two-stage premarital screening protocol demonstrated high sensitivity for detecting β-thalassemia carriers. The high prevalence of both the thalassemia trait and IDA in this young adult population underscores a substantial public health burden. The program successfully identified at-risk couples, enabling targeted genetic counseling and prenatal diagnosis to reduce the incidence of thalassemia major.

Introduction

β-thalassemia represents a significant public health challenge with widespread distribution across Africa, the Mediterranean region, the Middle East, including Iraq, the Indian peninsula, and Southeast Asia.[1] Globally, approximately 3% of the world population carries genes for β-thalassemia, with the highest prevalence rates documented in Cyprus, Italy (particularly Sardinia), and South Asia.[2,3] In Arab countries,

*** Corresponding contributor.**
Author's Address: *Department of Clinical Biochemistry, College of Medicine, University of Al-Qadisiyah, Iraq.*
E-mail address: krarr.haidar@yahoo.com
doi: 10.66285/JAMS.2026.072
E-ISSN: 2539-6056

β-thalassemia minor occurs at varying frequencies with carrier rates ranging from 1% to 11%, with particularly elevated frequencies observed in Jordan, Lebanon, Iraq, Egypt, Palestine, and other regional nations.[4,5]

In Iraq, the frequency of carriers for β-thalassemia is more evenly distributed throughout the country at rates that vary between 3.0% and 6.5%, about 15,000 thalassemia major or intermedia patients are registered with hemoglobinopathy.[4-6] T The Hereditary Blood Diseases Center of Babylon Teaching Hospital for Women and Children (BTHWC) is providing care for around 1,000 homozygous β-thalassemic patients. This is a huge patient load and emphasizes the need for large pragmatic intervention studies to be done in this area.[7]

The pathophysiology of thalassemia is characterized by deficient synthesis of either alpha or β hemoglobin chains, which manifests as alpha α- or β-thalassemia. Whilst alpha thalassemia is less common, β-thalassemia has a high prevalence in certain parts of the world. β-thalassemia manifested clinically in a continuum from thalassemia trait (thalassemia minor) to thalassemia major.[8,9] Thalassemia major patients (HbA2 2%, HbF 98%, HbA nil) have severe clinical manifestations and require regular blood transfusions, whereas thalassemia traits (HbA2>3.5%, HbF nil) are usually clinically unaffected and diagnosed on routine hematological examination.[10,11]

Multiple studies conducted throughout Iraq have demonstrated considerable regional variation in thalassemia prevalence across different provinces.[4-7] This common, incurable, autosomal recessive hemoglobinopathy causes substantial morbidity and mortality while imposing a considerable financial burden on society.[12] Importantly, β-thalassemia represents a preventable disease through the implementation of comprehensive screening programs.

Premarital screening programs have gained widespread acceptance since the early 1970s, with numerous countries implementing mandatory screening protocols.[13] In Iraq, the thalassemia prevention program was initiated in 2015 in the Babylon governorate.[7] This prevention strategy operates through the premarital screening of β-thalassemia carrier couples to encourage participation in genetic counseling and prenatal diagnosis.[4,5] Baghdad, a total of 12526 individuals (6,263 couples)

The screening protocol employs red blood cell indices, particularly mean corpuscular volume (MCV) and mean corpuscular hemoglobin (MCH), as initial population screening tools because of their suitability for large-scale epidemiological screening and incorporation into routine premarital examinations.[9] When both partners demonstrated an MCV<80 fL or MCH<27 pg (established cut-off values), hemoglobin A2 concentrations were measured using high-performance liquid chromatography.[10] Individuals with confirmed β-thalassemia minor characteristics (HbA2>3.5%) received genetic counseling referrals.[11] Partners with red cell indices below cut-off values but normal HbA2 concentrations are classified as suspected cases, requiring additional evaluation, including iron therapy for those with low serum ferritin levels, followed by repeat testing with or without molecular studies.[11,14]

Different cutoffs of MCH and MCV have been reported in hematologic references, probably due to population differences, including age and race.[15] Presently, MCV of 80 fL and MCH of 27 pg are used as cut-off values for the best sensitivity and specificity. The goal of premarital screening is to detect carriers of hemoglobin disorders to prevent the birth of a child with a severe form of the disease.

This research aims to investigate the prevalence of β-thalassemia minor in accordance with criteria by comprehensive assessment, including RBC indices, HPLC, and iron studies among 480 couples (male and female) from around primary health care centers in Babylon governorate, who presented for marriage at Hereditary Blood Diseases Center/ BTHWC.

The study was restricted to individuals aged 17-35 years because this age range represents the typical premarital age group in the Babylon governorate. In Iraq, the legal minimum age for marriage is 18 years (with judicial exceptions from age 15-17), and most individuals presenting for premarital screening fall within this range. Including older individuals would introduce confounding factors such as chronic disease-related anemia, age-related changes in hematological parameters, and reduced relevance to the premarital screening population. Furthermore, this age restriction ensures homogeneity of the study population and enhances the applicability of findings to the target demographic of premarital screening programs.

Materials and methods

Study design and participants

A cross-sectional study was conducted at the Hereditary Blood Diseases Center in BTHWC/Iraq. Recruitment period: 48 months (from January 1, 2016, to December 31, 2019). This study included a total of 960 (480 couples) persons. All were referred to the center from different premarital examination centers because of initial detection of blood indices abnormalities (Low MCV or Low MCH).

Inclusion and exclusion criteria

The inclusion criteria were as follows: 1) individuals of both sexes planning marriage, 2) age between 17 and 35 years, which represents the typical premarital age range in the Babylon governorate and ensures population homogeneity for screening purposes, and 3) initial screening results showing mean MCV<80 fL and/or MCH<27 pg. Individuals with a known history of hematological malignancies, chronic inflammatory diseases, liver or kidney disorders, or those who had

received a blood transfusion within the last three months were excluded from the study.

Sample size justification

The sample size was a convenience sample, encompassing all eligible couples referred to the center during the specified 48-month study period who met the inclusion criteria. Formal power calculations were not performed a priori.

Biological sample collection and processing

Trained phlebotomists collected peripheral venous blood samples (approximately 5 mL) from each participant. Collection was performed without fasting, primarily during the morning hours (8:00-12:00 pm). Blood was drawn into two vacuum tubes: A K_2EDTA tube for complete blood count (CBC) and hemoglobin electrophoresis, and a plain serum separator tube (SST™, BD Vacutainer®) for ferritin analysis.

The EDTA tube was gently inverted 8-10 times immediately after the collection. The SST tube was allowed to clot at room temperature for 30 minutes. Both tubes were transported to the laboratory at room temperature and processed within 2 hours of venipuncture. The SST tube was centrifuged at 3000 rpm for 10 min to separate serum. Serum aliquots for ferritin analysis were stored at -20 °C and analyzed in batches within one week of collection; no freeze-thaw cycles were permitted before analysis.

Laboratory assays

Complete blood count (CBC) and red cell indices were measured using an automated hematology analyzer (Mindray BC5000™, China). Hemoglobin electrophoresis quantification of HbA2 was performed using high-performance liquid chromatography (HPLC) on the VARIANT™ II Hemoglobin Testing System (Bio-Rad Laboratories, Hercules, CA, USA). Serum Ferritin was quantified by a chemiluminescent immunoassay (CLIA) on an Architect i2000SR immunoassay analyzer.

Calibration and quality control

All laboratory assays were performed in accordance with the manufacturer's protocol. Automated analyzers (Coulter LH 780, Architect i2000SR, and VARIANT II) were calibrated using their respective manufacturer-provided calibration materials before use and as required by the internal quality control protocols.

Internal quality control (IQC) was performed daily using commercial control materials of three levels (low, normal, and high) provided by the manufacturers (Bio-Rad Hemoglobin Control, Ref: 195-2051; Abbott Architect Ferritin Controls, Ref: 7K66-10). External quality assurance (EQA) was conducted through the Iraqi National Quality Control Program.

The inter-assay coefficient of variation (CV) for ferritin was <5%, and that for HbA2 quantification was <3%. All the assays were performed in a single run for each sample.

Ethical considerations

The study protocol was reviewed and approved by the Scientific and Ethical Committee of the University of Al-Mustaqbal, College of Medicine (Approval No. UM/CoM/Ethics/2015-038). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments. Written informed consent was obtained from all participants included in the study after explaining the nature and purpose of the study.

Statistical analysis

Data were collected and organized with Microsoft Excel and then analyzed by GraphPad Prism Software version 9.0. Study variables were summarized using descriptive statistics. For categorical data (the proportions of β-thalassemia traits, iron-deficiency anemia, and demographic characteristics), number (%) was used. Quantitative variables (hematological indices such as Hb, MCV, MCH, HbA2, and ferritin) were reported by means±SD. For inferential statistics, an independent sample t-test was used to determine the mean difference between continuous variables (e.g., hematological parameters) for two independent groups (for example, males vs females within β-thalassemia carrier analysis). The $p<0.05$ was regarded as statistically significant in all tests. Sensitivity and specificity were calculated to determine the diagnostic performance of the first screening phase (MCV and MCH cutoff values).

Result

The study was cross-sectional and included (480 males, 480 females) premarital individuals on a visit to the center due to abnormal red cell indices. From laboratory results, the study population was divided into three diagnostic groups: 447 (46.6%) with iron deficiency anemia (IDA), 193 (20.1%) as β-thalassemia carriers, and 320 (33.3%) with normal hematological parameters. In comparison to males (N=167), a higher prevalence of IDA was evident among females (N=280), and β-thalassemia trait was more common in males than in females (male: N=133; female: N=60). The hematological evaluation referred to normal reference values, and the anemia was defined as Hb<14 gm/dL in males and <12 gm/dL in females. Microcytosis was MCV<80 fL, and hypochromia was an MCH<27 pg for both genders. HbA2 level greater than 3.5% was considered the diagnostic cutoff for β-thalassemia minor, and ferritin less than 20 ng/mL as cut-off of iron deficiency.

The largest share (N=458, 47.7%) of the 960 participants belonged to the age group between 21 and 25 years, as presented in Table 1, which corresponds with the common age for marriage in this population. There were 280 (29.2%) and 222 (23.1%) participants aged 17-20 and 26-35 years, respectively. Distribution

of all three diagnostic categories in normal individuals, β-thalassemia carriers, and IDA followed a similar trend among different age groups, with the greatest absolute values prevailing among 21-25-year-olds, as shown in Table 1.

Table 1. Age distribution of participants by diagnostic group.

Age group (years)	Normal individuals	β-Thalassemia carriers	Iron deficiency anemia	Total
17-20	85	45	150	280
21-25	150	98	210	458
26-35	85	50	87	222
Total	320	193	447	960

The comparison of hematological parameters among four diagnostic groups: healthy, β-thalassemia trait, combined β-thalassemia trait with IDA, and IDA is shown in Table 2. The β-Thalassemia carriers expressed typical microcytic hypochromic anemia in that the MCV and MCH values were significantly decreased and showed diagnostically high (HbA2>3.5%). On the other hand, patients with IDA also had microcytic hypochromic indices (low MCV and MCH) but were characterized by significantly low serum ferritin concentrations and normal or decreased HbA2 percentage. The combined β-thalassemia and IDA group (N=155) demonstrated features of both conditions, including elevated HbA2 levels alongside depleted iron stores. All the parameters were found within the normal reference range in the healthy control group and served as the baseline. Of particular interest, females in the β-thalassemia carrier group had lower mean values of hemoglobin, MCV, MCH, and ferritin than males.

Table 2. Hematological parameters by diagnostic group.

Parameter	Gender	Normal individuals (N=320)	β-Thalassemia carriers (N=193)	Combined β-Thal+IDA (N=155)	IDA (N=447)
Hb (gm/dL)	Male	15±1.5	10.5	9.8±1.0	10.5±1.5
	Female	12±1.5	8.5-10	8.5±0.9	9.0±1.5
MCV (fL)	Male	85±2	65±2	62±3	75±2
	Female	83±2	60±2	58±3	70±2
MCH (pg)	Male	30±2	<27	21±2	20±2
	Female	30±2	<27	20±1.5	18±1
HbA2 (%)	Male	2.5±0.5	4.5±0.5	3.8±0.4	1.8±0.5
	Female	3±0.2	5±0.2	3.6±0.3	1.6±0.5
Ferritin (ng/mL)	Male	100±15	60±10	12±4	10±5
	Female	60±5	18±3	8±3	5.5±4

Comparative gender-based analysis of hematological parameters between β-thalassemia minor (BTM) carriers (N=193) is depicted in Table 3, which shows a significant difference among genders, i.e., males (N=133) and females (N=60). Males had higher mean Hb (10.5±0.8 gm/dL vs 9.3±0.8 gm/dL), greater MCV (65±2 fL vs 60±2 fL), greater MCH (23±2 pg vs 22±1.5 pg) and much higher levels of ferritin (60±10 ng/mL versus 18±3 ng/mL); whereas females had a slightly increased HbA2 percentage (5.0±0.2% versus 4.5±0.5%). All intersex contrasts were statistically significant (p<0.001).

Table 3. Comparison of key parameters between genders in the β-thalassemia carrier group.

Parameter	Mean±SD		*p* value
	Males (N=133)	**Females** (N=60)	
Age (years)	24.5±3.2	22.8±3.5	-
Hb (gm/dL)	10.5±0.8	9.3±0.8	<0.001
MCV (fL)	65±2	60±2	<0.001
MCH (pg)	23±2	22±1.5	<0.001
HbA2 (%)	4.5±0.5	5.0±0.2	<0.001
Ferritin (ng/mL)	60±10	18±3	<0.001

Table 4 shows the severity of IDA in the cohort according to the distribution of ferritin among 447 confirmed patients: severe deficiency (ferritin <15 ng/mL) was strongly predominant, representing 79.4% (N=355) of IDAs. There was a significant gender difference, with 230 females and 125 males in this serious category. The remaining 20.6% (N=92) were diagnosed as being moderately deficient (ferritin 15 to <30 ng/mL). The classification of iron deficiency severity was based on the World Health Organization (WHO) guidelines on the use of ferritin concentrations to assess iron status, which define iron deficiency as serum ferritin <15 µg/L in adults, and iron depletion as ferritin 15-30 µg/L.[16] No subjects with ferritin ≥30 ng/mL were observed in the screened case series, as all individuals with ferritin ≥30 ng/mL were classified as iron-replete rather than iron-deficient.

Table 4. Distribution of iron deficiency anemia severity based on ferritin levels.

Severity category	Ferritin Level (ng/mL)	Males	Females	Total	%
Severe deficiency	<15	125	230	355	79.4
Moderate deficiency	15 - <30	42	50	92	20.6
Iron-replete	≥30	0	0	0	0
Total		167	280	447	100

Table 5 presents a risk assessment for married couples from premarital screening results, classifying 480 couples (960 individuals) into five-level groups in accordance with the Hb profiles. Fifteen (3.1%) were couples, both carriers of β-thalassemia minor, which established a 25% risk per pregnancy for the couple to have progeny with thalassemia major and had to go through compulsory genetic counseling and prenatal detection. Among the rest of the couples, 200 had only one carrier partner, 223 couples had iron deficiency anemia, and 155 had a combined phenotype of carriers and abnormal respondents (each required special action, such as counseling and iron therapy).

Consanguinity data were collected for all 480 couples. Of these, 216 couples (45.0%) reported consanguineous marriages (defined as unions between second cousins or closer). Among the 15 high-risk couples (both carriers), 11 (73.3%) were consanguineous, which was significantly higher than the overall consanguinity rate in the study population (p<0.01). Among the 200 couples with one carrier partner, 88 (44.0%) were consanguineous. The high rate of consanguinity among at-risk couples is consistent with the known elevated prevalence of autosomal recessive disorders in consanguineous populations in Iraq.

Table 5. Risk assessment of married couples based on screening results.

Couple Type	Number of couples	Risk of having an affected child	Recommended action
Both normal	87	Very low	No action needed
One carrier, one normal	115	Carriers (no disease)	Genetic counseling
Both carriers	15	High (25% per pregnancy)	Mandatory counseling and PND offer
Both with iron deficiency	148	Low (but requires treatment)	Iron therapy
Carrier + iron deficiency	115	Variable	Iron therapy + HbA2 recheck + counseling
Total	480		

Note: *PND: prenatal diagnosis.*

The effectiveness of the two-step premarital screening program in detecting β-thalassemia minor is summarized in Table 6. Screening of 960 individuals using low MCV (<80 fL) and/or low MCH (<27 pg) as cut-off criteria indicated 100% sensitivity because all 193 carriers among the tested 960 individuals were identified. This step, however, had a low specificity (53.4%) due to the high number of false positives (N=447), predominantly caused by IDA. The second step test, including treatment with iron and the re-evaluation of indices to all who presented low indices at baseline was essential for refining the diagnosis, since it resolved microcytosis and hypochromia in a substantial number of cases (292 individuals) deeming 155 persons with persistently low indices as suspicious "true" carriers sent to have confirmatory Hb electrophoresis; thus improving program specificity.

Table 6. Effectiveness of primary screening (MCV/MCH) in detecting carriers.

Screening step	Number tested	Number referred	True carriers identified	False positives	Sensitivity (%)	Specificity (%)
Step 1: Initial screening (MCV<80 fL and/or MCH <27 pg)	960	640	193	447 (IDA cases)	100	53.4
Step 2: After iron therapy (re-check of indices)	447	155 (Remained low)	15 (Newly identified)	140	-	-
Overall program performance	960	795 (Total referrals)	193	-	100	70

Note: *False Positives were primarily due to iron deficiency.*

Discussion

The current study offers several informative and useful data concerning the preventive screening efficacy in relation to couples' premarital status, while investigating their implications within local and global prevention programs. The detected β-thalassemia trait prevalence of 20.1% in the studied individuals greatly surpasses reported rates among Iraqis (3-6.5%) and Middle Easterners at large (1-11%).[5] This relatively high detection rate is due to the focused screening strategy in which participants are selected on the basis of abnormal red blood cell indices, which will enrich the sample with possible carriers.[17]

The findings align with similar targeted screening approaches reported in other Middle Eastern populations.[18] A premarital screening study in Al Majma'ah, Saudi Arabia, identified the β-thalassemia trait in 3.4% of unselected couples,[19] while a study from Southwest Iran reported a 5.6% prevalence among general premarital populations.[20] All 581 couples live in Khuzestan Province, Southwest Iran. Individuals with mean corpuscular volume <80, mean corpuscular hemoglobin <27, hemoglobin A2 ≥3/5 were considered as β-thalassemia traits. Prevalence of minor β-thalassemia, α-thalassemia, SCA, iron deficiency anemia, and silent thalassemia were respectively identified in 995 (5.6% The marked difference between these general population studies and the current investigation underscores the effectiveness of the two-stage screening protocol employed in Babylon, where initial red cell index filtering significantly enhanced the detection efficiency.

Studies from other Iraqi provinces have demonstrated regional variations in the carrier rates. Hassan *et al*. reported a 3.5% prevalence of β-thalassemia trait in the Baghdad premarital screening program, Baghdad,[5] a total of 12,526 individuals (6,263 couples while Abbas *et al*. documented prevalence rates ranging from 3.0% to 6.5% across different Iraqi governorates.[4] In the current study, a higher detection rate likely reflects both the pre-screening strategy and potential geographic clustering of β-thalassemia carriers in the Babylon region, which is consistent with the established founder effect phenomena observed in isolated populations.

The observed gender disparity, with males exhibiting a higher carrier rate (27.7%) than females (12.5%), represents a notable finding that warrants careful interpretation. While β-thalassemia follows autosomal recessive inheritance patterns with an expected equal gender distribution, several factors may contribute to this apparent discrepancy. The significantly lower ferritin levels observed in female carriers (18±3 ng/mL) compared to males (60±10 ng/mL) suggest that concurrent iron deficiency may mask the characteristic hematological features of β-thalassemia in women.[21]

The significant difference in the prevalence of β-thalassemia between males and females observed in this study deserves further elaboration. Since β-thalassemia is an autosomal recessive disorder, the true carrier frequency should be approximately equal between sexes. The apparent male predominance (27.7% vs 12.5%) is most likely an artifact of the screening methodology rather than a true biological difference. Several mechanisms may account for this observation. Firstly, iron deficiency, which was significantly more prevalent and severe among females (82.1% of female IDA cases had ferritin <15 ng/mL), can suppress HbA2 levels below the diagnostic threshold of 3.5%, thereby masking the β-thalassemia trait in iron-deficient women.[22] Secondly, the overlapping hematological features of IDA and β-thalassemia trait (both presenting with microcytic hypochromic anemia) make it more difficult to distinguish between the two conditions in females who are more frequently iron-deficient. Thirdly, some female carriers with concurrent severe iron deficiency may have been misclassified as having IDA alone, as their HbA2 levels may have been falsely normalized. This diagnostic challenge has been reported in multiple studies from the Middle East and South Asia, where iron deficiency is endemic among women of reproductive age.[21,22] These findings emphasize the critical importance of iron repletion before definitive β-thalassemia screening in populations with high rates of iron deficiency, particularly among women.

This phenomenon has also been observed in other populations. Jameel *et al*. demonstrated that iron deficiency can suppress HbA2 levels below the diagnostic threshold of 3.5%, potentially leading to

under-detection of the β-thalassemia trait in iron-deficient individuals.[22] The high prevalence of severe iron deficiency among females in the current study (82.1% of female IDA cases had ferritin <15 ng/mL) supports this hypothesis and aligns with the regional patterns of iron deficiency in young women across the Middle East. [23]

The two-stage screening protocol demonstrated excellent sensitivity (100%) for detecting β-thalassemia carriers, confirming the effectiveness of MCV and MCH cut-off values as primary screening tools. These findings corroborate the results of large-scale studies conducted in other populations. Heng *et al*. analyzed 6,779 β-thalassemia carriers in Guangdong province, China, and reported that combined MCV (<80 fL) and MCH (<27 pg) screening reduced false-negatives to only 0.22%.[24] Similarly, Sahiratmadja *et al*. demonstrated that an MCV <80 fL and/or an MCH<27 pg provided optimal first-round screening sensitivity in Indonesian populations.[25]

The observed specificity of 53.4% in the initial screening step reflects the expected trade-off between sensitivity and specificity in population-screening programs. The high false-positive rate, primarily attributed to IDA (46.6% of screened individuals), aligns with established screening paradigms where initial high sensitivity is followed by confirmatory testing to improve specificity.[26] This approach has proven cost-effective in multiple healthcare settings, as demonstrated in prevention programs across Cyprus, Sardinia, and Greece.[27]

The finding that 46.6% of the screened individuals had IDA highlights a significant public health concern in the study population. This prevalence substantially exceeds the rates reported in other regional studies. Owaidah *et al*. documented an IDA prevalence of 10.7% among Saudi university students, while regional estimates for the Middle East indicated that IDA affects approximately 32.5% of pregnant women and 28% of adolescents.[28] The severity distribution observed in the current study, with 79.4% of IDA cases classified as severe deficiency (ferritin <15 ng/mL), indicates a critical level of iron depletion that requires immediate public health interventions.

The intersection of iron deficiency and β-thalassemia traits presents diagnostic challenges that are well documented in the literature. Sun et al. outlined key discriminating features, noting that β-thalassemia traits typically present with higher RBC counts (>5M/μL), lower MCV (<74 fL), and normal iron studies, whereas IDA is characterized by reduced RBC counts (<4M/μL) and depleted iron stores.[21] The current study protocol of iron therapy followed by repeat testing effectively addressed this diagnostic challenge, improving the overall program specificity from 53.4% to 70%.

The identification of 15 couples (3.1%) at high risk of having children with thalassemia major represents a significant achievement in disease prevention. Each at-risk couple has a 25% probability of having an affected child, making genetic counseling and prenatal diagnosis essential interventions. International experience has demonstrated the effectiveness of such prevention strategies. Cao *et al*. reported that comprehensive prevention programs in Sardinia achieved an 85% reduction in thalassemia major births from 1:250 to 1:1,660 live births.[27] Similar success has been documented in Cyprus, Iran, and Turkey, where mandatory premarital screening programs achieved substantial reductions in affected births despite limited success in preventing at-risk marriages.[29]

The cost-effectiveness of prevention programs has been well established. Kuwaiti studies have demonstrated that preventing one case of β-thalassemia major costs approximately 100 USD, substantially less than the annual treatment costs for affected patients.[30] Similar economic analyses from Greece and Cyprus indicate that prevention program costs equal those of treating one patient for one year, whereas the annual screening program expense equals one week of treatment costs for the entire affected population.[2,31,32]

The effectiveness of pre-marital screening programs relies on substantial cultural and community involvement. The Kurdistan, Iraq studies indicated significant challenges to program success, with only 10.5% of at-risk couples opting to cancel plans for marriage and 5.7% discussing termination in the case of affected fetuses.33 These results indicate the nuanced impact of medical advice and cultural values on the effects of the program.

Religious and cultural influences had a major effect on the acceptance of the program in the Middle East. A systematic review by Saffi and Howard of Middle Eastern programs found that, except for Iran, Turkey, and Iraq, most countries did not achieve 65% at-risk marriage cancellation rates; those three countries also experienced decreases in affected births through prenatal diagnosis and therapeutic abortion.[29] Identification of at-risk couples as performed in the current study forms a basis for selective culture-specific genetic counseling programs, which should be applicable and ethically valid.

The high rate of consanguinity observed in this study (45.0% overall, 73.3% among high-risk couples) is consistent with previously reported consanguinity rates in Iraq, which range from 24% to 71% across different regions.[34] Consanguineous marriages significantly increase the probability of offspring being homozygous for autosomal recessive alleles, including β-thalassemia mutations. The notably higher consanguinity rate among both-carrier couples (73.3%) compared to the overall study population (45.0%) further supports the well-established association between consanguinity and increased risk of autosomal recessive disorders. These findings underscore the importance of incorporating consanguinity assessment into premarital screening programs and providing targeted genetic counseling to

consanguineous couples, particularly in populations with high carrier rates for hemoglobinopathies.

Limitations and future directions. There are several limitations of this study to be accounted for when interpreting the results. The use of convenience sampling and a referral-based recruitment method may lead to selection bias that could impact generalizability to the wider Babylon population. Hence, lack of molecular characterization with many β-thalassemia mutations is a major drawback, since it is important to know the exact patterns and prevalence of the various mutations in order to make prevention programs and genetic counseling broad-based. Molecular analysis must be included in future studies to ascertain common mutations and customize preventive strategies.

Conclusion

This comprehensive premarital screening study demonstrated the effectiveness of systematic β-thalassemia trait detection in a high-prevalence population. The two-stage screening protocol achieved excellent sensitivity while managing concurrent iron-deficiency anemia. The findings underscore the critical importance of addressing iron deficiency as a public health priority while focusing on genetic disease prevention. Recognition of at-risk couples allows for the implementation of focused genetic counseling and prenatal diagnosis, as part of comprehensive measures aiming at the control of thalassemia in Iraq. The findings of the study would further justify and strengthen the continuation of premarital screening and counseling programs in Iraq, with a need for greater public awareness, better genetic counseling services, as well as strategies to alleviate the large burden of IDA in their target population. Further enhanced capacity in the program is to incorporate molecular diagnostic testing and thus provide a more in-depth level of genetic risk assessment for participating couples.

Ethical approval

The study protocol was reviewed and approved by the Scientific and Ethical Committee of the University of Al-Mustaqbal, College of Medicine (Approval No. UM/CoM/Ethics/2015-038). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments. Written informed consent was obtained from all participants included in the study after explaining the nature and purpose of the study.

Use of artificial intelligence

The artificial intelligence was applied to improve sentence structure and clarity.

Funding

No funds were received to fulfil this work.

Conflict of interest

The authors declare no competing interests.

CRediT authorship contribution statement

Widad Hamza Shekair: conceptualization, methodology, validation, formal analysis, resources, data curation, visualization, writing: original draft, review and edit, project administration, funding acquisition; **Zainab W. Al-Maaroof:** conceptualization, methodology, formal analysis, resources, visualization, resources, writing: review and edit; **Krarr Haider Haddawi:** investigation, writing: original draft, visualization, supervision; **Shaymaa Ali Abed Aljawazery:** formal analysis, investigation, resources, writing: original draft, review and edit, supervision; resources, writing: original draft preparation, visualization, supervision; **Suhad Aziz Mustafa Al-Khyate:** visualization.

Acknowledgements

The authors wish to thank the head and staff at the Hereditary Blood Diseases Center in BTHWC for their great efforts and contributions to this study. We would also like to acknowledge all technicians in the laboratory for their excellent work during sample processing and analysis. We would like to acknowledge all participants from the marital screening program, who cooperated in this study and made it possible.

References

[1] Abd Al-Abbass HA, Obaid AF, Makki Al-Hindy H, Abdulrasol ZA, Jasim Shlash AM. Epidemiological characteristics and disease complications in thalassemia syndrome patients in Babylon, Iraq. J Babol Univ Med Sci. 2024; 26(1): 1-10. https://www.researchgate.net/profile/Hayder-Makki-5/publication/380178836.

[2] Bellis G, Parant A. Beta-thalassemia in Mediterranean countries. Findings and outlook. Investig Geográficas. 2022; 77: 129-38. doi.org/10.14198/INGEO.19079.

[3] Kattamis A, Forni GL, Aydinok Y, Viprakasit V. Changing patterns in the epidemiology of β-thalassemia. Eur J Haematol. 2020;105(6): 692-703. doi.org/10.1111/ejh.13512.

[4] Abbas RA, Hassan RH, Taghlubee IM, Mohammed SI, Mohammed HH, Hasan HH, et al. Prevalence and molecular characterization of β-thalassemia in Kirkuk Province of Northern Iraq. Hemoglobin. 2024; 48(5): 308-3. doi.org/10.1080/03630269.2024.2418507.

[5] Hassan MK, Abbas RA, Hassan RA, Taghlubee IM, Abd Al Majeed SS, Khaleel GA, et al. Prevalence and spectrum of β-thalassemia mutations in Baghdad, Iraq: Data from the premarital screening program. Hemoglobin. 2025;49(1): 31-7. doi: 10.1080/03630269.2024.2446360.

[6] Al-Najafi WK. Premarital screening program for

hemoglobinopathies in Karbala, Iraq. Karbala J Med. 2020;13(1)): 2293-300. https://journals.uokerbala.edu.iq/index.php/kj/article/view/737.

[7] Hanoon BD, Khalf SA, Khalaf IA. Premarital screening program in Al-Nuaman Teaching Hospital. Iraqi J Med Sci. 2021; 19(2): 1-5. https://iraqijms.net/index.php/jms/article/view/788.

[8] Al-Maaroof ZW, Al Sharifi LM. Associations between serum ferritin levels, Age-stratified hematological parameters, and hemoglobin A2 levels in Iraqi Beta thalassemia carriers. J Appl Hematol. 2025; 16(2): 111-7. doi: 10.4103/joah.joah_49_25.

[9] Bordbar E, Taghipour M, Zucconi BE. Reliability of different RBC indices and formulas in discriminating between β-thalassemia minor and other microcytic hypochromic cases. Mediterr J Hematol Infect Dis. 2015; 7(1): e2015022. doi: 10.4084/MJHID.2015.022.

[10] Kablan NM, El Hamshary MO, Nasr GM, Hamza MT, Salam TH, Hemeida AA. Hemoglobin A2 cut off values in Egyptian cohort as a marker of β-thalassemia carriers. J Biosci Appl Res. 2015; 1(4): 154-9. https://jbaar.journals.ekb.eg/article_106021.

[11] Khera R, Singh T, Khuana N, Gupta N, Dubey AP. HPLC in characterization of hemoglobin profile in thalassemia syndromes and hemoglobinopathies: a clinicohematological correlation. Indian J Hematol Blood Transfus. 2015; 31(1): 110-5. doi: 10.1007/s12288-014-0409-x.

[12] Al-Maaroof ZW, Al-Taee ZM. Assessment of anemia prevalence in urban populations: A community-based study. J Appl Hematol. 2025; 16(1): 63-7. doi: 10.4103/joah.joah_128_24.

[13] Al-Allawi NA, Al-Dousky AA. Frequency of haemoglobinopathies at premarital health screening in Dohuk, Iraq: implications for a regional prevention programme. East Mediterr Heal J. 2010; 16(4): 1-5. https://pubmed.ncbi.nlm.nih.gov/20795420/.

[14] Al-Amodi AM, Ghanem NZ, Aldakeel SA, Ibrahim Al Asoom L, Rafique Ahmed N, Almandil NB, et al. Hemoglobin A2 (HbA2) has a measure of unreliability in diagnosing β-thalassemia trait (β-TT). Curr Med Res Opin. 2018; 34(5): 945-51. doi: 10.1080/03007995.2018.1435520.

[15] Samanta A, Chaudhuri PK, Das U, Bhattacharyya N. A new approach based on erythrocyte indices to exclude normal populations from chromatography based thalassemia screening programs with very Hhgh fidelity. Iran J Blood Cancer. 2021; 13(4): 107-18. http://ijbc.ir/files/site1/user_files_6-10-667-1-c721e4c.pdf.

[16] Organization WH. WHO guideline on use of ferritin concentrations to assess iron status in populations [Internet]. World Health Organization; 2020. https://iris.who.int/bitstream/handle/10665/331505/9789240000124.

[17] Hashemizadeh H, Noori R. Premarital screening of beta thalassemia minor in North-East of Iran. Iran J Pediatr Hematol Oncol. 2013; 3(1): 210-5. PMID: 24575266

[18] Khan AM, Al-Sulaiti AM, Younes S, Yassin M, Zayed H. The spectrum of beta-thalassemia mutations in the 22 Arab countries: a systematic review. Expert Rev Hematol. 2021; 14(1): 109-22. doi: 10.1080/17474086.2021.1860003

[19] Mir SA, Alshehri BM, Alaidarous M, Banawas SS, Dukhyil AAA Bin, Alturki MK. Prevalence of hemoglobinopathies (β-thalassemia and sickle cell trait) in the adult population of Al Majma'ah, Saudi Arabia. Hemoglobin. 2020; 44(1): 47-50. doi.org/10.1080/03630269.2020.1729175.

[20] Nezhad FH, Nezhad KH, Choghakabodi PM, Keikhaei B. Prevalence and genetic analysis of α- and β-thalassemia and sickle cell anemia in Southwest Iran. J Epidemiol Glob Health. 2018; 8(3): 189-95. doi.org/10.2991/j.jegh.2018.04.103

[21] Sun A, Chang JYF, Jin YT, Chiang CP. Differential diagnosis between iron deficiency anemia and thalassemia trait-induced anemia. J Dent Sci. 2023; 18(4): 1963-4.doi: 10.1016/j.jds.2023.07.036.

[22] Jameel T, Baig M, Ahmed I, Hussain MB, Alkhamaly MBD. Differentiation of beta thalassemia trait from iron deficiency anemia by hematological indices. Pakistan J Med Sci. 2017; 33(3): 665-9. doi: 10.12669/pjms.333.12098.

[23] Mohammad SM, Algahtani MS, Alsharif AA, Alshehri RH, Ghali KN Bin, Al-Almaei HMF, et al. Awareness of iron deficiency anemia among the adult population in Riyadh, Saudi Arabia. Int J Med Dev Ctries. 2020; 4(12): 2259. Available from: https://ijmdc.com/fulltext/51-1603909600.

[24] Heng GU, Wang YX, Du MX, Xu SS, Zhou BY, Li MZ. Effectiveness of using mean corpuscular volume and mean corpuscular hemoglobin for beta-thalassemia carrier screening in the Guangdong population of China. Biomed Environ Sci. 2021; 34(8): 667-71. doi: 10.3967/bes2021.094.

[25] Sahiratmadja E, Maskoen AM, Reniarti L, Prihatni D. Erythrocyte indices MCV and/or MCH as first round screening followed by Hb-analysis for β-thalassemia carrier state. Indones Biomed J. 2022; 14(3): 282-8. doi.org/10.18585/inabj.v14i3.1960.

[26] Sari DP, Wahidiyat PA, Setianingsih I, Timan IS, Gatot D, Kekalih A. Hematological parameters in individuals with beta thalassemia trait in South Sumatra, Indonesia. Anemia. 2022; 2022: 3572986. doi: 10.1155/2022/3572986.

[27] Cao A, Kan YW. The prevention of thalassemia. Cold Spring Harb Perspect Med. 2013; 3(2): a011775. doi: 10.1101/cshperspect.a011775.

[28] Owaidah T, Al-Numair N, Al-Suliman A, Zolaly M, Hasanato R, Al Zahrani F, et al. Iron deficiency and Iron deficiency anemia are common

epidemiological conditions in Saudi Arabia: Report of the national epidemiological survey. Anemia. 2020; 2020: 6642568. doi: 10.1155/2020/6642568.

[29] Saffi M, Howard N. Exploring the effectiveness of mandatory premarital screening and genetic counselling programmes for β-thalassaemia in the Middle East: A Scoping Review. Public Health Genom. 2015; 18(4): 193-203. doi: 10.1159/000430837.

[30] Rouh AlDeen N, Osman AA, Alhabashi MJ, Al Khaldi R, Alawadi H, Alromh MK, et al. The prevalence of β-thalassemia and other hemoglobinopathies in Kuwaiti premarital screening program: An 11-year experience. J Pers Med. 2021; 11(10): 980. doi: https://www.mdpi.com/2075-4426/11/10/980.

[31] Soteriades ES, Angastiniotis M, Economidou EC, Farmakis D, Avraam D, Eleftheriou A. The disease burden of β-thalassaemia revisited. Hematology. 2025; 30(1): 2551450. doi.org/10.1080/16078454.2025.2551450.

[32] Pitsillidou O, Petrou P, Postma MJ. Can Cyprus afford luspatercept a budget impact analysis of the reimbursement of luspatercept for the management of thalassaemia in Cyprus. Pharmaco Economics-Open. 2024; 8(3): 471-80. doi: 10.1007/s41669-024-00482-x.

[33] Rebeen Abdulsalam Mohammed Sediqm SASAA. Attitude and knowledge of participating couples towards the premarital hemoglobinopathy screening program in Duhok–Iraq. AMJ. 2024; 9(4): 84-93. doi.org/10.56056/amj.2024.302.

[34] Hawi ZA, Pugsley K, Namipashaki A, Samarrai W, Hawi Z. A Review of Consanguinity in the Iraqi Population: A Call to Action to Prevent a Future Health Catastrophe. Ibn AL-Haitham J Pure Appl Sci . 2024; 37(2 SE-Biology): 64-70. doi. org/10.30526/37.2.3472.

Journal of Associated Medical Sciences 2026; 59(2): 438-446

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Efficacy of digital gaming vs traditional Indian gaming on memory and attention in patients with persistent post-concussion syndrome: A feasibility study

J. Mohanasundaram, Vignesh Srinivasan*, Prathap Suganthirababu, Harishraj A., Madeshwaran S.

Department of Neuroscience, Saveetha College of Physiotherapy, Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai, Tamil Nadu, India.

ARTICLE INFO

Article history:
Received 18 December 2025
Accepted as revised 22 April 2026
Available online 29 April 2026

Keywords:
Persistent post-concussion syndrome, health, digital games, traditional Indian games, quality of life.

ABSTRACT

Background: Persistent post-concussion syndrome (PPCS) often presents long-lasting cognitive impairments such as memory and attention deficits, impacting daily life. Traditional cognitive rehabilitation faces challenges in patient engagement and personalization. Digital cognitive training games and culturally relevant traditional Indian games offer promising yet under-explored options for rehabilitation in PPCS patients.

Objectives: This research study aims to compare the effectiveness of digital gaming and traditional Indian games in enhancing memory and attention among individuals with PPCS.

Materials and methods: An experimental study was conducted with 10 PPCS patients aged 18-45 years, divided equally into two groups. Group A (N=5) received a digital gaming intervention, while group B (N=5) engaged in traditional Indian games. Both interventions consisted of 45-minute sessions, five days per week, over six weeks. Cognitive performance was evaluated using the Digit Span Test and Trail-Making Tests A and B. However, the sample size was small, within-group differences were analyzed using the Wilcoxon signed-rank test, and between-group comparisons were performed using the Mann–Whitney U test, with statistical significance set at $p<0.05$.

Results: Post-intervention analysis revealed that both groups experienced statistically significant gains in memory and attention performance ($p<0.05$). However, Group A exhibited significantly greater gains than group B in post-Digit Span Test scores and Trail-Making Test performance, indicating superior efficacy of digital gaming.

Conclusion: Digital gaming was more effective than traditional Indian games in improving attention and memory in PPCS patients. These findings support adopting technology-based cognitive training in rehabilitation while considering cultural relevance and patient preferences.

* **Corresponding contributor.**
Author's Address: *Department of Neuroscience, Saveetha College of Physiotherapy, Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai, Tamil Nadu, India.*
E-mail address: vignesh.scpt@saveetha.com
doi: 10.66285/JAMS.2026.073
E-ISSN: 2539-6056

Introduction

Persistent post-concussion syndrome (PPCS) is a multifactorial condition in which individuals continue to experience symptoms long after sustaining a first mild traumatic brain injury (mTBI).[1] Although most people recover within weeks, approximately 15-30% develop prolonged symptoms such as headaches, dizziness, emotional disturbances, and prominent cognitive impairments, particularly in attention and memory.[2-4] However, around 10-15% experience truly

persistent symptoms extending beyond the typical recovery period.[5] Neurocognitive research consistently demonstrates that individuals with PPCS show impairments in sustained and divided attention, working memory, processing speed, and episodic memory. Neuropsychological studies report poorer performance on attention and memory tasks compared to healthy controls, reflecting ongoing dysfunction within fronto-parietal attention networks and hippocampal–prefrontal circuitry.[6,7] These deficits interfere with daily functioning, academic performance, and occupational reintegration, underscoring the need for targeted interventions addressing attention and memory. Emerging evidence from brain physiology, including measures of brain tissue oxygenation, further highlights the importance of early cognitive and functional restoration.[8]

Cognitive rehabilitation is a cornerstone of PPCS management; however, traditional paper-and-pencil tasks and therapist-guided drills often lack engagement, adaptability, and ecological validity.[9] With the growth of digital health technologies, interactive and adaptive platforms now offer enriched, multisensory cognitive training environments that leverage neuroplasticity to promote functional reorganization of attention and memory networks after brain injury.[10] Digital cognitive training platforms, such as Lumosity and CogniFit provide gamified, adaptive modules with real-time feedback targeting attention, working memory, executive function, and processing speed. Individualized difficulty adjustment and performance-based progression support sustained engagement and experience-dependent neural change.[11] Clinical studies and meta-analyses in neurological populations, including traumatic brain injury, stroke, and multiple sclerosis, report improvements in attention, working memory, executive functioning, and global cognition.[12] Therapists also report that digital and virtual rehabilitation tools are feasible and easy to integrate into clinical practice.[13]

In contrast, traditional Indian games provide culturally familiar, low-cost, and socially engaging cognitive stimulation. Games such as Pallanguzhi (Mancala) and Snakes and Ladders engage sustained attention, numerical processing, working memory updating, planning, rule monitoring, and strategic decision-making. These processes map directly onto attention and memory domains commonly impaired in PPCS. Evidence from board games and structured leisure activities indicates improvements in memory, attentional control, and executive functioning across age groups, with traditional games proposed as therapeutic tools for special populations.[14,15]

Previous studies show that both digital and non-digital game-based cognitive training effectively improve attention and memory. Systematic reviews and meta-analyses of digital interventions report significant gains in attention, working memory, and executive functioning following structured, adaptive gameplay.[16] Similarly, evidence from board-game-based and other non-digital interventions demonstrates improvements in attentional control, memory performance, and executive functioning.[14]

The rationale for employing game-based interventions in PPCS is grounded in neuroplasticity and cognitive rehabilitation theory, game-based interventions repeatedly engage sustained attention, working memory updating, response inhibition, and rapid information processing, promoting experience-dependent neural reorganization.[17] Digital games offer adaptive, individualized progression, while traditional games provide ecologically valid and socially motivating contexts that enhance adherence.

Despite growing interest, comparative evidence in PPCS remains limited. This study evaluates digital cognitive programs versus traditional Indian games for improving attention and memory, informing the use of culturally grounded and technology-based approaches, individually or in combination, to enhance cognitive recovery in PPCS.

Materials and methods

An experimental study was conducted among individuals diagnosed with PPCS in a private clinical setting. Ten participants meeting the inclusion and exclusion criteria were selected through convenience sampling, informed of the study procedures, and provided written consent. Inclusion criteria were PPCS of at least three months duration, age 18-45 years, either gender, presence of memory or attention complaints, capacity to provide informed consent, and ability to follow instructions. Exclusion criteria included major neurological or psychiatric disorders, moderate to severe TBI (GCS<13), cognitive rehabilitation within the previous three months, and significant visual, motor, or physical impairments limiting participation. Diagnosis of PPCS was established by a consulting neurologist in accordance with the World Health Organization ICD-10 criteria (F07.2 - Post-concussional syndrome), based on a documented history of mtbi and persistence of post-concussive symptoms for more than three months. Identification of memory and attention complaints was conducted through a structured clinical interview and symptom checklist administered by the clinician, focusing on difficulties in sustained attention, concentration, working memory, mental fatigue, forgetfulness in daily activities, and reduced task efficiency. These subjective complaints were corroborated through baseline neuropsychological screening using the Digit Span Test (DST) and Trail-Making Test (TMT A & B), ensuring that participants demonstrated objective evidence of attention and/or memory impairment before enrollment.

Baseline cognitive performance was assessed prior to intervention using standardized neuropsychological measures, including the Digit Span Test (DST) and Trail-Making Test (TMT A & B), to ensure group comparability and control for premorbid cognitive variability.

Educational level was documented as a proxy for cognitive reserve, given its association with attention and working memory. Years of formal education ranged from 12 to 18 years (mean=15.1 years) and were comparable between groups, minimizing confounding effects related to premorbid cognitive ability. Occupational status (student or employed) was documented for all participants. Although occupation may influence cognitive performance, subgroup analysis was not conducted due to the small sample size. The mean age of participants in group A was 26.8±4.1 years, and in group B was 27.4±3.6 years, with no statistically or clinically meaningful difference between groups.

Participants were randomly assigned to group A (Digital gaming group, N=5) or group B (Traditional Indian gaming group, N=5). The combination of objective baseline assessment, documentation of cognitive reserve variables, and randomized group allocation strengthened the attribution of post-intervention changes to the interventions.

Intervention

Group A (Digital gaming group) participated in digital cognitive training such as Lumosity and CogniFit, targeting executive function, working memory, attention, and processing speed. Each session lasted 45 minutes, five days a week for six weeks. Sessions began with brief relaxation exercises to optimize focus. Participants then completed a rotation of cognitive games to maintain engagement and provide varied stimulation. Attention tasks included symbol matching, which required identifying identical symbols to strengthen selective attention, and visual tracking, where participants followed moving targets among distractors to improve sustained attention. Working memory tasks involved sequence recall, number pattern memorization, and spatial grid navigation, each progressively challenging memory updating and visuospatial processing. Game difficulty automatically adjusts based on performance, promoting continuous cognitive advancement. Sessions concluded with short relaxation exercises to reduce cognitive fatigue. Performance data and engagement were tracked after every session to ensure consistency and track improvement (Figure 1).

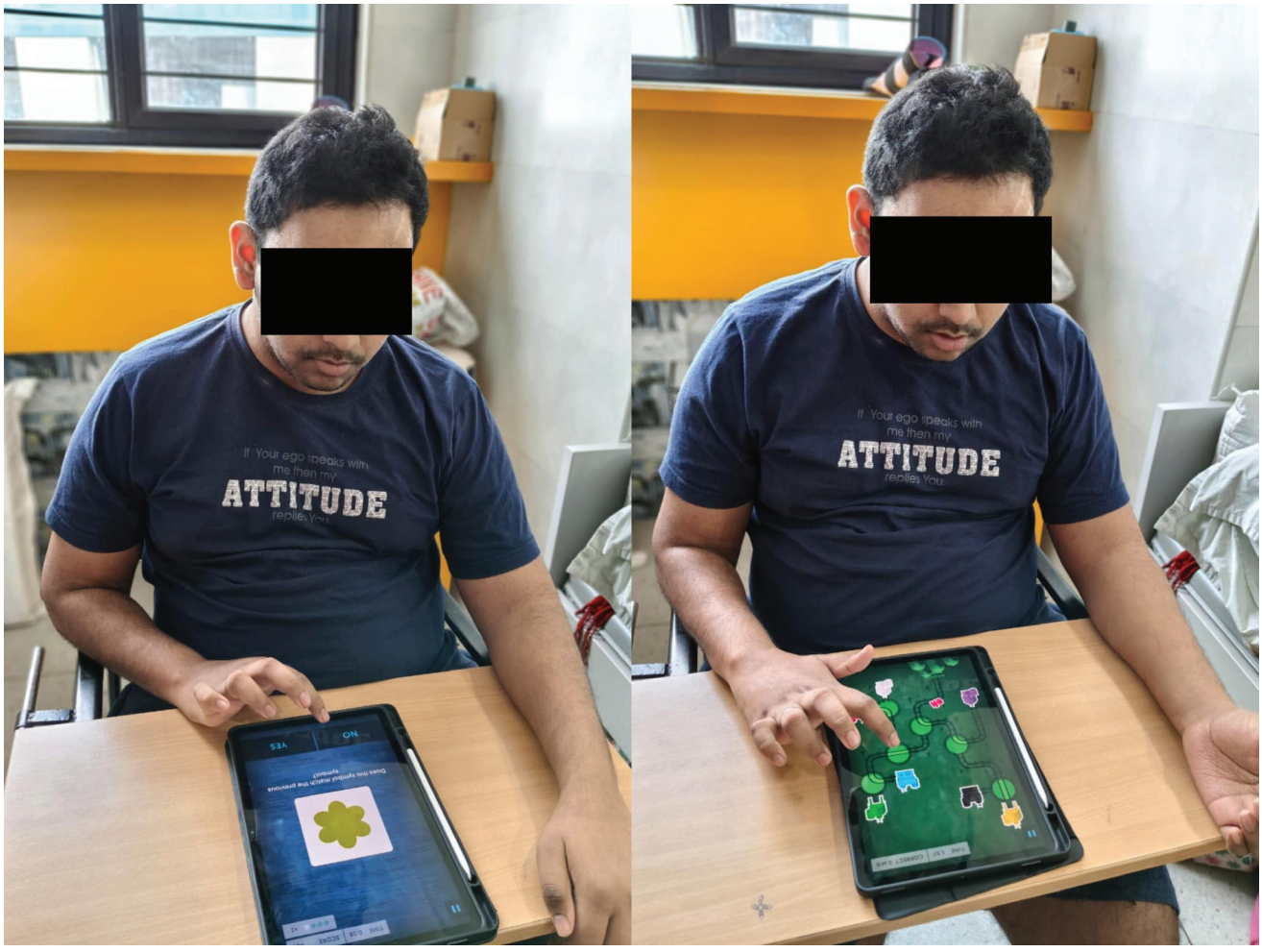


Figure 1. *Digital games.*

Group B (Traditional Indian gaming group) participated in culturally familiar traditional Indian games (Pallanguzhi (Mancala) and Snakes and Ladders) selected for their cognitive stimulation value. Pallanguzhi, played on a 14-pit board with shells, required counting, strategic planning, and recalling previous moves, thereby stimulating memory, mental arithmetic, and focused attention.[15] Snakes and Ladders involved dice-based number recognition, sequencing, and decision-making, with the unpredictability of snakes and ladders enhancing attention and impulse control. Verbal prompts-maintained engagement, encouraged interaction, and clarified rules. Game complexity increased gradually by adding strategic variations. Participants in both groups were advised to avoid other cognitive training activities during the six-week intervention to maintain consistency (Figure 2 and 3).



Figure 2. *Pallanguzhi (Mancala).*



Figure 3. *Snakes and ladder.*

Outcome measurements

1. Digit Span Test (DST): The part of a WAIS-IV that assesses attention span, immediate recall, and working memory through the Digit Span Forward and Backward tasks. Scores increase with correct responses as sequence length progresses, taking about 5-10 minutes. Higher scores indicate better cognitive capacity, with normative adult values ranging from 12 to 18. The test shows strong reliability (r=0.79-0.82).
2. Trail-Making Test (TMT A & B): The TMT evaluates processing speed, visual attention, and executive functioning. TMT-A requires connecting numbers; TMT-B alternates numbers and letters. Scoring is based on completion time and errors. Normal times: TMT-A: 20-40 seconds; TMT-B: 60-90 seconds. Reliability ranges from r=0.79 to 0.89.

Results

The study included 10 participants, with 5 assigned to the experimental group and 5 to the control group. The sex distribution showed more males (N=6) than females (N=4); the experimental group had more females (N=3) than males (N=2), while the control group had more males (N=4) than females (N=1). Participants were aged 18-28 years (N=5), 29-39 years (N=2), and >39 years (N=3); the mean age of participants in group A was 26.8±4.1 years, and in group B was 27.4±3.6 years, with no statistically or clinically meaningful difference between groups. Educational levels were evenly distributed overall (school N=4, college N=4, uneducated N=2), with the experimental group having more school-level participants and the control group having more college-level participants. Most participants were employed (N=7), with three unemployed, and employment status was included as a baseline characteristic due to its potential influence on time availability and engagement in both digital and traditional game activities.

Normality of data distribution was assessed using the Shapiro-Wilk test. Pre- and post-intervention Digit Span and Trail-Making Test scores were normally distributed in both groups (p>0.05). However, due to the small sample size, non-parametric tests were used for inferential analysis. Within-group comparisons were conducted using the Wilcoxon signed-rank test to evaluate pre and postintervention changes in each group. Between-group comparisons of posttest scores were performed using the Mann-Whitney U test. An alpha level of p<0.05 was considered statistically significant.

Within-group analyses revealed significant post-intervention improvements in both groups on the Digit Span Test and Trail-Making Test (A & B ratio). Group A showed a significant increase in Digit Span scores from 6±1.58 to 10.6±0.89 (z=-2.032, p=0.042), while group B improved from 5.8±0.83 to 8.4±1.14 (z=-2.070, p=0.038) as illustrated in Table 1, indicating enhanced memory and attention in both groups, with greater gains in group A. Similarly, significant improvements were observed in attention and executive function, with group A demonstrating a reduction in Trail-Making Test (A & B ratio) scores from 2.66±0.14 to 2.32±0.14 (z=-2.032, p=0.042) and group B showing a decrease from 2.68±0.12 to 2.50±0.08 (z=-2.121, p=0.034) (Table 2), reflecting improved cognitive flexibility and processing speed. Between-group comparisons of post-intervention outcomes showed that group A achieved significantly higher Digit Span scores than group B (10.6±0.89 vs 8.4±1.14, U=2.0, z=-2.193, p=0.014) and demonstrated superior performance on the Trail-Making Test (A & B ratio) (2.32±0.14 vs 2.50±0.08; U=4.5, z=-1.671, p=0.047) (Table 3). The graphical representations (Figures 4 and Figure 5) corroborated these findings, illustrating significant pre- to post-intervention improvements in both groups, with group A showing a greater magnitude of improvement across memory, attention, and executive function outcomes.

Table 1. Pre-test and post-test scores within group A and B on the Digital Span Test.

Group	Pre-intervention (Mean±SD)	Post-intervention (Mean±SD)	z-value	*p* value
Group A	6.0±1.581	10.6±0.894	-2.032	0.042
Group B	5.8±0.83	8.4±1.14	-2.070	0.038

Note: p values represent within-group comparisons using the Wilcoxon signed-rank test. Statistical significance was set at p<0.05.

Table 2. Pre- and post-test scores within group A and B on the Trail-Making Test.

Group	Pre-intervention (Mean±SD)	Post-intervention (Mean±SD)	z-value	*p* value
Group A	2.66±0.137	2.32±0.144	-2.032	0.042
Group B	2.68±0.115	2.5±0.079	-2.121	0.034

Note: The *p* value represents within-group comparisons using the Wilcoxon signed-rank test. Statistical significance was set at $p<0.05$.

Table 3. Between-group comparison of Digit Span Test and Trail-making Test scores for groups A and B.

Outcome measure	Group A (Mean±SD)	Group B (Mean±SD)	U value	z-value	*p* value
Digit span test	10.6±0.894	8.4±1.14	2.0	-2.193	0.014
Trail making test	2.32±0.144	2.5±0.079	4.5	-1.671	0.047

Note: The *p* value represents within-group comparisons using the Wilcoxon signed-rank test. Statistical significance was set at $p<0.05$.

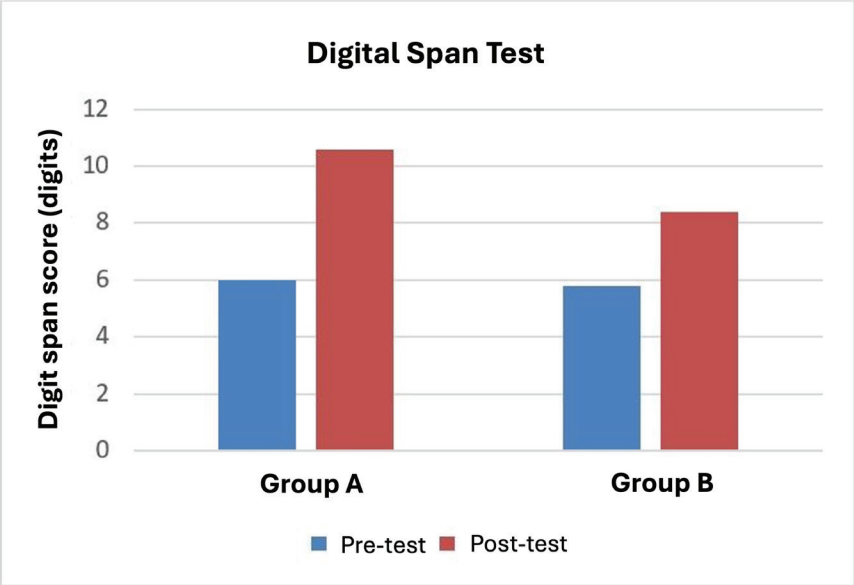


Figure 4. *Comparison of group A and group B Digit Span pre-test and post-test.*

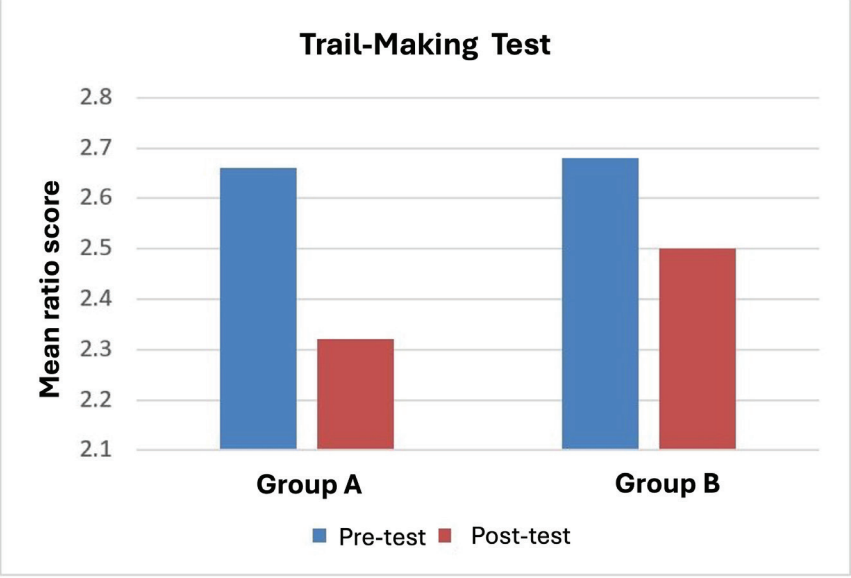


Figure 5. *Comparison of group A and group B Trail-Making pre-test and post-test.*

Discussion

The present study compared the efficacy of digital gaming interventions with traditional Indian games in improving memory and attention among individuals with PPCS. Results revealed significant cognitive gains in both groups; however, the digital gaming group showed superior improvements in DST and TMT scores. These findings support growing evidence that technology-driven cognitive rehabilitation is an effective and engaging approach for restoring neurocognitive function following mTBI.

The superior effects of digital gaming may be attributed to adaptive difficulty modulation, immediate performance feedback, high-frequency repetition, and multisensory engagement, which together maintain an optimal challenge level and promote experience-dependent neuroplasticity.[18] Digital platforms such as Lumosity and CogniFit dynamically adjust task complexity based on user performance, sustaining working memory and attentional load more precisely than fixed-rule traditional games. These design features likely contributed to greater gains in short-term memory, processing speed, and executive control, as reflected in DST and TMT improvements.[19,20]

The findings align with prior evidence showing that game-based cognitive interventions improve attention and memory in individuals with cognitive deficits, including those with traumatic brain injury and post-concussion symptoms.[20] Meta-analytic and clinical trial evidence support the efficacy of digital cognitive training for enhancing attention, working memory, and executive functioning through structured, repetitive, and adaptive practice.[21]

Traditional Indian games such as Pallanguzhi and Snakes and Ladders also produced meaningful cognitive gains by engaging sustained attention, rule monitoring, numerical processing, planning, and working memory updating. Their culturally familiar and socially interactive nature may enhance motivation and feasibility, particularly in low-resource or community-based rehabilitation settings. This aligns with Moser *et al.*, who emphasized the need for customized and contextually relevant rehabilitation strategies for PPCS.[22] Culturally grounded activities may also enhance engagement and reduce psychological resistance, especially among individuals less comfortable with technology. Shen *et al*. highlighted that successful cognitive rehabilitation relies not only on task specificity but also on ecological validity, engagement, and sustained participation.[19] Traditional games meet these criteria through intrinsic motivation and social interaction, promoting gradual cognitive reinforcement.

Participant characteristics may have influenced engagement and outcomes. Younger participants may have shown higher engagement with digital platforms due to greater familiarity with technology, while comparable educational levels between groups likely minimized confounding related to cognitive reserve. Sex- and occupation-related influences could not be examined due to the small sample size. Larger trials should evaluate whether age, sex, education, and occupational cognitive demands moderate responsiveness to digital versus traditional game-based interventions.

As a pilot feasibility study, this work aimed to assess acceptability and preliminary efficacy rather than draw definitive conclusions. The small sample size warranted the use of non-parametric analyses and one-tailed, hypothesis-driven testing to detect preliminary efficacy signals; however, significance was not retained with two-tailed testing, likely due to limited power. These methods were appropriate for exploratory pilot research and intended to inform effect size estimation and protocol refinement for a larger trial. Future studies should recruit adequately powered samples, use two-tailed testing, implement randomized controlled designs, incorporate neuroimaging to verify cortical changes, and include longer intervention periods with follow-up assessments to strengthen rigor and confirm the robustness, durability, and functional relevance of cognitive gains from both digital and traditional game-based interventions.

Limitations

The outcome measures employed were selected for their established reliability and clinical relevance in assessing memory and attention; however, future studies may incorporate a broader battery of neuropsychological assessments**,** including measures of executive function, processing speed, and quality of life to provide a more comprehensive cognitive profile. Although digital gaming interventions offer the advantage of remote and home-based delivery, in-person supervision was intentionally maintained in this study to ensure protocol adherence, standardized session delivery, and accurate performance monitoring. Future trials may explore home-based digital intervention models with remote supervision to enhance accessibility and scalability.

Conclusion

This study provides that digital gaming is more effective than traditional Indian games in enhancing memory and attention in PPCS patients. While traditional games offer valuable supplemental benefits, technology-driven cognitive rehabilitation shows superior impact. These findings support patient-centered, innovative approaches and highlight the need for larger randomized trials.

Ethical approval

The study was approved by the ISRB of Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai (ISRB No: 029/05/2025/ISRB/PGSR/SCPT).

Funding

No funding was obtained for this study.

Conflicts of interest

All the authors in this research declare no conflict of interest.

CRediT authorship contribution statement

J. Mohanasundaram: conceptualization, methodology, supervision; **Vignesh Srinivasan**: conceptualization, methodology, supervision, writing: review and edit; **Prathap Suganthirababu**: validation, writing: review and edit; **Harishraj A**: data curation, investigation, writing: original draft; **Madeshwaran S**: data curation, investigation. All authors read and approved the final manuscript.

Acknowledgements

I want to sincerely thank the department as well as everyone else who supported and encouraged me to set the foundation for this research. Each participant in this study has my sincere gratitude because their participation allowed this project to be completed.

References

[1] Mavroudis I, Chatzikonstantinou S, Petridis F, Palade OD, Ciobica A, Balmus IM. Functional overlay model of persistent post-concussion syndrome. Brain Sci. 2023; 13(7): 1028. doi: 10.3390/brainsci13071028.

[2] Carroll L, Cassidy JD, Peloso P, Borg J, Von Holst H, Holm L, et al. Prognosis for mild traumatic brain injury: results of the WHO collaborating centre task force on mild traumatic brain injury. J Rehabil Med. 2004; 36(0): 84-105. doi: 10.1080/16501960410023859.

[3] Santhana Lakshmi S, Patra D, Annadura B, Das P. The effect of long-term use of virtual reality training on cognitive deficits in post-traumatic brain injury patients: case report. Texila Int J Public Health. 2025; 25(2): Art.019. doi: 10.21522/tijph.2013.se.25.02.art019.

[4] Kumar A, Shabil M, Sah S. Comment on "depression after traumatic brain injury: A systematic review and meta-analysis." Am J Emerg Med. 2024; 86: 170. doi: 10.1016/j.ajem.2024.10.018.

[5] Suwaryo PAW, Kadir F, Omar A, Singh SKD. Prevalence of persistent post-concussion syndrome in adults after mild traumatic brain injury. Nurse Media J Nurs. 2024; 14(2): 199-209. doi: 10.14710/nmjn.v14i2.56529.

[6] Pauhl AN, Kells AM, MacNeill CLW, MacKenzie HM, Christie AD. Effects of persistent post-concussion symptoms on neurocognitive, motor, and neurophysiological outcomes. Brain Inj. 2026; 40(4): 260-9. doi: 10.1080/02699052.2026.2614067.

[7] Oldenburg C, Lundin A, Edman G, Boussard CND, Bartfai A. Cognitive reserve and persistent post-concussion symptoms-A prospective mild traumatic brain injury (mTBI) cohort study. Brain Inj. 2015; 30(2): 146-55. doi: 10.3109/02699052.2015.1089598.

[8] Sathishkumar K. Letter to the editor: Brain tissue oxygen partial pressure monitoring and prognosis of patients with traumatic brain injury: a meta-analysis. Neurosurg Rev. 2024; 47(1): 378. doi: 10.1007/s10143-024-02628-1.

[9] Wolinsky FD, Mahncke HW, Weg MWV, Martin R, Unverzagt FW, Ball KK, et al. The ACTIVE cognitive training interventions and the onset of and recovery from suspected clinical depression. J Gerontol B Psychol Sci Soc Sci. 2009; 64B(5): 577-85. doi: 10.1093/ geronb/gbp061.

[10] Nie L, He J, Wang J, Wang R, Huang L, Jia L, et al. Environmental enrichment for stroke and traumatic brain injury: mechanisms and translational implications. Compr Physiol. 2023; 14(1): 5291-323. doi: 10.1002/cphy.c230007.

[11] Lampit A, Hallock H, Valenzuela M. Computerized cognitive training in cognitively healthy older adults: a systematic review and meta-analysis of effect modifiers. PLoS Med. 2014; 11(11): e1001756. doi: 10.1371/journal.pmed.1001756.

[12] Svaerke K, Pyke SB, Tjoernlund M, Humle F, Mogensen J. Effects of computer-based cognitive rehabilitation on working memory in patients with acquired brain injury in the chronic phase, a pilot-study. Brain Inj. 2022; 36(4): 503-13. doi: 10.1080/02699052.2022.2034965.

[13] Vasanthi RK, Li BCJ, Purushothaman VK, A K, Chinnusamy S, Karuppaiyan RK, et al. Perceived usefulness and ease of using virtual reality during physiotherapy - a cross-sectional survey from physiotherapists perspective. Electron J Gen Med. 2023; 20(6): em530. doi: 10.29333/ejgm/13519.

[14] Verghese J, Lipton RB, Katz MJ, Hall CB, Derby CA, Kuslansky G, et al. Leisure activities and the risk of dementia in the elderly. N Engl J Med. 2003; 348(25): 2508-16. doi: 10.1056/nejmoa022252.

[15] Rahman M, Raja K, Rashid M, Kumar J. Listing of Indian folk games for potential therapeutic benefits in children with neurodevelopmental disability. Games Health J. 2020; 9(6): 453-60. doi: 10.1089/g4h.2019.0191.

[16] Chi K, Chen J, Zhou S, Han Z. The effectiveness of digital cognitive intervention in patients with traumatic brain injury: systematic review and meta-analysis. Front Neurol. 2025; 16:1651443. doi: 10.3389/fneur.2025.1651443.

[17] Välimäki M, Mishina K, Kaakinen JK, Holm SK, Vahlo J, Kirjonen M, et al. Digital gaming for improving the functioning of people with traumatic brain injury: randomized clinical feasibility study. J Med Internet Res. 2018; 20(3): e77. doi: 10.2196/jmir.7618.

[18] Nithianantharajah J, Hannan AJ. Enriched environments, experience-dependent plasticity and disorders of the nervous system. Nat Rev Neurosci. 2006; 7(9): 697-709. doi: 10.1038/nrn1970.

[19] Shen Y, Jiang L, Lai J, Hu J, Liang F, Zhang X, et al. A comprehensive review of rehabilitation approaches

for traumatic brain injury: efficacy and outcomes. Front Neurol. 2025; 16: 1608645. doi: 10.3389/fneur.2025.1608645.

[20] Lynch JM, Anderson M, Benton B, Green SS. The gaming of concussions: a unique intervention in postconcussion syndrome. J Athl Train. 2015; 50(3): 270-6. doi: 10.4085/1062-6050-49.3.78.

[21] Rioux M. A novel intervention for persistent memory symptoms after concussion: a pilot randomized controlled trial [dissertation]. Vancouver: University of British Columbia; 2025.

[22] Moser N, Popovic MR, Kalsi-Ryan S. Complexity of post-concussion syndrome assessment and management: a case for customizing rehabilitation. Biomed Eng Online. 2025; 24(1): 48. doi: 10.1186/s12938-025-01380-x.

Author guidelines

Instructions for authors

Manuscripts can be submitted through the online system via website https://www.tci-thaijo.org/index.php/bulletinAMS/

General principles

The *Journal of Associated Medical Sciences* is a scientific journal of the Faculty of Associated Medical Sciences, Chiang Mai University. The articles submitted to the journal that are relevant to any of all aspects of Medical Technology, Radiologic Technology, Occupational Therapy, Physical Therapy, Physiotherapy, Communication Disorders, and other fields related to the health sciences are welcome. Before publication, the articles will undergo a system of assessment and acceptance by at least three experts specializing in the relevant discipline. All manuscripts submitted to the Journal of Associated Medical Sciences should not have been previously published or are under consideration for publication elsewhere. All publications are protected by the *Journal of Associated Medical Sciences*' copyright.

Generative AI and AI-assisted technologies for the authors

JAMS's policy refers only to manuscript preparation, not to using AI and AI-assisted tools to analyze and draw insight from data as a part of the research.

It is recommended that the authors use AI and AI-assisted tools to improve English writing for the readability and language of the manuscript. The authors should carefully review and edit the result with human oversight and control since AI may generate authoritative-sounding output incorrectly or bias.

Before submission, The authors must ensure the work is original and does not infringe third-party rights.

Authorship implies responsibilities and tasks that can only be attributed to and performed by humans. By this, AI and AI-assisted technologies must not be listed as authors, co-authors, or cited AI.

The authors must be responsible and accountable for the content of the manuscript. Moreover, the authors should disclose in the manuscript the use of AI and AI-assisted technologies, and a statement must appear in the published article since it will support transparency and trust among the authors, readers, reviewers, editors, and contributors.

The author must declare of AI-generate or AI-assist by including source of AI used. If failed to do so, the paper will be retracted as AI related retraction.

Using AI and AI-assisted tools in creating the figures, images, and artwork

AI or AI-assisted tools cannot be used to create figures or pictures or alter images in manuscripts. The tools can be applied only for adjusting the color's brightness, contrast, or balance without obscuring or eliminating any information.

The exception applies only when AI or AI-assisted tools are used or are part of the research designs, materials, and methods. If they are used, the authors must explain how they are used in creating the images or the alteration process. Moreover, the model or tool, version and extension numbers, and manufacturer must be indicated. Authors may use AI or AI-assisted tools to help create graphical abstracts, provided they strictly verify that the content aligns accurately with the abstract and is of high definition.

Manuscript catagories

- **Review** Article must not exceed 20 journal pages (not more than 5,000 words), including 6 tables/figures and references (maximum 75, recent and relevant).
- **Original Articles** must not exceed 15 journal pages (not more than 3,500 words), including tables, figures, and references (maximum 40, recent and relevant).
- **Short Communications** (including technical reports, notes, and letters to the editor) must be at most 5 journal pages (not more than 1,500 words), including 2 tables/figures and references (maximum 10, recent and relevant).

Note: *Before any submission, the author must register based on the institute's official email address only.*

Manuscript files

To submit your manuscripts, you will need the following files:

1. A Title page file with name and ORCID of all authors and corresponding author(s)*
2. Main document file with abstract, keywords, main article, and references.
3. Figure files.
4. Table files.
5. Any extra files, such as supplemental files or Author Biographical notes.
6. A graphical abstract that represents the submitted manuscript. The author may search for those examples in the journal's current issues. Permission to reprint table(s) or figure(s) as a graphical abstract from other sources must be obtained from the original publishers and authors and submitted with the typescript. For clear print quality, the graphical abstract file must be within the following requirements:
   - at least 300 dpi (dots per inch)
   - the CMYK color mode is required for the high-quality presentation

- fits within an A4 size landscape layout
- JPEG or PNG formatted files.

Please find the Manuscript template provided here.

Manuscript Format

1. **Language**: English, Aptos 10 for text and 7 for all symbols. PLEASE be informed that the Journal only accepts the submission of English manuscripts.
2. **Format:** One-side printing, double spacing. Use the standard program and fonts and add page and line numbers for all pages.
3. **A Title Page:** Include the article title, names of all authors and co-authors, name of the corresponding author, and acknowledgements. Prepare according to the following contents.

- *Title of the Article:* Concise and Informative. Titles are often used in information retrieval systems. Avoid abbreviations and formulate where possible.
- *Author names and affiliation:* Where the family name may be ambiguous (e.g. a double name), please indicate this. Below the names, present the authors' affiliation addresses (where the work was done). Indicate all affiliations with a superscript number immediately after the author's name and in front of the appropriate address. Provide the full postal address of each affiliation, including the province, country, and, if available, the e-mail address of each author.
- *Corresponding author:* Indicate who will handle correspondence at all stages of refereeing and publication. Also, post-publication, ensure that telephone and fax numbers (with postal area code) are provided in addition to the e-mail address and the complete postal address. The corresponding author must keep contact details up to date.
- *ORCID of all authors:* By the end of each author's affiliation addresses, the ORCID must be provided with the full link as an example; https://orcid.org/xxxx-xxxx-xxxx-xxxx
- *Acknowledgements:* Acknowledgements will be collated in a separate section at the end of the article before the references in the copyediting stage. Please, therefore, include them on the title page; list here those individuals who provided help during the research (e.g., providing language help, writing assistance, or proofreading the article, etc.)

4. **Main Article Structure:** The manuscripts should be arranged in the following headings: Title, Abstract, Introduction, Materials and methods, Results, Discussion and Conclusion, and References. Prepare according to the following contents:

1) **Title of the article:** The title must be included in the main article.
2) **Abstract,** at most 400 words. The abstract must be structured with the below headings in separate paragraphs:
   - Background
   - Objectives
   - Materials and methods
   - Results
   - Conclusion, and
   - Keywords (3-5 keywords should be included)
3) **Introduction:** State the objectives of the work and provide an adequate background, avoiding a detailed literature survey or a summary of the results.
4) **Materials and methods:** Provide sufficient detail to allow the work to be reproduced. A reference should indicate published methods; only relevant modifications should be described. Ensure that each table, graph, or figure is referred to in the text. According to the policy of ethical approval, authors must state the ethical approval code and provide informed consent for human subject research (if any). For animal research, authors must include a statement or text describing the experimental procedures that affirms all appropriate measures (if any) in this section.

- ***Reporting Guidelines***

  It is important that the manuscript gives clear and complete information in the methods section. For Health Research, the Journal requests that the author provides reporting guidelines for the main study type. The EQUATOR Network Resource Centre (Enhancing the Quality and Transparency Of health Research) which is an international initiative that aims to improve the quality of research publications. It provides a comprehensive list of reporting guidelines available for many different study designs. A list full of all of the reporting guidelines endorsed by the EQUATOR Network can be found here, in addition, the supporting resource can be found here. Some of the reporting guidelines for common study designs are:

  - CONSORT for reporting randomized trials
  - STROBE for observational studies
  - PRISMA for systematic reviews
  - SPIRIT for Study protocols
  - STARD for Diagnostic/prognostic studies
  - CARE for Case reports
  - AGREE for Clinical practice guidelines
  - SRQR for Qualitative research

- SQUIRE for Quality improvement studies
- CHEERS for Economic evaluations
- ARRIVE for Animal pre-clinical studies

5) **Data analysis**: Provide sufficient detail to allow the study of the research's measurement outcomes (if any).

6) **Results:** Results should be clear and concise. Present the new results of the study, such as tables and figures mentioned in the article's main body and numbered in the order in which they appear in the text or discussion.

7) **Discussion:**This should explore the significance of the results of the work, not repeat them. A combined Results and Discussion section is often appropriate. Avoid extensive citations and discussion of published literature.

8) **Limitation** *(optional)***:** Discuss any limitations of the study, such as sample size, study duration, or uncontrolled variables, if applicable.

9) **Conclusion:** The study's main findings may be presented in a short Conclusions section, which may stand alone or form a subsection of a "Discussion" or "Results and Discussion."

10) **Ethical approval:** Ethical clearance for research involving human and animal subjects. The authors must state the ethical approval, date, and/or duration, and name of the institution or board of approval.

11) **Funding:** List all sources of financial support, including grant numbers. If no funding was received, please state: "This research received no external funding."

12) **Conflict of interest:** All authors must declare any financial and personal relationship with others or organizations that could inappropriately influence (bias) their work. If there is no interest to declare, please state, "The authors declare no conflict of interest."

13) **CRediT authorship contribution statement:** Describe each author's contributions using the CRediT (Contributor Roles Taxonomy), such as: Conceptualization, Methodology, Data curation, Writing - original draft, Writing - review & editing, etc.

14) **Acknowledgements:** Acknowledge individuals or institutions that supported the work but do not qualify as authors, such as advisors, data collectors, translators, or non-funding institutional support.

15) **References:** Vancouver's style

Tips: *If your manuscript contains mathematical equations or scientific formulas, you may utilize the ThaiJO Equation Composer available at https://equation.tci-thaijo.org/ that detailed instructions and usage guidelines can be found directly on the website.*

5. **Artwork Requirements**

1) Each table, graph, and figure should be self-explanatory and should present new information rather than duplicating what is in the text. Prepare one page for each and submit it separately as a supplementary file(s).

2) Save the figures as high-resolution JPEG or TIFF files.

Note: *Permission to reprint table(s) and figure(s) from other sources must be obtained from the original publishers and authors and submitted with the typescript.*

Ensuring a double-blinded peer review

To ensure the integrity of the double-blinded peer review for submission to this journal, every effort should be made to prevent the identities of the authors and reviewers from being known to each other. The document's authors have deleted their names from the main text, with "Author" and year used in the references and footnotes instead of the authors' name, article title, etc. After the journal is accepted, the author's name and affiliation and the corresponding author's name must be included in the document and re-submitted in the copyediting stage.

Proof correction

The Editor-in-Chief and production team are responsible for preparing the reprint for online publication. The corresponding author will shortly be informed by email of the proof of the final reprint paper for approval for publication. **Please be advised that the corresponding author checks email frequently.** The journal strongly recommends providing an **institute/formal email address** for submission.

The proof/edit process will take one week. If action is not taken within the specified time, the article will be published online as accepted proof.

Page charge

No page charge.

References format

1. **References style:** using the ***Vancouver*** referencing style (see example below).
2. **In-text citation:** Indicate references by number(s) in the order of appearance in the text with superscript format. Reference numbers must be placed immediately after the punctuation (with no spacing). The actual authors can be referred to, but the reference number(s) must always be given. When multiple references are cited at a given place in the text, use a hyphen (with no spacing) to join the inclusive first and last numbers. Use commas (with spaces) to separate non-inclusive

numbers in multiple citations, e.g. (2-5,7,10). Do not use a hyphen if there are no citation numbers between inclusive statements, e.g. (1-2); use a comma instead (1,2).

3. **References list:** The number of references (numbers in square brackets) in the list must be in the order in which they are mentioned in the text. In case of reference sources from non-English languages, translate the title to English and retain "in Thai" in the parentheses, i.e. (in Thai).

4. Please note that if references are not cited correctly, the manuscript may be returned for amendment before it is passed on to the editor for review.

Examples of references list

Multiple Authors: List up to the first 6 authors/editors, and use "et al." for any additional authors.

Journal Articles (in print): If the reference source contains a Digital Object Identifier (DOI), retain doi: at the end of the reference. (Link to CrossRef.) Vancouver Style does not use the full journal name; only the commonly used abbreviation: "Physical Therapy" is cited as "Phys Ther." (Link to Journal title of abbreviation) As an option, if a journal carries continuous pagination throughout a volume (as many medical journals do) the month and issue number may be omitted. Allow one space after the semi-colon and colon, and then end each reference with a full stop after the page number.

- Pachori P, Gothalwal R, Gandhi P. Emergence of antibiotic resistance Pseudomonas aeruginosa in intensive care unit; a critical review. Genes Dis. 2019; 6(2): 109-19. doi: 10.1016/j.gendis.2019.04.001.
- Hung Kn G, Fong KN. Effects of telerehabilitation in occupational therapy practice: A systematic review. Hong Kong J Occup Ther. 2019; 32(1): 3-21. doi: 10.1177/1569186119849119.
- Wijesooriya K, Liyanage NK, Kaluarachchi M, Sawkey D. Part II: Verification of the TrueBeam head shielding model in Varian VirtuaLinac via out-of-field doses. Med Phys. 2019; 46(2): 877-84. doi: 10.1002/mp.13263.
- Velayati F, Ayatollahi H, Hemmat M. A systematic review of the effectiveness of telerehabilitation interventions for therapeutic purposes in the elderly. Methods Inf Med. 2020; 59(2-03): 104-9. doi: 10.1055/s-0040-1713398.
- Junmee C, Siriwachirachai P, Chompoonimit A, Chanavirut R, Thaweewannakij T, Nualnetr N. Health status of patients with stroke in Ubolratana District, Khon Kaen Province: International Classification of Functioning, Disability and Health-based assessments. Thai J Phys Ther. 2021; 43(1): 45-63 (in Thai).

Journal Articles (on the Internet)

- Siegel PM, Bojti I, Bassler N, Holien J, Flierl U, Wang X, et al. A DARPin targeting activated Mac-1 is a novel diagnostic tool and potential anti-inflammatory agent in myocarditis, sepsis and myocardial infarction. Basic Res Cardiol [Internet]. 2021; 116(1): 1-25. Available from: https://doi.org/10.1007/s00395-021-00849-9.

Journal Articles (from an online database)

- Jackson D, Firtko A, Edenborough M. Personal resilience as a strategy for surviving and thriving in the face of workplace adversity: A literature review. J Adv Nurs [serial online]. 2007; 60(1): 1-9. doi: 10.1111/j.1365-2648.2007.04412.x.

Book / Chapter in an Edited Book References

Please be informed that references to books and chapters in edited books should not be included in the research article but in other manuscript categories.

- Grove SK, Cipher DJ. Statistics for nursing research: A workbook for evidence-based practice. 3rd St. Louis, Missouri: Elsevier; 2019.
- Haznadar M, Editor. Cancer metabolism: Methods and protocols. New York: Humana Press; 2019. doi: 10.1007/978-1-4939-9027-6.
- Perrin DH. The evaluation process in rehabilitation. In: Prentice WE, Editor. Rehabilitation techniques in sports medicine. 2nd Edition. St Louis, MO: Mosby Year Book; 2020: pp 253-78.

E-book

- Dehkharghani S, Editor. Stroke [e-book]. Brisbane (AU): Exon Publications; 2021 [cited 2021 Jul 31]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK572004/ doi: 10.36255/exonpublications.stroke.2021.
- Tran K, Mierzwinski-Urban M. Serial X-Ray radiography for the diagnosis of osteomyelitis: A review of diagnostic accuracy, clinical utility, cost-effectiveness, and guidelines [e-book]. Ottawa (ON): Canadian Agency for Drugs and Technologies in Health; 2020 [cited 2021 Jul 31]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK562943/

Dissertation/Thesis

- On-Takrai J. Production of monoclonal antibody specific to recombinant gp41 of HIV-1 subtype E [Term paper]. Faculty of Associated Medical Sciences: Chiang Mai University; 2001 [in Thai].
- Seale AC. The clinical and molecular epidemiology of streptococcus agalactiae in Kenya: maternal colonization and perinatal outcomes [Dissertation on the Internet]. [Oxford (England)]: University Oxford; 2015 [cited 2015 Jul 28]. Available from: http://ora.ox.ac.uk/objects/uuid:6e7d952a-dc5b-4af0-b0bb-f2ae2184eed0.

Conference Proceedings

- Lake M, Isherwood J, Clansey. Determining initial knee joint loading during a single limb drop landing: reducing soft tissue errors. Proceedings of 34thInternational Conference of Biomechanics in Sport; 2016 Jul 18-22; Tsukuba, Japan, 2016. Available from: https://ojs.ub.uni-konstanz.de/cpa/article/view/7126.
- Ellis MD, Carmona C, Drogos J, Traxel S, Dewald JP. Progressive abduction loading therapy targeting flexion synergy to regain reaching function in chronic stroke: preliminary results from an RCT. Proceedings of the 38thAnnual International Conference of the IEEE Engineering in Medicine and Biology Society; 2016: 5837-40. doi: 10.1109/EMBC.2016.7592055.

Organization as Author / Government Document

- Australian Government, Department of Health. Physical activity and exercise guidelines for all Australian. 2021 [updated 2021 May 7; cited 2021 Jul 15]. Available from: https://www.health.gov.au/health-topics/physical-activity-and-exercise/physical-activity-and-exercise-guidelines-for-all-australians.
- Department of Health. Situation survey on policy and implementation of physical activity promotion in schools for first year 2005. Nonthaburi: Ministry of Public Health; 2005. [in Thai].
- Department of Local Administration, Ministry of Interior Affairs. Standard of Sports Promotion. Bangkok. 2015: 7-9. [in Thai].
- World Health Organization. WHO guidelines on physical activity and sedentary behaviour. Geneva: World Health Organization; 2020. License: CC BY-NC-SA 3.0 IGO.
- World Health Organization. The epidemiology and impact of dementia: current state and future trends. 2015 [cited 2021 Mar 8]. Available from: http://www.who.int/mental_health/neurology/dementia/dementia_thematicbrief_epidemiology.pdf.


OURNAL OF ASSOCIATED
MEDICAL SCIENCES
AMS