JOURNAL OF ASSOCIATED MEDICAL SCIENCES

JAMS

JOURNAL OF ASSOCIATED MEDICAL SCIENCES

THE OFFICIAL PEER-REVIEWED ONLINE JOURNAL

Volume 59 Number 3 September-December 2026 E-ISSN: 2539-6056

Journal of Associated Medical Sciences

Aims and scope

The Journal of Associated Medical Sciences is the official scientific publication of the Faculty of Associated Medical Sciences, Chiang Mai University, Thailand. Our primary mission is to provide an international platform for researchers and professionals to disseminate and discuss significant findings, inventions, and innovations in the following areas:

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 sponsored 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 — Thaksin 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 — Drugs for Neglected Diseases Initiative (DNDi), Switzerland.
marclallemant@gmail.com

Masahiro Hosoda — Hirosaki University, Japan.
m_hosoda@hirosaki-u.ac.jp

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
Natthayarat Suwanwarit

Webpage Administrative Staff
Nopporn Phuangsombat
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. 3 : September-December 2026

Content

1 Development of the Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OT-COM-CA) for older adults with cognitive impairment: A prototype mobile application

Pachpilai Chaiwong[1], Hataichanok Apikomonkon[1], Anuchart Kaunnil[1], Sulalita Hongnueng[2], Peeraya Munkhetvit[1]*

10 [Retracted] Altered visfatin and adiponectin levels and their interaction with HEGF in the metabolic dyslipidemia of overt hypothyroidism: A case-control study

Murtadha Kareem Najim and Abtisam Faris Al-Shukry*

23 Preliminary study: Effect of breathing exercise on hemodynamics and heart rate variability in hypertension

Sirinut Chaiduang[1] and Jirawat Wattanapanyawech[2]*

31 Effectiveness and satisfaction of the elastic ankle support devices for walking and mobility among healthy people

Sugalya Amatachaya[1,2], Kanchalika Ngamying[1], Pornpawee Thawthaisong1, Thaksin Chanata[1,2], Wilairat Namwong[1,2], Arpassanan Wiyanad[1,2], Thiwabhorn Thaweewannakij[1,2]*

42 Analysis of associations between personal attributes, emotional state, and functional performance after stroke

Euna Ko[1] and Wongjong Yu[2]*

55 The interplay between serum Interleukin-33 and vitamin D3 in childhood asthma: Diagnostic utility and severity predictors in an Iraqi community

Zahraa Hazim Aziz and Thikra Abdullah Mahmood*

72 Predicting dyscalcemia using machine learning models based on routine laboratory data

Busadee Pratumvinit[1,2], Kittisak Thawnashom[3,4], Bunjira Makond[5]*

85 Comparison of radiation and light field congruence between a developed collimator test tool and a standard collimator test tool in the general X-ray units of Lamphun Province hospitals

Kampol Taywadithep

92 Effect of prospective memory training on instrumental activities of daily living skills among individuals with mild dementia

Punitha P and Yamuna T*

Content

99 Complications of hyaluronic acid fillers in HIV-positive patients: A narrative review

Kaewkan Thasana-Ekachit and Arisa Kaewkes*

107 Work performance assessment in Thai homeless shelters: An exploratory factor analysis

*Uthaikan Thanapet, Watthanaree Ammawat, Maliwan Rueankam, Winai Chatthong, Supalak Khemthong**

116 Wideband acoustic immittance (WAI) normative data in Thai normal hearing adults

Thanawan Wongkor[1], Nichtima Chayaopas[1], Supawan Laohasiriwong[1], Pittayapon Pitathawatchai[2], Sanathorn Chowsilpa[3]*

124 The re-emerging shadow: A systematic review of scrub typhus in India

Abhishek Mehta[1], Aditya Goel[2], Chandravarman Singh[3]*

132 Association of serum Interleukin-13 and arginase-1 with disease severity and control in patients with asthma: A biomarker-based study

Hayder Yaseen Ali[1], Haider Abdulhameed Alqaraghuli[2], Raid J. Mohammed AL-TIMIMI[1]*

142 Effectiveness of individualized multicomponent intervention for sensory food aversion feeding disorder in children: A case series

Kowsalya M[1], Sujata Missal[2], Sugi S[2]*

151 Diagnostic accuracy of urine AlereLAM and FujiLAM assays for diagnosing active tuberculosis in people living with HIV: A systematic review and meta-analysis

Andiesta Asriyah Mariam Saman[1], Puspa Wardhani[2,3,4,5], Aryati[2,3,4], Lukman Nur Rahman[1]*

168 Serum Galectin-3 as a diagnostic and discriminatory biomarker for chronic kidney disease: A case-control study

Asmaa. H. Nasser and Entedhar Rifaat Sarhat

177 Effect of a behavioral strategy program on quality of life: A quasiexperimental study in women with cognitive impairment receiving chemotherapy

M. Arun Kumar and Jamima Jamesraj

Content

185 Evaluation of changes in diabetes distress, resilience, and self-care behavior following Occupational Therapy Diabetes Self-Management (OTDSM) among adults with type 2 diabetes: A preliminary Quasi-experimental study

Kamala Arumugam, Loganathan Shanmugam, Raghuram Pattusamy, Darsana Ganesan, Jaivignesh Shanmugam*

196 Design and preliminary study of VP1-encoding mRNA-lipid nanoparticles for Enterovirus 71

Parnuphan Panyajai[1], Rujiraporn Pitaksalee[1], Apichai Prachasuphap[1], Kodcharad Jongpitisub[1], Pantida Treeyoung[1], Sirawit Wet-osot[1], Nalinee Saengtong[1], Nadthanan Pinyosukhee[1], Rattanawadee Wichajarn[1], Kritsamon Sophondilok[1], Anicha Luengchaichaweng[1], Panadda Dhepakson[2]*

205 Percentile-based infrared thermography for community screening of foot thermal abnormalities

Sirasith Uat-hao[1], Kemika Wongsatan[1], Runchana Kantanat[2], Kanyawee Sathapong[2], Kittichai Wantanajittikul[1], Chatchanok Udomtanakunchai[1]*

215 Exploring the antioxidant and prebiotic potential of fermented Thai jasmine rice beverage: An in vitro study

Ran Kitkangplu[1], Krongpong Monpengpinij[2], Kongkiat Kespechara[3], Woottichai Nachaiwieng[1,4], Niwed Kullawong[1,4]*

224 Deep learning approaches for CT image reconstruction: Advancements and clinical implications

Rita Yadav, Rajesh Kumar, Kavita Singh*

238 Proof-of-concept development of a cost-effective PACE-based assay for CYP2C19 genotyping

*Chonticha Sirikul, Auttachai Saejia, Nampeung Anukul**

247 Diagnostic potential of circulating Meteorin-like protein in Graves' disease with subclinical hyperthyroidism subgroup analysis

Sajjad Kadhim Ashour[1], Mohammed Imran Hamzah[1], Mahmood Shakir Khudhair[2]*

256 Method comparison and clinical classification agreement between Freelite and N-Latex serum-free light chain assays for clinical laboratory implementation

Dararat Dankai[1], Phichayut Phinyo[2], Wiyada Dankai[3]*

268 Avocado juice (Persea americana) with an abundance of bioactive substances on coupling with probiotics revealed potential health benefits

Kanchi Ravi Padma[1], Sankari Mohan[1], Lingareddy Nikitha Reddy[1], Dhanasekar Chitra[2]*

Content

285 Genotyping-based prediction of Diego blood group phenotypes: Implications for Thai blood donors

Sarisa Chidtrakoon[1], Wiradee Sasikarn[2], Tanaporn Choychimplee[2], Kamphon Intharanut[2], Oytip Nathalang[2]*

294 Robusta peaberry coffee attenuates inflammation, regulates apoptosis, and improves testicular cell integrity in diabetic rats

Mochamad Rizal Maulana[1], Teguh Budiharjo[1], Djoko Priyatno[1], Burhannudin Burhannudin[2], Puji Lestari[2], Sofhia Azqilla Nur[1], Yulian Nursaputra Habibina[1], Eko Naning Sofyanita[1]*

305 Prevalence of self-stigma and psychological distress among parents of children with autism spectrum disorder: A cross-sectional study

Harshavardhana Anandakumar, Raghuram Pattusamy[*], Loganathan Shanmugam, Deepika Pakkiri

313 Reduction of decoder convolutional blocks in Attention Gate U-Net++ to enhance computational efficiency in automatic HR-CTV delineation on CT-based cervical cancer brachytherapy images

Thinnagit Srikhot[1], Chatrawut Pattaweerakul[2], Kamonchanok Nobphuek[2], Titipong Kaewlek[1,3]*

325 Effectiveness of a tele-pulmonary rehabilitation program in individuals at risk of chronic obstructive pulmonary disease from artificial intelligence predictions

Busaba Chuatrakoon[1], Siriphun Kongsawasdi[1], Kittichai Wantanajittikul[2], Sompong Sriburee[2], Supatcha Konghakote[1]*

334 Prediction of coronavirus disease severity based on increased levels of microvesicles: A systematic review and meta-analysis

Nastasya Nunki[1,2], Yetti Hernaningsih[3,4,5], Puspa Wardhani[3,4,6], Citrawati Dyah Kencono Wungu[6,7], Agil Saputra[8], Narazah Mohd Yusoff[3,9]*

344 Development of a PETG-based 3D-printed lumbar spine phantom for bone mineral density assessment

*Thunyarat Chusin, Pakaphorn Kongfun, Suchanya Chanapim, Suphichaya Thanashappokin, Thanyawee Pengpan, Kingkarn Aphiwatthanasumet**

357 A review article: DIRFloortime® on improving social communication skills in Thai children with autism spectrum disorder

Natwipa Wanicharoen

364 Influence of Eu2O3 incorporation on the gamma-ray shielding performance of germanoborate glasses

Natthakridta Chanthima[1,2], Supakit Yonphan[1,2], Chalermpon Mutuwong[1,2], Wuttichai Chaiphaksa[1,2]*

Journal of Associated Medical Sciences 2026; 59(3): 1-9

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 the Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA) for older adults with cognitive impairment: A prototype mobile application

Pachpilai Chaiwong[1], Hataichanok Apikomonkon[1], Anuchart Kaunnil[1], Sulalita Hongnueng[2], Peeraya Munkhetvit[1]*

[1]Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]B.Care Medical Center, Sai Mai, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 17 November 2025
Accepted as revised 30 April 2026
Available online 14 May 2026

Keywords:
Occupational therapy outcome measure, cognitive assessment, older adults, computer-based assessment, cognitive Impairment

ABSTRACT

Background: Cognitive impairment is common in older adults and increases with age. It affects independence in daily activities, such as financial management, meal preparation, medication management, and driving. In healthcare, outcome measurement tools are essential for evaluating a patient's status, establishing baseline performance, guiding intervention planning, and monitoring intervention effectiveness. Computer-based assessments are feasible and well-accepted by patients, providing immediate and accurate scoring, real-time feedback, and automated reporting of results. However, in the Thai context, the availability of standardized assessment tools specifically designed for older adults with cognitive impairment remains limited. Consequently, occupational therapists require comprehensive and contextually appropriate instruments to effectively guide cognitive interventions throughout the continuum of care. To address this gap, a new assessment tool was developed and implemented as an Android tablet–based application to facilitate evidence-based practice.

Objectives: To develop the Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA) for older adults with cognitive impairment and evaluate its content validity and feasibility.

Materials and methods: This study employed a research and development (R&D) design to establish the content and workflow structure of OTCOM-CA. Following the content validity evaluation, the application's workflow and interface were designed in accordance with established user experience (UX) and user interface (UI) principles and subsequently validated. A pilot study involving 10 healthy older adults was conducted to evaluate the clarity of the workflow and the appropriateness of the language used within the application.

Results: Content validity evaluation of the OTCOM-CA's structure, workflow, and screen design showed acceptable ratings for all items (IOC=0.66-1.00). Most participants understood the workflow and test instructions.

Conclusion: OTCOM-CA demonstrated acceptable content validity. The prototype Android version shows a good workflow and language validation. Once revisions are made, its psychometric properties should be evaluated.

* **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: 10.66285/JAMS.2026.074
E-ISSN: 2539-6056

Introduction

Older adults constitute a substantial and growing proportion of the population in many countries worldwide. In 2020, nearly 1 billion people worldwide were aged 65 years and older.[1] In 2024, over 20%

of Thailand's population was older adults, making it an aged society.[2] Cognitive impairments, such as dementia, are prevalent among older adults and negatively impact their functional abilities and independence. In 2021, almost 700 thousand older adults in Thailand had dementia,[3] and the prevalence of mild cognitive impairment (MCI), which can progress to more severe forms like dementia, has increased fivefold among the Asian elderly.[4] The 2030 Agenda for Sustainable Development emphasizes that all segments of society should not be left behind, urging governments to prioritize the health and well-being of individuals with cognitive impairment.[1]

Occupational therapists (OTs) are essential in preserving the quality of life of older adults.[5] Evidence-based practice requires standardized cognitive outcome measures to prove the quality of OT services for clients with cognitive impairments.[6] In 2022, Chaiwong *et al*. developed the Occupational Therapy Cognitive Outcome Measure (OTCOM) for older adults with cognitive impairments. The measure was developed based on information processing theory, Toglia's Dynamic Interactional Approach, the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, relative cognitive assessment tools, clinical experience, and previous research. This initiative was driven by the lack of standardized cognitive outcome measures in occupational therapy in Thailand. It is a dynamic and comprehensive paper-and-pencil cognitive assessment tool. It showed good internal consistency (α=0.88), excellent inter-rater and intra-rater reliability (ICC=0.99), good concurrent validity with the Thai Cognitive-Perception Test (Thai-CPT) and the Dynamic Loewenstein Occupational Therapy Cognitive Assessment-Geriatric Version (DLOTCA-G), good construct validity, and sensitivity to change. However, the OTCOM requires further revision in some aspects, such as rescaling the scoring for the Orientation subtest, before it can be implemented in clinical practice.[7] Moreover, paper-and-pencil measures require more resources and offer less flexibility.[8] Computerized data collection is well accepted, efficient, and practical, producing immediate, accurate, and readily comparable results.[9] Because computerized assessment tools can provide real-time feedback and electronic outputs across multiple platforms (desktops, smartphones, and tablets), a tablet-based format was selected.[8] Tablets are compact, versatile, and affordable, and are increasingly used by medical professionals for data collection, diagnostics, and treatment.[10, 11] Moreover, studies have reported that older adults are satisfied with using tablets in medical care, and have recommended encouraging their use in healthcare services.[12] In addition, Android tablets account for nearly 80% of the market share.[13] Therefore, the OTCOM was revised and converted into an Android tablet application to improve its usefulness and applicability.

Therefore, the primary objective of this study was to develop a comprehensive and valid computer-based cognitive outcome assessment tool. The secondary objectives were to evaluate the content validity, workflow, user interface, and user experience of the developed assessment tool, and to pilot its use to assess the usability and acceptability of the new computer-based cognitive assessment tool among older adults. The prototype of the computer-based occupational therapy cognitive outcome measure for the elderly with cognitive impairment, named "Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA)," was achieved after the completion of the study.

Materials and methods

This study is an initial part of the development of OTCOM-CA. A research and development (R&D) design was adopted to develop a prototype of the OTCOM-CA. This study was divided into two phases. In the first phase, the content and workflow structure of the OTCOM-CA application were developed, followed by an examination of its content validity. In the second phase, the prototype application was developed and pilot tested to evaluate language validity. The details of each phase are as follows:

Phase 1: Development of OTCOM-CA

Firstly, the structure of the Occupational Therapy Cognitive Outcome Measure (OTCOM) [7] was reviewed and revised, considering findings from a previous study. Secondly, the revised structure was adapted for use as a mobile application (app). This modification involved several aspects, including rescaling, adjustments to administration procedures, and the addition of new sub-items. The modified version is called the Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA). Thirdly, the content validity of the OTCOM-CA was examined by three experts: an occupational therapist with five years of experience in cognitive assessment, a psychiatrist with more than five years of experience in cognitive intervention and using cognitive tests, and an occupational therapy lecturer with at least five years of experience in cognitive assessment and intervention with older adults with cognitive impairment. Fourthly, the OTCOM-CA was subsequently developed as a mobile application for Android tablets. Accordingly, the Android application's workflow was planned, and each response screen was designed in line with user experience (UX) and user interface (UI) principles. Finally, the workflow and all response screens were evaluated for content validity by three experts: two programmers with more than 5 years of experience developing Android applications and an occupational therapist with 5 years of experience in cognitive assessment. Content validity was determined by the index of item objective congruency (IOC) formula.[15,16]

Phase 2: Pilot use of OTCOM-CA

After revising the workflow and each response screen in response to experts' suggestions, the prototype application for the OTCOM-CA was developed. The pilot was then used with ten healthy older adults to examine the workflow and language validation. In this step, healthy older adults were asked to provide recommendations for the user experience (UX). The demographic characteristics of the participants and pilot use information were analyzed using descriptive statistics.

Instruments

The Montreal Cognitive Assessment (MoCA) Thai version 8.3 and the Montreal Cognitive Assessment Basic (MoCA-B) Thai version 2014 were used, depending on the participant's education; one of the two cognitive screening tools was used. The MoCA was administered to participants with five or more years of education, while the MoCA-B was used for those with fewer than five years of education. These tools have sensitivity and specificity ratings from 81% to 99%. A score above 24 implies the absence of cognitive impairment.[14] The participants in this phase were healthy older adults who scored above 24 on the screening tool.

Results

The OTCOM-CA comprises two parts: a self-report part and a performance-based part (Table 1). The former part includes 29 subjective questionnaires. The latter part includes eight domains (28 subtests), static cognitive assessment: orientation, attention, memory, perceptual-motor function, language, executive functions, social cognition, and self-awareness. Among these 8 domains, 5 (16 subtests) are dynamic measures, including a variety of step-by-step mediations. The content validity of the self-report part was acceptable, with an Index of Item-Objective Congruence (IOC) ranging from 0.67 to 1.00 across 24 questions. But all static subtests of the performance-based part demonstrate acceptable content validity (IOC=0.67-1.00). Only the dynamic subtest of the performance-based part had unacceptable content validity (IOC=0.33). Therefore, a word in the mediation step that could potentially cause misunderstandings was removed, following the expert's suggestion. After the revision, all questions in both the self-report and performance-based parts have acceptable content validity (IOC=0.67-1.00).

Table 1. Content and structure of the Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA).

Part 1: Self-report			**Questionnaires**		
Section 1: Self report of overall cognition and memory			2 questions		
Section 2: Self report of each component			27 questions		
Part 2: Performance-based					
Domain (8 domains)	**Subtest**	**Scale score**	**Test type**		**Mediation** (level)
			Static	Dynamic	
Orientation	Orientation of person	5 Interval scale	✓		-
	Orientation of time	5 Interval scale	✓		-
	Orientation of place	5 Interval scale	✓		-
	Orientation of situation	2 Interval scale	✓		-
Attention	Selective attention A	8 Interval scale	✓	✓	3
	Selective attention B	16 Interval scale	✓	✓	3
	Processing speed	31 Interval scale	✓		-
Memory	Immediate recall memory	6 Interval scale	✓		-
	Recognition memory	10 Interval scale	✓		-
	General semantic memory	5 Interval scale	✓	✓	2
	Delayed recall memory	5 Interval scale	✓	✓	2

Table 1. Content and structure of the Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA) (*continued*).

Part 2: Performance-based

Domain (8 domains)	Subtest	Scale score	Test type		Mediation (level)
			Static	Dynamic	
Perceptual-motor function	Geometric copy (Visuoconstruction)	5 Ordinal scale	✓	✓	2
	Pantomime praxis	10 Interval scale	✓	✓	2
	Imitation praxis	10 Interval scale	✓	✓	4
	Use of objects praxis	5 Interval scale	✓	✓	2
Language	Naming	26 Ordinal scale	✓		-
	Semantic fluency	10 Ordinal scale	✓		-
	Receptive language	5 Interval scale	✓		-
Executive functions	Flexibility	13 Ordinal scale	✓	✓	4
	Abstract thinking	5 Interval scale	✓	✓	2
	Categorization (grouping)	7 Ordinal scale	✓	✓	4
	Categorization (reason)	6 Ordinal scale	✓	✓	2
	Sequencing A	6 Ordinal scale	✓	✓	4
	Sequencing B	6 Ordinal scale	✓	✓	4
	Problem solving	12 Ordinal scale	✓	✓	2-3
Social cognition	Recognition of emotions	5 Interval scale	✓	✓	2
	Theory of mind	5 Interval scale	✓		-
Self-awareness	Awareness of cognitive decline	10 Ordinal scale	✓		-
	Awareness of memory	10 Ordinal scale	✓		-

The workflow was written on paper (Figure 1) and then translated into a flow in the Figma program. Two hundred and twenty-eight responding screens were created. Each response screen was designed using Canva (Figure 2-4). They were then tested for content validity in three aspects: the appropriateness of the workflow, the screen design as user experience (UX), and the screen design as user interface (UI). All aspects have an acceptable content validity (IOC=0.66-1.00). Experts suggested that some steps in the workflow should be clarified or reorganized. In addition, some buttons should be added or relocated, and some screens should be re-colored or redesigned to improve the UX/UI.

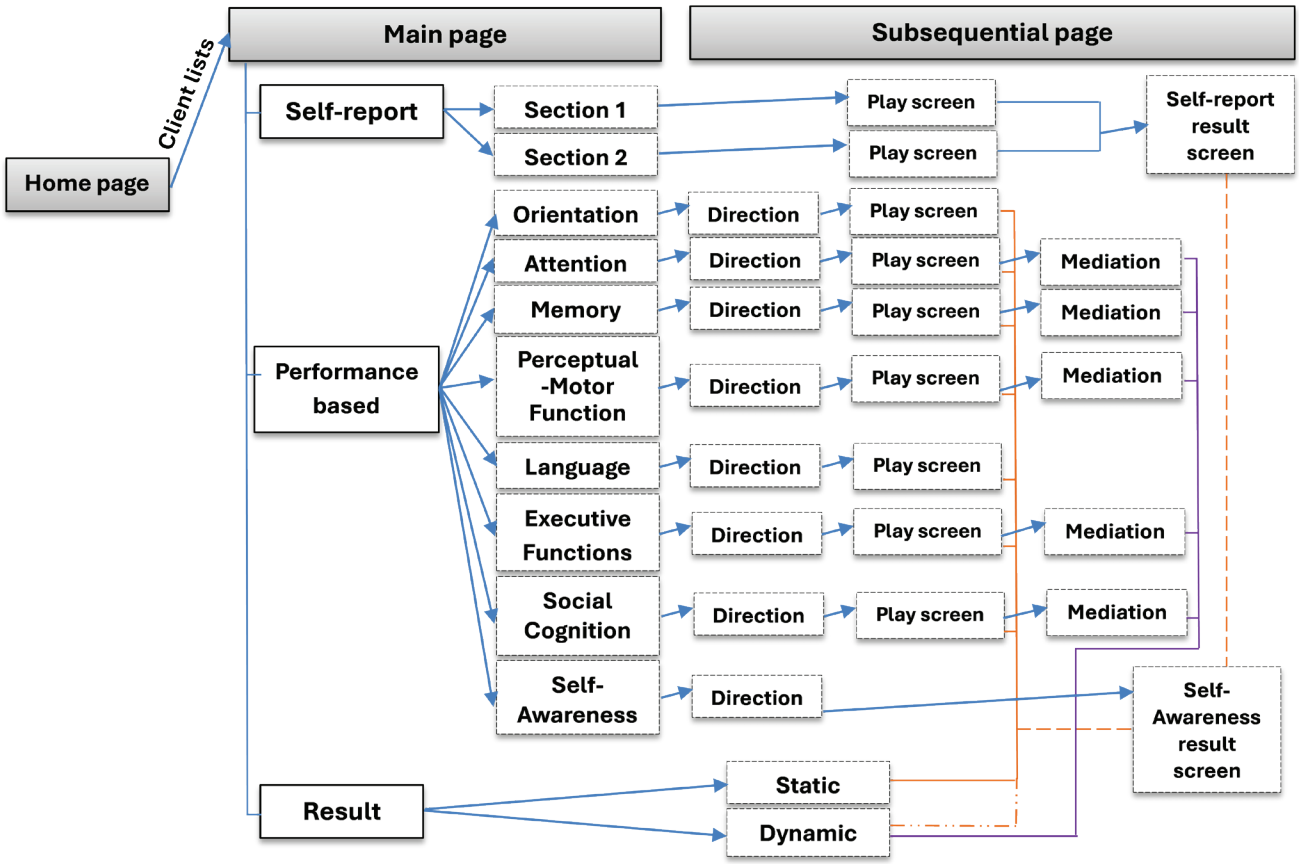


Figure 1. *Workflow of the Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA).*

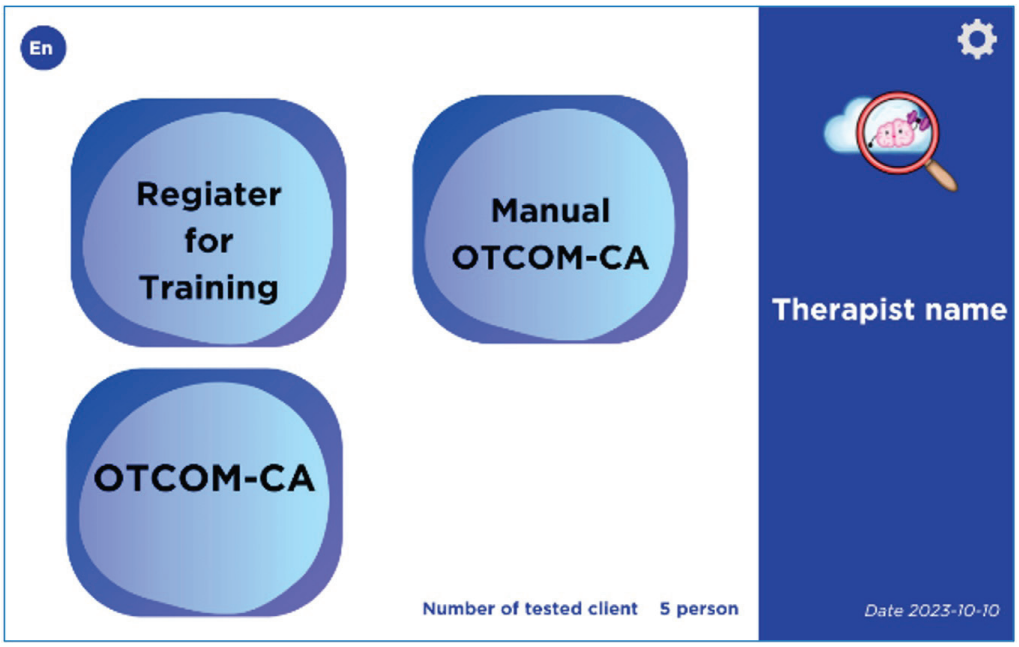


Figure 2. *Example of home page screen of the Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA).*

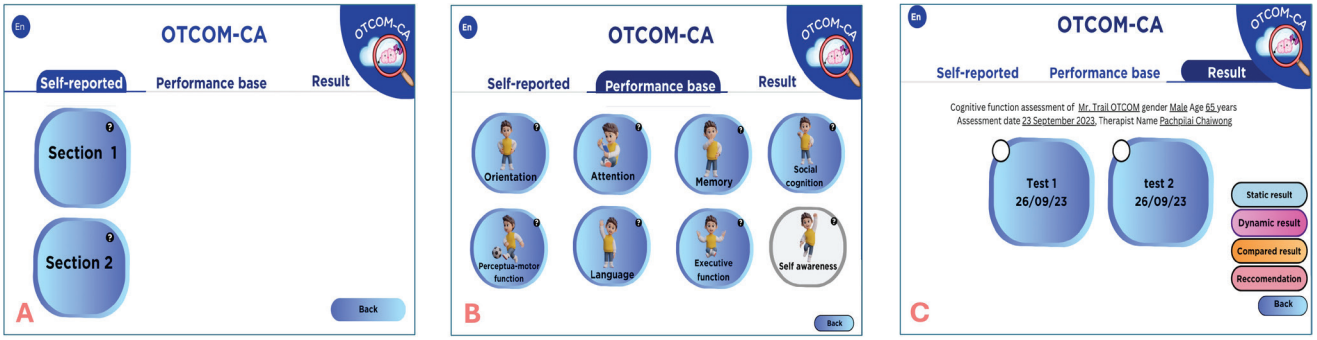


Figure 3. *Example of the main page screen of the Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA). A: Self-report part main page; B: Performance-based part main page; C: Result main page.*

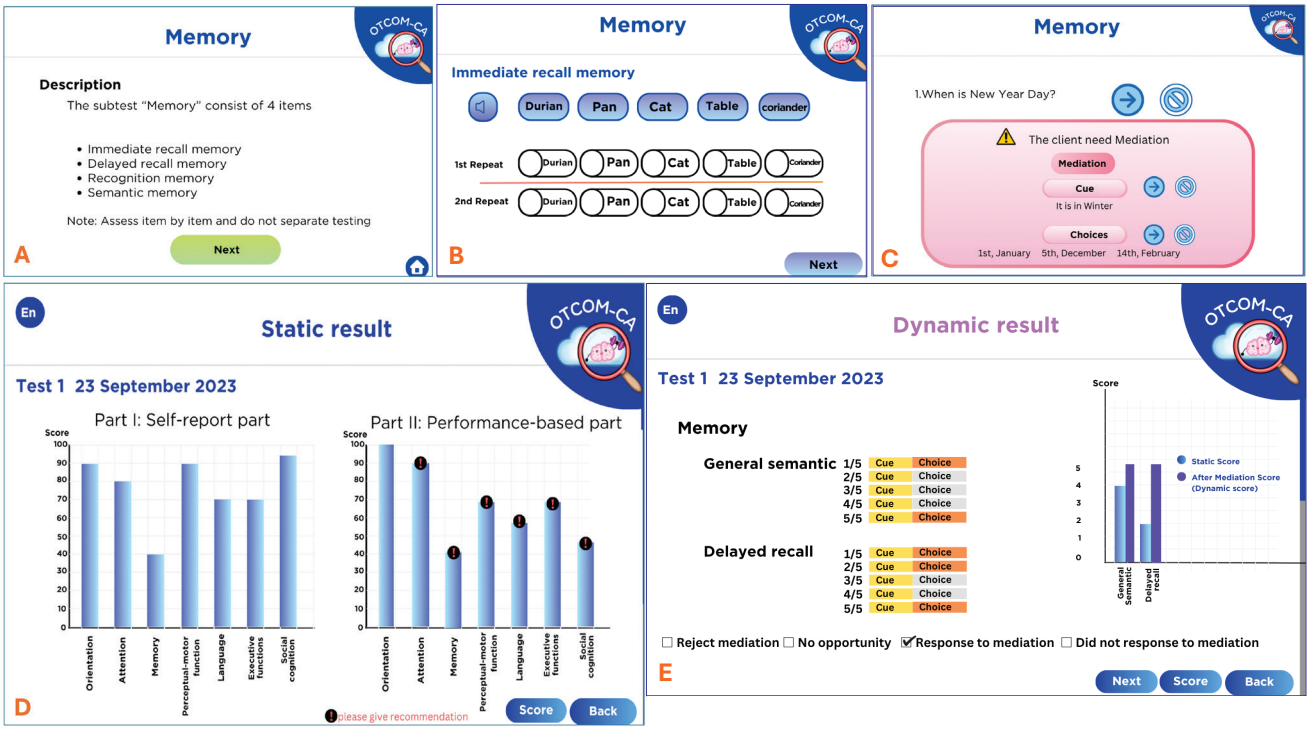


Figure 4. *Example of subsequent page screens of the Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA). A: description page of memory subtest, B: test page of immediate recall memory item, C: mediation page of semantic memory item, D: static result page, E: dynamic result page.*

In the pilot study, most participants were female (N=8, 80%), had fewer than five years of education (n=5, 50%), and had a mean cognitive screening test score of 26.8. (Table 2).

Table 2. Demographic characteristics of participants in pilot use (N=10)

Demographic		f	%	Mean±SD	Min-Max
Gender	Male	2	20.00		
	Female	8	80.00		
Age (years)	60-65	5	50.00		
	66-70	4	40.00	66.40±3.92	62-74
	71-75	1	10.00		
Education (grade)	1-4	5	50.00		
	5-8	0	0.00		2-18
	9-12	1	10.00		
	13+	4	40.00		
Cognitive screening				26.8±1.62	25-30

For the self-report part, most participants (37.90%) understood all 11 items immediately or after additional explanation. However, approximately 10-40% of participants required further examples for other items. For the static performance-based section, 53.8% of the participants understood all 14 test instructions either immediately or after additional explanation, whereas approximately 10-30% required additional examples for some instructions. The main difficulties were related to identifying the location of on-screen buttons and objects, unclear images, small image size, misspelled text, inconsistent on-screen display time for response options, and occasionally unclear explanations provided by the researchers. Some participants also noted that the proverbs used in the 'Abstract Thinking' subtest might be difficult for older adults with lower educational levels. In addition, some participants were unfamiliar with using a stylus and

suggested adding an extra practice example for certain tests. For the dynamic performance-based part, it was found that nearly half of the participants (40.00%) did not achieve a full score in the static part; as a result, they required 'Mediation' in the dynamic part, except for the 'General semantic memory' and the 'Use of objects Praxis' subtest. Most participants required mediation at all levels for the 'Delayed Recall Memory' subtest. Nearly all participants understood mediation either immediately or after further explanation; a small number required repeated explanations due to unclear researcher instructions, unfamiliarity with the proverb in an item of the 'Abstract Thinking' subtest, and a need for additional practice examples.

Discussion

This study applied a research and development design. The content and structure of the Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA) were developed for this study. After the content validity examination, the workflow of the OTCOM-CA Android application was planned, and each responding screen in the workflow was designed based on user experience (UX) and user interface (UI) principles. The workflow and response screens were then validated. Subsequently, a pilot study was conducted with ten healthy elderly individuals to assess the workflow and language validation.

The content validity evaluation of the OTCOM-CA's structure, workflow, and screen design demonstrated acceptable ratings for all items, with an Index of Item-Objective Congruence (IOC) ranging from 0.66 to 1.00. This outcome can be attributed to the well-designed structure of the OTCOM-CA, which was based on the previous standard version. Although some items were revised to include additional examples or explanations in the directions, these revisions were made at the experts' suggestion. The workflow and screen design can be attributed to the researcher's meticulous process, which involved creating a prototype, followed by iterative improvements and a final design phase. The structure of the workflow and screen design adhered to the UX principles established by Falmouth University and UI principles by Figma.[17,18] The use of Canva, an application known for its diverse, contemporary templates and illustrations for presentation creation,[19] contributed to the logical and appropriate linking of screens. This facilitated a clear and comprehensive view of the application's functionality. The application's screen design featured aesthetically pleasing visual elements, with clear fonts, sizes, and colors for images, buttons, and text, making it easy on the eyes and simple to read. These elements effectively highlighted important components. Furthermore, components such as buttons, images, text, and symbols were strategically positioned, making the interface intuitive and user-friendly, and promoting continuous use.

During the pilot use, most participants understood both the workflow and the test instructions. This is because of the well-developed process, as previously mentioned. However, there were some recommendations from the participants, including adding explanations for the button, picture size, and duration of the object's popping up. As a result, the pictures and buttons should be resized by at least 50 pixels to achieve an appropriate size for a mobile application designed for older adults.[20] The size and placement of buttons and display images should be adjusted to emphasize important elements and be strategically positioned for simplicity and usability. Moreover, additional test examples should be added. These features promote continuous use, prevent confusion during operation, and enhance the UX-UI experience.[18] Further, the researchers recommend that future users of the OTCOM-CA complete certification training to explain the instructions, ensuring that clients fully understand them before conducting the actual assessment. It was found that each multiple-choice subtest had an unequal duration. This results in test bias, causing participants to be confused about how to perform the test. The next version of OTCOM-CA should revise the pop-up duration to eliminate this test bias. Other psychometric properties will then be examined.

Limitations

The presentation of the workflow through the Figma application, although useful for illustrating the overall process, was highly complex and not easily accessible. This complexity makes it difficult for experts to review and evaluate the workflow. Consequently, the researcher developed an additional large-format cardboard workflow board to facilitate a clearer visualization of the overall structure for expert assessment.

Conclusion

The Occupational Therapy Cognitive Outcome Measure-Computer-Based Assessment (OTCOM-CA) has an acceptable content validity. The prototype Android version has a good workflow and language validation process. After revisions have been made, the psychometric properties should be evaluated.

Ethical approval

Ethical approval was obtained by the Ethics Committee at the Faculty of Associated Medical Sciences, Chiang Mai University (approval#AMSEC- 67EX-019). All participants signed a consent form after a verbal and written explanation of the study was provided.

Funding

This research was financially supported by Chiang Mai University and the Faculty of Associated Medical Sciences, Chiang Mai University, Thailand.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

CRediT authorship contribution statement

Pachpilai Chaiwong: conceptualization (lead), data curation (equal), formal analysis (equal), funding acquisition (equal), investigation (equal), methodology (equal), project administration; **Hataichanok Apikomonkon**: methodology (equal), writing: original draft preparation; **Anuchart Kaunnil**: formal analysis (equal), methodology (equal), writing: review and edit; **Sulalita Hongnueng**: conceptualization (supporting), data curation (equal), investigation (equal); **Peeraya Munkhetvit**: formal analysis (equal), funding acquisition (equal), methodology (equal), supervision.

Acknowledgements

The researchers extend their gratitude to all participants for their time and involvement in this study.

References

[1] United Nations Department of Economic and Social Affairs. World population aging 2020 highlights: living arrangements of older persons. New York: Population Division; 2020.

[2] Department of Older Persons. Elderly statistics 2024 [Internet]. 2024 [cited 2026 Apr 30]. Available from: https://www.dop.go.th/en/statistics_page?cat= 1&id=2559

[3] Foundation of Thai Gerontology Research and Development Institute (TGRI). Situation of the Thai older persons 2021. Nakhon Pathom: Institute for Population and Social Research, Mahidol University; 2022.

[4] Eshkoor SA, Hamid TA, Mun CY, Ng CK. Mild cognitive impairment and its management in older people. Clin Interv Aging. 2015; 10: 687-93. doi: 10.2147/CIA.S73922.

[5] O'Sullivan G, Hocking C. Translating action research into practice: seeking occupational justice for people with dementia. OTJR Occup Particip Health. 2013; 33(3): 168-76. doi: 10.3928/15394492-20130614-02.

[6] Lincoln N, das Nair R. Outcome measurement in cognitive neurorehabilitation. In: Cognitive neurorehabilitation: evidence and application [Internet]. 2nd ed. New York: Cambridge University Press; 2010 [cited 2026 Apr 30]. pp 91-105. doi: 10.1017/CBO9780511782015.007.

[7] Chaiwong P, Rattakorn P, Sungkarat S, Kattiya N, Tuttle D, Munkhetvit P. Psychometric properties of a new occupational therapy cognitive outcome measure for Thai older adults with cognitive impairments. J Assoc Med Sci. 2023; 56(1): 26-34. doi: 10.14456/jams.2023.4.

[8] Cantillon P, Irish B, Sales D. Using computers for assessment in medicine. BMJ. 2004; 329(7466): 606-9. doi: 10.1136/bmj.329.7466.606.

[9] Wilson AS, Kitas GD, Carruthers DM, Reay C, Skan J, Harris S, et al. Computerized information-gathering in specialist rheumatology clinics: an initial evaluation of an electronic version of the Short Form 36. Rheumatology (Oxford). 2002; 41(3): 268-73. doi: 10.1093/rheumatology/41.3.268.

[10] National Statistical Office. The 2024 household survey on the use of information and communication technology (Quarter 3). Bangkok: Ministry of Digital Economy and Society; 2024.

[11] Hogue RJ, editor. Tablet use within medicine. In: IADIS International Conference on Mobile Learning; 2013 Mar 14-16; Lisbon, Portugal.

[12] Ramprasad C, Tamariz L, Garcia-Barcena J, Nemeth Z, Palacio A. The use of tablet technology by older adults in health care settings-is it effective and satisfying? A systematic review and meta-analysis. Clin Gerontol. 2019; 42(1): 17-26. doi: 10.1080/07317115.2018.1457030.

[13] Statista. Market share of mobile operating systems worldwide from 2009 to 2024, by quarter 2025 [Internet]. 2025 [cited 2026 Apr 30]. Available from: https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/

[14] Nasreddine Z. MoCA clinic data 2024 [Internet]. 2024 [cited 2026 Apr 30]. Available from: https://www.mocatest.org/moca-clinic-data/

[15] Furr RM. Scale construction and psychometrics for social and personality psychology. New Delhi: Sage Publications; 2011. doi: 10.4135/9781446287910.

[16] DeVellis RF. Scale development: theory and applications. 4th Ed. Thousand Oaks (CA): SAGE Publications; 2016.

[17] Figma. Seven essential UI design principles + how to use them [Internet]. 2024 [cited 2026 Apr 30]. Available from: https://www.figma.com/resource-library/ui-design-principles/

[18] Hazzard L. The 6 key UX design principles to follow [Internet]. Falmouth: Falmouth University; 2024 [cited 2026 Apr 30]. Available from: https://www.falmouth.ac.uk/news/6-key-ux-design-principles-follow

[19] Teaching and Learning Innovation Center (TLIC). Knowing Canva [Internet]. Chiang Mai: Chiang Mai University [cited 2026 Apr 30]. Available from: https://docs.tlic.cmu.ac.th/canva-for-education

[20] Ketchaikosol V. A study of action-button size on mobile applications for elderly [master's thesis]. Bangkok: Bangkok University; 2016.

Journal of Associated Medical Sciences 2026; 59(3): 10-22

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



Altered visfatin and adiponectin levels and their interaction with HEGF in the metabolic dyslipidemia of overt hypothyroidism: A case-control study

Murtadha Kareem Najim* and Abtisam Faris Al-Shukry

Department of Medical Laboratory Technology College of Health and Medical Technology Al-Furat Al-Awsat Technical University, Kufa, Iraq.

RETRACTED

ARTICLE INFO

Article history:
Received 31 October 2025
Accepted as revised 20 April 2026
Available 14 May 2026

Keywords:
Adiponectin, Visfatin, hypothyroidism, metabolic dysfunction, metabolic biomarker

ABSTRACT

Background: Hypothyroidism is a common endocrine disorder associated with significant metabolic dysregulation, including dyslipidemia. Adipokines such as visfatin and adiponectin are key regulators of metabolism, but their specific roles and potential as markers of metabolic dysfunction in overt, treatment-naive hypothyroidism require further elucidation.

Objectives: This study aimed to investigate the association between visfatin and adiponectin levels with hypothyroidism, assess their potential utility for metabolic risk assessment, and explore their associations with metabolic parameters and human epidermal growth factor (HEGF).

Materials and methods: This case-control study included 71 hypothyroid patients and 71 frequency-matched euthyroid controls. All patients were newly diagnosed and therapy-naive. Serum levels of adipokines (visfatin, adiponectin), HEGF, thyroid hormones, and lipids were assessed. Statistical analyses included ROC curves, Spearman's correlation, multivariate regression, and ANCOVA to adjust for BMI and age.

Results: After adjusting for BMI and age, hypothyroid patients had significantly lower visfatin (adjusted LS mean: 11.04±1.56 vs 19.36±1.59 ng/mL, $p<0.001$) and adiponectin (10.77±2.04 vs 23.61±2.09 ng/mL, $p<0.001$) compared to controls and exhibited atherogenic dyslipidemia (all $p<0.001$). A strong positive correlation was observed between visfatin and adiponectin ($\rho=0.533$ and $p<0.001$). Regression analysis revealed significant interactions in the hypothyroid group: visfatin×HEGF was associated with an increased LDL-C ($\beta=0.1306$, $p=0.0188$), while adiponectin×HEGF was associated with decreased LDL-C (=-0.0919, $p=0.0062$). Adiponectin showed a good screening for metabolic risk in hypothyroidism (AUC=0.80) compared to visfatin (AUC=0.68).

Conclusion: Hypothyroidism is associated with altered adipokine levels and a specific interplay with HEGF that influences LDL-C. Adiponectin shows promise as adjunctive marker for assessing metabolic dysfunction severity in hypothyroidism, and the HEGF-adipokine network represents a potential pathophysiological mechanism worthy of further investigation.

Introduction

Hypothyroidism is one of the most common endocrine conditions associated with an increased risk of hyperlipidemia, hypertension, atherosclerosis, and low-grade inflammation.[1] Fatigue, lethargy, weight gain, and sensitivity to cold are cardinal features of hypothyroidism; however, these symptoms are indicators, and serum thyroid function tests are often

* **Corresponding contributor.**
Author's Address: *Department of Medical Laboratory Technology College of Health and Medical Technology Al-Furat Al-Awsat Technical University, Kufa, Iraq.*
E-mail address: murtadha.najeem.chm@student.atu.edu.iq
doi: 10.66285/JAMS.2026.075
E-ISSN: 2539-6056

used to diagnose the condition based on biochemical criteria.[2] Thyroid hormones affect many aspects of lipid metabolism, including synthesis, mobilization, and degradation.[3] However, the precise mechanisms linking thyroid dysfunction to metabolic disturbances, particularly through adipokine-mediated pathways, remain incompletely understood.

Thyroid hormones act through both genomic and non-genomic pathways to modulate the expression and activity of key enzymes and receptors. Crucially, it upregulates the activity of 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR), cholesteryl ester transfer protein (CETP), hepatic lipase (HL), and lecithin-cholesterol acyltransferase (LCAT), furthermore, simultaneously increasing the expression of LDL receptors (LDL-R) via the activation of sterol regulatory element-binding protein 2 (SREBP-2), enhancing cholesterol clearance.[4] Thyroid hormone is critical for enhancing β-oxidation via the induction of hepatocyte autophagy.[5] Consequently, in primary hypothyroidism the level of thyroid hormones is reduced or deficient; thus, many enzymes and receptors that are involved in the metabolism of lipids are decreased. TSH stimulates lipolysis in adipocytes via phosphorylation of perilipin and hormone-sensitive lipase (HSL), increasing circulating FFAs.[6] In hepatocytes, TSH promotes triglyceride synthesis through the THR/cAMP/PKA/PPARγ and SREBP1C pathways. It also stimulates the production of apolipoprotein B.[7] These disturbances create a pro-atherogenic lipid profile, but the contribution of adipose tissue-derived factors-particularly adipokines-to this metabolic dysregulation is an emerging area of interest.

Adipokines can affect numerous processes, including the regulation of energy and appetite, lipid and glucose metabolism, insulin sensitivity, endothelial cell function, inflammation, angiogenesis, blood pressure, hemostasis, and the onset of atherosclerosis and metabolic syndrome.[9] Owing to their pro- and anti-inflammatory properties, adipokines play a vital role in coordinating the immune response and systemic metabolism.[9] The balance between pro- and anti-inflammatory adipokines is maintained in individuals with normal metabolic functions. However, when adipose tissue accumulates in obese individuals, this balance shifts towards pro-inflammatory mediators.[10] Given the well-established link between hypothyroidism and altered adipose tissue function.

Visfatin, a pro-inflammatory adipokine, is most likely generated and released by white adipose tissue, namely the visceral white adipose tissue (vWAT). Visfatin exists in two forms: intracellular and extracellular visfatin, also known as nicotinamide phosphoribosyltransferase (NAMPT). Both intracellular (iNAMPT), which is crucial for the biosynthesis of nicotinamide adenine dinucleotide (NAD^+), and extracellular (eNAMPT) pathways are linked to numerous hormone-like signaling pathways and intracellular signaling cascades.[11] According to previous studies, plasma visfatin levels are high in both hyperthyroid and hypothyroid conditions, but significantly normalize once euthyroid is achieved.[12] This implies that visfatin can dynamically react to changes in thyroid hormone levels. Research suggests that visfatin may help stabilize the effects of thyroid hormones without affecting anthropometric or inflammatory factors.[12]

Adiponectin is an adipokine secreted by white adipose tissue and has pleiotropic effects on several physiological processes.[13] After therapy, adiponectin levels in hypothyroidism patients have been shown to drop noticeably, which may indicate a connection between thyroid function and adiponectin alterations.[14] Due to possible feedback pathways between adiponectin and thyroid hormones, adiponectin may be a candidate for monitoring metabolic disturbances associated with hypothyroidism. This variability highlights the complexity of adipokine regulation and the need for studies that account for potential modulating factors. HEGF and its receptor (EGFR) have been implicated in various metabolic processes beyond cell growth. Preclinical studies suggest that EGFR signaling can influence adipocyte differentiation and may modulate hepatic lipid metabolism.[15] Moreover, there is emerging evidence linking thyroid status to the expression of epidermal growth factors.[16,17] Thyroid hormones modulate both the production of EGF and the expression of its receptors in various tissues. T3 has been shown to potentially down-regulate EGF receptor mRNA, thereby altering cellular sensitivity to growth factors.[18] Conversely, long-term exposure to EGF can attenuate thyroid hormone responses by decreasing the concentration of specific nuclear thyroid hormone receptors.[19] This interplay suggests that EGF may mediate some of the developmental and metabolic actions of thyroid hormones. We hypothesized that HEGF may be involved in the relationship between adipokines and dyslipidemia in the hypothyroid state. Adipokines are thought to mediate metabolic disorders associated with hypothyroidism, such as dyslipidemia and insulin resistance. Elevated TSH levels are linked to enhanced insulin resistance, and alterations in adipokines may be associated with these effects.[20] Clinical investigations have connected adipokine level alterations to lipid disorders among individuals with hypothyroidism, indicating that they play a role in their pathogenesis. Despite growing evidence, the direct causal involvement of adipokines in metabolic diseases associated with thyroid dysfunction remains unclear. In contrast, adiponectin and visfatin levels in hypothyroidism reflect systematic variance, which is complicated by an overlap with obesity, inflammation, insulin resistance, and possible differences in local versus systemic adipokine effects.[12]

To address these gaps, this study focuses on two main objectives. Primary (confirmatory) aims were 1) to compare circulating levels of visfatin and adiponectin

RETRACTED

between patients with overt, treatment-naive hypothyroidism and age- and sex-matched euthyroid controls, and 2) to characterize the lipid profile (total cholesterol, LDL-C, HDL-C, triglycerides, VLDL) in hypothyroid patients compared to controls. Secondary (exploratory) aims were 1) to investigate whether HEGF levels differ between groups and whether HEGF interacts with visfatin and adiponectin to associate with lipid parameters, particularly LDL-C, and 2) to assess the potential of adipokines as adjunctive markers for metabolic dysfunction.

Materials and methods

Study population

This case-control study recruited participants (aged 16 to 65 years) who were then classified into two groups based on their thyroid function status: a hypothyroid group (N=71) and a euthyroid control group (N=67). The hypothyroid group (N=71) consisted of patients newly diagnosed with primary hypothyroidism (confirmed by elevated TSH (>5 mIU/L) and low T3 and T4 levels), who were recruited before the initiation of levothyroxine replacement therapy. The euthyroid control group (N=67) was recruited from individuals attending routine health check-ups, who were confirmed to have TSH and thyroid hormone levels within the reference range (TSH between 0.5-5.0 mIU/L and total T4 between 60-135 nmol/L) and had no personal or family history of thyroid disease, autoimmune disorders, or use of medications known to affect thyroid function. The controls were frequency-matched to the case group by age and sex.

Biochemical assays

Serum adiponectin (SunLong Biotech, SL0068Hu), visfatin (SL0325Hu), and Human Epidermal Growth Factor HEGF (SL0666Hu) were quantified using sandwich ELISA. The intra- and inter-assay coefficients of variation (CVs) were <10% and <12 % for all assays, respectively. Thyroid hormones (TSH, T3, T4) were measured using a Chemiluminescent Immunoassay (CLIA) on an SNIBE Maglumi 800 analyzer, and lipids (Total Cholesterol, BIOLABO 80106; Triglycerides, LP80019) were assayed by enzymatic methods. HDL was measured directly; LDL and very low-density lipoprotein (VLDL) levels were calculated using the Friedewald formula (LDL=Total Cholesterol-HDL-(Triglycerides/5), and VLDL=Triglycerides/5).

Statistical framework

All statistical analyses were performed using GraphPad Prism 10.5 for descriptive statistics, non-parametric tests, correlation analyses, and ROC curve analysis. Data were assessed for normality using the Shapiro-Wilk test. Normally distributed data are represented as mean±SD. Non-normally distributed data are presented as median (range). Categorical variables are presented as counts (percentages). Group comparisons for continuous variables were made using the independent t-test or Mann-Whitney U test. To account for significant differences in BMI between groups, we additionally performed Analysis of Covariance (ANCOVA) with group as the fixed factor and BMI and age as covariates. Adjusted least-squares means (LS means) with standard errors were calculated, along with adjusted mean differences (β), 95% confidence intervals, and partial eta-squared (η^2) effect sizes. To control for multiple comparisons, False Discovery Rate (FDR) adjustment was applied using the Benjamini-Hochberg procedure. The Chi-square test was used for categorical variables. Correlations were assessed using Spearman's rank correlation coefficient (ρ) adjusted for potential confounders (BMI and age) in the correlation between visfatin and adiponectin, a partial correlation analysis was conducted. The strength of correlations was interpreted as: |ρ| <0.3 (weak), 0.3-0.5 (moderate), >0.5 (strong). The diagnostic performance of biomarkers was evaluated using the Receiver Operating Characteristic (ROC) curve analysis. The area under the curve (AUC) is reported with 95% confidence intervals (CI). The optimal cut-off value was determined by maximizing the Youden Index. Pairwise comparisons of AUCs were performed using DeLong's test.

To investigate effect modification, multiple linear regression models were constructed with LDL-C as the dependent variable. Models included main effects and their two-way interaction terms, with adjustment for age, sex, and BMI. A $p<0.05$ was considered significant.

Results

Baseline characteristics and thyroid hormone profiles

The clinical characteristics of the study participants are shown in Table 1. The control group was frequency-matched to the hypothyroid group for age and sex, as confirmed by the non-significant differences in these variables ($p>0.05$), However, Body Mass Index (BMI) was significantly higher in the hypothyroid group (25.95±4.63 kg/m^2) compared to the control group (24.20±3.36 kg/m^2, $p=0.027$). Therefore, all subsequent group comparisons were adjusted for BMI and age using ANCOVA to isolate the effects of thyroid status from those of adiposity.

RETRACTED

Table 1. Demographic and clinical characteristics of participants.

Variable	Hypothyroid patients (N=71)	Healthy controls (N=67)	*p* value*	95% CI
Age (years) Median (IQR)	43 (35-52.5)	36(33-46.2)	0.078	
Sex			0.857	
• Male	15	15		
• Female	56	52		
BMI (Mean±SD)	25.95 ±4.63	24.20±3.36	0.027	0.198-3.287

Note: *Data are presented as median (IQR) for non-normally distributed variables (age), mean±SD for [illegible]ally distributed variables (BMI), and counts (%) for categorical variables (sex). The p values were calculated using Mann-Whitney U test for age, the unpaired t-test for BMI, and the Chi-square test for sex.*

RETRACTED

The detailed distribution of thyroid hormone levels is provided in Table 2. As expected, the hypothyroid group had significantly higher median [IQR] of TSH levels (22.00 [11.00-43.00] mIU/L) and significantly lower Mean±SD of T4 levels 31.38±19.49 nmol/L [illegible] compared to [illegible] control group (TSH median [IQR]: (1.80 [1.2[illegible]3.00] mIU/L), T4 [illegible]n±SD 10[illegible]6±17.86 nmol/L, all $p<0.0001$) [illegible]nfirming the distinct biochemical profile[illegible] of the two groups. T3 results are shown in Table [illegible]

Table 2. Comparison of thyroid hormone profiles between the cont[illegible] and hypothyroidism group.

Hormone	Healthy co[illegible]ols (N=67) Median [illegible]QR] or Mean±SD	[illegible]ypothyroid p[illegible]ients (N=71) [illegible]edian [illegible]QR] or Mean±SD	[illegible]ue
TSH (mIU/L)	1.80 (1.23-3.00)	[illegible]0 (11.00-43.00)	<0.0001
T4 (nmol/L)	102.6±[illegible]86	31.38±19[illegible]	<0.0001
T3 (n[illegible]/L)	[illegible](1.00-[illegible])	0.96[illegible]0-1.20)	<0.0001

[illegible]: Data are present[illegible]as me[illegible]±SD or [illegible]dian (IQR). TSH: thyro[illegible]stimulating hormone, T3: triiodothyronine, T4: thvroxine. The [illegible].0001 for all [illegible]mparisons between grou[illegible]Mann-Whitney U test).

[illegible]sex-strati[illegible]ana[illegible]sis [illegible]nfirmed that the significant reduct[illegible]ns in vi[illegible]ti[illegible]nd a[illegible]nectin were pr[illegible]nt in both males [illegible]nd fe[illegible]les with hypoth[illegible]oidism compared to their [illegible]x-matched control[illegible](all $p<0.05$, Table 3). Vi[illegible]atin levels in males s[illegible]wed a median of 13.0 ng/mL ([illegible]4-33.7) in healthy control versus 10.5 ng/mL (7.2-16.[illegible] in hypoth[illegible]oid patients ($p=0.0268$); in females, values [illegible]f vi[illegible]tin were 15.4 ng/mL (6.6-88.9) versus 10.25 ng/mL (6.1-28.1) ($p=0.0002$), in controls and patients respectively. Similarly, adiponectin levels were reduced in patients: males had 14.7 ng/mL (6.8-36.1) versus 9.1 ng/mL (5.8-25.8) ($p=0.0464$); females showed 19.2 ng/mL (6.0-135.8) versus 8.6 ng/mL (2.3-80.7) ($p<0.0001$). In contrast, no significant differences in adipokine levels were observed between males and females within either the hypothyroid or control groups ($p>0.7$, our unpublished data), indicating that the hypothyroid state, not biological sex, was the primary factor associated with adipokine dysregulation. Sex-stratified analysis using unadjusted data showed similar patterns in both males and females. After adjusting for BMI and age in the main ANCOVA models (Table 4), the overall differences between groups remained significant, though the study was not powered for sex-stratified adjusted analyses.

Table 3. Sex-stratified analysis of adipokine levels.

Biomarker	Sex	Healthy controls (N=67) Median (range)	Hypothyroid patients (N=71) Median (range)	*p* value
Visfatin (ng/mL)	Male (N=15/15)	13.0 (7.4-33.7)	10.5 (7.2-16.3)	0.0268
	Female (N=52/56)	15.4 (6.6-88.9)	10.25 (6.1-28.1)	0.0002
Adiponectin (ng/mL)	Male (N=15/15)	14.7 (6.8-36.1)	9.1 (5.8-25.8)	0.0464
	Female (N=52/56)	19.2 (6.0-135.8)	8.6 (2.3-80.7)	<0.0001

Note: *Data are presented as median (range). The p values were calculated using the Mann-Whitney U test with $p<0.05$ compared to same-sex controls. No significant differences were observed between males and females within the same group (all $p>0.05$).*

Comparative analysis of biomarkers and lipid profiles

To account for the significant difference in BMI between groups, we performed ANCOVA adjusting for BMI and age. The fully adjusted results are presented in Table 4. After adjustment, the hypothyroid group maintained significantly lower levels of both visfatin and adiponectin. Hypothyroid patients had an adjusted least-squares (LS) mean visfatin concentration of 11.04±1.59 ng/mL compared to 19.36±1.59 ng/mL in controls (adjusted mean difference [β]=-8.32, FDR-adjusted p<0.001). Similarly, adiponectin remained significantly lower in the hypothyroid group (10.77±2.04 vs. 23.61±2.09 ng/mL; β=-12.84, FDR-adjusted p<0.001). The partial η^2 values (0.092 and 0.124, respectively) indicate moderate effect sizes, suggesting that thyroid dysfunction itself, rather than differences in adiposity, is a primary driver of these adipokine alterations. Interestingly, HEGF levels were modestly but significantly elevated in the hypothyroid group after adjustment (47.96±2.10 vs. 41.73±2.14 pg/mL, β=6.23, FDR-adjusted p=0.043). The lipid profile remained markedly atherogenic in hypothyroid patients even after controlling for BMI and age. Compared to controls, the hypothyroid group showed significantly higher adjusted LS means for total cholesterol (202.84±6.58 vs 148.63±6.73 mg/dL, β=54.22, FDR-adjusted p<0.001), triglycerides (186.45±8.34 vs 106.75±8.53 mg/dL, β=79.69, FDR-adjusted p<0.001), LDL-C (137.89±6.10 vs 92.12±6.24 mg/dL, β=45.7[illegible], FDR-adjusted p<0.001), and VLDL (37.30±1.67 vs 21.35±1.71 mg/dL, β=15.95, FDR-adjusted p<0.001), together with significantly reduced HDL-C (27.65±1.10 vs 35.15±1.12 mg/dL, β=-7.50, FDR-adjusted p<0.001).

Table 4. ANCOVA-adjusted comparison of biomarkers between hypothyroid and control groups.

Parameter	Control group LS Mean±SE⁺ (N=67)	Hypothyroid group LS Mean±SE⁺ (N=71)	Adjusted mean difference (β)*	t-value*	p-value*	Partial $\eta^{2‡}$	FDR adjusted p-value†
Adipokines and growth factors							
Visfatin (ng/mL)	19.36±1.59	11.04±1.56	-8.319	2.266	<0.001	0.092	<0.001
Adiponectin (ng/mL)	23.61±2.09	10.77±2.04	-12.838	2.965	<0.001	0.124	<0.001
EGF (pg/mL)	41.73±2.14	47.96±2.10	6.229	3.046	0.043	0.030	0.043
Lipid profile							
Total cholesterol (mg/dL)	148.63±6.73	202.84±6.58	54.218	9.565	<0.001	0.195	<0.001
Triglycerides (mg/dL)	106.75±8.53	186.45±8.34	79.694	12.117	<0.001	0.245	<0.001
HDL-C (mg/dL)	35.15±1.12	27.65±1.10	-7.500	1.593	<0.001	0.143	<0.001
LDL-C (mg/dL)	92.12±6.24	137.89±6.10	45.768	8.859	<0.001	0.167	<0.001
VLDL (mg/dL)	21.35±1.71	37.30±1.67	15.951	2.423	<0.001	0.246	< 0.001

Note: *LS Mean±SE: Least-squares mean and standard error from ANCOVA, adjusting for age and BMI. *β, SE, and p value. Adjusted mean difference, its standard error, and raw significance for the group effect. ‡Partial η^2: Effect size (proportion of variance associated with thyroid status independent of covariates). †FDR-adjusted p value: False discovery rate q-value via Benjamini-Hochberg procedure (significance at<0.05), LS: least-squares, SE: standard error, FDR: false discovery rate, TSH: thyroid-stimulating hormone, HEGF: human epidermal growth factor, HDL-C: high-density lipoprotein cholesterol, LDL-C: low-density lipoprotein cholesterol, VLDL: very low-density lipoprotein, BMI: body mass index, ANCOVA: analysis of covariance.*

Correlation analysis

After adjusting for BMI and age using partial correlation, a strong positive association was observed between circulating visfatin and adiponectin concentrations (Spearman's p=0.533, p<0.001) in the hypothyroid group (Table 5). This alignment suggests concurrent alteration of these adipokines during thyroid dysfunction. No significant correlations were detected between TSH and any of the investigated adipokines or HEGF (all p>0.05). HEGF itself showed no significant correlation with other biomarkers (all p≥0.071), suggesting its signaling pathways operate independently within the adipokine-thyroid axis in this context.

RETRACTED

Table 5. Correlation matrix of study variables in the hypothyroid group (N=71).

Variable	TSH	Visfatin	HEGF	Adiponectin	Age
TSH	-				
Visfatin	0.140 (0.245)	-			
HEGF	-0.145 (0.220)	0.176 (0.141)	-		
Adiponectin	-0.030 (0.805)	0.533* (<0.001)	0.107 (0.376)	-	
Age	0.060 (0.618)	-0.126 (0.295)	-0.167 (0.163)	0.010 (0.934)	

Note: *Values represent Spearman's correlation coefficient (ρ) with the corresponding p value in parentheses.*

As shown in Table 6, TSH demonstrated significant positive correlations with atherogenic lipid parameters, total cholesterol (ρ=0.243, p=0.041), and LDL-C (ρ=0.298, p=0.012). A distinct inverse relationship emerged between visfatin and triglyceride-rich lipoproteins, including triglycerides (ρ=-0.271, p=0.021) and VLDL (ρ=-0.270, p=0.022). This inverse association suggests a potential link between lower visfatin levels and higher triglyceride-rich lipoproteins in hypothyroidism. Adiponectin showed similar, near-significant inverse trends with triglycerides (ρ=-0.229, p=0.054) and VLDL (ρ=-0.226, p=0.057).

Table 6. Correlations of biomarkers and lipid markers in hypothyroid patients (N=71).

Variables	Cholesterol	Triglyceride	HDL-C	LDL-C	VLDL
TSH	ρ 0.243	0.028	-0.015	0.298	-0.027
	p value 0.041*	(0.814)	(0.898)	0.012*	(0.822)
Visfatin	ρ -0.133	-0.271	-0.182	0.008	-0.270
	p value (0.269)	0.022*	(0.128)	(0.945)	0.022*
HEGF	0.155	0.067	0.031	0.119	0.067
	p value 0.196	(0.575)	(0.795)	(0.321)	(0.574)
Adiponectin	ρ -0.151	-0.229	0.025	-0.094	-0.226
	p value (0.208)	(0.054)	(0.832)	(0.432)	(0.057)
Age	ρ 0.251	0.140	0.063	0.187	0.138
	p value 0.034*	(0.243)	(0.599)	(0.117)	(0.248)

RETRACTED

Diagnostic performance (ROC analysis)

The discriminative ability of adiponectin and visfatin for differentiating hypothyroid patients from euthyroid controls is presented in Table 7. Adiponectin demonstrated the highest overall discriminative ability, with an Area Under the Curve (AUC) of 0.80 (95% CI: 0.72-0.87, *p*<0.0001) (Figure 1). At its optimal cut-off value of <14.50 ng/mL, determined by maximizing the Youden Index, adiponectin achieved a sensitivity of 81.69% and a specificity of 68.66%. Visfatin exhibited more modest discriminative power with an AUC of 0.68 (95% CI: 0.59-0.77, *p*<0.001). At its optimal cut-off of 12.35, visfatin achieved a sensitivity of 75.00% and a specificity of 62.69%.

Table 7. ROC analysis for adipokine discriminative ability

Biomarker	AUC	Optimal Cut-off	Sensitivity	95% CI	Specificity	95% CI	Youden Index
Adiponectin	0.80	<14.50	81.69%	71.15%-88.98%	68.66%	56.80%-78.49%	50.35
Visfatin	0.68	12.35	75%	0.6345-0.8332	62.69%	0.5072-0.7328	37%

Note: *AUC values are interpreted as: 0.90-1.00=excellent; 0.80-0.90=good, 0.70-0.80=fair, 0.60-0.70=poor; <0.60=fail. Adiponectin demonstrated good discriminative ability (AUC=0.80), while visfatin showed poor discriminative ability (AUC=0.68).*

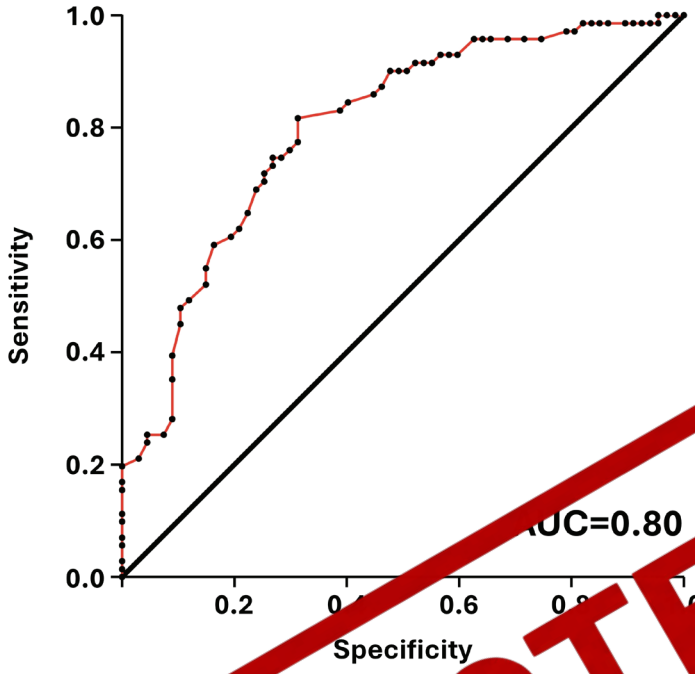


Figure 1. *ROC Curve for Adiponectin. The area under the curve (AUC) is 0.80 (p<0.0001). The optimal cut-off value is <14.50 ng/mL, yielding a sensitivity of 87% and a specificity of 68.7%.*

RETRACTED

Multivariate regression and interaction analysis

To investigate the interplay between adipokines and HEGF on lipid metabolism, this study conducted multiple linear regression models with LDL-C as the dependent variable, adjusting for age, sex, and BMI. The results are detailed in Table 8. Significant interaction effects between specific adipokines and HEGF on LDL-C levels were observed exclusively in the hypothyroid group and not in the euthyroid controls. The Visfatin×HEGF interaction term demonstrated a positive β-coefficient (β=0.1306, p=0.0188), indicating that the combination of higher levels of both factors associated with increased LDL-C concentrations. Conversely, the Adiponectin×HEGF interaction term showed a negative β-coefficient (β=-0.0919, p=0.0062), suggesting that this combination is associated with reduced LDL-C levels. Visfatin also independently and positively predicted adiponectin levels across all subjects (β=0.0328, p<0.0001), reinforcing the strong association observed in the correlation analysis. The absence of these significant interactions in the control group (all p>0.05) underscores the disease-specific nature of this novel signaling network in hypothyroidism (Table 8).

Table 8. Multiple linear regression and interaction analyses for adipokines and lipid parameters

Analysis type	Group	Predictor/interaction	Outcome	β-coefficient (SE)	p value	95% CI for β
Multivariate regression	Combined†	TSH	Adipokines		NS	-
Linear regression	Combined†	Visfatin	Adiponectin	0.328 (0.096)	<0.0001	0.121 to 0.045
Interaction analysis	Combined†	Visfatin×HEGF	Adiponectin	-	NS	-
Interaction analysis	Hypothyroidism	Visfatin×HEGF	LDL-C	0.1306 (0.052)	0.0188	0.028 to 0.233
Interaction analysis	Hypothyroidism	Adiponectin×HEGF	LDL-C	-0.0919 (0.032)	0.0062	-0.155 to -0.029
Interaction analysis	Control	Visfatin×HEGF	LDL-C	-	NS	-
Interaction analysis	Control	Adiponectin×HEGF	LDL-C	-	NS	-

Note: *Values represent unstandardized regression coefficients (β) with standard errors (SE) in parentheses, p values, and 95% confidence intervals (CI) for statistically significant associations. All models were adjusted for age, sex, and BMI unless otherwise specified. Interaction terms were created by multiplying the centered values of the two variables of interest. Statistical significance was set at p<0.05. NS (non-significant) indicates p≥0.05. †Combined group includes both hypothyroid patients and euthyroid controls (N=138), with adjustment for group status in the model.*

RETRACTED

Discussion

This case-control study examined the associations of visfatin and adiponectin with metabolic dysfunction in overt, treatment-naive hypothyroidism. Our main observations are threefold. Firstly, patients with hypothyroidism had significantly lower circulating levels of both visfatin and adiponectin than euthyroid controls, accompanied by a pronounced atherogenic lipid profile (elevated total cholesterol, LDL-C, triglycerides, and VLDL, and reduced HDL-C). Secondly, we observed a strong positive correlation between visfatin and adiponectin specifically in the hypothyroid group, contrasting with the inverse relationship typically reported in obesity and diabetes. Thirdly, exploratory interaction analyses revealed disease-specific associations between HEGF and both adipokines in relation to LDL-C metabolism: the visfatin×HEGF interaction was positively associated with LDL-C, while the adiponectin×HEGF interaction was negatively associated with LDL-C. These findings provide new insights into the complex adipokine-HEGF-lipid axis in hypothyroidism.

Our observation of reduced adiponectin in hypothyroidism aligns with some reports but contrasts with others that found increased or unchanged levels.[21-23] Similarly, our finding of decreased visfatin is consistent with Bekaya *et al.*,[24] but differs from Caixàs *et al.*[25] who reported elevated in thyroid dysfunction. These discrepancies might be the consequence of different study populations, methods, and definitions of hypothyroidism (overt vs subclinical). In milder or subclinical hypothyroidism, the body may mount a compensatory or stress-related response, leading to an initial increase in certain adipokines like visfatin. This biphasic response to elevation in early or mild dysfunction, followed by suppression in advanced or severe dysfunction, may reconcile the conflicting literature. According to Sawicka *et al.*, visfatin levels in hypothyroid patients are influenced by both the presence of autoimmune diseases and the levels of free thyroid hormones.[25]

In hypothyroidism, changes in visfatin and adiponectin levels are associated with metabolic dysfunction (dyslipidemia), which is shown by elevated total cholesterol, triglycerides, LDL, VLDL, and reduced HDL levels in hypothyroid individuals. Low visfatin levels could adversely affect NAD^+ dependent energy metabolism, aggravating lipid buildup, whereas reduced adiponectin impairs high-density lipoprotein(HDL) production and fatty acid oxidation.[11] Adiponectin decreases plasma triglyceride (TG) level by promoting the breakdown of triglyceride-rich lipoproteins by upregulating the expression of very low density lipoprotein (VLDL) receptors and skeletal muscle lipoprotein lipase (LPL), which in turn results in very low density lipoprotein and triglyceride catabolism.[26] This mechanism aligns directly with our findings in the hypothyroid group; the ANCOVA-adjusted results (Table 4) confirmed that hypothyroid patients have significantly higher triglycerides and VLDL alongside reduced visfatin and adiponectin, independent of BMI and age. The negative correlation we observed between adiponectin and both TG and VLDL further supports this relationship, although it did not reach statistical significance, possibly due to sample size limitations. Similarly, visfatin showed a significant negative correlation with both TG and VLDL ($p<0.05$, Table 6), suggesting that reduced levels of both adipokines may contribute to the atherogenic dyslipidemia characteristic of hypothyroidism.

TSH was correlated significantly but with weak positive correlations with LDL-C (ρ=0.298, p=0.012) and total cholesterol (ρ=0.243, p=0.041), but not with adipokines (Table 6). Our finding is consistent with many studies, such as that reported by Wang *et al.*,[27] who found that lipid abnormalities emerge even with TSH values at the upper end of the normal range, becoming more pronounced in subclinical and overt hypothyroidism. The biochemical mechanism involves decreased LDL receptor activity, resulting in the reduced catabolism of LDL and intermediate-density lipoproteins(28). Moreover, thyroid hormones influence cholesterol production through hepatic HMG-CoA reductase regulation.[29] The absence of direct correlations between TSH and adipokines suggests that the adipokine alterations in hypothyroidism may not be directly driven by TSH itself, but rather by broader metabolic disturbances associated with the hypothyroid state, including changes in adipose tissue biology and insulin sensitivity. In contrast to the negative association commonly found in patients with obesity and type 2 diabetes.[30] Our hypothyroid group showed a significant positive correlation between visfatin and adiponectin levels (ρ=0.533, $p<0.001$). This is counterintuitive, as visfatin is largely considered pro-inflammatory while adiponectin is anti-inflammatory; they are often inversely correlated.[31] However, coordinated reduction and strong correlation suggest a potential common upstream regulator. This finding suggests that in the unique metabolic environment of hypothyroidism, these two adipokines may be subject to common regulatory influences rather than existing in a simple pro-inflammatory or anti-inflammatory.

Notably, this regulator is unlikely to be TSH itself, given the absence of significant correlations between TSH and either adipokine. We hypothesize that the profound metabolic derangements of overt hypothyroidism, including altered adipose tissue biology, insulin sensitivity, and energy metabolism, may shift the relationship between visfatin and adiponectin towards coordinated regulation. Similar coordinated responses have been observed in other distinct physiological states, such as normal pregnancy.[32] The BMI-adjusted nature of our analysis suggests that this finding is not merely a consequence of adiposity differences between groups.

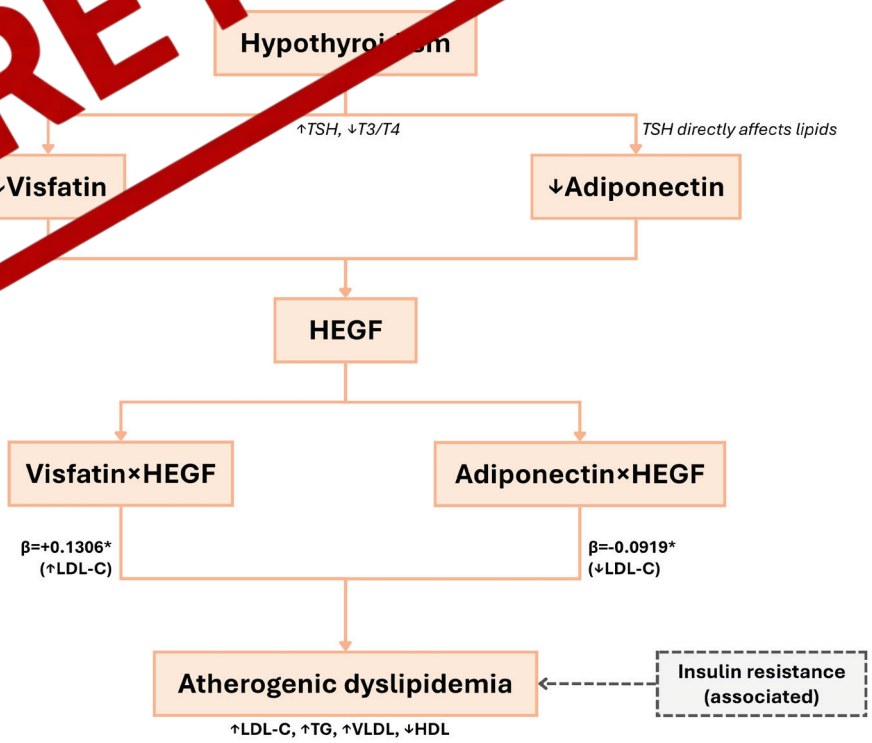
RETRACTED

Adiponectin emerged as the highest discriminative ability among the biomarkers studied, with an area under the curve (AUC) of 0.80 (Table 7 and Figure 1). Evaluating the diagnostic utility of these adipokines, our results demonstrated a moderate sensitivity of 81.7% and a modest specificity of 68.7%. While the sensitivity suggests these markers are reasonably capable of identifying metabolic alterations in hypothyroid patients, the specificity of 68.7% warrants cautious clinical interpretation. Crucially, this specificity translates to a false-positive rate of approximately 31.3%. In a practical clinical setting, utilizing adiponectin as a standalone diagnostic tool would result in nearly one-third of euthyroid individuals being incorrectly flagged for metabolic dysregulation. It should be utilized as part of a broader, comprehensive screening panel to aid in metabolic risk stratification alongside standard thyroid and lipid profiles. Its value likely lies in its ability to integrate information about thyroid dysfunction with downstream metabolic sequelae, including insulin resistance and inflammation.[33,34]

Regression analysis revealed significant interaction effects between specific adipokines and HEGF on LDL-C levels exclusively in the hypothyroid group. A statistically significant positive interaction was observed between visfatin and HEGF (β=0.1306, *p*=0.0188), indicating that high levels of this combination are associated with increased LDL concentrations. Conversely, a negative interaction was identified between adiponectin and HEGF (β=-0.0919, *p*=0.0062), signifying that this combination was correlated with reduced LDL-C levels (Table 8). A key finding of our study is the significant, disease-specific interaction between adipokines and HEGF on LDL-C levels. The significant interaction terms suggest that HEGF may influence the relationship between adipokines (both visfatin and adiponectin) and LDL-C metabolism, but only in the context of hypothyroidism. This raises the possibility that HEGF or the pathways it represents may function as a contextual factor in this setting.[35] In the euthyroid state, adipokine-HEGF interactions are quiescent. In the hypothyroid state, however, this axis becomes active, with HEGF potentially influencing both the visfatin-associated and adiponectin-associated effects on LDL-C.

Based on these findings, we propose a schematic mechanism summarizing the complex interplay between thyroid status, adipokines, and HEGF in hypothyroidism-associated dyslipidemia (Figure 2). In this proposed model, the hypothyroid state initiates reductions in both visfatin and adiponectin, which, together with HEGF, influence LDL-C metabolism through opposing interaction effects. This framework integrates our key observations and provides a visual hypothesis for how HEGF may act as a contextual influencing of adipokine function in the hypothyroid state, ultimately contributing to the observed atherogenic lipid profile.

RETRACTED



*p<0.05; interaction effects from multiple linear regression (adjusted for age, sex, BMI).

Figure 2. *Proposed schematic mechanism of adipokine-HEGF interactions in hypothyroidism-associated dyslipidemia.*

Limitations

There are several limitations of this study. Firstly, the case-control design precludes any conclusions about causality. Secondly, the BMI difference between groups to address this, we performed ANCOVA adjusting for BMI and age in all group comparisons (Table 4), and the significant differences persisted after adjustment. However, residual confounding cannot be entirely excluded, and future studies with BMI-matched controls would provide stronger evidence. Third, unmeasured confounders (insulin resistance, inflammation, lifestyle factors). Fourthly, the lack of interaction analyses for total cholesterol, triglycerides, and VLDL. Fifthly, the single-center design and modest sample size. Sixthly, the modest specificity of adiponectin and its implications; and finally, the exploratory nature of the HEGF findings and the need for validation in prospective studies.

Conclusion

In conclusion, this study demonstrates that overt, treatment-naive hypothyroidism is associated with significant metabolic dysregulation characterized by atherogenic dyslipidemia and marked reductions in serum visfatin and adiponectin levels even after controlling for differences in BMI and age. Furthermore, our findings highlight a potential interplay between altered HEGF expression and adipocyte dysfunction in the hypothyroid state, though these findings should be considered hypothesis-generating and require validation. However, given the modest sensitivity and modest specificity of visfatin and adiponectin should not be considered standalone diagnostic tools for metabolic complications.

Rather, a more cautious interpretation suggests they may serve as valuable adjunctive biomarkers, aiding in the broader metabolic risk stratification of hypothyroid patients. Future longitudinal and interventional studies are warranted to fully elucidate the prognostic value of these adipokines and to clarify the exact mechanistic pathways linking thyroid status, HEGF, and lipid metabolism.

Ethical approval

This study was approved by the Research Committee of the Thi-Qar Health Directorate, Training and Human Development Center (Approval No. 2025/20, dated January 14, 2025). The Committee confirmed compliance with ethical and scientific standards mandated by the Iraqi Ministry of Health and the Declaration of Helsinki. Written informed consent was obtained from all participants before enrollment. The fieldwork was authorized at Al-Rifae Teaching Hospital (Ref. No. 1642, dated January 14, 2025).

Conflict of interests

The authors have no conflicts of interest to declare relevant to the content of this article.

RETRACTED

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Murtadha Kareem Najim: conceptualization, methodology, investigation, formal analysis, writing: original draft, review and edit; **Abtisam Faris Al-Shukry**: investigation, resources, supervision, writing: review and edit.

Acknowledgements

We thank the staff of Al-Rifae Teaching Hospital for technical support.

References

[1] Verma DP, Chaudhary C, Singh A, Savyani KK, Gupta KK, Usman K, et al. Hypothyroidism in metabolic syndrome. Ann Med. 2024; 56(4): 717-2. doi: 10.4 8/aa aam_25_24. Epub 024 14.

[2] Taylor PN, Medici MM, Hubalewska-Dydejczyk A, Boelaert K. Hypothyroidism. Lancet (London, England). 2024; 404(1046): 1347-64. doi:10.1016/S0140-6736(24)016 3.

[3] Jonklaas J. Hypothyroidism, lipids, and lipidomics. Endocrine. 024; 84(2): 293-300. doi:10.1007/s12020-0 -03420-9 .

[4] Duntas LH, Brenta G. A renewed focus on the association between thyroid hormones and lipid metabolism. Front Endocrinol (Lausanne). 2018; 9: 511. doi: 10.3389/fendo.2018.00511.

[5] Sinha RA, You SH, Zhou J, Siddique MM, Bay BH, Zhu X, et al. Thyroid hormone stimulates hepatic lipid catabolism via activation of autophagy. J Clin Invest. 2012; 122(7): 2428-38. doi: 10.1172/JCI60580

[6] Gagnon A, Antunes TT, Ly T, Pongsuwan P, Gavin C, Lochnan HA, et al. Thyroid-stimulating hormone stimulates lipolysis in adipocytes in culture and raises serum free fatty acid levels in vivo. Metabolism. 2010; 59(4): 547-53. doi: 10.1016/j.metabol.2009.08.018.

[7] Yan F, Wang Q, Lu M, Chen W, Song Y, Jing F, et al. Thyrotropin increases hepatic triglyceride content through upregulation of SREBP-1c activity. J Hepatol. 2014; 61(6): 1358-64. doi: 10.1016/j.jhep.2014.06.037.

[8] Beukhof CM, Massolt ET, Visser TJ, Korevaar TIM, Medici M, de Herder WW, et al. Effects of thyrotropin on peripheral thyroid hormone metabolism and serum lipids. Thyroid. 2018; 28(2): 168-74. doi: 10.1089/thy.2017.0330.

[9] Kirichenko T V, Markina Y V, Bogatyreva AI, Tolstik T V, Varaeva YR, Starodubova A V. The role of adipokines in inflammatory mechanisms of obesity. Int J Mol Sci. 2022; 23(23): 14982. doi: 10.3390/

ijms232314982.
[10] Yao J, Wu D, Qiu Y. Adipose tissue macrophage in obesity-associated metabolic diseases. Front Immunol. 2022; 13: 977485. doi: 10.3389/fimmu.2022.977485.
[11] Dakroub A, Nasser SA, Kobeissy F, Yassine HM, Orekhov A, Sharifi-Rad J, et al. Visfatin: An emerging adipocytokine bridging the gap in the evolution of cardiovascular diseases. J Cell Physiol. 2021; 236(9): 6282-96. doi: 10.1002/jcp.30345.
[12] Caixàs A, Tirado R, Vendrell J, Gallart L, Megía A, Simón I, et al. Plasma visfatin concentrations increase in both hyper and hypothyroid subjects after normalization of thyroid function and are not related to insulin resistance, anthropometric or inflammatory parameters. Clin Endocrinol (Oxf). 2009; 71(5): 733-8. doi:10.1111/j.1365-2265.2009.03546.x.
[13] Brochu-Gaudreau K, Rehfeldt C, Blouin R, Bordignon V, Murphy BD, Palin MF. Adiponectin action from head to toe. Endocrine. 2010; [illegible](1): 11-32. doi: 10.1007/s12020-009-9278-8.
[14] Ozdemir D, Dagdelen S, Usm[illegible] A. Serum adiponectin levels and changes in glucose metabolism before and after treatment for thyroid dysfunction. Intern Med. 2015; 54(15): [illegible]9-57. doi: 10.2169/internalmedicine.54.06[illegible]
[15] Bhushan B, Apte[illegible]ichalopoulos GK, [illegible]ole of hepatocyte growth factor receptor in liver injury and lipid metabolism: Emerging new roles for an old receptor. Chem Biol Interact. 2020; 324: 109090. doi: 10.1016/j.cbi.2020.109090.
[16] Vi[illegible]awakarma J, Gupta K, [illegible]ira J, Garg A, Malik [illegible], Kashyap A, et al. Hypothyroidism induces motor deficits via altered cerebellar HB-EGF/EGFR and autophagy. J Endocrinol. 2023; 257(1). doi: 10.1530/JOE-22-0338.
[17] Schiff BA, McMurphy [illegible], Jasser SA, Younes MN, Doan D, Yeatba[illegible]OG, et al. Epidermal growth factor receptor (EGFR) is overexpressed in anaplastic thyroid cancer, and the EGFR inhibitor gefitinib inhibits the growth of anaplastic thyroid cancer. Clin cancer Res an Off J Am Assoc Cancer Res. 2004; 10(24): 8594-602. doi: 10.1158/1078-0432.CCR-04-0690.
[18] Mukku VR. Regulation of epidermal growth factor receptor levels by thyroid hormone. J Biol Chem. 1984; 259(10): 6543-7. PMID: 6327683.
[19] Kaji H, Hinkle PM. Epidermal growth factor decreases thyroid hormone receptors and attenuates thyroid hormone responses in GH4C1 cells. Endocrinology. 1987; 120(2): 537-43. doi: 10.1210/endo-120-2-537.
[20] Ekinci I. Metabolic Syndrome In Thyroid Disease. In: Metabolic syndrome: A comprehensive update with new insights. Bentham Science Publishers; 2025. pp 402-20.
[21] Kar K, Sinha S. Variations of adipokines and insulin resistance in primary hypothyroidism. J Clin Diagnostic Res. 2017; 11(8): BC07-9. doi: 10.7860/JCDR/2017/26666.10345.
[22] Lozanov LB, Gorcheva D, Lozanov BS, Argirova R, Koleva V, Kostova B, et al. Role of Thyroid deficiency on adiponectin, leptin, and metabolic status in visceral obesity: A cross-sectional study. Horm Metab Res. 2017; 49(9): 667-72. doi: 10.1055/s-0043-115532.
[23] Kokkinos S, Papazoglou D, Zisimopoulos A, Papanas N, Tiaka E, Antonoglou C, et al. Retinol binding protein-4 and adipo[illegible]n levels inthyroid overt and subclinic[illegible]dysfun[illegible]ion. Exp Clin Endocrinol Diabe[illegible]. 2015; 124([illegible] 87-92. doi: 10.1055/s-00[illegible]1564199.
[24] Ozkaya M, Sahin M, Cakal E, Yuzbasi[illegible]lu F, Sezer K, [illegible]c M, et al. Visfati[illegible]tas[illegible]n concen[illegible]ations in patients with hype[illegible]yro[illegible]ism a[illegible]d hypoth[illegible]roidism before and a[illegible]er contr[illegible] of thyroid fu[illegible]ction. J Endoc[illegible]ol [illegible]. 200[illegible]2(5): 435-9. doi: 10.1[illegible]7/[illegible]03346[illegible]2.
[25] [illegible]icka-C[illegible]aj N, [illegible]ybek-Kocik A, Kli[illegible]wicz A, Kloska M, [illegible]ańkowska-Wierzbic[illegible]D, Sowiński J, et [illegible] Circulating visfatin in hypothyroidism is associated with free thyroid hormones and antithyroperoxidase antibodies. Int J Endocrinol. 2016; 2016(1): [illegible]2469. doi:10.1155/2016/7402469.
[26] Qiao L, Z[illegible]C, van der Westhuyzen DR, Shao J. Adip[illegible]ectin reduces plasma triglyceride by increasing VLDL triglyceride catabolism. Diabetes. 2008; 57(7): 1824-33. doi: 10.2337/db07-0435.
[27] Wang F, Tan Y, Wang C, Zhang X, Zhao Y, Song X, et al. Thyroid-stimulating hormone levels within the reference range are associated with serum lipid profiles independent of thyroid hormones. J Clin Endocrinol Metab. 2012; 97(8): 2724–31. doi: 10.1210/jc.2012-1133.
[28] Rizos C V, Elisaf MS, Liberopoulos EN. Effects of thyroid dysfunction on lipid profile Open Cardiovasc Med J. 2011; 5: 76. doi: 10.2174/1874192401105010076.
[29] Zhang X, Song Y, Feng M, Zhou X, Lu Y, Gao L, et al. Thyroid-stimulating hormone decreases HMG-CoA reductase phosphorylation via AMP-activated protein kinase in the liver. J Lipid Res. 2015; 56(5): 963-71. doi: 10.1194/jlr.M047654.
[30] Zakaria I, Rizk MN, Rakha M, Shehaby A. Visfatin and adiponectin as early markers of atherosclerosis in type 2 diabetes mellitus. Egypt J Intern Med. 2013; 25: 143-8. doi: 10.7123/01.EJIM.0000433121.61723.90.
[31] Fantuzzi G. Adiponectin and inflammation: Consensus and controversy. J Allergy Clin Immunol. 2008; 121(2): 326-30. doi: 10.1016/j.jaci.2007.10.018.
[32] Souvannavong-Vilivong X, Sitticharoon C, Klinjampa R, Keadkraichaiwat I, Sripong C, Chatree S, et al. Placental expressions and serum levels of

RETRACTED

adiponectin, visfatin, and omentin in GDM. Acta Diabetol. 2019; 56: 1121-31. doi: 10.1007/s00592-019-01355-0.

[33] Abdalla MMI. Therapeutic potential of adiponectin in prediabetes: strategies, challenges, and future directions. Ther Adv Endocrinol Metab. 2024; 15: 20420188231222372. doi: 10.1177/20420188231222371.

[34] Al-rubaye AAH, Mohsin AA, Jasim WE. Circulating myokines and apelin-13 levels as predictive biomarkers for newly diagnosed type 2 diabetes mellitus. Indian J Biochem Biophys. 2025; 62: 798-805. doi: 10.56042/ijbb.v62i7.17204.

[35] Tomas A, Futter CE, Eden ER. EGF receptor trafficking: consequences for signaling and cancer. Trends Cell Biol. 2014; 24(1): 26-34. doi: 10.1016/j.tcb.2013.11.002.



Journal of Associated Medical Sciences 2026; 59(3): 23-30

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



Preliminary study: Effect of breathing exercise on hemodynamics and heart rate variability in hypertension

Sirinut Chaiduang[1] and Jirawat Wattanapanyawech[2*]

[1]Department of Physical Therapy, Faculty of Medicine, Prince of Songkla University, Songkhla Province, Thailand.
[2]Department of Physical Therapy, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 19 January 2026
Accepted as revised 29 April 2026
Available online 18 May 2026

Keywords:
Breathing exercise, hypertension, hemodynamics, heart rate variability parameter.

ABSTRACT

Background: Various types of breathing exercises (BE) have been investigated, mostly focusing on their effects on blood pressure (BP) in hypertension (HT). However, each study has had a different BE program, and strong evidence based on a standard protocol is lacking.

Objective: The present study was designed to compare the effects of BE across different breathing patterns within the same protocol on hemodynamic and heart rate variability (HRV) parameters.

Materials and methods: The study was a randomized crossover design. 20 hypertensive patients performed deep and slow, pursed-lip, and bee-humming breathing in a random sequence of the three breathing patterns. Hemodynamics and HRV parameters were measured before and immediately after exercise.

Results: At a rate of six breaths/min for ten minutes, there was a significant effect of time on SBP ($F(1,19)=5.17$, $p<0.05$). The post hoc Bonferroni test indicated a significant reduction in SBP during pursed-lip breathing compared to baseline ($p<0.01$). In addition, no significant differences were found in DBP, HR, and HRV parameters. Moreover, pursed-lip breathing showed the largest change in mean difference with 95% CI in SBP (-3.9 (-6.6, -1.1)). While deep and slow and bee-humming breathing were (-3.5(-7.4, 0.5)) and (-0.4(-4.4, 3.6)), respectively.

Conclusion: Among the three breathing exercises, only pursed-lip breathing significantly reduced SBP after a single 10-minute session at six breaths per minute, while no significant changes were observed in DBP, HR, or HRV parameters. These findings suggest that pursed-lip breathing may be the most effective technique for short-term BP reduction in hypertensive patients under standardized conditions.

Introduction

High BP, or HT, is a condition where the force of blood against the walls of arteries is consistently too high. BP is regulated by the autonomic nervous system (ANS), a function of sympathetic and parasympathetic activity. High BP is an autonomic imbalance characterized by increased sympathetic activity, with or without reduced parasympathetic activity.[1-3] Persistent elevated BP is a significant risk factor for stroke and cardiovascular disease, which leads to morbidity and mortality.[4] Antihypertensive drugs can reduce these risks. Unfortunately, either side effects or expenses are

* **Corresponding contributor.**
Author's Address: *Department of Physical Therapy, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand.*
E-mail address: Jirawat.W@chula.ac.th
doi: 10.66285/JAMS.2026.076
E-ISSN: 2539-6056

associated with the medications. Non-pharmacological approaches such as lifestyle modification, including dietary control and exercise, are recommended to control BP.[3,5]

BE can be a valuable addition to the overall strategy for managing high BP. BE is a contemporized technique, and it is interesting that it reduces sympathetic activation and increases baroreflex sensitivity.[2] There are several types of BE in the physical therapy field,[6-8] and various studies have claimed the positive effects of BE on BP regulation.[9-13] Deep, slow breathing can decrease BP in hypertensive patients after individual rates and depths are set for slow, rhythmic breathing.[14] Pursed-lip breathing was also reported to significantly lower BP in HT conditions. The participants counted two and four during inspiration and expiration, respectively. This technique involves either baroreflex or mindfulness relief.[15] We were also interested in one of the popular yoga breathing techniques in this study. Four to six breaths per minute (breaths/min) of bee-humming breathing showed positive effect in hypertensive patients and shifted the ANS toward a parasympathetic tone. The possible mechanism involves the baroreceptor and vibration, which enhance parasympathetic activation.[16] However, the standard time duration and rate of the BE are inconsistent.

Recent literature on deep and slow breathing suggests a breathing rate of less than 10 or 6 breaths/min for 10 minutes in a single session.[11] We also want to know whether this protocol is as effective as BE. In addition, the literature on the short-term effects of BE on BP and HRV should be expanded. Therefore, this study aims to compare the effect of deep and slow, pursed-lip and bee-humming breathing at a similar rate and duration (six breaths/min for ten minutes) on hemodynamic parameters, including BP and HR. Moreover, HRV parameters were rarely observed, as HRV is a physiological measure of variability between heartbeats and reflects ANS function in hypertensive patients.[1,17,18]

Materials and methods

Study design

The study was a randomized crossover design, randomizing the breathing pattern sequence. Participant recruitment began on 3rd December 2022 and concluded on 7th May 2023. The study has been executed in accordance with the Belmont Report 1979, the Declaration of Helsinki 2013, and the Council for International Organizations of Medical Sciences (CIOM) 2016 for human experiments. The committee approved the protocol and methodology. Previous studies have not reported any adverse effects related to the breathing patterns. The exercise session began after the participants provided informed consent (COA No.216/65).

Patient population

Both men and women aged 40-59 years old were eligible for enrollment in the study.[19] Other inclusion criteria were the essential hypertensive patients who received anti-hypertensive medications. The participants were willing to participate in the protocol. If the participants were unable to perform BE, they were excluded. Other exclusion criteria included smoking, secondary HT, chronic lung disease (asthma, chronic obstructive lung disease, or emphysema), and menopausal syndrome.

Experimental protocol

After meeting the inclusion and exclusion criteria and providing informed consent, participants were instructed to have at least six hours of adequate sleep and to refrain from alcohol, caffeinated beverages, and spicy foods for at least six hours prior to their appointment. They were also asked to take their antihypertensive medications as usual and to wear comfortable clothing on the day of the experiment.

Initially, the participants were assessed for general characteristics and health status. Subsequently, the participants were asked to firmly tie the chest strap (Polar H10) around their chest. The Polar application (HRV logger) was paired to a smartphone via Bluetooth. Then, the participants were asked to sit in a comfortable position (head, neck, shoulders, and hands relaxed). Next, the automatic sphygmomanometer was placed on the upper arm. Finally, the researcher taught each participant the breathing pattern. To ensure that all participants deeply inhale, slow vital capacity (SVC) was monitored during the BE session using a spirometer.

After 30 minutes of quietly sitting, an automatic sphygmomanometer measured the baseline hemodynamic parameters. The HRV parameters were obtained from the Polar logger application. Then, the participants were asked to randomly perform three BE patterns: deep and slow breathing, pursed-lip breathing, and bee-humming breathing, accompanied by a video guide. Then, post-test measurements were obtained. Between each breathing pattern, the participants rested for 30 minutes as a washout period between intervention before starting the next BE (Figure 1).

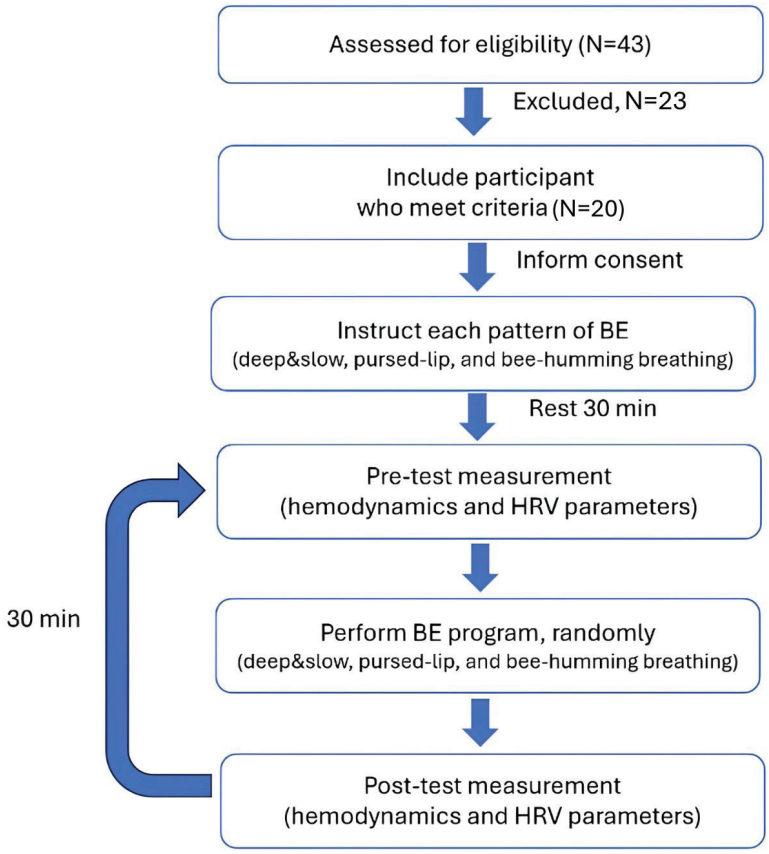


Figure 1. *Flow diagram of the experiment protocol.*

Interventions

Deep and slow breathing

The participant inhaled deeply and slowly through both nostrils for four seconds. Then, exhaled through both nostrils for six seconds. After 1 minute of deep, slow breathing, breathing control was performed at 12 breaths/min by breathing through both nostrils (inhale:exhale, 2:3), with the hand placed on the abdomen to feel the movement. This cycle was repeated for ten minutes (Figure 2A).

Pursed-lip breathing[15]

The participant inhaled deeply and slowly through both nostrils for four seconds and exhaled through the mouth while making an "O shape" with the lips for six seconds. After one minute of pursed-lip breathing, a minute of breathing control was performed. This cycle was repeated for ten minutes (Figure 2B).

Bee-humming breathing[16]

The participant inhaled through both nostrils for four seconds, followed by exhalation and the production of a humming "bee" sound ("MMMM") for six seconds. After one minute of bee-humming breathing, a minute of breathing control was performed. This cycle was repeated for ten minutes (Figure 2C). A beep sound from the video guide was used to help the patient control their breathing rate at the end of the session.

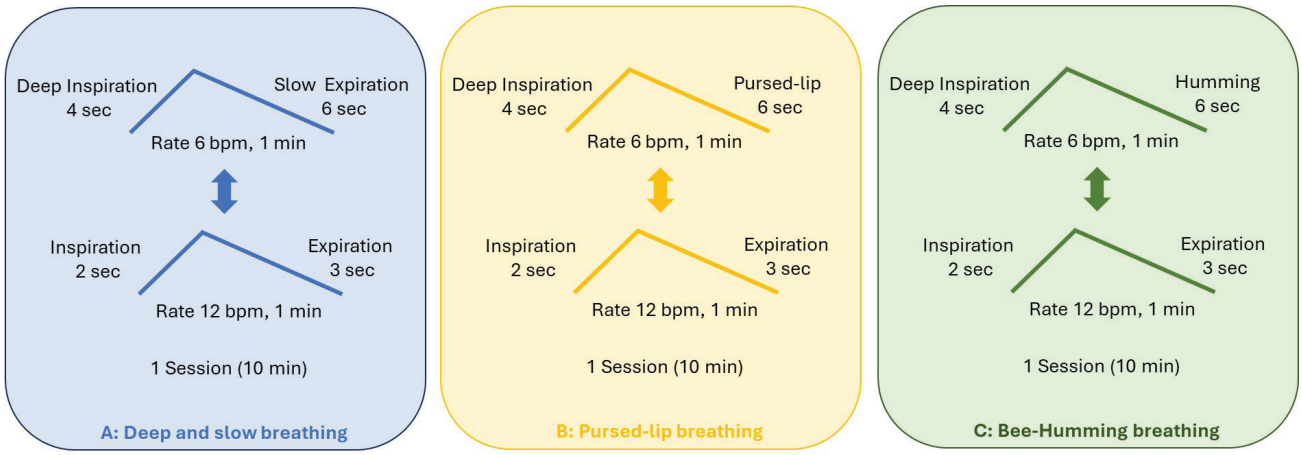


Figure 2. *BE protocol. A: deep and slow breathing, B: pursed-lip breathing, C: bee-humming breathing.*

Statistical analysis

The study outcomes were the percentage of sex, the number of antihypertensive drugs, the type of antihypertensive drugs, and other diseases. Besides, the mean±SEM of age, weight, height, body mass index (BMI), BP, HR, year since diagnosis, and HRV parameters were reported, as the SEM reflects the precision of the estimated mean and indicates how accurately the sample mean represents the population mean. Data analyses were performed using the statistical package (version 29.0; SPSS Inc, Chicago, IL). Descriptive statistics were computed to describe the sample. A Shapiro-Wilk test was conducted to ensure the distributions of the study variables. To assess whether hemodynamic or HRV parameters were modulated following the BE session, a two-way repeated-measures ANOVA was employed for parametric data, whereas nonparametric data were analyzed using the Friedman test. A p-value<0.05 was considered statistically significant. Based on the mean and SD from the previous study, the G*Power calculation indicated a minimum sample size of 20 participants.

Results

A total of 43 hypertensive patients were screened for the study. 20 fulfilled the eligibility criteria. They were instructed in three breathing patterns: deep and slow, pursed-lip, and bee-humming. The study design was a randomized crossover. Participants performed 10 minutes of BE, rested for half an hour, then repeated the process for all breathing patterns. The order of breathing patterns was randomly allocated to participants. The hemodynamics and HRV parameters were collected at the end of the exercise session. No participants were excluded during the protocol.

Descriptive data

Table 1 presents the general characteristics of the 20 hypertensive participants. The sample included equal numbers of men and women, with a mean age of 48.1±1.3 years and a mean BMI of 26.1±0.9 kg/m^2. Most participants (80%) were prescribed a single antihypertensive drug, predominantly calcium channel blockers (62.5%). BP at baseline was well-controlled, with a mean SBP of 133.9±3.6 mmHg and DBP of 87.6±2.5 mmHg. Two participants were accompanied by HT with hyperlipidemia, and one participant with diabetes mellitus (DM).

Table 1. General characteristics.

Characteristics	Mean±SEM	N (%)
Gender		
Male		10 (50.0)
Female		10 (50.0)
Age (years)	48.1±1.3	
Weight (kg)	70.3±3.1	
Height (cm)	163.7±1.5	
BMI (kg/m^2)	26.1±0.9	
SBP (mmHg)	133.9±3.6	

Table 1. General characteristics. (continue)

Characteristics	Mean±SEM	N (%)
DBP (mmHg)	87.6±2.5	
HR (bpm)	76.6±2.2	
Years since diagnosis	5.2±0.8	
Number of antihypertensive drugs		
• 1		16 (80.0)
• 2		4 (20.0)
Types of antihypertensive drugs		
• ACE Inhibitor		3 (12.5)
• Beta-blocker		1 (4.17)
• Calcium channel blocker		15 (62.5)
• Angiotensin II receptor blocker		5 (20.83)
Other diseases		
• Hyperlipidemia		2 (10.0)
• Diabetes mellitus		1 (5.0)

Note: *Value is mean±SEM or number (%), BMI: body mass index, SBP: systolic blood pressure, DBP: diastolic blood pressure, HR: heart rate, ACE Inhibitor: angiotensin converting enzyme inhibitor.*

Effect of breathing exercise on hemodynamics

An automatic BP monitor observed hemodynamics for three to four hours for all breathing patterns (Table 2). The analysis to assess the effect on SBP, DBP, and HR was performed. Two-way analysis of variance showed a significant effect of time on SBP ($F_{(1,19)}=5.17$, $p<0.05$). The post hoc Bonferroni test indicated that levels of SBP significantly decreased at the end of the 10-minute session ($p<0.01$) in pursed-lip breathing. However, there was no significant Intervention×Time interaction effect or intervention effect on SBP. No significant intervention, time, or Intervention×Time effect emerged for DBP and HR.

Table 2. Effect of breathing exercise on hemodynamics.

Parameters	Breathing technique					
	Deep and slow breathing		Pursed-lip breathing		Bee -humming breathing	
	Before	After	Before	After	Before	After
SBP (mmHg)	134.8±4.0	131.3±3.6	134.9±3.2	131.0±3.0*	133.0±4.0	132.6±3.5
DBP (mmHg)	89.5±2.9	88.2±2.8	88.1±2.6	89.0±2.6	88.1±2.6	88.8±2.8
HR (bpm)	74.7±1.8	73.9±1.8	73.3±2.3	73.3±2.0	74.6±2.0	74.8±2.0

Note: **p<0.05 indicates significant. Value is mean±SEM, SBP; systolic blood pressure, DBP; diastolic blood pressure, HR; heart rate.*

Effect of breathing exercise on HRV parameters

HRV parameters were uncertain (Table 3). The HRV parameters did not meet the assumption of normality; therefore, nonparametric tests were conducted. The test was analyzed to compare the effect of BE on HRV parameters. There was no significant difference in HRV parameters.

Table 3. Effect of breathing exercise on HRV parameters.

Parameters	Breathing technique					
	Deep and slow breathing		Pursed-lip breathing		Bee -humming breathing	
	Before	After	Before	After	Before	After
RR interval	778.59±25.03	777.92±21.00	787.67±25.09	790.68±22.06	789.66±24.82	796.83±23.88
SDNN	43.22±4.48	47.04±6.15	41.31±6.53	43.30±4.94	46.33±6.94	48.48±7.81
RMSSD	33.18±4.83	30.65±5.13	32.52±4.96	33.52±7.15	38.08±7.10	40.06±11.58
LF	0.03±0.00	0.03±0.00	0.03±0.01	0.03±0.01	0.03±0.01	0.02±0.00
HF	0.03±0.01	0.05±0.01	0.04±0.01	0.04±0.01	0.04±0.01	0.04±0.01
LF/HF ratio	1.18±0.15	0.88±0.08	1.09±0.19	0.94±0.15	1.23±0.24	0.72±0.09

Note: *Value is mean±SEM, SDNN: standard deviations of all the NN intervals, RMSSD: the root mean square of successive differences between normal heartbeats, LF: low frequency, HF: high frequency, LF/HF ratio: low frequency per high frequency ratio.*

Discussion

The study's objective aimed to compare the effect of BE (deep and slow, pursed-lip, and bee-humming breathing) on hemodynamics and HRV parameters in hypertensive patients. The time analysis found a short-term reduction in SBP, but it had no significant effect on DBP or HR. While there was no significant effect on HRV parameters.

SBP was significantly reduced at the end of the 10-minute session due to the effect of time on pursed-lip breathing. Like the previous study, which showed a significant reduction in SBP after performing ten minutes of BE.[20] Moreover, the investigation from Mitsungnern and colleagues revealed the statistical significance of SBP reduction in pursed-lip breathing exercises.[15] In addition, the investigation of our study further reported the value of the mean difference with 95% CI of SBP, including pursed-lip (-3.9(-6.6, -1.1)), deep and slow (-3.5(-7.4,0.5)), and bee-humming breathing (-0.4(-4.4, 3.6)). In this study, the mean difference in SBP reduction from pursed-lip breathing approached the minimal clinically important difference (MCID) threshold. Since SBP was the strongest predictor of stroke morbidity risk, a 4-mmHg reduction in average SBP is estimated to reduce stroke mortality by about 5.8-8.9%.[21] Therefore, although the present study found no significant intervention-by-time effect, the magnitude of reduction observed with pursed-lip breathing may still have potential clinical relevance and warrants further investigation in longer-term interventions.

Related to the study of BE on arterial baroreflex in hypertensive patients, the investigation noted that lowered BP during breathing practice is associated with increased vagal tone.[2] The following could be the mechanism of action for BE. High lung volume during inspiration expands the lung and compresses the heart chamber, thereby decreasing venous return. During expiration, the decrease in lung volume enhances venous return to the heart, resulting in an increased stroke volume. The subsequent rise in blood pressure stimulates baroreceptors, which inhibit sympathetic vasoconstrictor activity and enhance vagal innervation of the heart. Consequently, both blood pressure and heart rate decrease.[22] In addition, the pursed-lip creates back pressure, producing a small amount of positive end-expiratory pressure (PEEP), supporting the patency of the airways and alveoli in gas exchange and ventilation. Briefly, when blood gas levels are optimized, chemoreceptor activity is reduced. Since the baroreflex and chemoreflex systems interact to regulate autonomic balance, a diminished chemoreceptor drive allows baroreceptor-mediated parasympathetic activation to become more prominent. This shift leads to decreased sympathetic tone and ultimately lowers SBP.[22-24] While the postulate mechanism of bee-humming breathing reveals that the vibration of humming sound acts like an auditory stimulus that may stimulate the vagus nerve, contributing to parasympathetic tone. [16] However, the mechanism of bee-humming breathing lacks the mechanical effect of PEEP seen in pursed-lip breathing. This may explain why bee-humming breathing did not result in a significant reduction in SBP in the present study. Therefore, further research is needed.

Moreover, HRV parameters showed inconsistency across individual measures. Pursed-lip breathing appeared to increase RR interval, SDNN, and RMSSD, and slightly modulated the LF/HF ratio toward parasympathetic dominance. However, these changes were not statistically significant. According to the mechanisms discussed above regarding autonomic nervous system regulation, pursed-lip breathing may be more effective in influencing HRV responses. Although all three breathing techniques may share similar vagal stimulation effects, the scattered data observed with deep and slow breathing and bee-humming breathing limited the ability to draw definitive conclusions about autonomic modulation. Therefore, a single session may be insufficient to detect subtle autonomic changes. Future studies should include longer intervention durations and more sensitive measures to better differentiate the effects of each technique.

Limitations

For the parameters for which there was no significant finding, a previous study mentioned that ANS activity is unlikely to occur even if BE is performed in a stress state.[25] This study set up a spirometer to monitor breathing depth during BE. This discomfort may lead to stress responses. Participants may have consumed high-sodium foods, which may affect BP.[14] And the participants were properly taking anti-hypertensives that may attenuate the effect of BE.[16] Further study should aim to minimize stress during BE protocols. The biofeedback might be suitable for evaluating chest movement during BE. Food intake is recommended to control or retrain a high salt diet. The absence of a non-intervention or sham control group also limits the ability to attribute the observed effects solely to the breathing exercises. We recommend that future studies include a control group with matched time and conditions to better isolate the specific physiological effects of each breathing technique. Nevertheless, this randomized clinical trial demonstrates BE, which are easy to understand, quick to learn, and require no medical equipment. These findings underscore the need for further research to control potential confounding factors such as stress and diet, and to explore the use of biofeedback in evaluating BE.

Conclusion

This study found that among the three breathing exercises tested, only pursed-lip breathing effectively alleviates SBP after a single 10-minute session at six breaths per minute. Although deep and slow breathing and bee-humming breathing showed trends of improvement, the changes were not statistically significant. No significant effects were found in DBP, HR, or HRV parameters, possibly due to short-term intervention, antihypertensive medication use, or confounding factors such as stress or diet. Nonetheless, the modest but significant SBP reduction from pursed-lip breathing suggests it may be a simple and practical complementary approach for hypertensive patients, warranting further investigation in longer-term studies.

Ethical approval

This study was conducted in accordance with the ethical principles of the Belmont Report (1979), the Declaration of Helsinki (2013), and the Council for International Organizations of Medical Sciences (CIOMS) guidelines (2016) for research involving human participants. The study protocol and methodology were approved by the institutional ethics committee (COA No. 216/65). Written informed consent was obtained from all participants.

Funding

This research received no external funding.

Conflict of interest

The authors report no conflict of interest.

CRediT authorship contribution statement

Sirinut Chaiduang: conceptualization, methodology, investigation, data curation, formal analysis, writing: original draft and visualization, project administration; **Jirawat Wattanapanyawech**: supervision, validation, methodology, resources, writing: review and edit.

Acknowledgements

We would like to express our deep gratitude to the scholarship from the Graduate School, Chulalongkorn University, to commemorate the 72nd anniversary of His Majesty King Bhumibol Adulyadej. We sincerely thank all hypertensive patients who participated in this study. Finally, we thank Dr. James M. Brimson (Research Unit for Innovation and International Affairs, Faculty of Allied Health Sciences, Chulalongkorn University) for proofreading and editing the English of this manuscript.

References

[1] Bolin LP, Saul AD, Bethune Scroggs LL, Horne C. A pilot study investigating the relationship between heart rate variability and blood pressure in young adults at risk for cardiovascular disease. Clin Hypertens. 2022; 28(1): 1-8. doi: 10.1186/s40885-021-00185-z.

[2] Joseph CN, Porta C, Casucci G, Casiraghi N, Maffeis M, Rossi M, et al. Slow breathing improves arterial baroreflex sensitivity and decreases blood pressure in essential hypertension. Hypertension. 2005; 46(4): 714-8. doi: 10.1161/01.HYP.0000179581.68566.7d.

[3] Oparil S, Acelajado MC, Bakris GL, Berlowitz DR, Cífková R, Dominiczak AF, et al. Hypertension. Nat Rev Dis Primers. 2018; 4: 18014. doi: 10.1038/nrdp.2018.14.

[4] Brook RD, Julius S. Autonomic imbalance, hypertension, and cardiovascular risk. Am J Hypertens. 2000; 13(S4): 112S-22S. doi: 10.1016/S0895-7061(00)00228-4.

[5] Grossman E, Grossman A, Schein MH, Zimlichman R, Gavish B. Breathing-control lowers blood pressure. J Hum Hypertens. 2001;15(4):263-9. doi: 10.1038/sj.jhh.1001147.

[6] Solomen S, Aaron P. Breathing techniques-A review. Int J Phys Educ Sports Health. 2015; 2: 237-41.

[7] Innocenti D. An overview of the development of breathing exercises into the specialty of physiotherapy for heart and lung conditions. Physiotherapy. 1996; 82(12): 681-93. doi: 10.1016/S0031-9406(05)66377-8.

[8] Hoffman LA, Owens GR, Rogers RM, Mazzocco MC. Breathing exercises in chest physical therapy. Chest. 1987; 92(1): 191. doi: 10.1016/S0012-3692(16)37676-0.

[9] Brandani JZ, Mizuno J, Ciolac EG, Monteiro HL.

The hypotensive effect of Yoga's breathing exercises: A systematic review. Complement Ther Clin Pract. 2017; 28: 38-46. doi: 10.1016/j.ctcp.2017.05.002.

[10] Kuppusamy M, Kamaldeen D, Pitani R, Amaldas J, Shanmugam P. Effects of Bhramari Pranayama on health - A systematic review. J Tradit Complement Med. 2017; 8(1): 11-6. doi: 10.1016/j.jtcme.2017.02.003.

[11] Yau KK-Y, Loke AY. Effects of diaphragmatic deep breathing exercises on prehypertensive or hypertensive adults: A literature review. Complement Ther Clin Pract. 2021; 43: 101315. doi: 10.1016/j.ctcp.2021.101315.

[12] Chaddha A, Modaff D, Hooper-Lane C, Feldstein DA. Device and non-device-guided slow breathing to reduce blood pressure: A systematic review and meta-analysis. Complement Ther Med. 2019; 45: 179-84. doi: 10.1016/j.ctim.2019.03.005.

[13] Koul BR, Pity KP, Kota V. A systematic review and meta-analysis of effectiveness of breathing exercises on blood pressure of patients with hypertension. Int J Sci Eng Manag. 2020; 5(12): 51-6.

[14] Zahroh C, Nurjanah S, Widyarti NKS. Abdominal breathing affects blood pressure in hypertension sufferers. J Public Health Afr. 2019; 10(S1): 23-5. doi: 10.4081/jphia.2019.1174.

[15] Mitsungnern T, Srimookda N, Imoun S, Wansupong S, Kotruchin P. The effect of pursed-lip breathing combined with number counting on blood pressure and heart rate in hypertensive urgency patients: A randomized controlled trial. J Clin Hypertens (Greenwich). 2021; 23(3): 672-79. doi: 10.1111/jch.14168.

[16] Ghati N, Killa AK, Sharma G, Karunakaran B, Agarwal A, Mohanty S, et al. A randomized trial of the immediate effect of Bee-Humming Breathing exercise on blood pressure and heart rate variability in patients with essential hypertension. Explore. 2021; 17(4): 312-9. doi: 10.1016/j.explore.2020.03.009.

[17] Nuttaset M, Apiwan M, Rattana V, Jittima S, Naruemon L. Heart rate variability assessment and clinical uses. J Thai Rehabil Med. 2018; 28: 32-6. (inThai).

[18] Rajendra Acharya U, Paul Joseph K, Kannathal N, Lim CM, Suri JS. Heart rate variability: a review. Med Biol Eng Comput. 2006; 44(12): 1031-51. doi: 10.1007/s11517-006-0119-0.

[19] Fryar CD, Ostchega Y, Hales CM, Zhang G, Kruszon-Moran D. Hypertension prevalence and control among adults: United States, 2015-2016. NCHS Data Brief. 2017; (289): 1-8.

[20] Telles S, Yadav A, Kumar N, Sharma S, Visweshwaraiah NK, Balkrishna A. Blood pressure and purdue pegboard scores in individuals with hypertension after alternate nostril breathing, breath awareness, and no intervention. Med Sci Monit. 2013; 19: 61-6. doi: 10.12659/MSM.883743.

[21] Umemura S, Arima H, Arima S, Asayama K, Dohi Y, Hirooka Y, et al. The Japanese society of hypertension guidelines for the management of hypertension (JSH 2019). Hypertens Res. 2019;42(9):1235-481. doi: 10.1038/s41440-019-0284-9.

[22] Klabunde RE. Cardiovascular physiology concepts. 2nd ed. Philadelphia: Lippincott Williams & Wilkins; 2012.

[23] Nguyen JD, Duong H. Pursed-lip breathing. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 [cited 2026 May 5]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK545289/

[24] Malone DJ, Adler J. The patient with respiratory failure-preferred practice pattern 6F. In: Irwin S, Tecklin JS, editors. Cardiopulmonary physical therapy. 4th ed. Saint Louis: Mosby; 2004. p. 372-99.

[25] Kai S, Nagino K, Aoki T, Imura T, Kiyoshima K, Satake Y, et al. Cardiac autonomic nervous system activity during slow breathing in supine position. Rehabil Res Pract. 2021; 2021: 6619571. doi: 10.1155/2021/6619571.

Journal of Associated Medical Sciences 2026; 59(3): 31-41

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 and satisfaction of the elastic ankle support devices for walking and mobility among healthy people

Sugalya Amatachaya[1,2*], Kanchalika Ngamying[1], Pornpawee Thawthaisong[1], Thaksin Chanata[1,2], Wilairat Namwong[1,2], Arpassanan Wiyanad[1,2], Thiwabhorn Thaweewannakij[1,2]

[1]School of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen Province, Thailand.
[2]Improvement of Physical Performance and Quality of Life (IPQ) research group, Khon Kaen University, Khon Kaen Province, Thailand.

ARTICLE INFO

Article history:
Received 8 February 2026
Accepted as revised 14 May 2026
Available online 19 May 2026

Keywords:
Gait analysis, physical therapy, clinical biomechanics, rehabilitation

ABSTRACT

Background: A plastic non-hinged ankle-foot orthosis (AFO) is a standard ankle support device (ASD) commonly prescribed for individuals with walking impairments in general clinical settings of a developing country. Although it promotes toe clearance during the swing phase, its inherent rigidity can limit ankle rocker during the stance phase, thereby facilitating the development of compensatory strategies that affect walking and mobility in these individuals.

Objectives: To develop an alternative practical ASD assisting movements not only the swing but also the stance phases and compare their effects to a non-hinged AFO and no-ASD conditions on walking and mobility of healthy participants.

Materials and methods: This study developed elastic ASDs from a 3- to 4-inch bandage, called a long ankle sling, and elastic bands with buttonholes, called elastic ankle support equipment (EASE). Their effects and satisfaction were compared with those of no-ASD and non-hinged AFO in 18 healthy participants. Outcomes were measured as spatiotemporal gait variables, mobility using the timed up and go test (TUG), and satisfaction using a self-report questionnaire.

Results: Spatiotemporal and TUG data of the elastic ASD conditions showed no significant differences from the no-ASD condition, but significantly better than those of the AFO condition. The elastic ASDs met the criteria for dimensions, weight, safety, and effectiveness (p<0.001). The EASE was also comfortable and easy to adjust, whereas the AFO was durable (p<0.01).

Conclusion: The present study suggests the potential clinical benefit of elastic ASDs as they did not hinder the walking and mobility of healthy participants, i.e., allowed healthy participants to maintain walking ability and mobility comparable to unassisted walking. Given its ease of adjustment and application, the EASE is a particularly promising and user-friendly ASD for implementation in various clinical and community settings.

Introduction

Adequate ankle control is essential for normal walking across both the stance and swing phases. During the stance phase, ankle control is essential for body advancement as an inverted pendulum movement at the ankle, namely ankle rocker. During the swing phase, ankle control is also needed to clear the toes and move forward without the risk of stumbling or falling.[1,2] Impaired ankle control, regardless of its cause, can

* **Corresponding contributor.**
Author's Address: *School of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen Province, Thailand.*
E-mail address: samata@kku.ac.th
doi: 10.66285/JAMS.2026.077
E-ISSN: 2539-6056

significantly retard walking ability, which could further distort mobility, daily activities, and independence of the individuals.[1-3] As a result, an ankle-foot orthosis (AFO) is frequently applied to promote ankle control in patients.[3-6] There are many types of AFOs, both non-hinged and hinged styles, available nowadays. A non-hinged AFO is lightweight, but it is usually made as a single piece to hold the ankle in a single position. On the contrary, a hinged, both passive and active, AFO allows some degrees of ankle movement, but it is criticized for its weight (at least approximately 1 kg) and size.[3]

The sensorimotor impairments affecting walking ability likely result in the ankle being in plantarflexion with inversion affecting balance control and risk of fall of the patients.[1,2,7] Consequently, a plastic non-hinged AFO is commonly prescribed as a standard ankle support device (ASD) to aid the ankle in a neutral position and promote foot clearance of patients in general clinical settings of a developing country.[3-6] Previous studies also reported that a rigid, non-hinged AFO could enhance ankle and postural stability, thereby promoting balance, walking speed, step symmetry, and walking efficiency for individuals with movement impairments.[8-11]

Nevertheless, the rigid nature of a non-hinged AFO could restrict the ankle rocker, i.e., the change of the ankle joint from plantarflexion to dorsiflexion, during the stance phase.[7,12] This restriction could facilitate individuals' development of a compensatory strategy, i.e., forward trunk lean, to enable body-weight shifting over the small base of support during a single-limb support period of the stance phase. Such compensation shifts the body's center of mass anteriorly to the knee joint, which subsequently pushes the knee backward, resulting in knee hyperextension. These compensatory strategies could obviously distort patients' gait patterns.[1,2,7,13] Consequently, the researchers believed that the development of a practical, lightweight ASD that could assist ankle control in both the swing and stance phases would benefit the walking and mobility of individuals with movement impairments. However, in the initial phase after development, a new device should assess its potential benefits and satisfaction in healthy people without sensorimotor deficits to provide foundational information for further clinical application. Therefore, this study developed a practical and lightweight ASD that could assist ankle movements during both the stance and swing phases. Subsequently, the study compared the effects of the new ASD to a standard non-hinged AFO and no-ASD on walking ability and mobility of healthy individuals. The participants' satisfaction with the ASDs in this study was also assessed. The researchers hypothesized that the new ASD would assist ankle control during both the stance and swing phases. Thus, walking and mobility of healthy participants were comparable to the unassisted control condition (no-ASD) and superior to the data while walking with a standard non-hinged AFO. The present findings are intended to suggest potential clinical application of a new ASD for individuals with walking impairments in general clinical and home-based settings.

Materials and methods

Study design and protocols

This research and development (R&D) design study was conducted to develop and preliminarily assess the potential clinical benefit and satisfaction of a new practical ASD. The research protocols were divided into 2 phases: innovation development and outcome assessments. Details of each phase are as follows.

Phase I*: Innovation development*

The innovation was developed following the R&D process, starting with a literature review for needs analysis, followed by design and development, validation, and evaluation.[3,14,15]

Literature review for needs analysis:

In this step, the researchers reviewed existing literature and analyzed target users, i.e., individuals with walking impairments, to identify the problems and justify the need for a new ASD. From the literature, we found that normal body advancement during walking occurs at the ankle joint of the stance limb after the foot is placed flat on the floor, as the so-called ankle rocker, whereby the ankle must change from plantarflexion to dorsiflexion to create an inverted pendulum movement of the body. These movements occur simultaneously with the knee changing from full extension to slight flexion to lower the upward trajectory of the body center of mass for effective leg length and energy optimization, and to act as a shock absorber during the stance phase. During the swing phase, joints of the lower extremity must transition from extension to flexion to enable toe clearance and minimize the risk of stumbling and falling.[1,2,13,16] However, walking impairments are frequently caused by sensorimotor abnormality, resulting in the ankle likely being in plantarflexion with inversion, while the knee is in hyperextension patterns.[1,2,7,13] In general clinical settings, a non-hinged AFO is commonly prescribed to support the ankle in a neutral position and enable toe clearance during the swing phase.[4-6] However, the rigid nature of a non-hinged AFO could limit ankle rocker and facilitate individuals in developing compensatory strategies, i.e., forward trunk lean and knee hyperextension, which further distort the patients' gait patterns.[1-3,7,13]

Design and development:

Based on the preliminary review information, the researchers designed a framework (Figure 1) and searched for appropriate materials to provide the supporting and assisting forces for effective ankle control during the stance and swing phases. Finally, the researchers decided to use an elastic material and designed the device to maximize the lever arm length to achieve an effective assisting force during the stance and swing phases. Then, the researchers created an initial prototype from a 3- to 4-inch bandage, tailored to individuals' sizes, measuring 4.5 m long, namely a long ankle sling (Figure 2). The distal end of the long ankle sling was placed at the base of the toes, while the proximal end was wrapped around the top of the shank, at the level just below the knee joint. By doing this, it would help to maximize the lever arm lengths for biomechanical assistance during the stance phase (proximal end) and the swing phase (distal end). In addition, the two bands of the long ankle sling enable adjustment of force to achieve proper ankle alignment (inversion or eversion) based on the patient's needs.

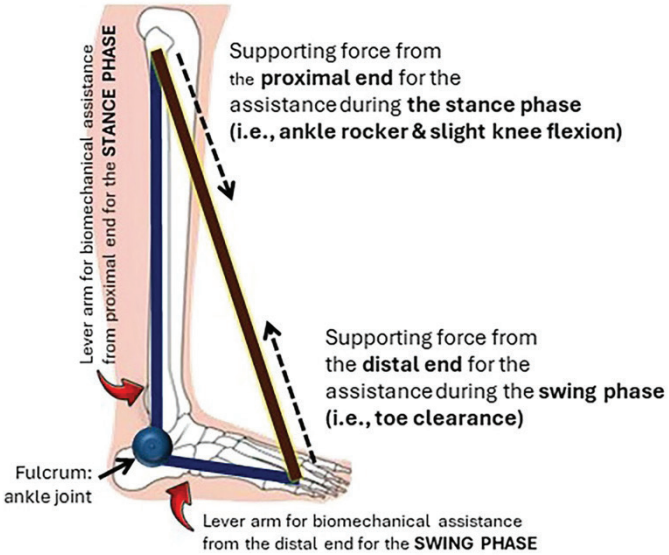


Figure 1. *The concepts for the development of elastic ankle support devices in this study.*

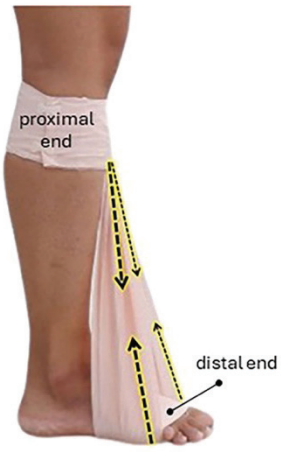


Figure 2. *Long ankle sling.*

Validation and evaluation:

After the development, the long ankle sling was assessed for its potential clinical application and adverse effects (if any). The preliminary data suggested that the long ankle sling could assist ankle control in both the stance and swing phases, but it must be applied by a skilled expert to generate appropriate supporting forces, making it challenging to implement in routine clinical practice. To address this issue, the researchers further developed an alternative ASD to facilitate adjustment and ensure proper ankle control. Finally, we decided to use elastic bands with buttonholes (with the width and length of 1.5 cm x 30 cm) that can be easily found in E-marketplaces, e.g., Shopee, Lazada, for creating a practical elastic ASD, named "elastic ankle support equipment (EASE)" (Figure 3A). The EASE was designed based on the concept of a long ankle sling, using four elastic bands and three plastic rivet buttons (0.5cm wide). One band and two buttons are used to make a circle wrapping around the forefoot, i.e., at the base of the toes (distal end). Another band is placed at the upper end of the shank, i.e., just below the knee joint (making the proximal end), with the other two elastic bands attached to a button behind the shank. After proper ankle alignment is set, each of the last two bands is drawn from the posterior aspect of the proximal end, crossing the anterior aspect of the shank to be attached to a button on the opposite side of the distal end at the base of the toes (Figure 3B).

After development, the researchers applied both the long ankle sling and EASE to preliminarily assess their potential clinical benefit and satisfaction in healthy individuals in this study.



Figure 3. *Elastic ankle support equipment (EASE). A: Equipment used for the EASE, B: The use of EASE for ankle support.*

Phase II*: Outcome assessments in healthy individuals*

Participants:

The study recruited a healthy cohort from several communities in a developing country using simple random sampling. The inclusion criteria were age 18 years and over, both male and female, and body mass index 18.5-29.9 kg/m^2. These individuals were excluded when they presented with any musculoskeletal and neurological conditions affecting walking and mobility, such as muscle weakness, numbness, impaired or loss of sensations, musculoskeletal pain with a visual analog scale score for pain of at least 5 out of 10, deformities of the spine and lower extremity, joint limitation, or an inability to understand and complete the study protocols. The sample size calculation for a major variable, i.e., walking speed, using data from 10 pilot participants, with a confidence level of 95% (significance level=0.05; Z_α=1.65) and a power of 90% (Z_β=1.28), indicated that the study required at least 16 participants.[17] All participants read and signed a written informed consent form approved by the Khon Kaen University Ethics Committees for Human Research (HE671050) before participation in the study.

Outcome measurements

Eligible participants were interviewed and assessed for their demographics. Subsequently, they were assessed for their walking ability and mobility under four conditions, including 1) no-ASD (control), 2) a plastic non-hinged AFO, 3) a long ankle sling, and 4) EASE, with the ASDs applied to the right leg (Figure 4), in a random order using a Latin square. For the AFO, this study used a commercial, non-hinged posterior leaf-spring AFO in a size that fit the participant's foot (Figure 4B) because it is commonly prescribed for individuals with walking impairments in general clinical settings in the country.[4-6] For the elastic ASDs, two well-trained researchers (KN & PT) used one of the five sets of bandages and EASE for all participants.

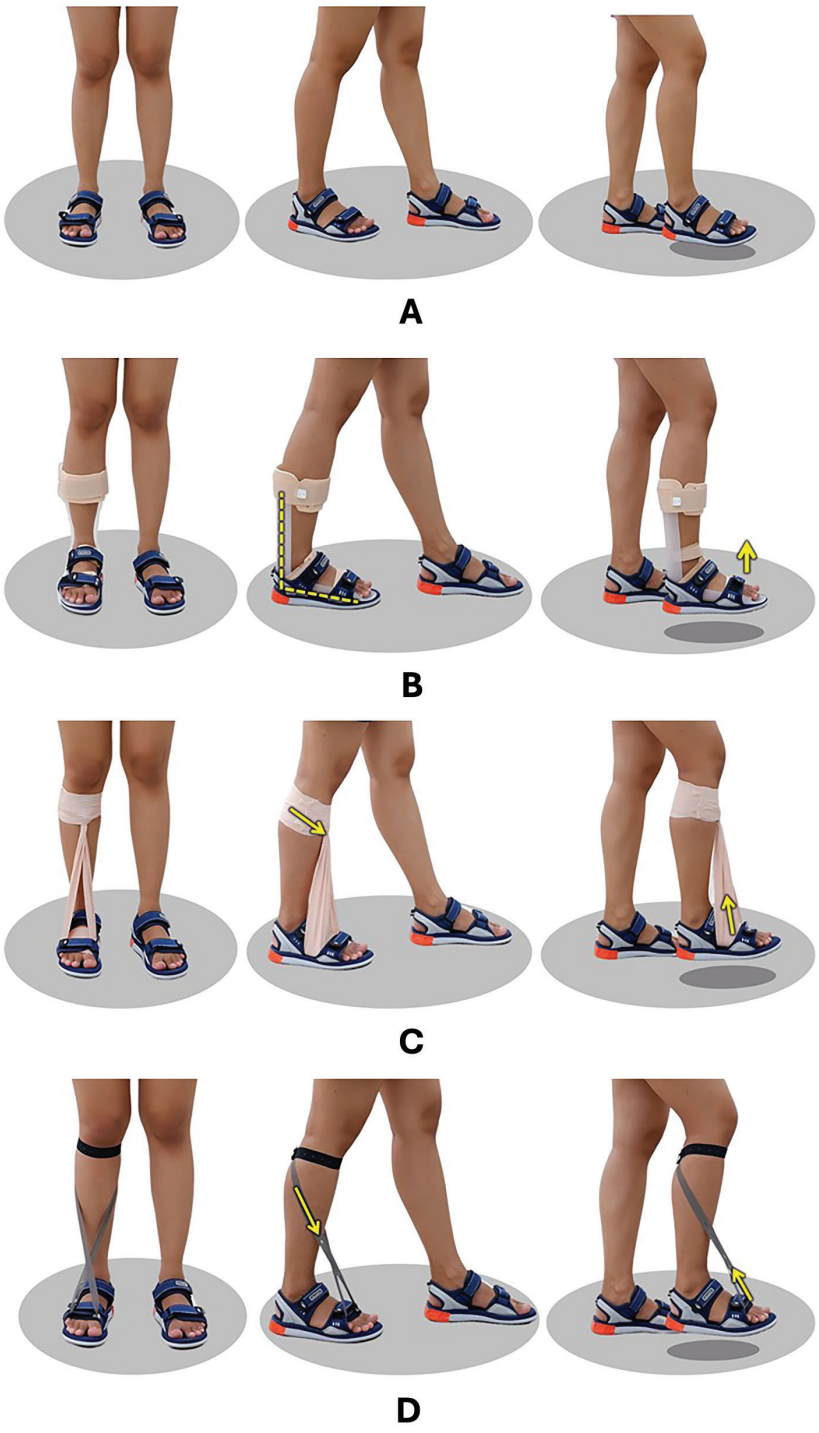


Figure 4. *Testing conditions in the study. A: without an ankle support device, B: with an ankle-foot orthosis, C: with a long ankle sling, D: with elastic ankle support equipment. Figures illustrate the front view, stance phase, and swing phases, respectively.*

In each condition, participants were assessed for walking and mobility by two excellent reliability raters (intraclass correlation coefficients, or ICCs, for intra- and inter-rater reliability >0.90). Each condition was assessed while participants wore sandal sports shoes prepared by the researchers to an appropriate size (Figure 4) to minimize the effects of different shoes on the outcomes. Prior to the assessments in each condition, participants were given familiarization trials. An examiner walked alongside or stayed beside the participants without interruption to ensure their safety and the accuracy of the outcomes, with sufficient rest periods between trials and tests as required. Thereafter, the participants completed a self-report questionnaire regarding their satisfaction with the three ASDs in the study. Details of the outcome measurements of the study are as follows:

Walking ability

The video-camera-based motion capture system with two-dimensional gait analysis and manual digitization was used as a practical, reliable method (ICC=0.95-0.99) for data collection in general settings.[18-20] The camera (GoPro Hero12 with a 120-Hz frame rate and full HD [1080p]) was mounted perpendicular to the 10-m walkway at 3.5 m, capturing data over 4 m in the middle of the walkway (a rhythmic phase). Prior to the data recording, the system was calibrated using a known-length object. Participants were attached with a colored tape at both heels for clear manual digitization. Then they walked at their preferred, fast, and safe speeds. Subsequently, the spatiotemporal data (i.e., step length, stride length, cadence, walking speed, and percent step length symmetry) over the two trials were analyzed using the Kinovea program as follows:[1,2,13,16]

- Step Length (cm): The distance measured in horizontal pixels between the heel of one foot and the heel of the other foot, scaled according to the calibration object.
- Stride Length (cm): The average summation of the right and left step lengths.
- Cadence (steps/minute): The number of steps taken in one minute.
- Walking speed (m/s): calculated using the formula $\frac{stride\ length \times cadence}{120}$
- Step length symmetry (%): calculated as $\frac{average\ data\ of\ the\ shorter\ step\ length}{average\ data\ of\ the\ longer\ step\ length} \times 100$

Mobility: The timed up and go test (TUG), a reliable measure (ICC=0.99) that assesses basic mobility activities in daily living, was used to reflect participants' mobility and dynamic balance abilities.[21,22] The participants timed their ability to stand up from a standard armchair, walk 3 m, turn around a traffic cone, walk back, and sit down on the chair at a fast, safe speed.[21-23] The average data over the two trials were reported.

Satisfaction: The Quebec user evaluation of satisfaction with assistive technology (QUEST, version 2.0), a standard, valid, and reliable (ICC=0.75-0.90) instrument, was used to assess participants' satisfaction.[6,24] Participants self-reported on the 8 items in the device domain. Each item was rated on a 5-point ordinal scale ranging from 1 (not satisfied at all) to 5 (very satisfied).[24,25]

Statistical analysis

The Shapiro–Wilk test was applied to assess the data distribution. Descriptive statistics were used to describe the participants' demographics and the study's findings. With the non-normal distribution, the data are presented using median (interquartile range). The Friedman test was utilized to analyze the differences among the four conditions. Then, the Wilcoxon signed-rank test was used to analyze all pairwise comparisons. The statistical significance level was set at $p<0.05$.

Results

Eighteen healthy participants (11 females), with an average age of 44.5±11.4 years, body mass index of 23.8±3.4 kg/m^2, and no significant musculoskeletal pain, completed the study. Their average preferred walking speed was 1.4 m/s, the fastest walking speed was 1.7 m/s, and the TUG data were 7.07 s (Table 1).

Table 1. Walking and mobility outcomes of the participants when assessed in four testing conditions.

	Variable	Ankle support device (ASD) conditions				*p* value[a]
		No-ASD	AFO	Long ankle sling	EASE	
Preferred walking speed	Right step length (cm)	0.71 (0.68:0.77)	0.69 (0.66:0.74)	0.71 (0.67:0.75)	0.71 (0.65:0.76)	0.091
	Left step length (cm)	**0.71**^{A*} (0.67:0.76)	0.68 (0.64:0.73)	**0.70**^{A*} (0.67:0.77)	**0.71**^{A*} (0.66:0.77)	0.002
	Stride length(m)	**1.42**^{A*} (1.35:1.54)	1.37 (1.31:1.47)	**1.41**^{A*} (1.35:1.51)	**1.42**^{A*} (1.32:1.55)	0.008
	Step length symmetry (%)	98.61 (95.55:99.05)	98.55 (94.02:99.08)	98.59 (94.78:99.34)	98.97 (94.78:99.34)	0.983
	Cadence (steps/minute)	**119**^{A*} (112:124)	116 (107:122)	**117**^{A*} (114:124)	117 (114:123)	0.022
	Walking speed (m/s)	**1.41**^{A*} (1.32:1.53)	1.32 (1.19:1.40)	**1.37**^{A*} (1.30:1.51)	**1.38**^{A*} (1.23:1.52)	0.012
Walking fastest speed	Right step length (cm)	**0.79**^{A*} (0.74:0.85)	0.76 (0.70:0.81)	**0.79**^{A*} (0.74:0.84)	**0.78**^{A*} (0.74:0.83)	0.003
	Left step length (cm)	**0.77**^{A*} (0.74:0.85)	0.75 (0.70:0.82)	**0.78**^{A*} (0.72:0.83)	**0.78**^{A*} (0.75:0.85)	0.012
	Stride length(m)	**1.56**^{A*} (1.50:1.69)	1.51 (1.41:1.64)	**1.57**^{A*} (1.48:1.66)	**1.56**^{A*} (1.48:1.67)	0.007
	Step length symmetry (%)	97.4 (97.43:99.49)	98.68 (96.33:98.85)	98.73 (97.04:99.04)	98.84 (95.98:98.29)	0.209
	Cadence (steps/minute)	**133**^{A*} (122:137)	128 (119:134)	**132**^{A*} (125:135)	**131**^{A*} (124:140)	<0.001
	Walking speed (m/s)	**1.73**^{A**} (1.56:1.90)	1.61 (1.43:1.71)	**1.73**^{A**} (1.56:1.77)	**1.70**^{A**} (1.56:1.78)	<0.001
Timed up and go test (s)		**7.07**^{A**} (6.38:7.68)	7.88 (7.14:8.54)	**7.15**^{A*} (6.81:8.28)	**7.10**^{A**} (6.62:7.74)	<0.001

Note: *Data are presented using median (interquartile range: IQR), [a]p values from the Friedman test, bold characters indicate the conditions with significant differences from the Wilcoxon signed rank test, whereby $^{*}p<0.05$, $^{**}p<0.001$, superscripts refer to the condition with statistically significant differences from the indicated condition, A: ankle-foot orthosis (AFO).*

Spatiotemporal data

At a preferred speed, the left step length, stride length, and walking speed of participants while walking with both elastic ASDs were comparable to those found while walking with no-ASD (p>0.05). However, the data for these conditions were significantly higher than those in the AFO condition (p<0.05, Table 1). In addition, their walking cadence was significantly higher when walking without an ASO and with a long ankle sling than in the AFO condition (p<0.05, Table 1).

When measured at the fastest and safest speeds, the findings showed the same trend as those for the preferred walking speed. However, the significant differences among the conditions were additionally found for the right step length (p=0.003) and cadence (p<0.001), whereby the data of no-ASD and elastic ASD conditions were significantly better than those of the AFO condition. However, there were no significant differences among the elastic ASDs and no-ASD conditions (p>0.05, Table 1). Notably, step-length symmetry showed no significant differences across the four conditions at both the preferred and fastest speeds.

Mobility

The TUG data demonstrated the same trend as those found for the fastest and safest speed. Participants took the longest time to complete the TUG while walking with an AFO (p<0.001, Table 1). On the contrary, the TUG data for the long ankle sling, EASE, and no-ASD conditions did not show significant differences among them (p>0.05).

Satisfaction

The QUEST scores of the long ankle sling and EASE were significantly better than the AFO in terms of dimensions, weight, safety and security, and comfort (p<0.001, Table 2). The EASE also had the highest QUEST scores for ease of adjustment and ease of use compared to the other ASDs. In contrast, the AFO received the highest scores for durability (p<0.001, Table 2). Participants also provided additional suggestions that the AFO was very hard and stiff, resulting in pain in their malleoli, the medial and lateral sides of the foot, and the heel. The hard and rigid natures of the AFO also limited their mobility, resulting in some participants feeling unsafe and insecure. In contrast, the long ankle sling was criticized in its clinical application for the appropriate supporting forces. The EASE was recommended to increase the size of elastic bands to promote an appropriate supporting force while walking.

Table 2. The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST, version 2.0) of the participants.

Items of the QUEST (version 2.0)	Types of ankle support device			*p* value[a]
	AFO	Long ankle sling	EASE	
1. The dimensions (size, height, length, and width) of your assistive device?	2.72 (1.23:4.21)	**4.22 (3.67:5.00)**^{A**}	**4.17 (3.46:5.00)**^{A**}	<0.001
2. The weight of your assistive device?	2.39 (1.30:3.48)	**4.17 (3.46:5.00)**^{A**}	**4.78 (4.35:5.00)**^{A**,L*}	<0.001
3. The ease in adjusting (fixing and fastening) the parts of your assistive device?	3.11 (1.79:4.43)	2.72 (1.65:3.79)	**4.61 (4.35:5.00)**^{A*,L**}	<0.001
4. How safe and secure your assistive device is?	2.67 (1.30:4.04)	**4.06 (3.26:5.00)**^{A*}	**4.39 (3.69:5.00)**^{A*}	<0.001
5. The durability (endurance and resistance to wear) of your assistive device?	**4.00 (3.19:5.00)**^{L**,E**}	3.78 (3.13:4.43)	3.78 (2.90:4.66)	<0.001
6. How easy is it to use your assistive device?	2.94 (1.55:4.33)	3.00 (1.94:4.06)	**4.72 (4.26:5.00)**^{A**,L**}	<0.001
7. How comfortable your assistive device is?	2.78 (1.67:3.89)	**3.83 (3.04:4.62)**^{A*}	**4.39 (3.69:5.00)**^{A**}	<0.001

Note: *Each item was scored from 1 (not satisfied at all) to 5 (very satisfied). The data are presented using median (interquartile range: IQR).* [a]*p value from the Friedman test. Bold characters indicate the conditions with significant differences from the Wilcoxon signed rank test, whereby* $^{*}p<0.05$, $^{**}p<0.001$. *Superscripts refer to the condition with statistically significant differences from the indicated condition, A: ankle-foot orthosis, L: long ankle sling, E: elastic ankle support equipment (EASE).*

Discussion

Healthy participants were presumed to exhibit a normal, unconstrained gait when walking without any ASD. Conversely, using an ASD is expected to alter natural walking patterns. Therefore, the no-ASD condition was established as a reference to evaluate the effects of the three ASDs in this study. The findings indicated no significant differences in the outcomes between the no-ASD and elastic ASD conditions, both the long ankle sling and EASE. The data for these conditions were also significantly better than those for the AFO condition, particularly when assessed at the fastest and safest speed (Table 1). Furthermore, the long ankle sling and EASE were rated favorably for their dimensions, weight, safety, and comfort. Conversely, the AFO caused discomfort, made participants feel unsafe, and caused pain in their feet.

All ASDs in this study were designed to support the foot, thereby enhancing toe clearance and facilitating forward limb progression during the swing phase.[4-6] Consequently, the right step length, i.e., the side with an ASD, showed no significant differences among the conditions when assessed at a preferred walking speed. Nonetheless, the long ankle sling and EASE were lightweight with elastic pulling force from the distal end (Figure 4C and Figure 4D) that additionally assisted the ground reaction force during the push-off phase, especially when walking at the fastest and safest speed. Thus, the participants could lift the foot from the floor and advance the lower limb forward more easily than in the AFO condition, which was criticized for its weight (Table 2). Consequently, the participants' step length and walking cadence while walking with elastic ASDs at the fastest and safest speed were significantly greater than those in the AFO condition (Table 1).

Subsequently, when the right foot was placed flat on the floor, the elastic pulling force from the proximal end of the tibia aided slight knee flexion and facilitated tibial advancement, i.e., ankle rocker, thereby creating an inverted pendulum movement of the body at the ankle joint. Such movements promoted a shift in the body's center of mass and line of gravity to be positioned over a small base of support during the single-limb support period on the foot with an elastic ASD. This mechanism helped lengthen the left step length (i.e., the other side without ASD) and improve balance while performing basic mobility tasks, including standing up, walking, turning around, and sitting down, as assessed using the TUG. Thus, the findings for these conditions showed no significant differences relative to the data from the no-ASD condition but were significantly better than those for the AFO condition. The increased step length and cadence under elastic ASD conditions were also accompanied by significant improvements in participants' stride length and walking speed compared to the AFO condition.[1,2,13,16]

In contrast, the weight of the non-hinged AFO confounded the ability to lift the foot up during the swing phase.[3-6,26] Consequently, participants reduced right step length and walking cadence compared to the other conditions, especially when assessed at the fastest and safest speed (p<0.05, Table 1). Moreover, the rigid nature of the non-hinged AFO imposed a resisting force on ankle joint rotation, limiting tibial advancement at the ankle joint and shifting bodyweight anteriorly over the base of support during the stance phase.[1,2,4-7] Such a manner impeded the other side's step length. Consequently, participants had the shortest left-step length when walking with the AFO. These effects subsequently reduced participants' walking speeds and mobility compared with the other conditions (p<0.05, Table 1). Similar changes, i.e., increased or decreased step lengths in both lower extremities while using each ASD, resulted in non-significant differences in percent step-length symmetry when assessed at either preferred or fastest walking speeds (p>0.05; Table 1).

Data from the QUEST 2.0 questionnaire additionally support the clinical application of the elastic ASDs. The participants expressed similar levels of satisfaction between the long ankle sling and the EASE in terms of dimensions, weight, safety and security, effectiveness, and comfort, except for the ease of adjustment of the long ankle sling. Conversely, they reported satisfaction with the AFO regarding the ease in clinical application and durability (Table 2). Some participants experienced pain in particular areas of the foot, felt unsafe, and insecure. Previous studies have also reported adverse events related to AFO use, including interference with gait performance and balance ability in individuals with walking deficits, along with increased skin irritation and rashes in the contact areas.[3-6,27-29] The present findings offer alternative elastic ASDs and their potential clinical application, especially the EASE, which is user-friendly, in promoting walking ability and mobility. However, this study has some limitations.

Limitations

This study contains some limitations. The present study assessed the effects of novel elastic ASDs in healthy individuals without sensorimotor deficits, as this is a critical first step toward providing foundational information on their potential utility in clinical populations. However, findings from healthy people may limit direct generalization to the target populations. The elastic ASDs may face durability issues or a potential decline in elasticity, which could compromise their performance and overall effectiveness over time. In addition, this study applied a standard, non-hinged AFO because it is commonly prescribed for individuals with walking deficits in general clinical settings of a developing country.[4-6] In fact, various types of AFOs are available, each with various degrees of joint limitation and weight.[3,26] The present findings may not fully support the clinical advantages of using a long ankle sling and EASE over other types of AFOs. Moreover, the findings were analyzed using the assisting forces provided by the elastic ASDs, without biomechanical

assessments or motion analysis. Therefore, a future study should include clinical populations and measure outcomes related to the durability of elastic ASDs and their effectiveness in more challenging environments, such as stair ascent and descent, as well as energy expenditure, to enhance ecological validity and comprehensive measurement.

Conclusions

Deficits in ankle control could impair walking and mobility. The present study demonstrates the potential clinical benefit of elastic ASDs as they did not hinder the performance of healthy individuals. Specifically, both the long ankle sling and EASE allowed healthy participants to maintain walking ability and mobility comparable to unassisted walking. Given its ease of adjustment and application, the EASE emerges as a particularly promising and user-friendly option for implementation in diverse clinical and community settings.

Ethical approval

All procedures of the study were in accordance with the standards of the Ethics Committee for Human Research (HE671050). Eligible participants completed a written informed consent form before participation in this study.

Funding

The Fundamental Fund (FF), Khon Kaen University, for the fiscal year 2026 from National Science, Research and Innovation Fund (NSRF), Thailand.

Conflict of interests

The authors declared no potential conflicts of interest.

Data availability statement

All relevant data are within this manuscript, and raw data are archived by the corresponding author.

CRediT authorship contribution statement

Sugalya Amatachaya: conceptualization, data curation, investigation, methodology, funding acquisition, project administration, visualization, resources, supervision, validation, writing: original draft, review and edit, approval of final manuscript; **Kanchalika Ngamying**: data curation, investigation, methodology, formal analysis, visualization, approval of final manuscript; **PornpaweeThawthaisong**: data curation, investigation, methodology, formal analysis, visualization, approval of final manuscript; **Thaksin Chanata**: data curation, investigation, methodology, formal analysis; **Wilairat Namwong**: conceptualization, investigation, funding acquisition, supervision, writing: original draft, review, and edit, approval of final manuscript; **Arpassanan Wiyanad**: conceptualization, funding acquisition, writing: original draft, review and edit, approval of final manuscript; **Thiwabhorn Thaweewannakij**: conceptualization, investigation, funding acquisition, supervision, writing: original draft, review and edit, approval of final manuscript.

Acknowledgements

This research has received funding support from the Fundamental Fund (FF), Khon Kaen University, for the fiscal year 2026 from the National Science, Research and Innovation Fund (NSRF), Thailand. The researchers also thank additional support from the Faculty of Associated Medical Sciences, Khon Kaen University. Thailand.

References

[1] Perry J and Burnfield JM. Gait analysis. Normal and pathological function 2nd Ed. California: Slack, 2010.

[2] Whittle MW. Gait analysis: an introduction. Churchill Livingstone, London, 2012

[3] Zhou C, Yang Z, Li K, Ye X. Research and development of ankle-foot orthoses: A Review. Sensors (Basel). 2022; 22(17): 6596. doi: 10.3390/s22176596.

[4] Dajpratham P, Pongakkasira C, Terachinda P, Phutthakumnerd W, Sirirak P. Patient compliance and factors related to usage of plastic ankle-foot orthoses by chronic stroke patients. ASEAN J Rehabil Med. 2021; 32: 29-33.

[5] Poolpipat P, Suksathien R. The use of plastic ankle-foot orthosis prescribed for patients with foot drop: The application and related Factors. ASEAN J Rehabil Med. 2023; 33(2): 81. doi

[6] Thitithunwarat N, Krityakiarana W, Kheowsri S, Jongkamonwiwat N, Richards J. The effect of a modified elastic band orthosis on gait and balance in stroke survivors. Prosthet Orthot Int. 2023; 47(5): 466-472. doi: 10.1097/PXR.0000000000000205.

[7] Amatachaya S. Abnormal gait. In Amatachaya S. Physical therapy on mobility and walking of patients with stroke. Khon Kaen University. 2025: 313-40.

[8] Esquenazi A, Ofluoglu D, Hirai B, Kim S. The effect of an ankle-foot orthosis on temporal spatial parameters and asymmetry of gait in hemiparetic patients. PM R. 2009; 1(11): 1014-8. doi: 10.1016/j.pmrj.2009.09.012.

[9] Franceschini M, Massucci M, Ferrari L, Agosti M, Paroli C. Effects of an ankle-foot orthosis on spatiotemporal parameters and energy cost of hemiparetic gait. Clin Rehabil. 2003; 17(4): 368-72. doi: 10.1191/0269215503cr622oa.

[10] Leung J, Moseley A. Impact of ankle-foot orthoses on gait and leg muscle activity in adults with hemiplegia: systematic literature review. Physiotherapy. 2003; 89: 39-55. doi: 10.1016/S0031-9406(05)60668-2.

[11] Cruz TH, Dhaher YY. Impact of ankle-foot-orthosis on frontal plane behaviors post-stroke. Gait Posture. 2009; 30(3): 312-6. doi: 10.1016/j.

gaitpost.2009.05.018.

[12] Yamamoto S, Fuchi M, Yasui T. Change of rocker function in the gait of stroke patients using an ankle foot orthosis with an oil damper: immediate changes and the short-term effects. Prosthet Orthot Int. 2011; 35(4): 350-9. doi: 10.1177/0309364611420200.

[13] Burnfield M. Gait analysis: Normal and pathological function. J Sports Sci Med. 2010; 9(2): 353.

[14] Afriani L, Mutmainnah, Sunarni. Understanding the design of research and development methods in the field of education. IJESS [Internet]. 2025; 6(1): 1-5.

[15] Umar U, Purwanto M, Al Firdaus M. Research and development: as the primary alternative to educational research design frameworks. JELL [Internet] 2023; 8(01): 73-2. doi: 10.37110/jell.v8i01.172.

[16] Amatachaya S. Normal gait. In Amatachaya S. Physical therapy on mobility and walking of patients with stroke. Khon Kaen University. 2025: 281-312.

[17] Portney LG. Foundations of clinical research: applications to evidence-based practice: FA Davis. 2020

[18] Nithiatthawanon T, Teesintanakorn A, Sanjai C, Hengchanoknun K, Saikaew N, Simarattanamongkhon S, et al. Concurrent validity of two-dimensional motion analysis using Kinovea for measuring spatiotemporal gait parameters in healthy individuals. Arch AHS. 2024; 36(3): 21-30.

[19] Promkeaw D, Mato L, Thaweewannaki T, Arrayawichanon P, Saensook W, Amatachaya S. Walking on different surfaces challenged ability of community-dwelling elderly. Arch AHS. 2018; 30(1): 39-46.

[20] Promkeaw D, Arrayawichanon P, Thaweewannakij T, Mato L, Amatachaya P, Amatachaya S. Various surfaces challenge gait characteristics of ambulatory patients with spinal cord injury. Spinal Cord. 2019; 57(9): 805-13. doi: 10.1038/s41393-019-0282-4.

[21] Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc. 1991; 39(2): 142-8. doi: 10.1111/j.1532-5415.1991.tb01616.x.

[22] Phonthee S, Amatachaya P, Sooknuan T, Amatachaya S. Stepping training with external feedback relating to lower limb support ability effectively improved complex motor activity in ambulatory patients with stroke: a randomized controlled trial. Eur J Phys Rehabil Med. 2020; 56(1): 14-23. doi: 10.23736/S1973-9087.19.05907-0.

[23] Thaweewannakij T, Suwannarat P, Mato L, Amatachaya S. Functional ability and health status of community-dwelling late age elderly people with and without a history of falls. Hong Kong Physiother J. 2015; 34: 1-9. doi: 10.1016/j.hkpj.2015.08.001.

[24] Aledda S, Galeoto G, Fabbrini G, Lucibello L, Tofani M, Conte A, Berardi A. A systematic review of the psychometric properties of Quebec user evaluation of satisfaction with assistive technology (QUEST). Disabil Rehabil Assist Technol. 2024; 19(4): 1228-35. doi: 10.1080/17483107.2022.2161648.

[25] Demers L, Weiss-Lambrou R, Ska B. Development of the Quebec user evaluation of satisfaction with assistive technology (QUEST). Assist Technol. 1996; 8(1): 3-13. doi: 10.1080/10400435.1996.10132268.

[26] Daryabor A, Arazpour M, Aminian G. Effect of different designs of ankle-foot orthoses on gait in patients with stroke: A systematic review. Gait Posture. 2018; 62: 268-79. doi: 10.1016/j.gaitpost.2018.03.026.

[27] Sankaranarayan H, Gupta A, Khanna M, Taly AB, Thennarasu K. Role of ankle foot orthosis in improving locomotion and functional recovery in patients with stroke: A prospective rehabilitation study. J Neurosci Rural Pract. 2016; 7(4): 544-9. doi: 10.4103/0976-3147.185507.

[28] Geboers JF, Drost MR, Spaans F, Kuipers H, Seelen HA. Immediate and long-term effects of ankle-foot orthosis on muscle activity during walking: a randomized study of patients with unilateral foot drop. Arch Phys Med Rehabil. 2002; 83(2): 240-5. doi: 10.1053/apmr.2002.27462.

[29] Totah D, Menon M, Jones-Hershinow C, Barton K, Gates DH. The impact of ankle-foot orthosis stiffness on gait: A systematic literature review. Gait Posture. 2019; 69: 101-11. doi: 10.1016/j.gaitpost.2019.01.020.

Journal of Associated Medical Sciences 2026; 59(3): 42-54

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



Analysis of associations between personal attributes, emotional state, and functional performance after stroke

Euna Ko[1] and Wongjong Yu[2*]

[1]Department of Rehabilitation Medicine, Chungdam Hospital, Seoul, Republic of Korea.
[2]Department of Physical Therapy, Eulji University, Seongnam, Republic of Korea.

ARTICLE INFO

Article history:
Received 29 December 2025
Accepted as revised 22 April 2026
Available online 20 May 2025

Keywords:
Stroke, emotional factors, functional performance, rehabilitation.

ABSTRACT

Background: Emotional problems such as depression, anxiety, and apathy frequently occur after stroke and can negatively impact the functional recovery process. However, most previous studies have focused on specific emotional factors or examined the relationship between emotional and functional variables in a predominantly one-directional manner. Because emotional management is also important in stroke rehabilitation, further research is needed to clarify the complex associations among emotional, functional, and personal variables.

Objectives: The purpose of this study was to examine the bidirectional associations between emotional factors and functional performance in stroke patients undergoing inpatient rehabilitation and to explore the role of personal factors in emotional status.

Materials and methods: The study was conducted with 102 stroke patients hospitalized in a rehabilitation hospital. Emotion was assessed using the Korean version of the Positive and Negative Affect Schedule (PANAS), the Beck Depression Inventory (BDI), the Hospital Anxiety and Depression Scale-Anxiety (HADS-A), and the Apathy Evaluation Scale (AES). Function was measured by cognitive function (MMSE), balance ability (BBS), gait ability (FAC), and activities of daily living (K-MBI). Personal factors included sex, age, duration of illness (DISEASE_DUR), and length of stay (LOS). Statistical analyses included simple and multiple regression.

Results: Simple regression analysis showed that better functional performance was associated with lower negative affect, depression, and anxiety. Conversely, higher negative affect, depression, and anxiety were associated with poorer cognitive function, balance, gait ability, and activities of daily living. Apathy showed a limited positive association with cognitive function. Among personal factors, positive affect decreased slightly with increasing age, and apathy was higher in patients with a shorter illness duration. Multiple regression analysis showed that most relationships disappeared, but depression remained independently associated with functional performance after adjustment for other emotional variables. Among functional domains, only activities of daily living independently explained emotional state.

Conclusion: Emotional state and functional performance in stroke patients were significantly associated, particularly with respect to negative emotions and depression. These findings suggest that a functional-focused rehabilitation approach alone may be insufficient, highlighting the potential need for rehabilitation strategies that integrate emotional interventions with functional training. This study can serve as foundational data for developing integrated rehabilitation models to promote recovery in stroke patients.

***Corresponding contributor.**
Author's Address: *Department of Physical Therapy, Eulji University, Seongnam, Republic of Korea.*
E-mail address: wjyu@eulji.ac.kr
doi: 10.66285/JAMS.2026.078
E-ISSN: 2539-6056

Introduction

The demand for stroke treatment and rehabilitation continues to increase, placing a significant burden on society.[1] Stroke patients with severe functional limitations who cannot be discharged directly from tertiary hospitals are provided with inpatient rehabilitation programs at secondary hospitals.[2] Most functional recovery after stroke occurs within the first three months, but complete recovery happens slowly over 6 to 12 months.[3] Therefore, stroke patients often require prolonged hospitalization for rehabilitation. In Korea, a nationwide study reported that the mean total length of hospital stays after stroke, from acute care admission to home discharge across multiple hospitals, was 115.6 days.[4]

Stroke patients experience disabilities due to paralysis and consequently report difficulties with walking.[5] Balance issues are a common problem after stroke and are associated with reduced activities of daily living and mobility, as well as an increased risk of falls.[6] Stroke rehabilitation, particularly physical therapy, focuses on restoring physical independence and functional abilities during daily living.[7] The ultimate goal of treatment is to improve walking function and restore balance and movement.

However, the problems caused by stroke are not limited to physical and functional issues. Depression, anxiety, and apathy are frequent neuropsychiatric outcomes following stroke.[8,9] Depression occurs in 18-39% of stroke patients,[10,11] anxiety in 18-27%,[12,13] and apathy in 8-36%.[8,14] The most common neuropsychiatric outcome after stroke is depression.[15] Depression is closely associated with brain structural and functional changes,[16] and there is a negative correlation between motor function and depression in the elderly.[17] Furthermore, post-stroke depression is associated with impaired functional outcomes[18], reduced rehabilitation outcomes,[19] and cognitive impairment.[20]

Yejin Lee *et al*. investigated the impact of emotional factors on social participation among 73 stroke patients. The results indicated that stroke severity and depression significantly influence social activities and participation.[21] Yu-Won Choe and Myoung-Kwon Kim's study examined the relationship between emotional factors and physical abilities in stroke patients, reporting significant associations between negative emotional states and poorer physical function variables.[22] Additionally, individuals with depression are reported to have lower levels of physical activity.[23]

A large-scale retrospective cohort study based on the National Health Database revealed that female stroke survivors have a 20% higher risk of depression than male stroke survivors.[24] A meta-analysis on risk factors for post-stroke depression indicated higher depression risk among women and younger age groups.[15] Analysis of the relationship between personality traits and emotional factors in subacute stroke patients within 3 months showed higher levels of emotional distress in women and older patients.[25] Furthermore, analysis of the mediating effects of age, sex, and duration of illness on the relationship between trunk control ability and functional recovery in stroke patients revealed that recovery levels were lower with increasing age and longer duration of illness.[26] That is, since personal characteristics are factors related to physical functional recovery, this can serve as evidence that, when expanded, personal factors influence physical function and thus may also affect emotions.

Numerous studies exist on the relationship between stroke and emotions, particularly depression. However, most existing research has focused primarily on depression and has examined associations between emotions and function from a predominantly one-directional perspective. Stroke patients undergoing inpatient rehabilitation generally show changes in functional performance over time, consistent with neuroplasticity principles.[27] However, a review of the existing literature suggests that emotional factors remain closely associated with functional outcomes in stroke patients, indicating that emotional aspects should also be considered alongside functional performance.[28] Therefore, further research exploring the associations among functional, personal, and emotional variables is warranted.

This study evaluates four emotional factors-positive and negative affect, depression, anxiety, and apathy-in stroke patients and analyzes their interrelationships with eight variables: functional factors (cognition, balance, gait, and activities of daily living) and personal factors (sex, age, duration of illness, and length of stay). It aims to assist various healthcare professionals, including physical therapists, in understanding stroke patients hospitalized for treatment and in establishing treatment directions. The hypotheses of this study were as follows: 1) emotional variables are significantly associated with functional variables, and 2) personal factors are significantly associated with emotional variables.

Materials and methods

Study participants

This study employed a cross-sectional observational design involving a simple cohort of stroke patients. Emotional assessments are conducted by distributing Korean-language evaluation forms to participants. One-on-one explanations by the researcher are provided as needed. Participants submit all completed evaluation forms to the researcher without omission. Functional assessment data is collected concurrently with the emotional assessments.

Participants are stroke patients hospitalized in a rehabilitation hospital who understand the evaluation methods and can provide informed consent. Participants were recruited using a purposive sampling approach by promoting participation to eligible stroke patients hospitalized at Hospital C in Seoul, South Korea, from August 2024 to May 2025. Researchers

distribute informational leaflets to participants explaining the study. For those who may have difficulty understanding the leaflet due to language barriers, researchers provide a one-on-one explanation. The research participant's consent is obtained directly from the subject. Upon consent, a handwritten signature is obtained on the consent form, and the evaluation proceeds. The final number of research subjects who participated is 102. The sample size was not determined by a priori power analysis but rather by the number of eligible stroke patients recruited during the study period at a simple rehabilitation hospital.

Evaluation items

Emotional factors

Positive and negative affect: Korean version of the Positive and Negative Affect Schedule (PANAS)

A scale to identify positive and negative affect in stroke patients, with an internal consistency reliability of 0.84. The Korean version of PANAS consists of 20 items (10 positive, 10 negative), each rated on a 5-point Likert scale (1 = Very slightly or not at all, 5 = Extremely.[29] The total score ranges from 10 to 50.

Depression: Korean version of the Beck Depression Inventory (BDI)

Comprising 21 items, the 0-3-point scale yields scores ranging from 0 to 63. Depression severity is categorized as mild (14-19), moderate (20-28), or severe (29-63). The Korean version of the BDI-II has demonstrated high internal consistency, with a Cronbach's alpha of 0.89.[30]

Anxiety: Korean version of the Hospital Anxiety and Depression Scale-Anxiety (HADS-A)

The assessment tool measures both anxiety and depression, but since it overlaps with the depression assessment used in this study, only the anxiety assessment was employed. It was developed in 1983 to measure the level of anxiety exhibited by patients visiting hospitals. This scale consists of 7 items. The score range on the 0-3-point scale is 0-21. Severity is categorized as mild (8-10), moderate (11-13), or severe (14-21). Anxiety was assessed using the Korean version of the Hospital Anxiety and Depression Scale-Anxiety subscale (HADS-A), which demonstrated good internal consistency (Cronbach's α=0.83).[31]

Apathy: Korean version of the Apathy Evaluation Scale (AES)

The 18 items related to apathy over the past 4 weeks are rated on a 4-point Likert scale (1-Not at all, 4-a lot). The score range is 18 to 72, with higher scores indicating a greater level of apathy symptoms. The Korean version of the Apathy Evaluation Scale (AES) demonstrated acceptable internal consistency, with a Cronbach's α of 0.76.[32]

Function factors

Cognition: Mini-Mental State Examination (MMSE)

The MMSE is a tool for assessing the subject's overall cognitive function level, covering domains such as orientation, memory, attention and calculation, language function, and spatial-temporal organization ability. The score range is 0-30 points, with a higher score indicating better cognitive function.[33,34]

Balance: Berg Balance Scale (BBS)

The BBS is a tool for assessing a subject's static and dynamic balance abilities. It consists of 14 items related to daily living activities, such as sitting, standing, changing position, and moving. Each item is scored from 0 to 4 points based on performance ability, with a total score range of 0 to 56 points. A higher score indicates better balance ability.[35]

Gait: Functional Ambulation Category (FAC)

FAC is a tool for assessing an individual's functional walking ability. It classifies walking ability into six levels from 0 to 5 based on the degree of assistance and supervision required during walking. A score of 0 indicates inability to walk, while a score of 5 indicates the ability to walk independently.[36]

Activities of daily living: Korean version of the Modified Barthel Index (K-MBI)

The K-MBI is an assessment tool for evaluating an individual's basic Activities of Daily Living (ADL). It consists of 10 items: eating, personal hygiene(grooming), dressing, mobility, bowel control, and bladder control. Each item is scored differently based on the level of performance, with a total score range of 0-100 points. A higher score indicates greater functional independence in daily living.[37]

Personal factors

The personal factors selected as independent variables in this study are sex, age, duration of illness (DISEASE_DUR), and length of stay (LOS).

Statistical analysis

Analysis was performed using the R statistical software (ver. 4.3.3, R Foundation for Statistical Computing, Vienna, Austria). Descriptive statistics were used to summarize the mean, standard deviation, minimum, and maximum values for all variables. To examine the relationships among emotional, functional, and personal variables, simple and multiple regression analyses were performed. Stepwise regression was used as an exploratory procedure to identify variables associated with emotional and functional outcomes, while multicollinearity and model assumptions were subsequently assessed. The level of statistical significance was set at $p<0.05$. Data were analyzed primarily using complete cases; however, one participant had missing data for length of

stay (LOS), and analyses involving LOS were conducted with N=101.

Simple Regression Analysis

To examine the associations among the study variables, separate simple regression models were fitted within the domains of emotional, functional, and personal variables.

Multiple Regression Analysis

Significant variables from the simple regression analyses were selected to construct multiple regression models. Separate models were constructed in which the emotion variable predicted the function, and in which the function variable predicted emotion. If a simple regression analysis yielded no more than two significant independent variables, multiple regression analysis was not performed, and the results were interpreted as is.

Testing Assumptions for Regression Analyses

To assess the suitability of the regression models, assumption tests were conducted for both simple and multiple regression analyses. The assumptions examined included linearity (RESET test), normality of residuals (Shapiro-Wilk test), homoscedasticity (Breusch-Pagan test), independence (Durbin-Watson test), and multicollinearity (Variance Inflation Factor <10).

Results

The descriptive statistics of all study variables, including the mean, standard deviation, minimum, and maximum values, are summarized in Table 1. Results of the regression analyses examining the associations between emotional variables and functional variables. Table 2 showed that cognitive function was significantly associated with lower negative affect and higher apathy. A better balance function was significantly associated with lower negative affect, depression, and anxiety. Better gait function was significantly associated with lower negative affect and depression. Better activities of daily living were significantly associated with lower negative affect, depression, and anxiety.

Table 1. Descriptive statistics of variables (N=102).

	Mean/N	SD	Min	Max
PANAS_P	25.45	8.90	10	43
PANAS_N	18.99	7.29	10	40
BDI	12.65	9.75	0	49
HADS	4.50	3.67	0	19
AES	50.64	8.88	30	66
MMSE	24.71	5.04	9	30
BBS	25.95	15.99	0	55
FAC	2.00	1.40	0	5
K-MBI	54.97	19.54	19	94
Male	44 (43.1%)	-	-	-
Female	58 (56.9%)	-	-	-
Age	65.13	12.94	27.0	86.0
DISEASE_DUR (months)	5.44	5.17	0.8	25.0
LOS**** (months)	4.29	5.11	0.2	22.6

Note: Gender is presented as N (%). All other variables are presented as mean, standard deviation, minimum, and maximum. One participant had missing data for length of stay (LOS), and analyses involving LOS were conducted with **** N=101.

Table 2. Simple regression analysis of emotion factors associated with function (N=102).

Dependent variable	Independent variable	β	SE	t	*p* value	R2
MMSE	PANAS_P	0.008	0.057	0.140	0.889	0.000
	PANAS_N	-0.160	0.067	-2.386	0.019*	0.054
	BDI	-0.067	0.051	-1.310	0.193	0.017
	HADS	-0.206	0.136	-1.521	0.131	0.023
	AES	0.141	0.055	2.573	0.012*	0.062
BBS	PANAS_P	0.086	0.180	0.477	0.635	0.002
	PANAS_N	-0.606	0.211	-2.874	0.005**	0.076
	BDI	-0.622	0.152	-4.097	<0.001**	0.144
	HADS	-1.001	0.424	-2.363	0.020*	0.053
	AES	0.213	0.179	1.191	0.236	0.014
FAC	PANAS_P	0.013	0.016	0.854	0.395	0.007
	PANAS_N	-0.050	0.019	-2.670	0.009**	0.067
	BDI	-0.048	0.014	-3.513	<0.001**	0.110
	HADS	-0.061	0.038	-1.619	0.109	0.026
	AES	0.023	0.016	1.473	0.144	0.021
K-MBI	PANAS_P	-0.052	0.220	-0.236	0.814	0.001
	PANAS_N	-0.929	0.251	-3.695	<0.001**	0.120
	BDI	-0.962	0.176	-5.473	<0.001**	0.231
	HADS	-1.579	0.508	-3.107	0.002**	0.088
	AES	0.337	0.217	1.550	0.124	0.023

***Note**: *p<0.05, **p<0.01.*

The results of the simple regression analysis examining the association between functional variables and emotional variables. Table 3 showed that negative affect was significantly associated with lower cognitive function, balance, gait, and activities of daily living. Depression was significantly associated with lower balance, gait, and activities of daily living, whereas anxiety was significantly associated with lower in balance and activities of daily living. Apathy was significantly associated only with higher cognitive function. Overall, higher functional levels tend to be associated with lower negative emotions, and higher cognitive function tends to be associated with higher apathy. To improve readability, the statistically significant results from Table 2 are additionally summarized in Figure 1 using standardized regression coefficients (β).

Table 3. Simple regression analysis of functional factors associated with emotion (N=102).

Dependent variable	Independent variable	B	SE	t	*p* value	R2
PANAS_P	MMSE	0.025	0.177	0.140	0.889	0.000
	BBS	0.026	0.056	0.477	0.635	0.002
	FAC	0.540	0.633	0.854	0.395	0.007
	K-MBI	-0.011	0.046	-0.236	0.814	0.001
PANAS_N	MMSE	-0.336	0.141	-2.386	0.019*	0.054
	BBS	-0.126	0.044	-2.874	0.005**	0.076
	FAC	-1.343	0.503	-2.670	0.009**	0.067
	K-MBI	-0.129	0.035	-3.695	<0.001**	0.120
BDI	MMSE	-0.251	0.192	-1.310	0.193	0.017
	BBS	-0.231	0.056	-4.097	<0.001**	0.144
	FAC	-2.308	0.657	-3.513	<0.001**	0.110
	K-MBI	-0.240	0.044	-5.473	<0.001**	0.231
HADS	MMSE	-0.110	0.072	-1.521	0.131	0.023
	BBS	-0.053	0.022	-2.363	0.020*	0.053
	FAC	-0.419	0.259	-1.619	0.109	0.026
	K-MBI	-0.056	0.018	-3.107	0.002**	0.088
AES	MMSE	0.439	0.171	2.573	0.012*	0.062
	BBS	0.066	0.055	1.191	0.236	0.014
	FAC	0.924	0.627	1.473	0.144	0.021
	K-MBI	0.070	0.045	1.550	0.124	0.023

Note: **p<0.05, **p<0.01.*

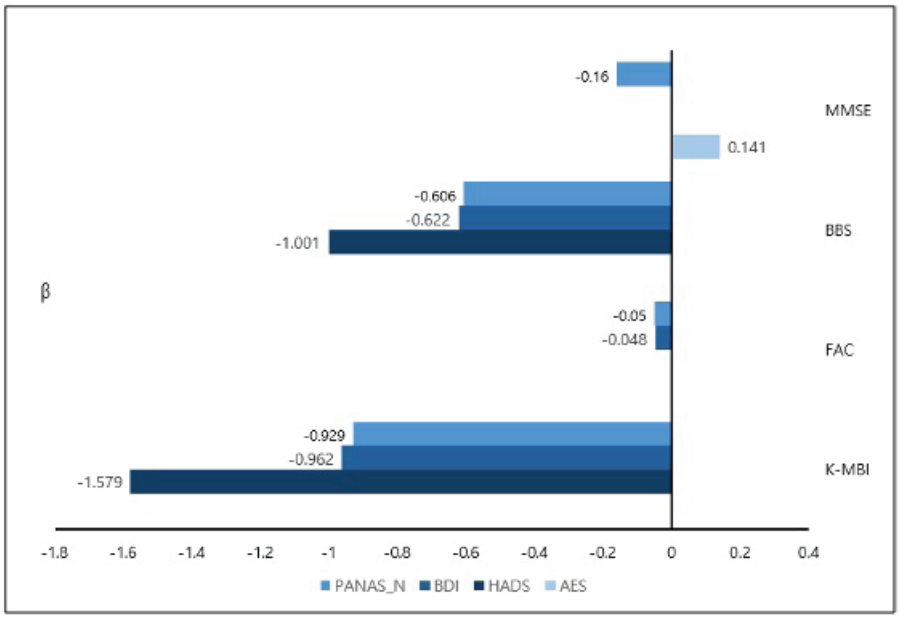


Figure 1. *Significant emotional factors associated with functional outcomes in simple regression analyses. Bars represent standardized regression coefficients (β) for statistically significant associations only.*

Taken together, no significant bidirectional association was observed between positive affect and functional variables. Negative affect, depression, anxiety, and apathy showed differential association patterns with functional variables (Table 3). Negative affect was significantly associated with all functional indicators, whereas depression was significantly associated with balance, gait, and activities of daily living. Anxiety showed more limited associations, being significant only for balance and activities of daily living. Apathy demonstrated a distinct association pattern, with significant associations only with cognition. The statistically significant results from Table 3 are likewise summarized in Figure 2 to facilitate visual comparison of the functional predictors of emotional outcomes.

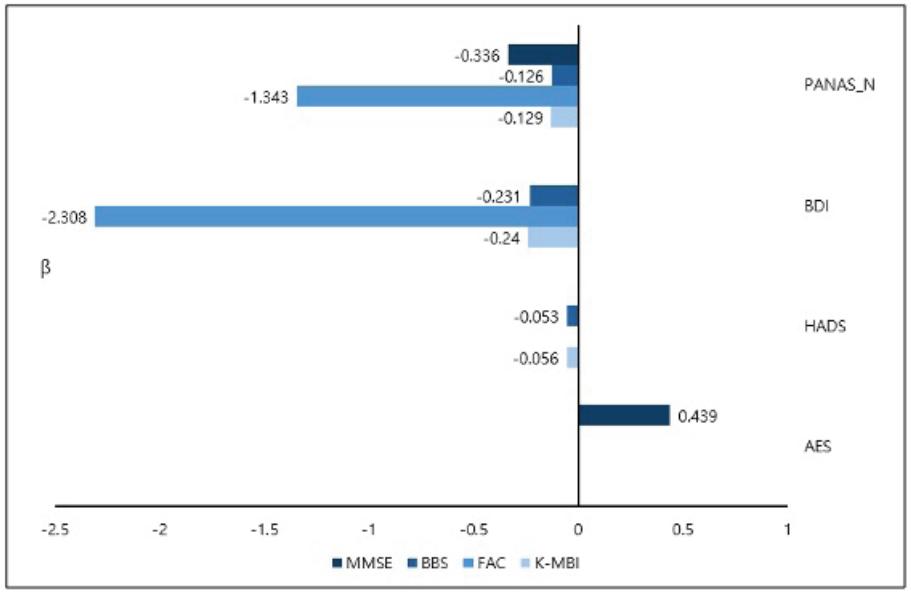


Figure 2. *Significant functional predictors of emotional outcomes in simple regression analyses. Bars represent standardized regression coefficients (β) for statistically significant associations only.*

The results of the simple regression analysis on the influence of personal variables and emotional variables. Table 4 showed no significant effects for sex or length of hospitalization. However, age showed a slight decrease in positive affect, and the duration of illness showed significantly higher apathy scores for shorter durations. This indicates that apathy scores tended to be lower among patients with a longer illness duration. In the simple regression analysis for personal variables, no more than two independent variables were significant, so they were not included in the multiple regression model. Because only two statistically significant associations were identified in Table 4, these results are presented separately in Figure 3.

Table 4. Simple regression analysis of personal factors associated with emotion (N=102).

Dependent variable	Independent variable	β	SE	t	*p* value	R2
PANAS_P	Sex	-2.432	1.771	-1.373	0.173	0.019
	Age	-0.134	0.067	-1.991	0.049*	0.038
	DISEASE_DUR	-0.150	0.171	-0.876	0.383	0.008
	LOS	-0.139	0.175	-0.791	0.431	0.006
PANAS_N	Sex	0.097	1.465	0.066	0.947	0.000
	Age	0.009	0.056	0.162	0.872	0.000
	DISEASE_DUR	0.065	0.141	0.463	0.644	0.002
	LOS****	0.007	0.143	0.046	0.963	0.000
BDI	Sex	-1.977	1.949	-1.014	0.313	0.010
	Age	0.055	0.075	0.732	0.466	0.005
	DISEASE_DUR	0.168	0.188	0.892	0.374	0.008
	LOS****	0.173	0.189	0.915	0.362	0.008
HADS	Sex	-0.799	0.733	-1.090	0.278	0.012
	Age	-0.017	0.028	-0.584	0.561	0.003
	DISEASE_DUR	-0.020	0.071	-0.280	0.780	0.001
	LOS****	-0.038	0.069	-0.541	0.590	0.003
AES	Sex	1.437	1.778	0.808	0.421	0.006
	Age	-0.108	0.068	-1.591	0.115	0.025
	DISEASE_DUR	-0.346	0.168	-2.059	0.042*	0.041
	LOS****	-0.327	0.172	-1.901	0.060	0.035

***Note**: *p<0.05, **p<0.01. ****N=101.*

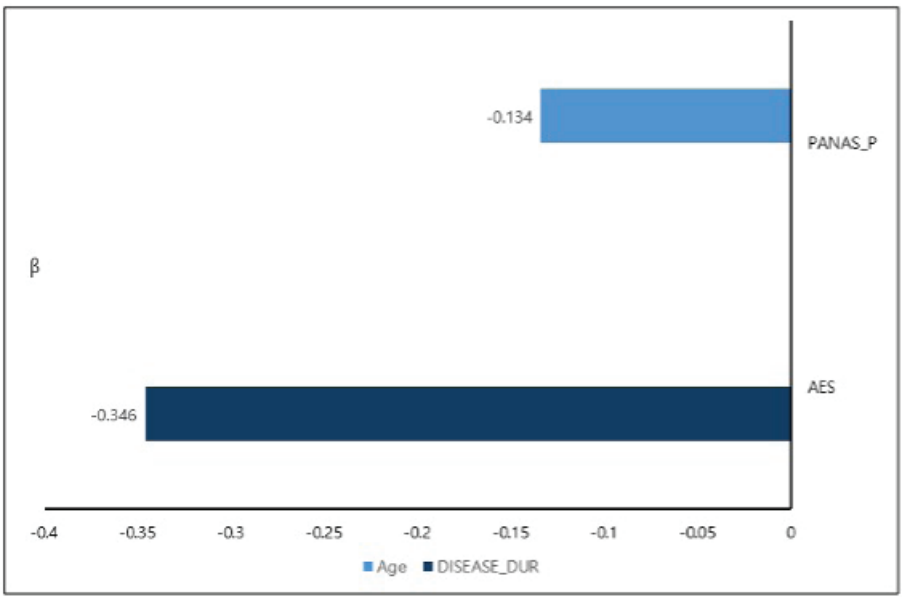


***Figure 3**. Significant associations of personal factors with emotional outcomes in simple regression analyses. Bars represent standardized regression coefficients (β) for statistically significant associations only.*

For the multiple regression analysis, only variables that were significant in two or more simple regression analyses were included in the model, and only statistically significant results are presented in Table 5. In the simple regression, multiple variables were mutually significant, but most disappeared in the multiple regression analysis. Among emotional variables, depression remained a significant predictor of balance, gait, and activities of daily living. Conversely, among functional variables, activities of daily living remained a significant predictor of depression. No significant predictors were identified for MMSE, negative affect, or anxiety in the multiple regression models.

Table 5. Statistically significant findings from the multiple regression analyses (N=102).

Dependent variable	Independent variable	B	SE	t	*p* value	R2
BBS	BDI	-0.606	0.219	-2.767	0.0068*	0.154
FAC	BDI	-0.039	0.016	-2.414	0.0176*	0.118
K-MBI	BDI	-0.929	0.253	-3.675	0.0004*	0.245
BDI	K-MBI	-0.280	0.081	-3.466	0.0008*	0.237

***Note**: Only statistically significant predictors from the multiple regression analyses are presented. No significant predictors were identified for MMSE, PANAS_N, or HADS in the multiple regression models. N=102.*

The assumptions for multiple regression were tested, and linearity and homoscedasticity were satisfied in most models. Multicollinearity was also found to be at an acceptable level (VIF<10). Although Durbin-Watson statistics were generally low, given that this study is based on cross-sectional data, violations of residual independence are not expected to significantly affect interpretation. Therefore, the regression models were considered acceptable for exploratory interpretation, although the results should be interpreted with caution.

Discussion

In the simple regression analyses, emotionalboth variables and functional variables showed significant associations, although the models' explanatory power was modest. Higher levels of negative affect, depression, and anxiety were consistently associated with poorer cognitive function, balance, gait ability, and activities of daily living. Conversely, lower functional performance was associated with higher levels of negative affect, depression, and anxiety. These findings suggest a bidirectional association between emotion variables and functional performance, with negative affect and depression emerging as the most consistently associated factors. This finding is consistent with prior research indicating that negative emotion affects functional levels in stroke patients.[22] Previous studies also indicated that self-regulation ability affects activities of daily living (ADL) performance. Taken together with the findings of this study, this suggests that rehabilitation medical institutions need to foster an environment that enables patients to maintain positive emotions and facilitates goal-directed, attentive focus.

However, apathy showed a distinct pattern, being significantly associated only with cognitive function and not with the other functional domains. Few studies have directly analyzed whether cognitive function and gait ability influence emotion. A study analyzing the effects of an early tailored nursing program in stroke patients found that the experimental group showed significantly improved cognitive function and significantly reduced depression and anxiety compared to the control group.[38] This suggests that improved cognitive function may influence emotional stability. This study provides empirical evidence supporting the possibility suggested by previous research.

Analysis of the effects of individual factors on emotional variables revealed that most individual factors did not significantly explain emotional state. This finding is consistent with previous research showing that post-stroke emotional disturbances are more strongly associated with disability, cognitive impairment, and stroke severity than with demographic characteristics such as age or sex.[39] In particular, disability and stroke severity have been identified as major predictors of post-stroke depression in systematic reviews.[40]

In this study, increasing age was associated with a slight decrease in positive affect; however, given the relatively narrow age distribution of the sample, this finding should be interpreted with caution. Previous research has suggested that this pattern may be related to reduced vitality and a more restricted expression of positive affect in older adults.[41] In the present study, sex was not significantly associated with emotional outcomes, in contrast to previous studies that reported a higher risk of depression among women.[15,24] Although prior research has reported that younger age may be associated with a higher risk of depression,[15] the current results indicate that demographic factors showed limited explanatory value in this sample.

The finding that apathy was higher in patients with a shorter illness duration may reflect temporal variation in post-stroke apathy rather than a simple linear change over time. Longitudinal research has shown that among

stroke survivors, 29% exhibited persistent apathy, 13% showed resolution over time, and 10% developed incident apathy between follow-up assessments.[42] These findings suggest that illness duration may be related to variation in apathy levels across different phases after stroke. However, in the present study, the explanatory power of illness duration was very limited (R^2=0.041), and this interpretation should therefore be made with caution. Duration of illness alone appears insufficient to predict apathy.

Furthermore, while simple regression analysis revealed a broad, bidirectional relationship between emotion and function, the multiple regression model showed a significant reduction in this association. Among the emotional factors explaining function, only depression remained independent, and among the functional factors explaining emotion, only activities of daily living remained. These results indicate that emotional and functional variables are correlated and share partially overlapping variance. That is, the relationship between emotion and function in stroke patients likely reflects a complex interplay among multiple interacting factors rather than a direct effect of a single factor.

Although regression analyses were conducted to explore relationships between emotional and functional variables, the cross-sectional nature of the data precludes conclusions about causality or directionality. Accordingly, the findings should be interpreted as associations rather than directional effects. In addition, reversing independent and dependent variables within the same dataset reflects correlational relationships rather than causal mechanisms. Future studies using longitudinal designs or more advanced analytical approaches, such as structural equation modeling or cross-lagged panel designs, are needed to clarify the directionality and underlying mechanisms of these relationships.

In summary, positive affect was not significantly associated with functional variables, whereas negative emotional variables were significantly associated with functional outcomes. These findings suggest meaningful associations between emotional and functional variables. However, given the cross-sectional design, they should be interpreted as correlational rather than directional. Therefore, medical institutions providing rehabilitation for stroke patients need to incorporate both emotional interventions or management and functional training when designing rehabilitation programs.

Individual factors did not, for the most part, significantly explain emotional states. However, given that emotional patterns may vary somewhat with age and illness duration, the results of this study suggest that emotional interventions could indirectly aid functional recovery and highlight the need for tailored psychological support. Xu *et al*. found that a multidimensional approach, including psychological counseling and emotional support, cognitive stimulation training, functional training guidance, family participation education, and motivational sessions as intervention elements in an early tailored nursing program for stroke patients, had a positive impact on emotional recovery and functional improvement[38]. In conclusion, a hospital-based, customized rehabilitation model that comprehensively addresses emotion, cognition, and function is likely to be effective for stroke patient recovery. Therefore, future research should focus on developing such a model in detail and designing a systematic program applicable in clinical settings.

Limitations

Limitations of this study include its cross-sectional design, which limits the causal interpretation, the single-institution setting, and the use of self-reported emotional scales. In addition, the use of exploratory regression approaches, including stepwise selection, and the relatively low explanatory power observed in several models may limit the robustness and generalizability of the findings. Future studies using longitudinal designs or more advanced analytical approaches, such as structural equation modeling, are needed to clarify potential directional relationships among these variables.

Conclusion

This study comprehensively analyzed the relationships among emotional, functional, and personal factors in stroke patients to explore their associations. The analysis revealed that higher levels of cognitive function (MMSE), balance (BBS), gait ability (FAC), and activities of daily living (K-MBI) were associated with lower levels of negative affect (PANAS_N), depression (BDI), and anxiety (HADS). Conversely, lower levels of negative affect and depression were associated with improved balance and activities of daily living, supporting the presence of reciprocal associations between emotional state and functional performance. These findings suggest the necessity for integrated management of emotional regulation and functional recovery in stroke patients.

Among individual factors, most showed no significant influence; however, differences in certain emotional responses were observed, such as a decrease in positive emotions with increasing age and a reduction in apathy with longer illness duration. This suggests that emotional states may be influenced by complex factors such as functional level and psychological adaptation, rather than by simple demographic characteristics.

In multiple regression analysis, most effects disappeared, leaving only depression and activities of daily living as independent predictors. This suggests that the relationship between emotion and function is not a simple one-to-one correspondence but rather a structure formed through the interaction of multiple factors. It implies that managing depression and

enhancing activities of daily living are particularly important intervention elements in rehabilitation therapy.

Based on these findings, this study emphasizes that a functional-focused approach alone is insufficient in the rehabilitation process for stroke patients; it must be complemented by emotional and psychological assessment and intervention. Particularly, among emotional factors, negative emotions and depression consistently showed a strong association across all functional performance domains. Therefore, early screening and systematic management of emotional status within healthcare institutions can serve as a key strategy to enhance the efficiency of functional recovery.

Future research requires analysis utilizing methods to more clearly identify causal pathways among function, emotion, and personal factors. Additionally, studies are needed to develop and validate an integrated model incorporating multidimensional elements such as psychological counseling, emotion regulation training, cognitive stimulation, family involvement, and motivational strategies when designing hospital-based rehabilitation programs. This study can serve as foundational data for developing protocols for such integrated rehabilitation approaches.

Ethical approval

This study was approved by the Institutional Review Board of Eulji University (IRB No. EUIRB2024-046).

Funding

This research received no external funding

Conflict of interest

The authors declare no conflict of interest in this study

CRediT authorship contribution statement

Euna Ko: conceptualization, methodology, data collection, data analysis, interpretation of results, writing: original draft preparation, review and editing; **Wongjong Yu**: supervision, research consultation, and critical review of the manuscript.

References

[1] Feigin VL, Norrving B, Mensah GA. Global burden of stroke. Circ Res. 2017; 120(3): 439-48. doi: 10.1161/CIRCRESAHA.116.308413

[2] Stineman MG, Bates BE, Kurichi JE, Kwong PL, Ripley DC, Vogel WB, et al. Initial physical grades and cognitive stages after acute stroke: who receives comprehensive rehabilitation services? PM R. 2013; 5(12): 1007-18. doi: 10.1016/j.pmrj.2013.07.008.

[3] Wade DT, Wood VA, Hewer RL. Recovery after stroke--the first 3 months. J Neurol Neurosurg Psychiatry.1985; 48(1): 7-13. doi: 10.1136/jnnp.48.1.7.

[4] Kang JH, Bae HJ, Choi YA, Lee SH, Shin HI. Length of hospital stay after stroke: A Korean nationwide study. Ann Rehabil Med. 2016; 40(4): 675-81. doi: 10.5535/arm.2016.40.4.675.

[5] Kim MK, Park YJS, Choi EH. Effect of mirror therapy combined with lower extremity muscle strength exercise on gait and balance of patients with chronic stroke. J Korean Soc Phys Med. 2018; 13(1): 81-8. doi: 10.13066/ksep.2018.13.1.81.

[6] Tyson SF, Hanley M, Chillala J, Selley A, Tallis RC. Balance disability after stroke. Phys Ther. 2006; 86(1): 30-8. doi: 10.1093/ptj/86.1.30.

[7] Sturm JW, Donnan GA, Dewey HM, Macdonell RA, Gilligan AK, Srikanth V, et al. Quality of life after stroke: the North East Melbourne Stroke Incidence Study (NEMESIS). Stroke. 2004; 35(10): 2340-5. doi: 10.1161/01.STR.0000141977.18520.3b.

[8] Suzuki A, Mutai H, Furukawa T, Wakabayashi A, Hanihara T. The prevalence and course of neuropsychiatric symptoms in stroke patients impact functional recovery during in-hospital rehabilitation. Top Stroke Rehabil.2022; 29(1): 1-8. doi: 10.1080/10749357.2020.1863261.

[9] Nemani K, Gurin L. Neuropsychiatric Complications after Stroke. Semin Neurol. 2021; 41(1): 85-100. doi: 10.1055/s-0040-1722223.

[10] Medeiros GC, Roy D, Kontos N, Beach SR. Post-stroke depression: A 2020 updated review. Gen Hosp Psychiatry.2020; 66: 70-80. doi: 10.1016/j.genhosppsych.2020.06.011.

[11] Isuru A, Hapangama A, Ediriweera D, Samarasinghe L, Fonseka M, Ranawaka U. Prevalence and predictors of new onset depression in the acute phase of stroke. Asian J Psychiatr. 2021; 59: 102636. doi: 10.1016/j.ajp.2021.102636.

[12] Cumming TB, Blomstrand C, Skoog I, Linden T. The high prevalence of anxiety disorders after stroke. Am J Geriatr Psychiatry. 2016; 24(2): 154-60. doi: 10.1016/j.jagp.2015.06.003.

[13] Knapp P, Dunn-Roberts A, Sahib N, Cook L, Astin F, Kontou E, et al. Frequency of anxiety after stroke: An updated systematic review and meta-analysis of observational studies. Int J Stroke. 2020; 15(3): 244-55. doi: 10.1177/1747493019896958.

[14] Carnes-Vendrell A, Deus J, Molina-Seguin J, Pifarre J, Purroy F. Depression and apathy after transient ischemic attack or minor stroke: Prevalence, evolution and predictors. Sci Rep. 2019; 9(1): 16248. doi: 10.1038/s41598-019-52711-2.

[15] Shi Y, Yang D, Zeng Y, Wu W. Risk factors for post-stroke depression: A meta-analysis. Front Aging Neurosci.2017; 9: 218. doi: 10.3389/fnagi.2017.00218.

[16] Schmaal L, Veltman DJ, van Erp TG, Samann PG, Frodl T, Jahanshad N, et al. Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. Mol Psychiatry.2016; 21(6): 806-12. doi: 10.1038/mp.2015.69.

[17] Kvael LAH, Bergland A, Telenius EW. Associations between physical function and depression in nursing home residents with mild and moderate dementia: a cross-sectional study. BMJ Open. 2017; 7(7): e016875. doi: 10.1136/bmjopen-2017- 016875.

[18] Robinson RG, Jorge RE. Post-stroke depression: A review. Am J Psychiatry. 2016; 173(3): 221-31. doi: 10.1176/appi.ajp.2016.15030363.

[19] Kutlubaev MA, Hackett ML. Part II: predictors of depression after stroke and impact of depression on stroke outcome: an updated systematic review of observational studies. Int J Stroke. 2014; 9(8): 1026-36. doi: 10.1111/ijs.12356.

[20] Shi Y, Lenze EJ, Mohr DC, Lee JM, Hu L, Metts CL, et al. Post-stroke depressive symptoms and cognitive performances: A network analysis. Arch Phys Med Rehabil. 2024; 105(5): 892-900. doi: 10.1016/j.apmr.2023.12.012.

[21] Lee Y, Nicholas ML, Connor LT. Identifying emotional contributors to participation post-stroke. Top Stroke Rehabil.2023; 30(2): 180-92. doi: 10.1080/10749357.2021.2010834.

[22] Choe YW, Kim MK. Could self-control and emotion influence physical ability and functional recovery after stroke? Medicina (Kaunas). 2021; 57(10): 1038. doi: 10.3390/medicina57101038.

[23] Schuch F, Vancampfort D, Firth J, Rosenbaum S, Ward P, Reichert T, et al. Physical activity and sedentary behavior in people with major depressive disorder: A systematic review and meta-analysis. J Affect Disord. 2017; 210: 139-50. doi: 10.1016/ j. jad.2016.12.011.

[24] Kim M, Lee YH. Gender Differences in the risk of depression in community-dwelling stroke survivors compared to the general population without stroke. Chonnam Med J. 2023; 59(2): 134-9. doi: 10.4068/ cmj.2023.59.2.134.

[25] Xiao W, Liu Y, Huang J, Huang LA, Bian Y, Zou G. Analysis of factors associated with depressive symptoms in stroke patients based on a national cross-sectional study. Sci Rep. 2024; 14(1): 9268. doi: 10.1038/s41598-024-60124-x.

[26] Gular K, Sivasubramanian V, Reddy RS, Tedla JS, Dixit S. The mediating effect of age, gender, and post-stroke duration on the association between trunk and upper limb recovery in subacute stroke population: A cross-sectional study with mediation analysis. Int J Environ Res Public Health. 2022; 19(23): 15644. doi: 10.3390/ijerph19 2315644.

[27] Aderinto N, AbdulBasit MO, Olatunji G, Adejumo T. Exploring the transformative influence of neuroplasticity on stroke rehabilitation: a narrative review of current evidence. Ann Med Surg (Lond). 2023; 85(9): 4425-32. doi: 10.1097/MS9.0000000000001152.

[28] Butsing N, Zauszniewski JA, Ruksakulpiwat S, Griffin MTQ, Niyomyart A. Association between post-stroke depression and functional outcomes: A systematic review. PLoS One. 2024; 19(8): e0309158. doi: 10.1371/journal.pone.0309158.

[29] Lee HH, Kim EJ, Lee MG. A validation study of Korean positive and negative affect schedule: The PANAS Scales. Korean J Clin Psychol. 2003; 22(4): 935-46. doi: 10.15842/kjcp.2003.22.4.014.

[30] Lee EH, Lee SJ, Hwang ST, Hong SH, Kim JH. Reliability and validity of the Beck Depression Inventory-II among Korean adolescents. Psychiatry Investig. 2017; 14(1): 30-6. doi: 10.4306/pi.2017. 14.1.30.

[31] Choi JH, Jeon S, Hong S, Kim A, Park JY, Yang HJ. The Reliability and validity of the Korean version of hospital anxiety and depression scale using Rasch Measurement Theory in patients with Parkinson's disease. J Korean Neurol Assoc. 2021; 39(4): 312-21. doi: 10.17340/jkna.2021.4.3.

[32] Lee YM, Park IH, Koo MS, Ko SH, Kang HM, Song JE. The reliability and validity of the Korean version of apathy evaluation scale and its application in patients with schizophrenia. J Korean Neuropsychiatr Assoc. 2013; 16(2): 80-5. doi: 10.12934/ jkpmhn.2013.22.2.80.

[33] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975; 12(3): 189-98. doi: 10.1016/0022-3956 (75)90026-6.

[34] Kim TH, Jhoo JH, Park JH, Kim JL, Ryu SH, Moon SW, et al. Korean version of mini mental status examination for dementia screening and its' short form. Psychiatry Investig. 2010; 7(2): 102-8. doi: 10.4306/pi.2010.7.2.102.

[35] Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health. 1992;83 Suppl 2:S7-11. (PMID: 1468055).

[36] Holden MK, Gill KM, Magliozzi MR, Nathan J, Piehl-Baker L. Clinical gait assessment in the neurologically impaired. Reliability and meaningfulness. Phys Ther. 1984; 64(1): 35-40. doi: 10.1093/ptj/ 64.1.35.

[37] Jung HY, Park BK, Shin HS, Kang YK, Pyun SB, Paik NJ, et al. Development of the Korean version of Modified Barthel Index (K-MBI): Multi-center study for subjects with stroke. Ann Rehabil Med. 2007; 31(3): 283-97.

[38] Xu L, An L, Lin Q, Liu M. Early personalized care in stroke patients improves the cognitive function and promotes the recovery process of disease. Int J Clin Exp Med. 2020; 13(4): 2466-73. (ISSN: 1940-5901).

[39] Ayerbe L, Ayis S, Wolfe CD, Rudd AG. Natural history, predictors and outcomes of depression after stroke: systematic review and meta-analysis. Br J Psychiatry. 2013; 202(1): 14-21. doi: 10.1192/ bjp.bp.111.107664.

[40] Ayerbe L, Ayis S, Rudd AG, Heuschmann PU, Wolfe CD. Natural history, predictors, and associations of depression 5 years after stroke: the South

London Stroke Register. Stroke. 2011; 42(7): 1907-11. doi: 10.1161/STROKEAHA.110.605808.

[41] Schweizer S, Stretton J, Van Belle J, Price D, Calder AJ, Cam CAN, et al. Age-related decline in positive emotional reactivity and emotion regulation in a population-derived cohort. Soc Cogn Affect Neurosci. 2019; 14(6): 623-31. doi: 10.1093/scan/nsz 038.

[42] Chen J, Li Y, Liu J, Zhang Y, Zeng Y, Chen M, et al. Symptom clusters of early-stage poststroke depression: A mixed-methods study. Nurs Open. 2021; 8(5): 2488-97. doi: 10.1002/nop2.894.

Journal of Associated Medical Sciences 2026; 59(3): 55-71

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



The interplay between serum Interleukin-33 and vitamin D3 in childhood asthma: Diagnostic utility and severity predictors in an Iraqi community

Zahraa Hazim Aziz* **and Thikra Abdullah Mahmood**

Department of Microbiology, College of Medicine, University of Kufa, Najaf Governorate, Iraq.

ARTICLE INFO

Article history:
Received 12 March 2026
Accepted as revised 15 May 2026
Available online 20 May 2026

Keywords:
Childhood asthma, interleukin-33, vitamin D deficiency, biomarkers, ROC curve, alarmin

ABSTRACT

Background: The epithelial alarmin Interleukin-33 (IL-33) and the immunomodulator vitamin D3 are implicated in asthma pathogenesis, yet their combined clinical utility and interplay in pediatric populations remain underexplored.

Objectives: This study aimed to investigate the diagnostic performance and predictive value of serum IL-33 and vitamin D3 in childhood asthma, and their association with disease severity.

Materials and methods: A case-control study was conducted involving 120 children (60 asthmatics and 60 age- and gender-matched healthy controls) in Karbala, Iraq. Serum levels of IL-33 and 25-hydroxyvitamin D3 were quantified using enzyme-linked immunosorbent assay (ELISA). Diagnostic utility was evaluated via Receiver Operating Characteristic (ROC) curve analysis, and independent predictors were identified using multivariate binary logistic regression.

Results: Asthmatic children exhibited significantly elevated IL-33 (395.56±218.75 vs 159.95±55.32 ng/L, p<0.001) and lower vitamin D3 (17.01±7.38 vs 28.20±10.60 ng/mL, p<0.001) than controls. A strong inverse correlation was observed between IL-33 and Vitamin D3 (r = -0.614, p<0.001), both of which showed a dose-dependent relationship with asthma severity. ROC analysis revealed exceptional diagnostic accuracy for IL-33 (AUC=0.933). Multivariate regression identified body mass index (aOR = 1.528, p=0.014) and IL-33 (aOR=1.015 per 1 ng/L increase, p=0.020) as significant independent predictors of asthma, yielding an overall model accuracy of 95.8%. Notably, male asthmatics exhibited significantly lower vitamin D3 levels than females (p<0.001).

Conclusion: The inverse relationship between elevated IL-33 and vitamin D3 deficiency underscores their synergistic role in asthma exacerbation. IL-33 serves as a highly accurate diagnostic biomarker, highlighting the potential for targeted anti-alarmin therapies and gender-stratified nutritional interventions in pediatric asthma management.

Introduction

Asthma is one of the most prevalent chronic respiratory diseases globally, imposing a substantial burden on healthcare systems and diminishing the quality of life for millions of individuals. According to the Global Burden of Disease (GBD) study and the World Health Organization (WHO), asthma affected an estimated 262 million people worldwide in 2019 and accounted for approximately 455,000 deaths.[1] The

* **Corresponding contributor.**
Author's Address: *Department of Microbiology, College of Medicine, University of Kufa, Najaf Governorate, Iraq.*
E-mail address: zahraah.aljanaby@student.uokufa.edu.iq
doi: 10.66285/JAMS.2026.079
E-ISSN: 2539-6056

prevalence of childhood asthma has been rising steadily, particularly in developing nations and the Middle East, driven by a complex interplay of genetic predisposition, urbanization, and environmental exposures such as passive smoking.[2] In Iraq, epidemiological studies have reported childhood asthma prevalence rates ranging from 7% to over 22% in certain urban centers, highlighting a critical regional health challenge.[3] Despite advancements in therapeutic strategies, a significant proportion of pediatric patients continue to experience recurrent exacerbations and suboptimal disease control, underscoring the urgent need to elucidate the precise molecular and immunological mechanisms driving asthma pathogenesis.

Recent immunological research has increasingly focused on the role of airway epithelial cells, not merely as a physical barrier, but as frontline regulators of innate immunity. Upon exposure to environmental triggers such as allergens, pollutants, or cold air, damaged epithelial cells rapidly release a group of cytokines known as "alarmins," which include Thymic Stromal Lymphopoietin (TSLP), Interleukin-25 (IL-25), and Interleukin-33 (IL-33).[4] Among these, IL-33, a member of the IL-1 cytokine superfamily discovered by Schmitz et al. in 2005, has emerged as a crucial mediator of type-2 airway inflammation.[5] IL-33 exerts its biological effects by binding to its specific receptor, ST2 (IL-1RL1), which is widely expressed on innate lymphoid cells type 2 (ILC2s), mast cells, Th2 cells, and eosinophils. This binding triggers a potent cascade of pro-inflammatory cytokines (IL-4, IL-5, and IL-13), leading to airway hyperresponsiveness, mucus hypersecretion, and airway remodeling.[6] Meta-analyses have demonstrated that serum IL-33 levels are significantly elevated in asthmatic children and correlate with disease severity, positioning IL-33 as both a promising biomarker and a potential therapeutic target.[7]

Concurrently, there is growing evidence implicating nutritional status, particularly vitamin D deficiency, in the pathogenesis and exacerbation of childhood asthma. Vitamin D is a potent immunomodulator that regulates both innate and adaptive immune responses.[8] Epidemiological data indicate a pandemic of vitamin D deficiency among children, particularly in the Middle East, where cultural practices, reduced outdoor physical activity, and limited sun exposure contribute to critically low serum 25-hydroxyvitamin D3 levels.[9] Systematic reviews have shown that vitamin D supplementation can significantly reduce the rate of asthma exacerbations requiring systemic corticosteroids.[10] However, the exact molecular mechanisms by which vitamin D deficiency exacerbates asthma remain incompletely understood.

A pivotal yet underexplored mechanistic link between vitamin D and airway inflammation involves regulation of the IL-33/ST2 axis. Experimental *in vitro* studies have demonstrated that vitamin D actively enhances the production of soluble ST2 (sST2), a decoy receptor that binds to free IL-33, thereby neutralizing its pro-inflammatory signaling.[11] Consequently, vitamin D deficiency may reduce sST2, allowing unchecked IL-33 activity and amplifying the inflammatory cascade. Despite this compelling biological rationale, there is a distinct knowledge gap in the clinical literature. A recent comprehensive systematic review of 71 studies investigating the association between vitamin D status and inflammatory biomarkers in asthma noted significant inconsistencies and a lack of sufficient clinical data specifically examining the simultaneous interaction between vitamin D and IL-33 in pediatric cohorts.[12] Furthermore, most existing studies have evaluated these biomarkers in isolation, and the combined diagnostic utility of serum IL-33 and vitamin D-particularly using rigorous Receiver Operating Characteristic (ROC) curve analysis and multivariate predictive modeling-has not been adequately established.

To address this knowledge gap, the present study was designed to investigate the dual profiles of serum IL-33 and vitamin D3 in children with asthma. The novelty of this research extends significantly beyond being the first study of its kind in an Iraqi pediatric population. Specifically, this study aims to: 1) provide clinical validation of the inverse correlation between a potent epithelial alarmin (IL-33) and an immunomodulatory nutrient (vitamin D3) across varying degrees of asthma severity, 2) establish precise, ROC-derived diagnostic cut-off values for both biomarkers to differentiate asthmatic children from healthy controls, and 3) construct a highly accurate multivariate logistic regression model to evaluate the independent predictive value of these biomarkers alongside anthropometric factors such as body mass index (BMI). By bridging the gap between *in vitro* mechanistic theories and *in vivo* clinical realities, this study seeks to provide new scientific insights that could inform targeted, biomarker-guided therapeutic and nutritional interventions for childhood asthma.

Materials and methods

Study design and setting

A case-control study was conducted over 7 months, from July 10, 2025, to February 15, 2026. The study enrolled children attending the outpatient clinics and emergency department of the Children Teaching Hospital in Karbala, Iraq. The research protocol was reviewed and officially approved by the Institutional Ethics Committee at the College of Medicine, University of Kufa, in strict adherence to the ethical principles outlined in the Declaration of Helsinki. Prior to enrollment, written informed consent was obtained from the parents or legal guardians of all participating children.

Study population and sampling

The study comprised 120 pediatric subjects, systematically divided into two age- and gender-matched

groups. The case group included 60 children diagnosed with asthma, while the control group consisted of 60 apparently healthy children with no personal or familial history of allergic or chronic respiratory diseases. The sample size was determined to be statistically adequate and consistent with similar regional case-control studies in Iraq that investigated immunological markers and vitamin D status in pediatric asthma. For instance, Al-Tuama *et al*. used a comparable cohort of 116 subjects (50 asthmatics and 66 controls) in Karbala,[13] Al-Hasnawi and Al-Hasnawi evaluated Interleukins in 149 pediatric subjects[14] and Al-Sharifi *et al*. investigated vitamin D deficiency in 100 Iraqi children.[15]

Inclusion and exclusion criteria

Asthmatic children were diagnosed and classified according to the Global Initiative for Asthma (GINA) guidelines.[16] Disease severity was stratified into mild, moderate, and severe based on clinical presentation, symptom frequency, and the requirement for pharmacological intervention. Children with other chronic systemic illnesses, autoimmune diseases, congenital anomalies, or those currently receiving immunomodulatory therapies (other than standard asthma medications) were excluded from the study to prevent confounding effects on inflammatory biomarkers.

Clinical and demographic data collection

A comprehensive questionnaire was administered via direct interviews with the parents/guardians to collect detailed demographic and clinical data. Anthropometric measurements, including weight and height, were recorded to calculate the BMI. Additional data gathered included passive smoking exposure, residential area (urban vs rural), household pet exposure, and a detailed family history of asthma or atopy. Clinical manifestations (e.g., shortness of breath, wheezing, cough) and specific environmental triggers (e.g., cold air, dust, food allergies) were systematically documented.

Blood sampling and preparation

Approximately 3-5 mL of venous blood was drawn from each participant under aseptic conditions. The whole blood samples were collected into sterile, plain gel tubes and allowed to clot at room temperature for 30 min. Subsequently, the samples were centrifuged at 3000 rpm for 15 min to separate the serum. The extracted serum was carefully aliquoted into sterile Eppendorf tubes and immediately cryopreserved at -20 °C until biochemical analysis, avoiding repeated freeze-thaw cycles that could degrade the target proteins.

Biochemical analysis (ELISA)

The quantitative determination of serum biomarkers was performed using highly sensitive, commercially available enzyme-linked immunosorbent assay (ELISA) kits, strictly following the manufacturer's standard operating procedures. Serum IL-33 and 25-hydroxyvitamin D3 were measured using commercially available ELISA kits according to the manufacturers' instructions. Serum IL-33 was quantified with the human IL-33 ELISA Kit (Cat. No. E0044Hu, Bioassay Technology Laboratory [BT LAB], Zhejiang, China), and serum 25-hydroxyvitamin D3 was assessed using the Human Vitamin D ELISA Kit (Cat. No. E1543Hu, BT LAB, Zhejiang, China). Both assays were based on a quantitative sandwich enzyme immunoassay technique using pre-coated microtiter plates, biotin-conjugated detection antibodies, and streptavidin-horseradish peroxidase, with optical density measured at 450 nm. The assay range and analytical sensitivity were 10-2000 ng/L and 5.36 ng/L for IL-33, and 0.5-200 ng/mL and 0.23 ng/mL for 25-hydroxyvitamin D3, respectively. Assay precision was acceptable for both biomarkers, with intra-assay and inter-assay coefficients of variation of <8% and <10%, respectively. Vitamin D status was classified as sufficient (≥30 ng/mL), insufficient (20.0-29.9 ng/mL), or deficient (<20 ng/mL) according to established pediatric endocrine guidelines.[17]

Statistical analysis

Data management and statistical analyses were executed using the Statistical Package for the Social Sciences (SPSS) software, version 26.0 (IBM Corp., Armonk, NY, USA). Continuous variables were assessed for normal distribution using the Shapiro-Wilk test. Normally distributed continuous data were presented as Mean±SD and compared using the independent samples t-test for two groups, or One-way Analysis of Variance (ANOVA) followed by Tukey's post hoc test for multi-group comparisons (e.g., asthma severity). Categorical variables were expressed as frequencies and percentages (N, %) and analyzed utilizing the Pearson Chi-Square test or Fisher's Exact test, as appropriate.

Bivariate correlations between continuous variables were evaluated using Spearman's rank correlation coefficient (r) given the data's nonnormal distribution. To assess the diagnostic utility of the biomarkers, Receiver Operating Characteristic (ROC) curve analysis was performed, calculating the Area Under the Curve (AUC), sensitivity, specificity, and optimal cut-off values determined via Youden's *J* statistic. A multivariate binary logistic regression model was constructed to identify independent predictors of childhood asthma, reporting adjusted Odds Ratios (aOR) and 95% Confidence Intervals (CI). The model's goodness-of-fit was verified using the Hosmer-Leme show test and Nagelkerke R^2. A two-tailed $p<0.05$ was considered statistically significant, while $p<0.001$ was deemed highly significant.

Results

Table 1 shows that children with asthma and controls were comparable in age and sex (p=0.718 and p=0.708, respectively). However, BMI was significantly higher in the asthma group (19.58±5.73 vs 16.61±2.18 kg/m^2, p=0.001), with overweight/obesity identified only among asthmatic children (15.0%). Passive smoking, positive family history of asthma/atopy, and urban residence were all significantly more frequent in asthmatic children than in controls (60.0% vs 28.3%, p<0.001, 53.3% vs 13.3%, p<0.001, and 70.0% vs 50.0%, p=0.040, respectively), while pet exposure was comparable between groups (p=0.265). All respiratory symptoms and triggers were significantly more common in the asthma group (all p<0.001). Asthma severity was mild in 33.3%, moderate in 40.0%, and severe in 26.7% of cases.

Table 1. Comparison of demographic and clinical characteristics between asthma patients and healthy controls.

Parameter	**Asthma** (N=60)	**Control** (N=60)	***p* value**
Age (years), Mean±SD	9.27±2.79	9.53±3.16	0.718 (NS)
• 5-8	26 (43.3%)	26 (43.3%)	
• 9-11	18 (30.0%)	15 (25.0%)	
• 12-15	16 (26.7%)	19 (31.7%)	
Gender, N (%)			0.708 (NS)
• Male	38 (63.3%)	35 (58.3%)	
• Female	22 (36.7%)	25 (41.7%)	
BMI (kg/m^2), Mean±SD	19.58±5.73	16.61±2.18	0.001**
• Underweight (<18.5)	35 (58.3%)	51 (85.0%)	
• Normal (18.5-24.9)	16 (26.7%)	9 (15.0%)	
• Overweight (25-29.9)	5 (8.3%)	0 (0.0%)	
• Obese (≥30)	4 (6.7%)	0 (0.0%)	
Passive smoking exposure	36 (60.0%)	17 (28.3%)	<0.001***
Pet exposure	12 (20.0%)	7 (11.7%)	0.265 (NS)
Urban residence	42 (70.0%)	30 (50.0%)	0.040*
Family history of asthma	32 (53.3%)	8 (13.3%)	<0.001***
Clinical manifestations, N (%)			
• Shortness of breath	54 (90.0%)	0 (0.0%)	<0.001***
• Wheezing	37 (61.7%)	0 (0.0%)	<0.001***
• Cough	44 (73.3%)	0 (0.0%)	<0.001***
Food allergy	27 (45.0%)	3 (5.0%)	<0.001***
Triggers, N (%)			
• Cold air	46 (76.7%)	0 (0.0%)	<0.001***
• Dust	45 (75.0%)	0 (0.0%)	<0.001***
Asthma severity, N (%)			
• Mild	20 (33.3%)	NA	
• Moderate	24 (40.0%)	NA	
• Severe	16 (26.7%)	NA	

Note: *NS: not significant at p>0.05), * p<0.05, **p<0.01; ***p<0.001, NA: not applicable. Continuous variables are expressed as Mean±SD and compared using the independent samples t-test. Categorical variables are expressed as N (%) and compared using the Chi-Square test or Fisher's Exact test.*

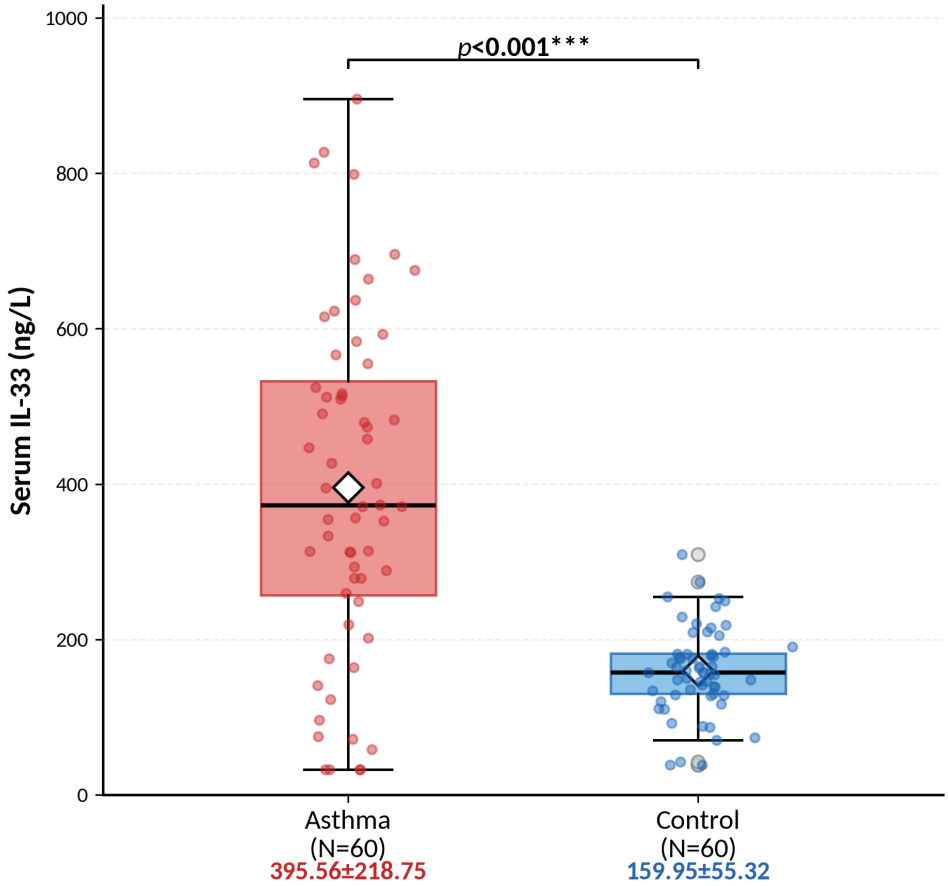


Figure 1. *Comparison of IL-33 levels between asthma patients (N=60) and healthy controls (N=60). Box plots show median (horizontal line), mean (diamond), interquartile range (box), and individual data points. The numbers represent Mean±SD.*

Comparison of serum biomarker levels between asthma patients and healthy controls

Table 2, Figures, and Figure 2 showed marked differences in serum biomarker levels between asthmatic children and controls. Serum IL-33 was significantly higher in the asthma group than in controls (395.56 218.75 vs 159.95±55.32 ng/L, $p<0.001$), whereas serum 25(OH)D3 was significantly lower (17.01±7.38 vs 28.20±10.60 ng/mL, $p<0.001$). These findings indicate an enhanced inflammatory profile and reduced vitamin D status in children with asthma.

Table 2. Serum biomarker levels in asthma patients and healthy controls.

Parameter	**Asthma** (N=60)	**Control** (N=60)	***p* value**
IL-33 (ng/L)	395.56±218.75	159.95±55.32	<0.001***
25-Hydroxyvitamin D3 (vitamin D3, ng/L)	17.01±7.38	28.20±10.60	<0.001***

Note: *Data are expressed as Mean±SD. A statistical comparison was performed using the independent-samples t-test. ***p<0.001 indicates a highly significant difference.*

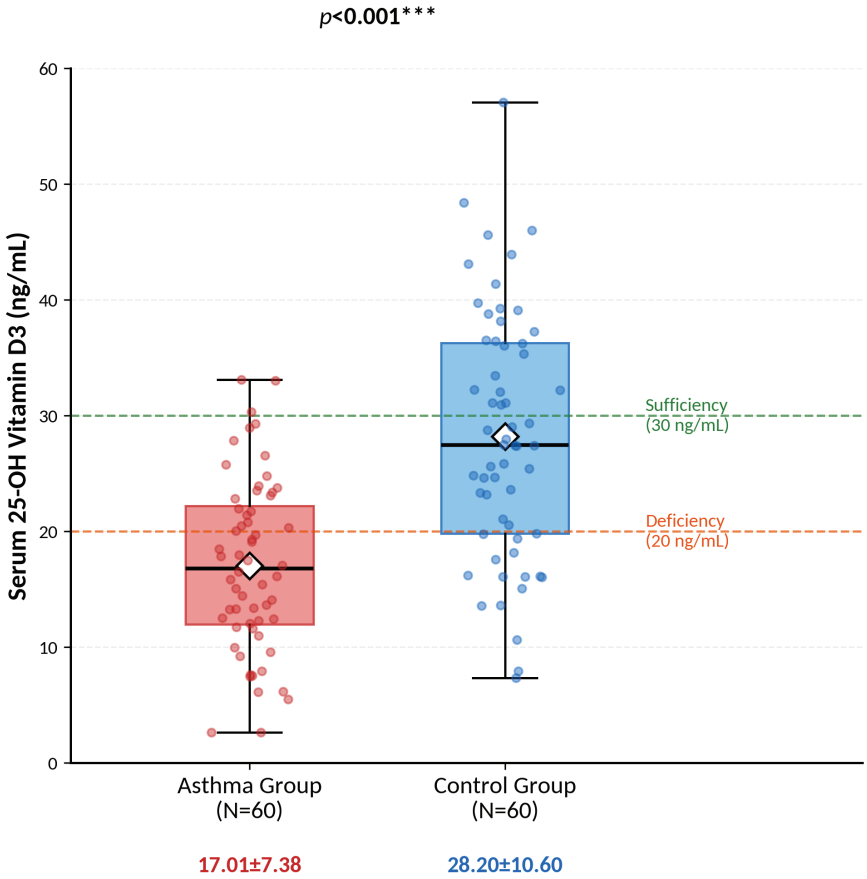


Figure 2. *Comparison of serum 25-hydroxyvitamin D3 levels between asthma patients (N=60) and healthy controls (N=60). Box plots show median (horizontal line), mean (diamond), interquartile range (box), and individual data points. Dashed lines indicate vitamin D deficiency (<20 ng/mL) and sufficiency (≥30 ng/mL) cutoffs. The numbers represent Mean±SD.*

Table 3, Figures 3, and Figure 4 showed significant associations between serum biomarkers and asthma severity. IL-33 increased progressively across mild, moderate, and severe asthma (276.2±87.6, 386.7±79.3, and 595.5±132.9 ng/L, respectively, $p<0.001$), while 25(OH)D3 decreased significantly with increasing severity (20.7±4.4, 17.5±5.6, and 11.0±3.8 ng/mL, respectively, $p<0.001$). Pairwise comparisons confirmed significant differences between severity groups for both markers, indicating that higher IL-33 and lower vitamin D levels were associated with more severe asthma.

Table 3. Comparison of serum biomarker levels among asthmatic children classified by disease severity.

Biomarker	**Mild** (N=20)	**Moderate** (N=24)	**Severe** (N=16)	***p* value**
Serum IL-33 (ng/L)	276.2±87.6A	386.7±79.3B	595.5±132.9C	<0.001***
25(OH)D3 (ng/mL)	20.7±4.4A	17.5±5.6B	11.0±3.8C	<0.001***

Note: *Data presented as mean±SD. Different letters (A/B/C) indicate significant pairwise differences (one-way ANOVA followed by Tukey's post hoc test, p<0.05); ***p<0.001 indicates a highly significant difference among severity groups, A, B, C: significant pairwise differences (one-way ANOVA followed by Tukey's post hoc test).*

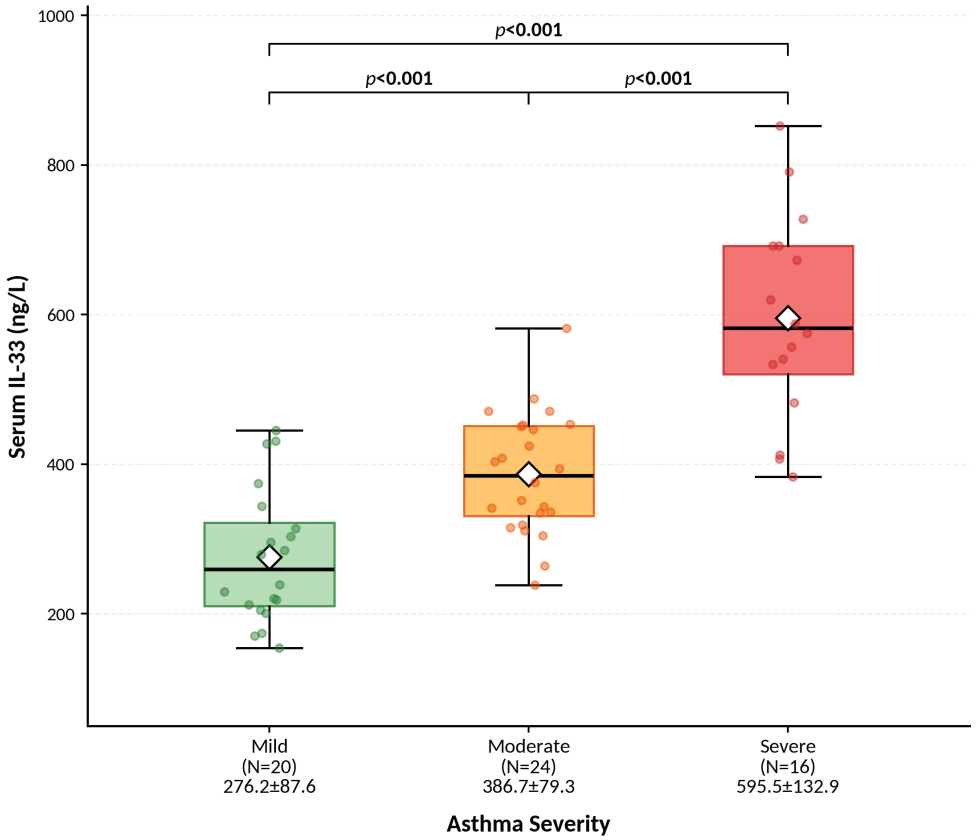


Figure 3. *Comparison of serum IL-33 levels (ng/L) among asthmatic children classified by disease severity. IL-33 levels increased progressively with severity (p<0.001 for all pairwise comparisons).*

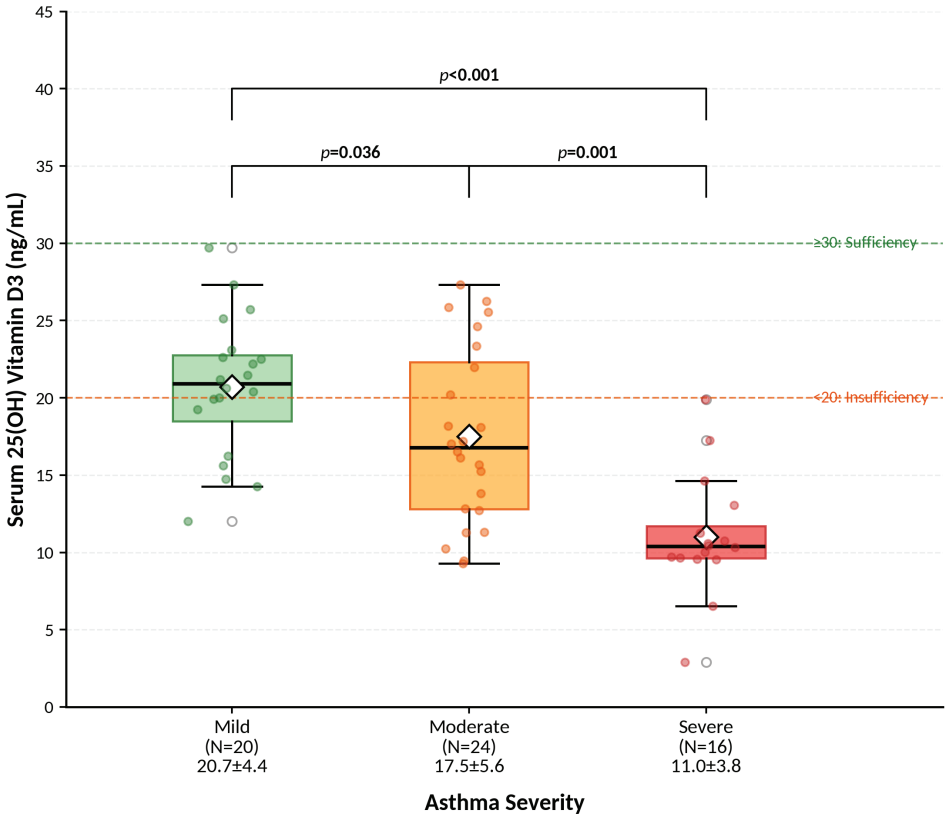


Figure 4. *Comparison of serum 25-hydroxyvitamin D3 levels (ng/mL) among asthmatic children classified by disease severity. Dashed lines indicate clinical thresholds for vitamin D insufficiency (<20 ng/mL) and sufficiency (≥30 ng/mL). Vitamin D3 showed an inverse relationship with severity (p=0.036 for Mild vs Moderate, p=0.001 for Moderate vs Severe, p<0.001 for Mild vs Severe).*

Spearman correlation matrix between serum biomarkers in asthmatic children

Table 4 and Figure 5 showed a significant inverse correlation between serum IL-33 and 25(OH)D3 levels in asthmatic children (r=-0.614, p<0.001). Higher IL-33 levels were associated with lower vitamin D concentrations, supporting an inverse relationship between inflammatory activity and vitamin D status in childhood asthma.

Table 4. Spearman's rank correlation coefficient (r) between inflammatory biomarkers and vitamin D3 among asthmatic children (N=60).

Biomarkers	Serum IL-33	Serum vitamin D3
Serum IL-33 (ng/L)	1.000	-0.614***
Serum vitamin D3 (ng/mL)	-0.614***	1.000

Note: *Data are presented as Spearman's rank correlation coefficient (r), ***p<0.001 indicates a highly significant correlation.*

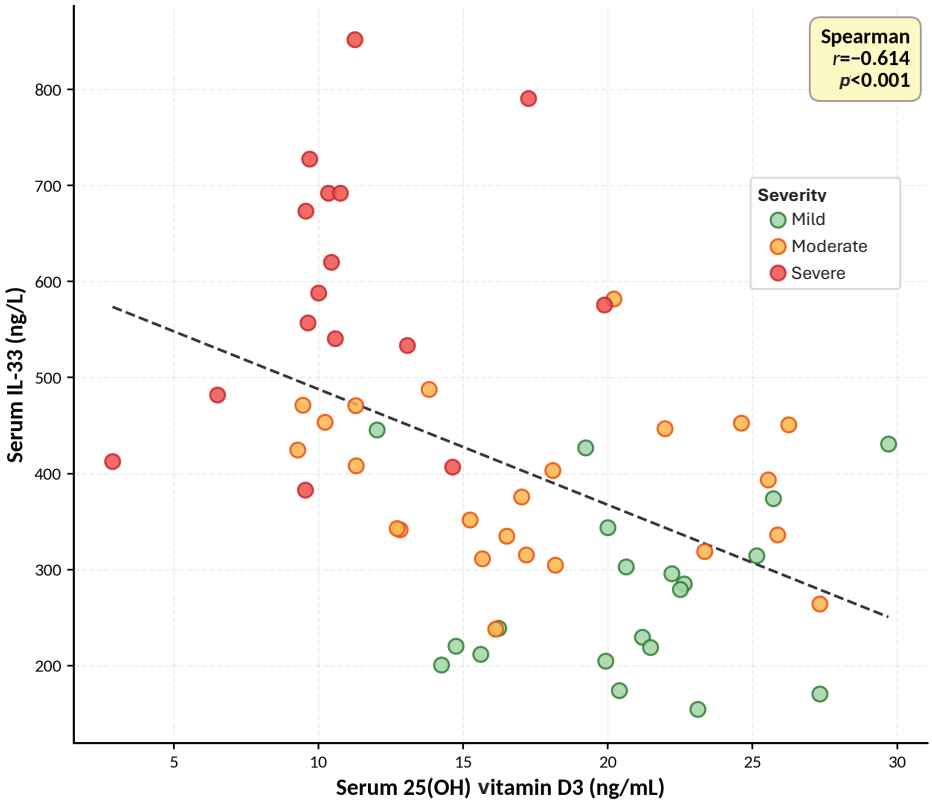


Figure 5. *Scatter plot showing the inverse correlation between serum 25(OH) vitamin D3 (ng/mL) and serum IL-33 (ng/L) in asthmatic children (N=60). Color-coded by disease severity. Spearman correlation coefficient r=-0.614, p<0.001. This negative correlation is consistent with vitamin D's immunomodulatory role in suppressing epithelial alarmin production.*

Table 5 and Figure 6 showed that both IL-33 and 25(OH)D3 significantly distinguished asthmatic children from healthy controls (both p<0.001). IL-33 demonstrated superior diagnostic performance, with an AUC of 0.933 (95% CI: 0.874-0.978) and an optimal cutoff of ≥239.4 ng/L, yielding 85.0% sensitivity and 93.3% specificity. In comparison, 25(OH)D3 showed a lower but still significant diagnostic value, with an AUC of 0.819 (95% CI: 0.734-0.890) and an optimal cutoff of ≤20.3 ng/mL, corresponding to 80.0% sensitivity and 75.0% specificity.

Table 5. Diagnostic performance of biomarkers in differentiating asthma patients from healthy controls (ROC curve analysis).

Diagnostic performance	Biomarker	
	Serum IL-33 (ng/L)	Serum vitamin D3 (ng/mL)
Cut-off	≥239.4	≤20.3
AUC (95% CI)	0.933 (0.874-0.978)	0.819 (0.734-0.890)
p value	<0.001***	<0.001***
Sensitivity (%)	85.0	85.0
Specificity (%)	93.3	75.0
PPV (%)	92.7	76.2
NPV (%)	86.2	78.9
Accuracy (%)	89.2	77.5
Youden's *J*	0.783	0.550

Note: *ROC: receiver operating characteristic, AUC: area under the curve, CI: confidence interval, PPV: positive predictive value, NPV: negative predictive value. Cut-off values were determined using Youden's J statistic. For vitamin D3, values below the cut-off indicate asthma, and for IL-33, values above the cut-off indicate asthma.*

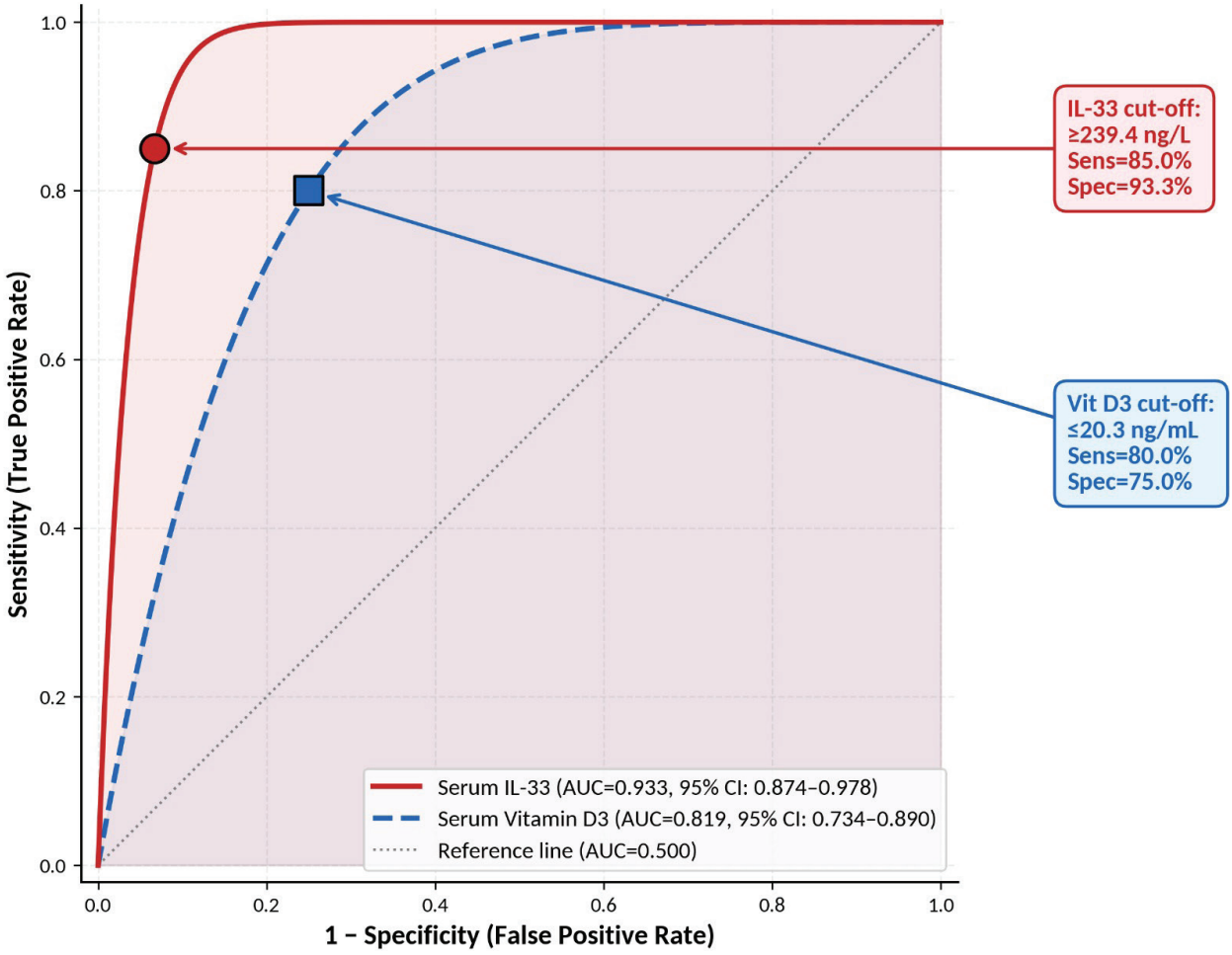


Figure 6. *ROC curve analysis of serum IL-33 and vitamin D3 for the diagnosis of childhood asthma.*

Binary logistic regression analysis of independent predictors of childhood asthma

Table 6 and Figure 7 showed that BMI and serum IL-33 were independent predictors of childhood asthma in the multivariable logistic regression model. Higher BMI was associated with increased odds of asthma (aOR=1.528, 95% CI: 1.089-2.144, *p*=0.014), and higher IL-33 was also significantly associated with asthma risk (aOR=1.015, 95% CI: 1.002-1.027, *p*=0.020). Although higher 25(OH)D3 was associated with lower odds of asthma, this association was not statistically significant (aOR=0.919, *p*=0.097). The model showed excellent performance, with good calibration, high discriminatory ability (AUC=0.972), and an overall accuracy of 95.8%.

Table 6. Binary logistic regression model for independent predictors of childhood asthma.

Panel A. Final parsimonious logistic regression coefficients.

Predictor	B	SE	Wald	aOR	95% CI	*p* value	Interpretation
BMI (kg/m^2)	0.424	0.173	6.02	1.528	1.089-2.144	0.014*	Each 1 kg/m^2 increase: 52.8% higher asthma odds.
Serum IL-33 (ng/L)	0.015	0.006	5.38	1.015	1.002-1.027	0.020*	Each 10 ng/L increase: 16.1% higher asthma odds.
Serum vitamin D3 (ng/mL)	-0.084	0.051	2.75	0.919	0.832-1.015	0.097	Protective trend; 8.1% lower odds per 1 ng/mL, NS.

Panel B. Model summary and goodness-of-fit.

Metric	Value	Interpretation
-2 Log-Likelihood	40.63	Lower values indicate better fit in comparative models.
Chi-square (model)	125.72, *p*<0.001	Overall model was highly significant.
Cox & Snell R^2	0.649	About 64.9% of variance explained.
Nagelkerke R^2	0.866	About 86.6% of variance explained.
Overall classification accuracy	95.8% (115/120)	Excellent overall classification performance.
Sensitivity	93.3% (56/60)	Strong case-detection ability.
Specificity	98.3% (59/60)	Excellent control identification.
AUC (ROC)	0.972	Outstanding discriminative ability.
Hosmer-Leneshow test	chi-square=8.95, *p*=0.176	Adequate calibration and good model fit.

Note: *B: regression coefficient, SE: standard error, AOR: adjusted odds ratio, CI: confidence interval. *p<0.05.*

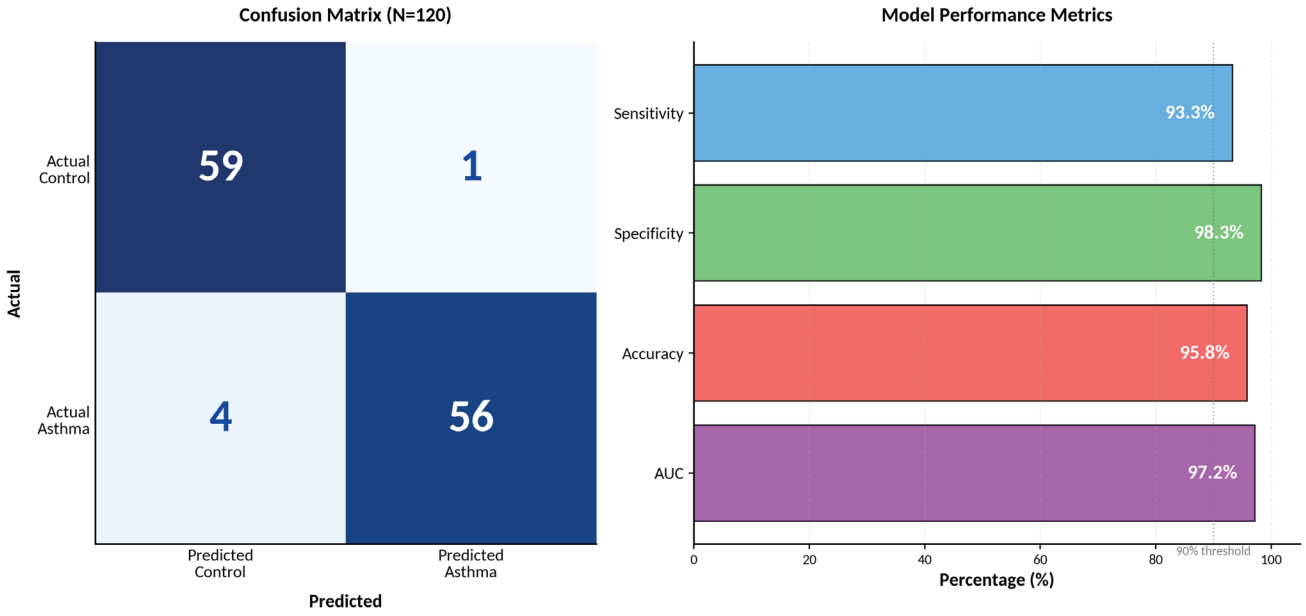


Figure 7. *Classification performance of the final multivariate logistic regression model. Left panel: confusion matrix, Right panel: model performance metrics (sensitivity, specificity, accuracy, AUC).*

Effect of gender on serum biomarker levels in asthmatic children

Table 7 showed a significant sex-related difference in serum 25(OH)D3 levels among asthmatic children, with males having lower levels than females (13.79±5.36 vs. 21.51±5.15 ng/mL, $p<0.001$), with a large effect size (r=0.609). In contrast, serum IL-33 levels were higher in males than females, but the difference was not statistically significant (428.91±156.41 vs 365.70±132.17 ng/L, $p=0.0858$), with a small effect size (r=0.222).

Table 7. Comparison of biomarker levels between male and female asthmatic children (N=60).

Parameter	**Males** (N=38)	**Females** (N=22)	p value	**Effect size** (r)	Significance
IL-33 (ng/L)	428.91±156.41	365.70±132.17	0.0858	0.222 (small)	NS
Vitamin D3 (ng/mL)	13.79±5.36	21.51±5.15	<0.001	0.609 (large)	***

Note: *Data presented as Mean±SD, NS: not significant, ***p<0.001 = highly significant.*

Discussion

The present study provides compelling evidence for the interplay between systemic inflammation, as reflected by elevated serum IL-33, and nutritional deficiency, manifested by reduced serum 25-hydroxyvitamin D3, in the pathogenesis and clinical severity of childhood asthma. To the best of our knowledge, this is one of the first studies conducted in Iraq to simultaneously investigate both biomarkers in a pediatric asthmatic population and to evaluate their combined diagnostic and predictive utility through ROC curve analysis and multivariate logistic regression modeling.

Demographic and environmental risk factors

The demographic analysis revealed that the asthma and control groups were well-matched for age and sex, thereby minimizing confounding and strengthening the internal validity of the comparative analyses. However, several modifiable and non-modifiable risk factors were significantly associated with the asthma phenotype in our cohort.

BMI was significantly higher in asthmatic children than in healthy controls (19.58±5.73 vs 16.61±2.18 kg/m^2, $p=0.001$), with 15.0% of the asthma group classified as overweight or obese, compared to none in the control group. This finding is consistent with the large-scale meta-analysis by Deng *et al.*, which demonstrated that overweight and obesity are significant risk factors for childhood asthma and wheeze across 18 studies involving 73,252 children.[18] Similarly, Azizpour *et al.* confirmed this association in their systematic review and meta-analysis of case-control studies, reporting a significant positive relationship between elevated BMI and asthma risk in children.[19] The mechanistic basis for this association involves adipose tissue-derived pro-inflammatory cytokines (adipokines), including leptin and TNF-alpha, which promote systemic inflammation and airway hyperresponsiveness, thereby predisposing obese children to asthma development and exacerbation.[20]

Notably, the relationship between BMI, IL-33, vitamin D, and asthma is recognized to vary across pediatric age strata and between sexes. Adiposity-related inflammation may exert distinct effects across early childhood, pre-pubertal, and adolescent periods, and pubertal changes substantially modulate vitamin D status, particularly in adolescent boys and in regions with high baseline deficiency, such as the Middle East.[9] A formal age- and gender-stratified evaluation of these biomarkers across the three age subgroups (5-8, 9-11, and 12–15 years) was beyond the statistical power of the present cohort and is acknowledged in the Limitations section as an important direction for future, adequately powered studies.

Passive smoking exposure was markedly more prevalent among asthmatic children (60.0%) compared to controls (28.3%, p<0.001). This observation aligns with the comprehensive meta-analysis conducted by He *et al.,* which demonstrated a robust positive association between both prenatal and postnatal secondhand smoke exposure and the occurrence of childhood asthma.[21] Environmental tobacco smoke contains numerous irritants and oxidative agents that damage the airway epithelium, enhance mucus secretion, and promote Th2-mediated inflammatory responses, all of which are central to asthma pathophysiology.[1]

A positive family history of asthma or atopy was reported in 53.3% of asthmatic children compared to only 13.3% of controls (p<0.001). This strong familial predisposition is well-documented in the literature. Ober and Yao reported heritability estimates for asthma ranging between 35% and 95%, underscoring the substantial genetic contribution to disease susceptibility.[22] Bjerg *et al*. further demonstrated in a large pediatric cohort that a family history of asthma and atopy significantly increases the risk of developing these conditions in offspring, even after adjusting for known environmental risk factors.[23]

Urban residence was significantly associated with asthma in our study (70.0% vs 50.0%, p=0.040). This finding is supported by the extensive review by Grant *et al*., which concluded that children living in urban areas are at higher risk of developing asthma due to increased exposure to traffic-related air pollution, indoor allergens, and reduced green space.[24] Aligne *et al*. similarly identified urban residence as an independent risk factor for childhood asthma after controlling for race, poverty, and other environmental variables.[25]

Serum IL-33 levels

The central finding of this study is the markedly elevated serum IL-33 levels in asthmatic children compared to healthy controls (395.56±218.75 ng/L vs 159.95±55.32 ng/L, p<0.001). This result agrees with several Iraqi studies. Shaban *et al*., in a study conducted in Iraq, reported that serum IL-33 was significantly up-regulated in asthma patients compared to controls, further supporting the role of this alarmin cytokine in the local asthmatic population.[26] Additionally, Al-Hasnawi and Al-Hasnawi investigated IL-25 and IL-33 as immunological markers in Iraqi pediatric asthma patients and confirmed the elevation of both Interleukins in the asthma group.[14]

Internationally, our findings are corroborated by the landmark study of Bahrami Mahneh et al. conducted in Iran, which demonstrated significantly higher serum IL-33 levels in 61 asthmatic children compared to 63 age- and gender-matched controls (15.17±32.3 vs 0.61±2.16 pg/ml, p=0.028).[27] Although the absolute concentrations differ due to variations in ELISA kits and assay methodologies, the directional trend is identical. Wang *et al*. further consolidated this evidence through a systematic review and meta-analysis, confirming that serum IL-33 is consistently elevated in children with asthma across multiple populations.[7]

Furthermore, the progressive increase in IL-33 levels corresponding to asthma severity observed in our study (mild: 276.2±87.6, moderate: 386.7±79.3, severe: 595.5±132.9 ng/L, p<0.001) mirrors the severity-dependent pattern reported by Bahrami Mahneh et al. (2016), who documented IL-33 concentrations of 9.92, 13.68, and 31.92 pg/ml in mild, moderate, and severe asthma, respectively (p=0.026).[27] Yang *et al*. also confirmed that serum IL-33 is elevated in asthmatic children and increases in proportion to the severity of pulmonary ventilation obstruction.[28] These converging findings across diverse populations strongly support the utility of IL-33 as a biomarker of disease activity and severity in childhood asthma.

Serum vitamin D3 levels

Serum 25-hydroxyvitamin D3 levels were significantly lower in asthmatic children compared to controls (17.01±7.38 vs 28.20±10.60 ng/mL, p<0.001), with the asthma group's mean falling below the clinical deficiency threshold of 20 ng/mL. This finding is particularly relevant in the Iraqi context. Al-Tuama *et al*., in a study conducted at the same hospital setting in Karbala, Iraq, reported a significant association between vitamin D deficiency and childhood asthma, with asthmatic children demonstrating substantially lower vitamin D levels than their healthy counterparts.[13] Al-Sharifi *et al*. similarly concluded that vitamin D deficiency is very common among Iraqi asthmatic children and that vitamin D inadequacy is predominant in asthmatic children living in Iraq.[15]

At the international level, our results are consistent with the large case-control study by Bener *et al*. in Qatar, which identified vitamin D deficiency as a strong predictor of asthma in children, with most asthmatic children exhibiting vitamin D deficiency compared to controls.[29] Esfandiar *et al*. also demonstrated a significant inverse relationship between vitamin D levels and asthma severity in children in an Iranian study.[30] The dose-dependent decline in vitamin D

levels with increasing asthma severity observed in our study (mild: 20.7±4.4, moderate: 17.5±5.6, severe: 11.0±3.8 ng/mL, $p<0.001$) is particularly noteworthy, as it indicates that severe asthmatic children are in a state of critical vitamin D deficiency, potentially exacerbating their inflammatory burden.

Inverse correlation between IL-33 and vitamin D: mechanistic insights

A key finding of this study is the strong inverse correlation between serum IL-33 and vitamin D3 levels (Spearman r=-0.614, $p<0.001$). This relationship provides important mechanistic insights into the immunopathology of childhood asthma. The biological plausibility of this inverse association is well-supported by the seminal work of Pfeffer *et al.,* who demonstrated that vitamin D enhances the production of soluble ST2 (sST2), a decoy receptor that sequesters IL-33 and inhibits its pro-inflammatory signaling.[11] Specifically, vitamin D selectively upregulates the expression of sST2 mRNA in human bronchial epithelial cells and $CD4^+$ T lymphocytes, and conditioned medium from vitamin D-treated epithelial cells significantly inhibited IL-33-induced mast cell activation.[11] This mechanism provides a direct molecular link between vitamin D deficiency and the unchecked elevation of IL-33 observed in our asthmatic cohort.

IL-33, a member of the IL-1 cytokine family, functions as an epithelial-derived alarmin that is released upon cellular damage or stress. As described by Cayrol and Girard, IL-33 activates group 2 innate lymphoid cells (ILC2s), inducing the production of large amounts of IL-5 and IL-13, which are central mediators of type 2 inflammation, eosinophilic infiltration, mucus hypersecretion, and airway remodeling.[6] Duchesne *et al*. further elaborated that IL-33 released from bronchial epithelial cells serves as a frontline mediator of the asthmatic inflammatory response, bridging innate and adaptive immunity.[4]

Rai *et al*. provided additional evidence that vitamin D acts as an immunomodulator that regulates IL-33 secretion and macrophage phenotypic polarization, suggesting that adequate vitamin D status may serve as a natural brake on the IL-33/ST2 inflammatory axis.[31] Bonanno *et al*. also reported an inverse correlation between 25-hydroxyvitamin D levels and IL-33 in children with allergic airway disease, further supporting the biological relevance of our findings.[32]

Diagnostic utility of IL-33 and vitamin D3

The ROC curve analysis revealed that serum IL-33 possesses excellent diagnostic accuracy for childhood asthma, with an AUC of 0.933 (95% CI: 0.874-0.978), a sensitivity of 85.0%, and a specificity of 93.3% at an optimal cut-off of ≥239.4 ng/L. This diagnostic performance substantially exceeds that reported by Shaban *et al*. in Iraq, who found a sensitivity of 60.6% and specificity of 59.5% at a cut-off of 68.3 pg/mL.[26] The superior performance in our study may be attributed to differences in ELISA kit sensitivity, patient selection criteria, and the larger proportion of moderate-to-severe cases in our cohort. Yang *et al*. demonstrated that combining IL-33 with other markers, such as FEF75% z-score and FeNO, can further improve diagnostic sensitivity to 90.1%, suggesting that multi-marker panels may offer even greater clinical utility.[28]

Serum vitamin D3 also demonstrated good diagnostic value (AUC=0.819, 95% CI: 0.734-0.890), with an optimal cut-off of ≤20.3 ng/mL closely aligning with the established clinical threshold for vitamin D deficiency (<20 ng/mL). This convergence between the statistically derived cut-off and the clinically recognized deficiency threshold enhances the translational relevance of our findings and supports the integration of vitamin D assessment into routine asthma evaluation protocols.

Independent predictors of childhood asthma

The multivariate binary logistic regression analysis identified BMI (aOR=1.528, 95% CI: 1.089-2.144, $p=0.014$) and serum IL-33 (aOR=1.015 per 1 ng/L increase, $p=0.020$) as significant independent predictors of childhood asthma. The identification of BMI as an independent predictor is consistent with the findings of Fang *et al*., who reported that a one-unit increase in BMI was associated with 4% higher odds of asthma in children,[33,] and with those of Li *et al*., who confirmed that high BMI independently predicts increased asthma occurrence.[34]

Although serum vitamin D3 showed a protective trend (aOR=0.919), it did not reach statistical significance in the multivariate model ($p=0.097$). This may be explained by the strong collinearity between vitamin D and IL-33, as evidenced by their robust inverse correlation (r=-0.614). When both variables are included simultaneously in a regression model, the dominant variance explained by IL-33 may attenuate the independent contribution of vitamin D. Nevertheless, the overall model demonstrated outstanding performance, with a Nagelkerke R^2 of 0.866, an overall classification accuracy of 95.8%, and an AUC of 0.972, confirming the excellent predictive capacity of the combined biomarker panel.

Gender disparities in biomarker levels

An important secondary finding of this study is the significant gender-based disparity in vitamin D levels among asthmatic children. Males exhibited substantially lower vitamin D concentrations compared to females (13.79±5.36 vs 21.51±5.15 ng/mL, $p<0.001$), with a large effect size (r=0.609). This finding is consistent with the study by Li *et al*., who reported that males with two or more allergic diseases had significantly lower vitamin D levels than their female counterparts.[35] Hollams *et al*. in a large Australian cohort, also found that males with inadequate vitamin D at ages

6 and 14 years had increased atopy and bronchial hyperresponsiveness.[36]

The gender disparity in vitamin D levels may be attributed to several factors, including differences in outdoor activity patterns, clothing habits (particularly in Middle Eastern societies where cultural practices may differ between genders), dietary intake, and hormonal influences on vitamin D metabolism. This finding has important clinical implications, suggesting that male asthmatic children may represent a particularly vulnerable subgroup requiring targeted vitamin D screening and supplementation strategies.

In contrast, serum IL-33 levels did not differ significantly between male and female asthmatic children (p=0.0858), with only a small effect size (r=0.222). This suggests that the inflammatory burden, as measured by IL-33, is primarily driven by disease-related factors rather than gender-specific biological differences.

Clinical implications

The findings of this study carry several important clinical implications. Firstly, the strong diagnostic performance of serum IL-33 (AUC=0.933) supports its potential utility as a non-invasive biomarker for diagnosing childhood asthma, particularly in settings where pulmonary function testing may be challenging in young children. Secondly, the significant inverse correlation between vitamin D and IL-33 suggests that vitamin D supplementation may be a viable adjunctive therapeutic strategy to modulate the inflammatory response in asthmatic children. Pfeffer and Hawrylowicz reviewed the evidence supporting a role for vitamin D in secondary prevention, with a focus on reducing asthma exacerbations and inflammation.[37] However, the VDKA randomized clinical trial by Forno *et al*. found that vitamin D3 supplementation did not significantly improve time to severe asthma exacerbation in children with low vitamin D levels, highlighting the need for further investigation into optimal dosing, timing, and patient selection for vitamin D interventions.[38] Thirdly, the identification of BMI as an independent predictor of asthma reinforces the importance of weight management programs as part of comprehensive asthma care in children. Lastly, the pronounced vulnerability of male asthmatic children to severe vitamin D deficiency warrants gender-stratified clinical assessments and targeted nutritional interventions.

Strengths and limitations

This study has several notable strengths. It is one of the first studies in Iraq to simultaneously evaluate both IL-33 and vitamin D in pediatric asthma within a single cohort, enabling direct correlation analysis. The comprehensive statistical approach, including ROC curve analysis with Youden's *J* statistic optimization and multivariate logistic regression with a detailed goodness-of-fit assessment, yields robust, clinically translatable findings. The inclusion of effect size calculations alongside p values enhances the interpretability of the results beyond mere statistical significance. Furthermore, the age- and sex-matching of the study groups minimizes potential confounding.

However, several limitations should be acknowledged. Firstly, the cross-sectional case-control design precludes establishing causal relationships between the biomarkers and asthma. Longitudinal prospective studies are needed to determine whether elevated IL-33 and reduced vitamin D precede asthma onset or are consequences of the disease process. Secondly, the sample size (N=120), while consistent with comparable regional studies, may limit the generalizability of the findings to the broader Iraqi pediatric population. Thirdly, the study did not assess other potential confounders such as dietary habits, sun exposure duration, physical activity levels, or concurrent medication use, which may influence both vitamin D status and inflammatory markers. Fourthly, although the cohort spanned three age subgroups (5-8, 9-11, and 12-15 years) and both genders, the study was not originally powered for age- and gender-stratified subgroup analyses of BMI, IL-33, and vitamin D within the asthma group. With per-subgroup sizes ranging from 16 to 38 asthmatic children, formal stratified statistical comparisons would be underpowered and at high risk of yielding unstable estimates. Accordingly, age- and sex-stratified evaluations of these biomarkers-including age-specific identification of subgroups with the most pronounced IL-33 elevation and vitamin D deficiency-represent an important direction for future, adequately powered, prospectively designed studies. Fifthly, soluble ST2 (sST2) was not measured in our cohort. Although the proposed mechanistic link between vitamin D deficiency and unrestrained IL-33 activity is grounded in the previously reported finding that vitamin D enhances sST2 production *in vitro,*[11] this pathway was inferred rather than directly tested in our participants. Future work should incorporate sST2 measurement from stored sera, where available, to directly validate the vitamin D → sST2 → IL-33 axis in this population. Sixthly, asthma severity was classified based on clinical presentation, symptom frequency, and pharmacological requirements in accordance with GINA guidelines without spirometric confirmation (FEV_1 and FEV_1/FVC ratio).[16] The absence of objective lung function testing may limit the precision of severity stratification and the robustness of dose-response analyses across severity groups. We acknowledge that integrating spirometry-based indices into future studies, particularly in older children able to perform reliable maneuvers, would strengthen severity classification and enable more refined biomarker-lung function correlations. Lastly, the single-center design may introduce selection bias, and multi-center studies across different Iraqi governorates would strengthen

the external validity of these findings.

Future directions

Future research should prioritize large-scale, multi-center prospective cohort studies to establish the temporal relationship between IL-33, vitamin D, and asthma development in Iraqi children. Randomized controlled trials evaluating the efficacy of vitamin D supplementation in reducing IL-33 levels and improving asthma outcomes in this population are warranted. Additionally, investigating the combined diagnostic utility of IL-33 with other biomarkers (such as FeNO, IgE, and eosinophil counts) through multi-marker panel approaches may further enhance diagnostic precision. Genetic studies exploring polymorphisms in the IL1RL1 (ST2) and VDR (vitamin D receptor) genes in the Iraqi population could provide valuable insights into the genetic susceptibility underlying the observed biomarker associations.

Finally, gender-stratified intervention studies targeting vitamin D supplementation in male asthmatic children are recommended to address the pronounced deficiency observed in this subgroup.

Conclusion

This study provides compelling clinical evidence elucidating the intertwined roles of the epithelial alarmin IL-33 and vitamin D3 in the pathogenesis and severity of childhood asthma. Our findings demonstrate a robust, inverse relationship between elevated systemic IL-33 levels and profound vitamin D3 deficiency, both of which correlate significantly with escalating disease severity. The exceptional diagnostic performance of IL-33, alongside the independent predictive value of both IL-33 and elevated BMI, underscores the multifaceted nature of asthma exacerbations. Furthermore, the pronounced vulnerability of male asthmatic children to severe vitamin D deficiency highlights the necessity for individualized clinical assessments. Ultimately, these results bridge the gap between *in vitro* immunological mechanisms and clinical phenotypes, suggesting that targeted interventions-such as biomarker-guided anti-IL-33 therapies combined with gender-stratified vitamin D supplementation and weight management-could offer a synergistic approach to improving disease control and mitigating the inflammatory burden in pediatric asthma.

Ethical approval

The study protocol was reviewed and approved by the Institutional Ethics Committee at the College of Medicine, University of Kufa, Iraq (No: MEC-103, Dated October,2025). The research was conducted in strict adherence to the ethical principles outlined in the Declaration of Helsinki. Written informed consent was obtained from the parents or legal guardians of all participating children prior to their inclusion in the study.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interests

The authors declare no competing interests.

Consent for publication

Not applicable. This manuscript does not contain any individual person's data in any form (including individual details, images, or videos).

Availability of data and materials

The datasets generated and/or analyzed during the current study are not publicly available due to institutional privacy policies regarding patient data, but are available from the corresponding author on reasonable request.

CRediT authorship contribution statement

Zahraa Hazim Aziz: conceptualization, methodology, investigation, data collection, formal analysis, writing: original draft and visualization; **Thikra Abdullah Mahmood:** supervision, validation, methodology, data interpretation, writing: review and edit, project administration.

Acknowledgements

The authors would like to express their profound gratitude to the medical and nursing staff at the Children Teaching Hospital in Karbala, Iraq, for their invaluable assistance during sample collection. We also extend our deepest appreciation to the children and their families who participated in this study.

References

[1] Lu W, Rylance S, Schotte K, Aarsand R, Lebedeva E, Bill W, et al. Tobacco and asthma: presenting the World Health Organization (WHO) tobacco knowledge summary. Subst Abuse Treat Prev Policy. 2025; 20(1): 34. doi.org/10.1186/s13011-025-00632-6.

[2] Taherian MR, Fatemian F, Halimi A, Soleimani Y, Jorjani G, Nozari P, et al. Prevalence of asthma among children and adolescents in WHO's Eastern Mediterranean Region: a meta-analysis of over 0.5 million participants. BMC Public Health. 2024; 24(1): 2148. doi.org/10.1186/s12889-024-19614-5.

[3] Mirzaei M, Karimi M, Beheshti S, Mohammadi M. Prevalence of asthma among Middle Eastern children: a systematic review. Med J Islam Repub Iran. 2017; 31: 9. doi.org/10.18869/mjiri.31.9.

[4] Duchesne M, Okoye I, Lacy P. Epithelial cell alarmin cytokines: frontline mediators of the asthma inflammatory response. Front Immunol. 2022; 13: 975914. doi.org/10.3389/fimmu.2022.975914.

[5] Schmitz J, Owyang A, Oldham E, Song Y, Murphy E, McClanahan TK, et al. IL-33, an interleukin-1-like

cytokine that signals via the IL-1 receptor-related protein ST2 and induces T helper type 2-associated cytokines. Immunity. 2005; 23(5): 479-90. doi.org/10.1016/j.immuni.2005.09.015.

[6] Cayrol C, Girard JP. IL-33: an alarmin cytokine with crucial roles in innate immunity, inflammation and allergy. Curr Opin Immunol. 2014; 31: 31-7. doi.org/10.1016/j.coi.2014.09.004.

[7] Wang Y, Wang L, Hua S. Interleukin-33 in children with asthma: a systematic review and meta-analysis. Allergol Immunopathol (Madr). 2017; 45(4): 387-92. doi.org/10.1016/j.aller.2016.12.003.

[8] Litonjua AA. Childhood asthma may be a consequence of vitamin D deficiency. Curr Opin Allergy Clin Immunol. 2009; 9(3): 202-7. doi.org/10.1097/ACI.0b013e32832b36cd.

[9] Al-Ghamdi MA, Lanham-New SA, Kahn JA. Differences in vitamin D status and calcium metabolism in Saudi Arabian boys and girls aged 6 to 18 years: effects of age, gender, extent of veiling and physical activity with concomitant implications for bone health. Public Health Nutr. 2012; 15(10): 1845-53. doi.org/10.1017/S136898001100344X.

[10] Jolliffe DA, Greenberg L, Hooper RL, Griffiths CJ, Camargo CA, Kerley CP, et al. Vitamin D supplementation to prevent asthma exacerbations: a systematic review and meta-analysis of individual participant data. Lancet Respiratory Medicine. 2017; 5(11): 881-90. doi.org/10.1016/S2213-2600(17)30306-5.

[11] Pfeffer PE, Chen YH, Woszczek G, Matthews NC, Chevretton E, Gupta A, et al. Vitamin D enhances production of soluble ST2, inhibiting the action of IL-33. J Allergy Clin Immunol. 2015; 135(3): 824-7.e3. doi.org/10.1016/j.jaci.2014.09.044.

[12] El Abd A, Dasari H, Dodin P, Trottier H, Ducharme FM. Associations between vitamin D status and biomarkers linked with inflammation in patients with asthma: a systematic review and meta-analysis of interventional and observational studies. Respiratory Research. 2024; 25(1): 344. doi.org/10.1186/s12931-024-02947-6.

[13] Al-Tuama NMJ, Almusawi ZM, Abood HAAN, Abutiheen A, Al-tuama KMJ. The association between vitamin D deficiency and childhood asthma. J Prev Diagn Treat Strateg Med. 2022; 1(2): 115-20. doi.org/10.4103/jpdtsm.jpdtsm_31_22.

[14] AL-Hasnawi ATN, AL-Hasnawi SMJ. Role of interleukin 25 and interleukin 33 as immunological markers in pediatric asthma. Eurasia J Biosci. 2020; 14: 4625-30. www.ejobios.org/article/role-of-interleukin-25-and-interleukin-33-as-immunological-markers-in-pediatric-asthma-7975.

[15] Al-Sharifi ZAR, Al-Ammar HAJ, Mahmood HG, Turki KM, Al-Karkhi II. Lack of vitamin D in Iraqi children with asthma. Biomed Pharmacol J. 2017; 10(1): 89-93. doi.org/10.13005/bpj/1085.

[16] Hoque F, Nayak R. Focused overview of the 2024 Global Initiative for Asthma guidelines. APIK J Int Med. 2025; 13(1): 4-12. doi.org/10.4103/ajim.ajim_45_24.

[17] Misra M, Pacaud D, Petryk A, Collett-Solberg PF, Kappy M; Drug and Therapeutics Committee of the Lawson Wilkins Pediatric Endocrine Society. Vitamin D deficiency in children and its management: review of current knowledge and recommendations. Pediatrics. 2008; 122(2): 398-417. doi.org/10.1542/peds.2007-1894.

[18] Deng X, Ma J, Yuan Y, Zhang Z, Niu W. Association between overweight or obesity and the risk for childhood asthma and wheeze: an updated meta-analysis on 18 articles and 73 252 children. Pediatr Obes. 2019; 14(9): e12532. doi.org/10.1111/ijpo.12532.

[19] Azizpour Y, Delpisheh A, Montazeri Z, Sayehmiri K, Darabi B. Effect of childhood BMI on asthma: a systematic review and meta-analysis of case-control studies. BMC Pediatr. 2018; 18(1): 143. doi.org/10.1186/s12887-018-1117-0.

[20] Reyes-Angel J, Kaviany P, Rastogi D, Forno E. Obesity-related asthma in children and adolescents. Lancet Child Adolesc Health. 2022; 6(10): 713-24. doi.org/10.1016/S2352-4642(22)00185-7.

[21] He Z, Wu H, Zhang S, Lin Y, Li R, Xie L, et al. The association between secondhand smoke and childhood asthma: a systematic review and meta-analysis. Pediatr Pulmonol. 2020; 55(10): 2518-31. doi.org/10.1002/ppul.24961.

[22] Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. Immunol Rev. 2011; 242(1): 10-30. doi.org/10.1111/j.1600-065X.2011.01008.x.

[23] Bjerg A, Hedman L, Perzanowski MS, Platts-Mills T, Lundbäck B, Rönmark E. Family history of asthma and atopy: in-depth analyses of the impact on asthma and wheeze in 7-to 8-year-old children. Pediatrics. 2007; 120(4): 741-8. doi.org/10.1542/peds.2006-3112.

[24] Grant TL, Wood RA. The influence of urban exposures and residence on childhood asthma. Pediatr Allergy Immunol. 2022; 33(5): e13784. doi.org/10.1111/pai.13784.

[25] Aligne CA, Auinger P, Byrd RS, Weitzman M. Risk factors for pediatric asthma: contributions of poverty, race, and urban residence. Am J Respir Crit Care Med. 2000; 162(3 Pt 1): 873-7. doi.org/10.1164/ajrccm.162.3.9908085.

[26] Shaban SA, Brakhas SA, Ad'hiah AH. Interleukin-33 is up-regulated in serum of asthma patients. AIP Conf Proc. 2022; 2437(1): 020015. doi.org/10.1063/5.0092770.

[27] Bahrami Mahneh S, Movahedi M, Aryan Z, Bahar MA, Rezaei A, Sadr M, et al. Serum IL-33 is elevated in children with asthma and is associated with disease severity. Int Arch Allergy Immunol. 2016; 168(3): 193-6. doi.org/10.1159/000442385.

[28] Yang Q, Zheng Y, Dong J, Xu Q, Li S, Li A, et al. Serum

interleukin-33 combined with FEF75% z-score and FeNO improves the diagnostic accuracy of asthma in children. J Pediatr (Rio J). 2024; 100(1): 81-7. doi.org/10.1016/j.jped.2023.05.008.

[29] Bener A, Ehlayel MS, Tulic MK, Hamid Q. Vitamin D deficiency as a strong predictor of asthma in children. Int Arch Allergy Immunol. 2012; 157(2): 168-75. doi.org/10.1159/000327389.

[30] Esfandiar N, Alaei F, Fallah S, Babaie D, Sedghi N. Vitamin D deficiency and its impact on asthma severity in asthmatic children. Ital J Pediatr. 2016; 42(1): 108. doi.org/10.1186/s13052-016-0318-z.

[31] Rai V, Agrawal DK. Immunomodulation of IL-33 and IL-37 with vitamin D in the neointima of coronary artery: a comparative study between balloon angioplasty and stent in hyperlipidemic microswine. Int J Mol Sci. 2021; 22(16): 8824. doi.org/10.3390/ijms22168824.

[32] Bonanno A, Gangemi S, La Grutta S, Malizia V, Riccobono L, Colombo P, et al. 25-hydroxyvitamin D, IL-31, and IL-33 in children with allergic disease of the airways. Mediators Inflamm. 2014; 2014: 520241. doi.org/10.1155/2014/520241.

[33] Fang C, Jiang Z, Su X, Fan W. The association between body mass index and asthma in children: a cross-sectional study from NHANES 1999 to 2020. Sci Rep. 2025; 15(1): 9448. doi.org/10.1038/s41598-025-89476-y.

[34] Li RJ, Wen YX. Association of body mass index with asthma occurrence and persistence in adolescents: a retrospective study of NHANES (2011–2018). Heliyon. 2023; 9(9): e19623. doi.org/10.1016/j.heliyon.2023.e19623.

[35] Li CJ, Chang LS, Guo MMH, Wang LJ, Kuo HC. Sex differences in vitamin D and behavioral profiles among children with allergic diseases. Food Science & Nutrition. 2023; 11(9): 5492-500. doi.org/10.1002/fsn3.3504.

[36] Hollams E, Hart P, Holt B, Serralha M, Parsons F, De Klerk N, et al. Vitamin D and atopy and asthma phenotypes in children: a longitudinal cohort study. Eur Respir J. 2011; 38(6): 1320-7. doi.org/10.1183/09031936.00032911.

[37] Pfeffer PE, Hawrylowicz CM. Vitamin D in asthma: mechanisms of action and considerations for clinical trials. Chest. 2018; 153(5): 1229-39. doi.org/10.1016/j.chest.2017.09.005.

[38] Forno E, Bacharier LB, Phipatanakul W, Guilbert TW, Cabana MD, Ross K, et al. Effect of vitamin D3 supplementation on severe asthma exacerbations in children with asthma and low vitamin D levels: the VDKA randomized clinical trial. JAMA. 2020; 324(8): 752-60. doi.org/10.1001/jama.2020.12384.

Journal of Associated Medical Sciences 2026; 59(3): 72-84

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



Predicting dyscalcemia using machine learning models based on routine laboratory data

Busadee Pratumvinit[1,2], Kittisak Thawnashom[3,4], Bunjira Makond[5*]

[1]*College of Digital Science, Prince of Songkla University, Hat Yai, Songkhla Province, Thailand.*
[2]*Department of Clinical Pathology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.*
[3]*Department of Medical Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*
[4]*Interdisciplinary Health and Data Sciences Research Unit, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*
[5]*Faculty of Commerce and Management, Prince of Songkla University, Trang Province, Thailand.*

ARTICLE INFO

Article history:
Received 13 January 2026
Accepted as revised 14 May 2026
Available online 21 May 2026

Keywords:
Calcium, hypocalcemia, machine learning, clinical laboratory techniques, utilization review

ABSTRACT

Background: Dyscalcemia is associated with significant clinical complications, and the demand for calcium testing is rising disproportionately.

Objectives: This study aimed to develop machine learning (ML) models to predict total calcium abnormalities using routinely available demographic and laboratory data to optimize laboratory test utilization.

Materials and methods: This retrospective study analyzed data from 7,951 patients at Siriraj Hospital between April 2023 and March 2024. Feature selection identified ten predictors, including patient status, absolute lymphocyte count, albumin, bicarbonate, hematocrit, hemoglobin, platelet count, potassium, red blood cell count, and total protein. Seven ML algorithms were developed and validated using an 80:20 training-testing split with 10-fold cross-validation and hyperparameter tuning. The Synthetic Minority Oversampling Technique was applied to address class imbalance. Additionally, model calibration was evaluated using Brier scores, and Platt scaling was applied to ensure the reliability of predicted probabilities.

Results: XGBoost achieved the highest AUC (0.84), indicating slightly better discriminative performance, and the highest specificity (0.86), reflecting a stronger ability to correctly identify normal cases. Random forest yielded the highest F1 score (0.71) and recall (0.76), indicating a better balance between precision and recall, with higher sensitivity for detecting abnormal cases. Following Platt scaling, the calibrated models achieved robust Brier scores (e.g., 0.1496 for random forest and 0.1497 for XGBoost), demonstrating highly accurate probability risk estimation. SHAP analysis identified albumin as the most influential feature, followed by total protein and hemoglobin.

Conclusion: ML models utilizing demographic and laboratory data can accurately predict plasma total calcium status, and probability calibration ensures reliable risk estimates. XGBoost and random forest demonstrated robust performance with complementary strengths. These models show promise as screening tools for identifying high-risk patients and optimizing rational laboratory utilization.

* **Corresponding contributor.**
Author's Address: Faculty of Commerce and Management, Prince of Songkla University, Trang Province, Thailand.
E-mail address: bunjira.m@psu.ac.th
doi: 10.66285/JAMS.2026.080
E-ISSN: 2539-6056

Introduction

Laboratory testing is essential for disease screening, diagnosis, and monitoring; however, inappropriate test utilization remains a significant challenge within healthcare systems.[1] Overutilization can lead to

unnecessary costs and downstream investigations, whereas underutilization may result in delayed or missed diagnoses. Rational laboratory utilization through data-driven approaches is critical for improving clinical efficiency and patient safety.

Dyscalcemia, encompassing both hypercalcemia and hypocalcemia, is a common laboratory abnormality associated with a broad spectrum of clinical manifestations, ranging from asymptomatic biochemical changes to severe neurological and cardiovascular complications.[2,3] Over the past decade (2013-2023), Siriraj Hospital observed a marked 41.3% increase in total calcium test volume, rising from 90,953 to 128,539 tests annually. This growth rate markedly outpaced that of other routine electrolytes, such as sodium and potassium, which increased by approximately 14% over the same period. This disproportionate increase in calcium testing highlights the growing clinical demand for calcium assessment and underscores the need for tools to identify patients at risk of abnormal calcium levels using routinely available clinical and laboratory data.

Machine learning (ML) has emerged as an important tool in healthcare informatics for extracting meaningful patterns from large-scale clinical and laboratory datasets. In laboratory medicine, ML models have been successfully applied to predict abnormal test results and to support clinical decision-making related to test utilization. Previous studies have demonstrated the feasibility of ML-based prediction for various laboratory parameters, such as liver function tests,[4] renal function markers,[5] ferritin,[6-10] parathyroid hormone-related peptide,[11] and ionized calcium.[12]

However, most calcium-related prediction studies rely on datasets that combine total calcium with additional parameters to predict ionized calcium, which may not reflect routine clinical workflows or represent rational laboratory utilization.[12-15] Although ionized calcium is considered the most accurate indicator of calcium status, its routine measurement is limited by higher cost and pre-analytical constraints. Consequently, total calcium measurement, with or without albumin-based correction, remains the most used approach in everyday clinical practice.

Accordingly, there is a need for predictive models that leverage patient demographic information and laboratory data to efficiently identify dyscalcemia in real-world clinical settings. Therefore, the objectives of this study were: 1) to apply machine learning models to predict abnormalities in plasma calcium levels using patient demographic characteristics and laboratory test results, and 2) to compare the predictive performance of different machine learning models in identifying dyscalcemia.

Materials and methods

Dataset

This study used demographic and laboratory testing data extracted from the electronic health record system of Siriraj Hospital, a 2,200-bed university hospital in Bangkok, Thailand. The study was conducted with approval from the Siriraj Institutional Review Board (COA No. Si 405/2024). The study included randomly selected patients aged 19 years and older who underwent total calcium testing between April 2023 and March 2024. Patients without any additional laboratory tests beyond total calcium were excluded.

Demographic data collected included age, sex, and patient status (outpatient or inpatient). The selection of laboratory variables was adapted from the study by Yang *et al.*, which developed machine learning models to predict parathyroid hormone-related protein (PTHrP), a biomarker associated with calcium metabolism.[11] In the present study, a modified set of variables was used by excluding laboratory tests that were unavailable in the central laboratory of Siriraj Hospital or not directly relevant to calcium regulation. The final dataset comprised 47 predictive features (3 demographic and 44 laboratory variables) and 1 target variable (total calcium).

Laboratory variables included hemoglobin, hematocrit, red blood cell count, mean corpuscular volume, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, red cell distribution width, white blood cell count, percentages of neutrophils, lymphocytes, monocytes, eosinophils, and basophils; absolute neutrophil count, absolute lymphocyte count, absolute monocyte count, absolute eosinophil count, absolute basophil count, platelet count, total calcium (target variable), albumin, phosphate, magnesium, sodium, potassium, chloride, bicarbonate, anion gap, alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, total bilirubin, total protein, glucose, blood urea nitrogen, creatinine, measured ionized calcium, ionized calcium adjusted to pH 7.4, measured pH, vitamin D, parathyroid hormone, thyroid-stimulating hormone, activated partial thromboplastin time, prothrombin time, and international normalized ratio.

Machine learning analysis

Data preprocessing

Patient status was encoded as outpatient = 1 and inpatient = 2. The target variable was defined as calcium status and was derived from plasma total calcium measurements. Outcomes were dichotomized as normal (class = 0) or abnormal (class = 1) using age-specific reference intervals. Reference intervals were obtained from the manufacturer's package insert of the total calcium assay. For individuals aged 19-60 years, the reference interval for total calcium was 8.6-10.0 mg/dL, for individuals aged 61-90 years, 8.8-10.2 mg/dL, and for individuals older than 90 years, 8.2-9.6 mg/dL.

Feature selection

Feature selection was performed using the Machine Intelligence Learning Optimizer [MILO] (MILO-ML, LLC, Sacramento, CA). The MILO feature selection

tool employs two different methods to assess the contribution of independent variables to the outcome: an ANOVA F-value approach and a Random Forest feature importances method. Variables were ranked by their scores from each method, and features in the top 25th percentile of either approach were retained for further analysis. The initial dataset consisted of 39,926 cases. Prior to feature selection, several data preprocessing steps were applied to ensure data quality and computational feasibility. Laboratory variables with fewer than 10,000 available observations (measured pH, measured ionized calcium, ionized calcium corrected to pH 7.4, international normalized ratio, thyroid-stimulating hormone, activated partial thromboplastin time, prothrombin time, vitamin D, and parathyroid hormone) were excluded, as their inclusion prevented successful model execution.

Feature importance was assessed by identifying the features ranked within the top 25% using an ANOVA F-statistic approach and the top 25% based on random forest feature importance, as implemented in MILO-ML. The combined feature set from both methods was retained for subsequent model development. These features included nine continuous variables-total protein (TP), albumin (Alb), hematocrit (Hct), hemoglobin (Hb), absolute lymphocyte count, red blood cell count (RBC), platelet count, bicarbonate, and potassium—and one categorical variable, patient status (Figure 1).

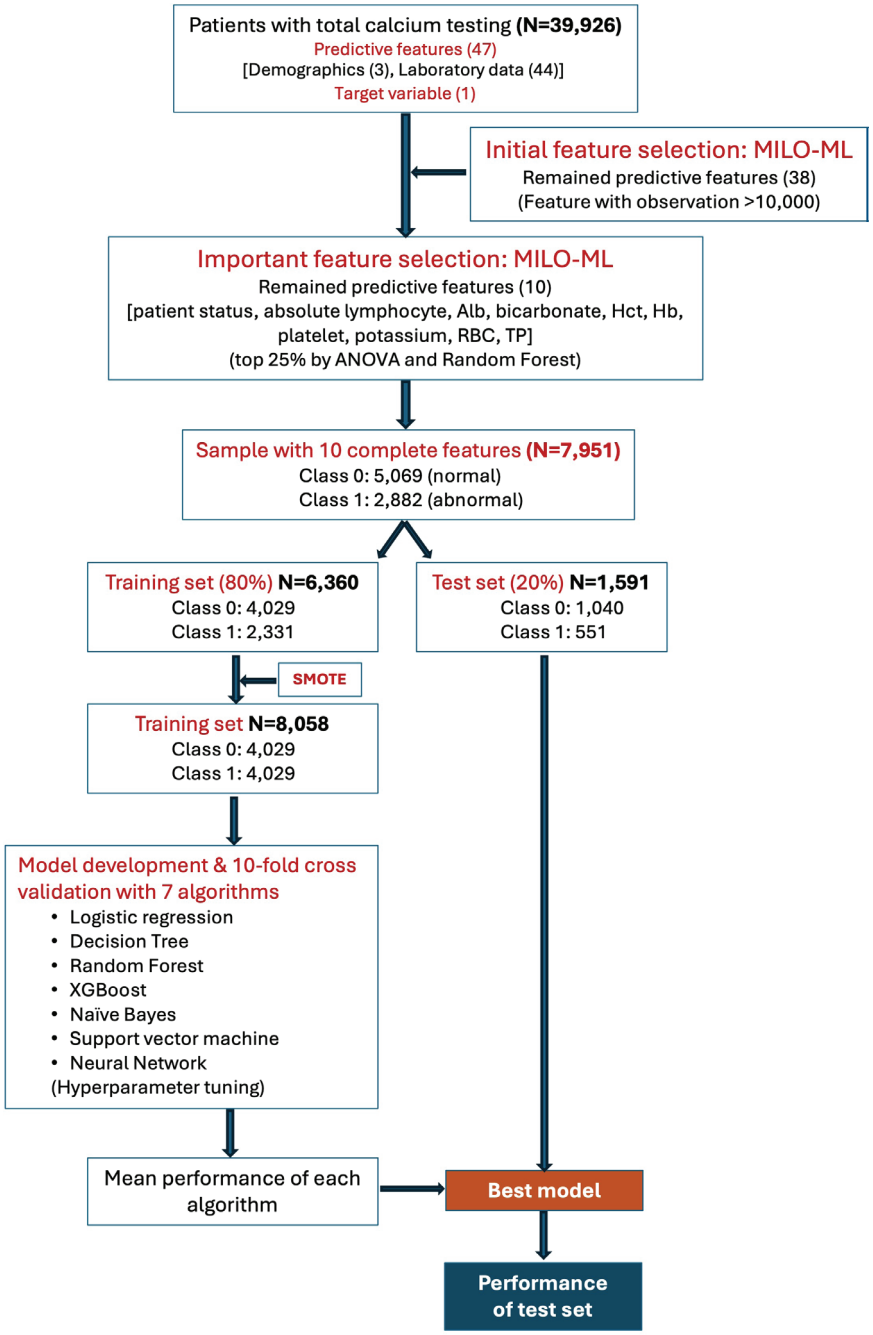


Figure 1. *Schematic diagram of model construction and validation.*

Missing data were assessed for all candidate variables prior to analysis. The proportion of missingness varied across laboratory parameters, with several variables exhibiting substantial incomplete observations. To ensure data integrity and compatibility with the machine learning pipeline, a complete case analysis was performed. Only samples with all 10 selected features available were included in the final dataset, reducing the dataset from 39,926 to 7,951 samples.

Model development and validation

The dataset was randomly split into a training set (80%) and a test set (20%), with the test set reserved exclusively for internal validation and not used during model training or hyperparameter tuning. To address class imbalance in the training dataset, the Synthetic Minority Over-sampling Technique (SMOTE) was applied only to the training data.[16] Prior to SMOTE, the class distribution in the training set was imbalanced (Class 0: 4,029; Class 1: 2,331). After applying SMOTE, the dataset was balanced with equal representation of both classes (Class 0: 4,029; Class 1: 4,029; total N=8,058).

Feature scaling was performed using standardization (z-score normalization) for models sensitive to feature magnitudes, including support vector machines, neural networks, and logistic regression models. Tree-based models (decision tree, random forest, XGBoost) and naïve Bayes classifiers were trained using unscaled features. All preprocessing steps were applied to the training data first, and the same transformations were subsequently applied to the test data to prevent data leakage.

Machine learning analyses were performed in Python using Google Colaboratory. Multiple supervised machine learning algorithms were developed and evaluated, including Logistic Regression, Decision Trees, Random Forests, Extreme Gradient Boosting (XGBoost), Naïve Bayes, Support Vector Machines, and Neural Networks. These algorithms were selected to represent a diverse range of linear, nonlinear, probabilistic, ensemble, and deep learning–based approaches.

Hyperparameter tuning for each model was performed using a grid search with k-fold cross-validation on the training dataset. Ten-fold cross-validation was used to optimize model parameters and assess training performance. The optimal hyperparameter set for each algorithm was selected based on cross-validated accuracy.

Following hyperparameter optimization, the best-performing configuration of each model was retrained using the entire resampled training dataset. Model performance was then evaluated on the independent test dataset, which served as an internal validation cohort (Figure 1).

Performance metrics included accuracy, precision (positive predictive value), recall (sensitivity), specificity, F1 score, and area under the receiver operating characteristic curve (AUC). Receiver operating characteristic curves were generated for each model to visualize discriminative performance. For the training dataset, performance metrics were summarized as mean and standard deviation across cross-validation folds. For the test dataset, point estimates of performance metrics were reported. Confusion matrices for each model were generated using predictions from the test dataset to illustrate classification performance.

In addition to discriminative metrics, the models predicted-probability accuracy was evaluated using the Brier score. Because SMOTE can distort probability estimates and lead to poor model calibration, Platt scaling (logistic calibration) was applied to correct these estimates. The calibrator was fitted using the testing dataset via a pre-fit cross-validation approach. Reliability diagrams (calibration curves) were generated to visualize the agreement between the predicted probabilities and the observed outcomes before and after calibration.

To enhance model interpretability, Shapley Additive Explanations (SHAP) were used to quantify the contribution of individual features to model predictions. SHAP values were calculated for the best-performing model among the seven algorithms.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics (IBM Corp., Armonk, NY, USA). Continuous variables were summarized as mean (standard deviation, SD) for approximately normally distributed data and as median (interquartile range, IQR) for non-normally distributed data. The categorical variable was expressed as a number and a percentage. Differences in laboratory parameters between the normal and abnormal calcium groups were compared using the independent samples t-test for normally distributed variables and the Mann–Whitney U test for non-normally distributed variables. The chi-square test was used to compare the proportions between the groups. A two-sided $p<0.05$ was considered statistically significant.

Results

Baseline characteristics of predictive variables stratified by calcium status

Among 7,951 patients, 5,069 were classified as having normal calcium levels and 2,882 as having abnormal calcium levels. Significant differences in clinical and laboratory parameters were observed between the two groups. Patients with abnormal calcium levels had significantly lower median (IQR) values for absolute lymphocyte count, albumin, bicarbonate, hematocrit, hemoglobin, platelet count, potassium, red blood cell count, and total protein than those with normal calcium levels (all $p<0.001$). The abnormal calcium group had a substantially higher proportion of inpatients than the normal calcium group (61.4% vs 37.0%), whereas the normal calcium group had a higher proportion of outpatients (63.0% vs 38.6%, $p<0.001$) (Table 1).

Table 1. Characteristics of predictive variables stratified by calcium status (normal vs abnormal).

Parameters	Normal calcium (N=5,069)	Abnormal calcium (N=2,882)	*p* value
Continuous variables	**Median (IQR)**	**Median (IQR)**	
Absolute lymphocyte (10^3/µL)	1.42 (1.12)	1.00 (1.00)	<0.001
Albumin (gm/dL)	3.90 (0.70)	3.10 (0.90)	<0.001
Bicarbonate (mmol/L)	26.00 (4.00)	24.00 (6.00)	<0.001
Hematocrit (%)	34.30 (10.00)	29.10 (8.10)	<0.001
Hemoglobin (gm/dL)	11.10 (3.20)	9.50 (2.70)	<0.001
Platelet count (10^3/µL)	222.00 (128.00)	178.00 (161.00)	<0.001
Potassium (mmol/L)	4.10 (0.60)	3.90 (0.80)	<0.001
Red blood cell count (10^6/µL)	3.96 (1.16)	3.42 (1.03)	<0.001
Total protein (gm/dL)	7.20 (1.00)	6.30 (1.40)	<0.001
Categorical variable	**Number (%)**	**Number (%)**	
Patient status			<0.001
Outpatient =1	3,195 (63.0%)	1,112 (38.6%)	
Inpatient = 2	1,874 (37.0%)	1,770 (61.4%)	

Hyperparameter tuning and model performance

The evaluated hyperparameters and the optimized values for each machine learning model are summarized in Table 2. Model performance in predicting total calcium status is presented in Tables 3 and 4 for the training and testing datasets. The random forest model achieved the highest AUC (0.89) and F1 score (0.81), followed closely by XGBoost (AUC=0.88; F1 score=0.81). Logistic regression, support vector machine, and neural network models demonstrated comparable moderate performance, with AUCs ranging from 0.83 to 0.84 and F1 scores between 0.76 and 0.77. The decision tree model exhibited the lowest performance (AUC=0.73; F1 score=0.73).

Table 2. The list of hyperparameters and optimized values of each machine learning model.

Model	Data preprocessing	Hyperparameter	Values tested	Selected value
Logistic regression	Scaled	C	log-scale from 10^{-4} to 10^{4} (20 values)	0.0127
		Penalty	L1, L2	L2
		Solver	liblinear	liblinear
		Maximum iterations	1,000; 2,000; 5,000	1,000
Decision tree	Unscaled	Maximum depth	3, 5, 7	7
		Minimum samples per split	2, 5, 10	2
Random forest	Unscaled	Number of estimators	100, 200, 300	100
		Maximum depth	3, 5, 7	7
XGBoost	Unscaled	Learning rate	0.1, 0.01	0.1
		Number of estimators	100, 200, 300	200
Naïve Bayes	Unscaled	Variance smoothing	$1e^{-9}$ to $1e^{-5}$	$1e^{-9}$
Support vector machine	Scaled	C	0.1, 1, 10	1
		kernel	linear	linear
Neural network	Scaled	Hidden layer size(s)	50, 100, 50–50	50
		Activation	ReLU, tanh	ReLU

Table 3. Performance evaluation of each model in the training dataset.

Models	Accuracy	Precision	Recall	Specificity	F1 score	AUC
Logistic regression	0.76 (0.009)	0.76 (0.020)	0.76 (0.012)	0.76 (0.022)	0.76 (0.009)	0.83 (0.010)
Decision tree	0.73 (0.015)	0.73 (0.027)	0.73 (0.022)	0.73 (0.023)	0.73 (0.020)	0.73 (0.015)
Random forest	0.81 (0.012)	0.81 (0.016)	0.80 (0.021)	0.82 (0.014)	0.81 (0.014)	0.89 (0.009)
XGBoost	0.81 (0.010)	0.82 (0.019)	0.80 (0.023)	0.83 (0.016)	0.81 (0.015)	0.88 (0.009)
Naïve Bayes	0.71 (0.014)	0.70 (0.020)	0.73 (0.022)	0.69 (0.018)	0.71 (0.017)	0.79 (0.011)
Support vector machine	0.78 (0.010)	0.79 (0.012)	0.75 (0.022)	0.80 (0.015)	0.77 (0.009)	0.84 (0.011)
Neural network	0.77 (0.010)	0.78 (0.022)	0.75 (0.020)	0.78 (0.024)	0.76 (0.008)	0.84 (0.011)

Note: *Data presented as mean (SD).*

Table 4. Performance evaluation of each model in the testing data.

Models	Accuracy	Precision	Recall	Specificity	F1 score	AUC
Logistic Regression	0.77	0.65	0.76	0.78	0.70	0.82
Decision Tree	0.76	0.64	0.69	0.79	0.66	0.79
Random Forest	0.79	0.67	0.76	0.80	0.71	0.83
XGBoost	0.79	0.72	0.66	0.86	0.69	0.84
Naïve Bayes	0.71	0.56	0.74	0.70	0.64	0.79
Support Vector Machine	0.77	0.65	0.76	0.78	0.70	0.83
Neural Network	0.77	0.65	0.73	0.79	0.69	0.83

In the testing dataset, XGBoost achieved the highest AUC (0.84), indicating a slightly superior discriminative ability, while random forest yielded the highest F1 score (0.71), reflecting the best balance between precision and recall. Logistic regression, support vector machine, and neural network models demonstrated comparable performance (AUC: 0.82-0.83; F1 score: 0.69-0.70) (Figure 2). Overall, XGBoost showed slightly better performance in four evaluation metrics, including the highest specificity, suggesting good discrimination of non-diseased individuals. Random forest showed slightly lower overall performance, but achieved higher recall, indicating superior identification of diseased individuals. The remaining models exhibited superior performance in fewer metrics and overall lower effectiveness. Detailed classification results from the confusion matrices for the testing dataset are presented in Figure 3.

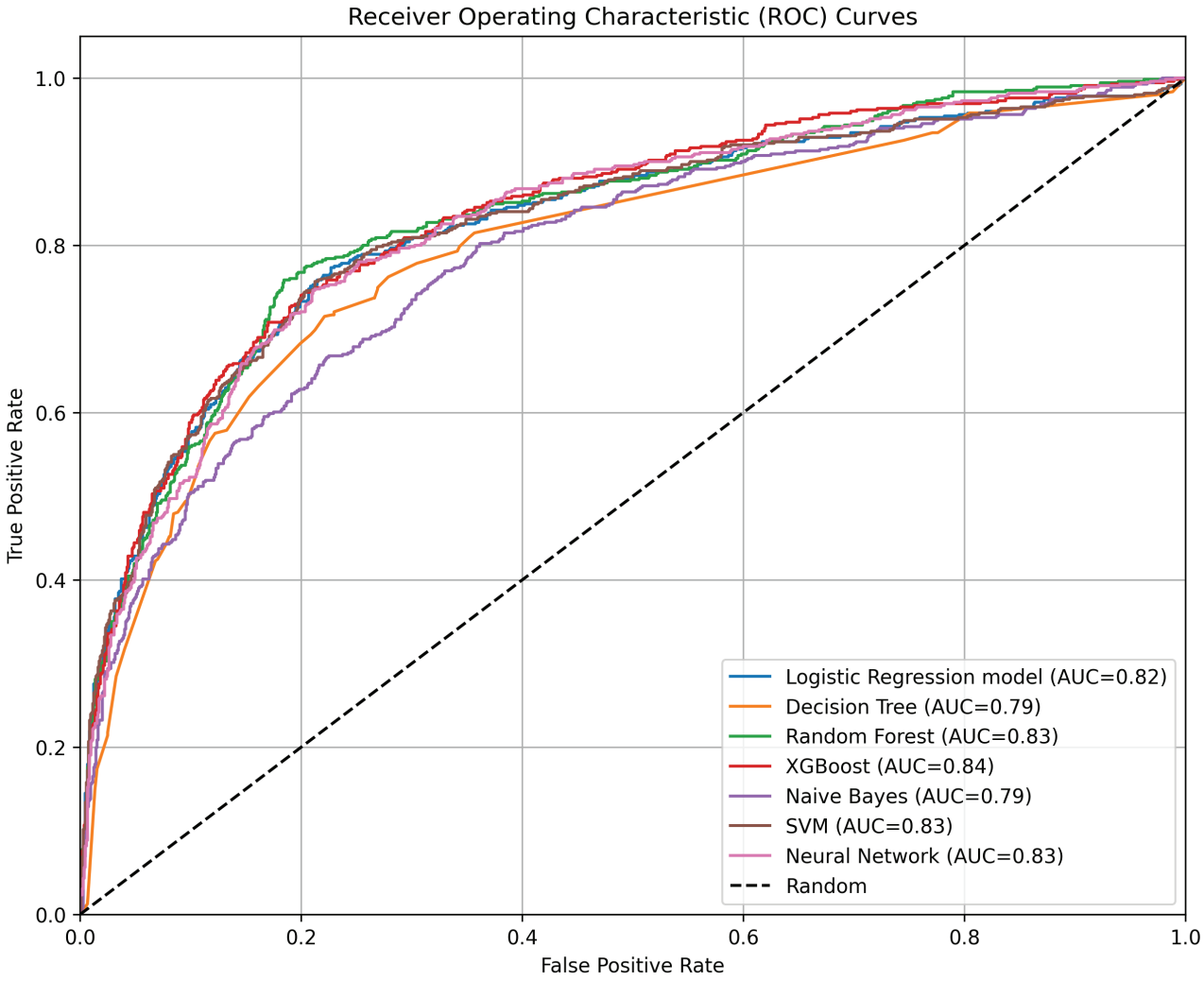


Figure 2. *Comparative performance analysis based on receiver operating characteristic (ROC) curves.*

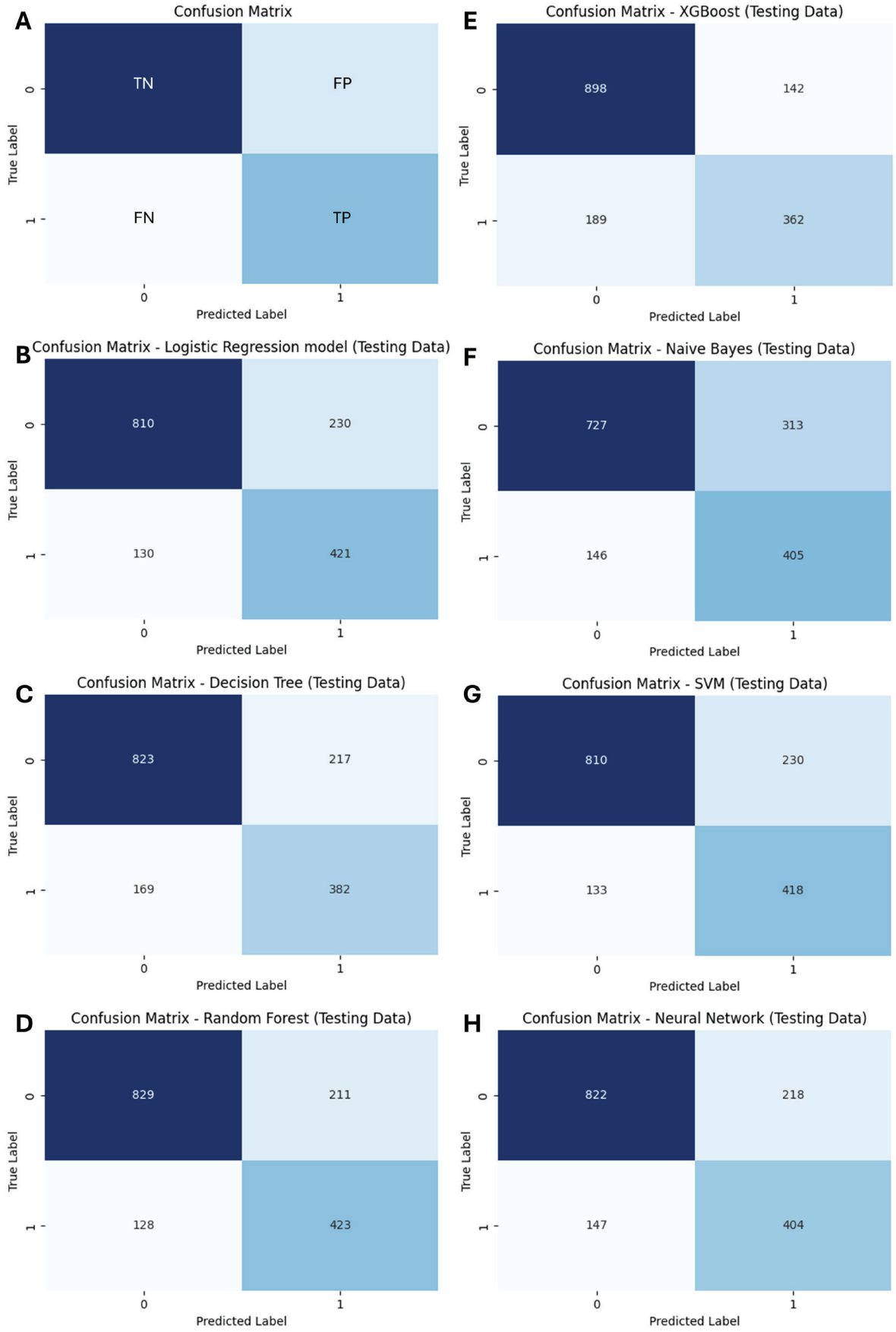


Figure 3. *Confusion matrices of the evaluated models. A: schematic representation of a confusion matrix illustrating true negatives (TN), false positives (FP), false negatives (FN), and true positives (TP), B: Logistic Regression, C: Decision Tree, D: Random Forest, E: XGBoost, F: Naïve Bayes, G: Support Vector Machine, H: Neural Network.*

To ensure the reliability of the predicted probabilities for clinical decision-making, model calibration was further evaluated using Brier scores. Recognizing that the SMOTE technique can distort probability estimates, we applied Platt scaling (logistic calibration) to the testing dataset. As anticipated, calibration substantially improved the accuracy of probability estimates across all algorithms, reflected by lower Brier scores. The random forest and XGBoost models achieved the best calibrated Brier scores of 0.1496 and 0.1497, respectively. The complete pre- and post-calibration Brier scores for all models are detailed in the supplementary Table S1, and the reliability diagrams (calibration curves) demonstrating the improved probability fits for the top-performing models are presented in the supplementary Figure S1.

Feature importance

SHAP summary plots were generated for both the random forest and XGBoost (Figure 4). The analysis revealed that albumin was the most influential feature for predicting calcium status in both models. Lower albumin values were generally associated with the positive class (class 1 = abnormal). Total protein and hemoglobin were the next most influential variables in both models, contributing moderately to class separation.

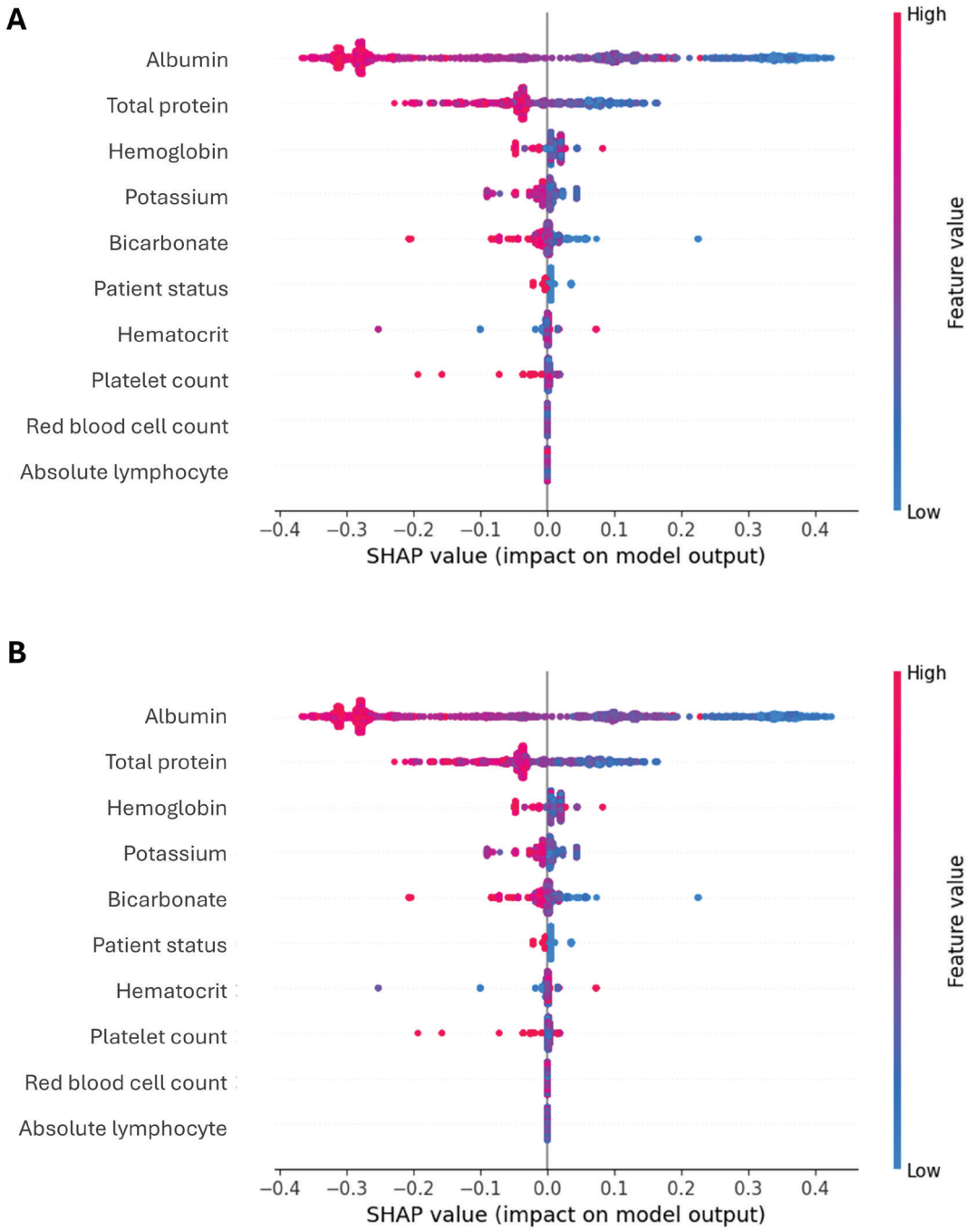


Figure 4. *SHAP value analysis for class 1 (positive). A: Random Forest, B: XGBoost.*

Discussion

This study demonstrates the feasibility and effectiveness of using ML to predict plasma total calcium abnormalities using only patient demographic characteristics and routinely available laboratory data. By leveraging a large dataset from a tertiary care center, we developed and validated seven ML models. The XGBoost model achieved the highest discriminative performance in the testing set (AUC=0.84) and showed particularly high specificity, indicating a strong ability to correctly identify patients with normal calcium status. This property is advantageous for laboratory test utilization, as improved specificity may reduce unnecessary follow-up testing and downstream investigations in low-risk patients. In contrast, the random forest achieved the highest recall, reflecting greater sensitivity for detecting abnormal calcium levels, which is suitable for screening purposes, where minimizing missed diagnoses is paramount. Taken together, these complementary performance profiles suggest that model selection can be aligned with specific clinical objectives, balancing the priorities of case detection and avoidance of over-investigation.

Beyond discriminative power, the reliability of a model's probabilistic output is paramount for clinical adoption. While our models demonstrated high AUC, the initial probability estimates were affected by the SMOTE resampling process, a phenomenon well-documented in machine learning literature. By applying Platt scaling, we successfully calibrated the models, as evidenced by the substantial reduction in Brier scores. This calibration ensures that the system's risk estimates accurately reflect the true probability of dyscalcemia, facilitating more reliable clinical decision-making when integrated into laboratory information systems.

Previous applications of ML in calcium metabolism have predominantly focused on predicting ionized calcium levels from total calcium and other parameters.[12-15] However, the present study addresses a different clinical need by identifying patients at risk of abnormal total calcium. To our knowledge, this is one of the first studies to focus specifically on predicting total calcium status using a broad panel of common hematology and chemistry markers.

The interpretability analysis using SHAP values provided physiological validation for our models. Albumin was identified as the most influential feature for predicting calcium status in both the random forest and XGBoost models, which is biologically consistent with albumin's known calcium-binding capacity. In plasma, approximately 40% of calcium is protein-bound, and approximately 80% of this fraction is bound to albumin.[17] Hypoalbuminemia is therefore a major determinant of reduced total calcium levels.[2,12] Payne et al. demonstrated a strong correlation between serum calcium and albumin (r=0.867), compared with a weaker correlation with total protein (r=0.682), and proposed an albumin-adjusted calcium formula that is widely used in clinical practice.[18] Hemoglobin also ranked among the top three predictors, consistent with previous studies reporting a positive correlation between hemoglobin and calcium levels.[19-21]

Based on the XGBoost model's performance, 1,087 of 1,591 cases (68%) in the test dataset were predicted as normal, suggesting that up to 68% of calcium tests could be avoided. During the study period (April 2023 to March 2024), a total of 133,436 calcium tests were performed at Siriraj Hospital. Applying this proportion, approximately 90,736 tests could potentially be avoided. Given a cost of 55 Thai Baht (THB) per test, this corresponds to an estimated annual cost saving of 4,990,480 THB (approximately 140,000 USD, assuming 1 USD ≈ 35 THB).

However, the model also yielded 189 false-negative cases (11.8%), indicating that some abnormal results may have been missed. Therefore, this model should be used as a decision-support tool rather than as a standalone diagnostic approach. In clinical practice, laboratory results should be interpreted in conjunction with patient history and clinical findings to ensure safe decision-making. These estimates are based on model performance in the test dataset, and actual cost savings may vary depending on model implementation and clinical practice patterns.

Limitations

Several limitations of this study should be acknowledged. First, the study is retrospective and single center, which may limit generalizability to other populations with different reference intervals or patient demographics. Second, data-related limitations should be considered. A complete-case analysis was performed, resulting in a substantial reduction in sample size from 39,926 to 7,951. Although this approach ensured data consistency, it may have introduced selection bias if the missingness was not completely at random. In addition, certain laboratory variables with substantial missingness were excluded prior to feature selection, which may have limited the inclusion of potentially relevant predictors. Third, the model development and evaluation were based solely on internal validation without external validation using independent datasets. This may limit the generalizability and robustness of the findings, as model performance may differ in other clinical settings or populations. In addition, temporal drift and stability of model performance over time were not formally evaluated, which may limit assessment of temporal generalizability. Fourth, the use of SMOTE to address class imbalance may introduce a risk of overfitting, although it was applied only to the training dataset. Additionally, we did not explore advanced oversampling variants (e.g., Borderline-SMOTE or SMOTE-ENN); investigating whether these techniques further refine decision boundaries is an area for future research. Finally, while the models predict total

calcium abnormalities, they do not distinguish among the underlying etiologies of dyscalcemia, which would require further clinical correlation.

Conclusion

Machine learning models using routine demographic and laboratory data can accurately predict plasma calcium status. Random forest and XGBoost demonstrated robust performance, with distinct strengths in sensitivity and specificity, respectively. Applying probability calibration ensures these models provide reliable risk estimates for clinical use. This ML-based screening approach could be integrated into the laboratory information system to flag high-risk patients without a recent calcium order or to predict normal results in low-risk patients, thereby optimizing laboratory test utilization.

Ethical approval

The study was conducted with approval from the Siriraj Institutional Review Board (COA no. Si 405/2024).

Funding

No funding was provided for this research.

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

Busadee Pratumvinit: conceptualization, methodology, formal analysis, investigation, data curation, writing original draft, review and edit; **Kittisak Thawnashom:** conceptualization, methodology, formal analysis, writing: review and edit; **Bunjira Makond:** conceptualization, methodology, formal analysis, writing: review and edit.

Acknowledgements

The authors gratefully acknowledge the Siriraj Informatics and Data Innovation Center (SiData+), Faculty of Medicine Siriraj Hospital, Mahidol University, for retrieving the data used in this study. The authors also acknowledge the use of ChatGPT (OpenAI, GPT-4.1-turbo) to improve the clarity and language of the manuscript; however, we maintain full responsibility for all content and interpretations presented in this work.

References

[1] Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. PLoS One. 2013; 8(11): e78962. doi: 10.1371/journal.pone.0078962.

[2] Walker MD, Shane E. Hypercalcemia: A Review. JAMA. 2022; 328(16): 1624-36. doi: 10.1001/jama.2022.18331.

[3] Bilezikian JP. Hypoparathyroidism. J Clin Endocrinol Metab. 2020; 105(6): 1722-36. doi: 10.1210/clinem/dgaa113.

[4] Lidbury BA, Richardson AM, Badrick T. Assessment of machine-learning techniques on large pathology data sets to address assay redundancy in routine liver function test profiles. Diagnosis (Berl). 2015; 2(1): 41-51. doi: 10.1515/dx-2014-0063.

[5] Pawuś D, Porażko T, Paszkiel S. Automated biomedical measurements analysis: Innovative models based on machine learning for predicting laboratory results in nephrology. Expert Syst Appl. 2025; 270: 126568. doi: 10.1016/j.eswa.2025.126568.

[6] Luo Y, Szolovits P, Dighe AS, Baron JM. Using machine learning to predict laboratory test results. Am J Clin Pathol. 2016; 145(6): 778-88. doi: 10.1093/ajcp/aqw064.

[7] Kurstjens S, de Bel T, van der Horst A, Kusters R, Krabbe J, van Balveren J. Automated prediction of low ferritin concentrations using a machine learning algorithm. Clin Chem Lab Med. 2022; 60(12): 1921-8. doi: 10.1515/cclm-2021-1194.

[8] Efros O, Soffer S, Mudrik A, Robinson R, Kenet G, Nadkarni GN, et al. Predictive machine-learning model for screening iron deficiency without anaemia: a retrospective cohort study. BMJ Open. 2025; 15(8): e097016. doi: 10.1136/bmjopen-2024-097016.

[9] Varekha NV, Stuklov NI, Gordienko KV, Gimadiev RR, Shchegolev OB, Kislaya SN, et al. Development of a method for differential diagnosis of iron deficiency anemia and anemia of chronic disease based on demographic data and routine laboratory tests using machine learning technologies. Oncohem. 2025; 20(1): 171-81. doi: 10.17650/1818-8346-2025-20-1-171-181.

[10] Li W, Su CY, Meulenbeld A, Jagirdar H, Janssen MP, Swanevelder R, et al. Machine-learning models to predict iron recovery after blood donation: a model development and external validation study. Lancet Haematol. 2025; 12(6): e431-e41. doi: 10.1016/S2352-3026(25)00068-7.

[11] Yang HS, Pan W, Wang Y, Zaydman MA, Spies NC, Zhao Z, et al. Generalizability of a machine learning model for improving utilization of parathyroid hormone-related peptide testing across multiple clinical centers. Clin Chem. 2023; 69(11): 1260-9. doi: 10.1093/clinchem/hvad141.

[12] Bancal C, Salipante F, Hannas N, Lumbroso S, Cavalier E, De Brauwere DP. A new approach to assessing calcium status via a machine learning algorithm. Clin Chim Acta. 2023; 539: 198-205. doi: 10.1016/j.cca.2022.12.018.

[13] Kumari S, Nayak S, Mangaraj M. A new machine learning approach for actual calcium measurement. Indian J Clin Biochem. 2025; 40(2): 300-6. doi: 10.1007/s12291-024-01182-3.

[14] Hung WL, Huang CF, Tsai MH, Liou HH, Liu PY, Fang YW. A new predictive equation for estimating

serum ionized calcium levels in patients on chronic hemodialysis. Med Sci Monit. 2023; 29: e941321. doi: 10.12659/MSM.941321.

[15] Yap E, Ouyang J, Puri I, Melaku Y, Goldwasser P. Novel methods of predicting ionized calcium status from routine data in critical care: External validation in MIMIC-III. Clin Chim Acta. 2022; 531: 375-81. doi: 10.1016/j.cca.2022.05.003.

[16] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002; 16: 321-57. doi: 10.1613/jair.953.

[17] Fraser WD, Alter DN. Bone and mineral metabolism. In: Rifai N, Chiu RW, Young I, Burnham C-AD, Wittwer CT, editors. Tietz Textbook of Laboratory Medicine. 7th Ed. St. Louis: Elsevier; 2023. pp 766-e85.

[18] Payne RB, Little AJ, Williams RB, Milner JR. Interpretation of serum calcium in patients with abnormal serum proteins. Br Med J. 1973; 4(5893): 643-6. doi: 10.1136/bmj.4.5893.643.

[19] Wang Z, Chen X, Chen Y, Yang L, Wang H, Jiang W, et al. Association between admission serum calcium and hemoglobin in older patients with hip fracture: a cross-sectional study. Eur Geriatr Med. 2022; 13(2): 445-52. doi: 10.1007/s41999-021-00569-2.

[20] Haji S, Atta A, Sagir Ahmad J, Farrukh Zia T, Javed N, Muhammad N, et al. An analysis of anaemia and serum calcium levels in chronic kidney disease as an indicator of severity and progression of disease in patients of South Punjab, Pakistan: Anemia and calcium in CKD severity. J Health Rehabil Res. 2024; 4(3): 1-9. doi: 10.61919/jhrr.v4i3.1261.

[21] Şahin S, Yıldız N, Esenülkü G, Özkan Kart P, Kamasak T, Özkaya AK, et al. Serum calcium levels, erythrocyte indices, and associated factors in children with febrile seizures. Farabi Tıp Dergisi. 2025; 4(2): 17-26. doi: 10.59518/farabimedj.1590060.

Supplementary data

Table S1. Brier scores of machine learning models before and after probability calibration.

Machine learning model	Brier score (uncalibrated)	Brier score (calibrated)
Logistic Regression	0.1660	0.1525
Decision Tree	0.1732	0.1644
Random Forest	0.1589	0.1496
XGBoost	0.1512	0.1497
Naïve Bayes	0.2261	0.1769
Support Vector Machine	0.1655	0.1524
Neural Network	0.1635	0.1543

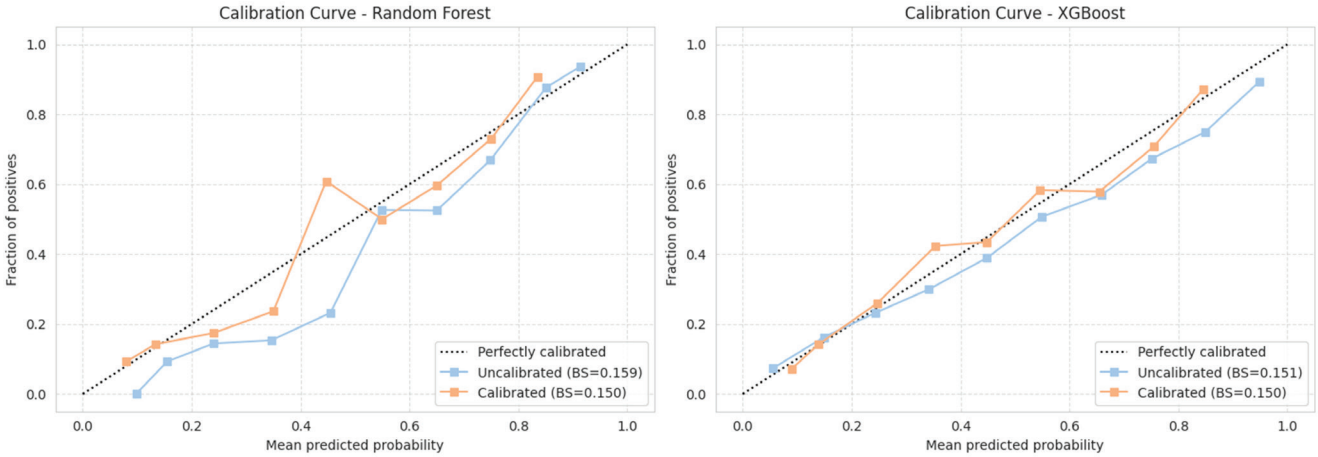


Figure S1. *Reliability diagrams (calibration curves) for the top-performing models: Random Forest (left) and XGBoost (right). The plots demonstrate the improvement in probability estimation after applying Platt scaling. The uncalibrated models (blue line with squares) show miscalibration typical of SMOTE-resampled training data. Following calibration (orange line with squares), the curves align more closely with the perfectly calibrated reference line (black dotted diagonal), which corresponds to a reduction (improvement) in the Brier Scores (BS).*

Journal of Associated Medical Sciences 2026; 59(3): 85-91

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 radiation and light field congruence between a developed collimator test tool and a standard collimator test tool in the general X-ray units of Lamphun Province hospitals

Kampol Taywadithep

Regional Medical Sciences Center 1 Chiang Mai, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 8 December 2025
Accepted as revised 18 May 2026
Available online 25 May 2026

Keywords:
Radiation and light field congruence, collimator, collimator test tool

ABSTRACT

Background: Assessment of radiation and light-field congruence is a key quality-control procedure for general X-ray machines. This process typically requires a standard collimator test tool, which can be costly. Therefore, the author developed an innovative collimator test tool as a cost-effective alternative for evaluating the congruence between the radiation field and the collimator light field.

Objectives: The present study aims to evaluate and compare radiation and light-field congruence measured using the developed collimator test tool and a standard tool on general X-ray units across eight hospitals in Lamphun Province.

Materials and methods: The developed collimator test tool was constructed from an etched stainless-steel plate measuring 20×25 cm, with scale resolutions of 2.5 mm and 5 mm. The standard collimator test tool (GAMMEX, Model 161B) is made from a brass plate of the same dimensions with a scale resolution of 5 mm. Measurements followed the 2023 diagnostic X-ray quality standards established by the Department of Medical Sciences, Ministry of Public Health. Radiation and light field borders were measured three times on each of four sides, yielding a total of 96 measurements. Error and percentage error were calculated, and a paired-sample t-test analysis was performed.

Results: Comparison of radiation and light-field congruence between the developed collimator and the standard collimator test tools revealed a maximum error of 0.5 mm, corresponding to a percentage error of 3%, which is below the acceptable limits of 1 mm or 5%. Statistical analysis using a paired-sample t-test ($p<0.05$) indicated no significant difference between measurements obtained from the two test tools ($p\geq0.05$).

Conclusion: The developed collimator test tool can be used as an alternative to the standard tool for radiation and light field congruence testing in general X-ray units. It offers a cost-effective option for self-quality control and monitoring in radiology departments while maintaining the same measurement accuracy.

* **Corresponding contributor.**
Author's Address: *Regional Medical Sciences Center 1 Chiang Mai, Chiang Mai Province, Thailand.*
E-mail address: kampol.t@dmsc.mail.go.th
doi: 10.66285/JAMS.2026.081
E-ISSN: 2539-6056

Introduction

Radiological diagnostic technology has been rapidly and continuously evolving to support medical diagnosis. Diagnostic X-rays are now an essential tool that helps doctors diagnose diseases and detect abnormalities in patients' organs more accurately and quickly. International and national radiation safety organizations regulate and supervise the use

of radiation to ensure compliance with international standards for radiation hazard prevention. To maximize the benefits of radiation in diagnosis while minimizing exposure, it is essential to implement appropriate protective measures. Awareness of radiation protection for patients can be achieved by establishing a quality management system to supervise operations and maintain equipment. Improving the quality of work and services, as well as controlling the radiation dose received by patients, ensures that exposure does not exceed necessary levels, thereby reducing the risk to both patients and radiation workers.[1-4]

In Thailand, the Department of Medical Sciences has a policy to enhance the quality and capabilities of diagnostic radiology laboratories by focusing on developing and supporting quality systems and standards in accordance with the Ministry of Public Health's regulations. To support public health, quality control of medical diagnostic X-ray machines is implemented, providing hospital radiology personnel with guidelines for their work. Efforts are made to improve the quality of the service system to ensure efficient services for the public. Additionally, the quality of services in diagnostic radiology laboratories is enhanced to maintain radiation safety and security in accordance with nationwide standards.[7]

Quality control of medical diagnostic X-ray machines, digital imaging equipment, and other accessories is essential to ensure that diagnostic radiology laboratories produce high-quality X-rays that are safe for both patients and operators. X-ray images are a critical component of medical diagnosis, helping doctors ensure accuracy and precision in treatment planning. High-quality imaging benefits practitioners by reducing the time required to obtain X-rays and supports the development of efficient diagnostic radiology services within hospitals, ultimately contributing to an effective public healthcare system.[9]

The Annual Reports from 2017 to 2024 of the Regional Medical Sciences Center 1, Chiang Mai, reported that up to 12.7 percent of general diagnostic X-ray machines failed to meet quality standards each year. One of the main causes was the misalignment between the radiation beam and the collimator light. Similarly, the Annual Report 2023 of the Bureau of Radiation and Medical Devices, Department of Medical Sciences, showed that out of 4,262 general diagnostic X-ray machines tested nationwide, 258 units (6.1 percent) did not meet the standards, primarily due to a mismatch between the X-ray beam and the collimator light.[10,11]

Quality standards for radiation and light-field congruence in medical diagnostic X-ray machines specify that the alignment error should not exceed 2 centimeters at a source-to-image distance (SID) of 100 centimeters, and that testing should be conducted every 6 months.[8,9] ICRP Publication 154 emphasizes the importance of proper collimation to reduce the total patient dose as part of a consistent quality control program. Misalignment between the radiation beam and the collimator light can increase the patient's radiation dose, resulting in unnecessary exposure.[2]

X-ray machine quality control standards are important for ensuring patient safety, minimizing radiation dose, and identifying test items that do not meet quality requirements. Problems related to the adjustment of X-ray machines have been observed during radiation and light-field congruence testing of general X-ray machines in Health Region 1. To deal with these issues, the author developed an innovative collimator test tool for assessing radiation and light field congruence, in collaboration with The Brick Fabrication Lab (THE BRICK FABLAB) at the Northern Sciences and Technology Park (STeP), Chiang Mai University.[12] This developed tool serves as an alternative for testing the congruence of general X-ray machines and can be used to monitor and adjust collimators in diagnostic radiology laboratories in Lamphun Province, other provinces in Health Region 1, or nationwide.

Therefore, the purpose of this study was to evaluate and compare radiation and light-field congruence measurements obtained with the developed collimator test tool against those obtained with the standard collimator test tool.

Materials and methods

Development of a collimator test tool

A collimator test tool was developed using an etched stainless-steel plate measuring 20×25 cm, with scale resolutions of 2.5 and 5 mm, as shown in Figure 1.

Standard collimator test tool

The standard tool (GAMMEX 161B) used a brass plate measuring 20×25 cm with a 5 mm scale resolution (Figure 1).

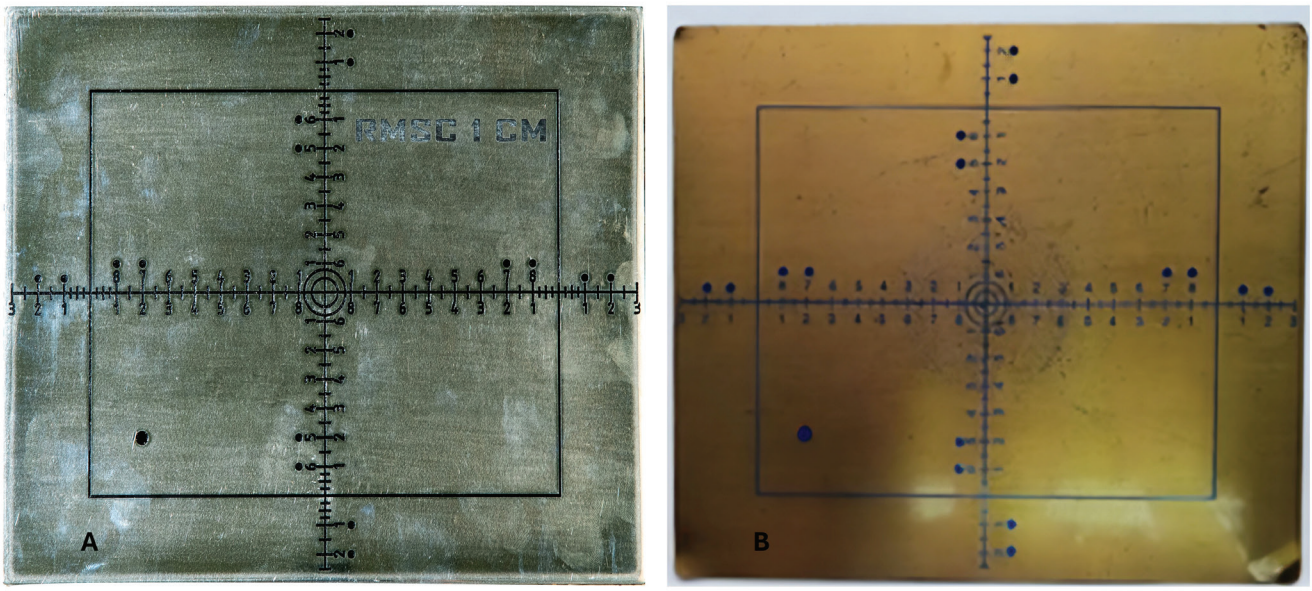


Figure 1. *The collimator test tool images. A: The developed collimator test tool, B: The standard collimator test tool.*

Test method

Radiation and light-field congruence testing between the developed collimator test tool and the standard collimator test tool on general X-ray machines was conducted at eight hospitals.[1,6,7] A total of 96 measurements was collected from eight general X-ray units in this preliminary investigation. The radiation and light field congruence test began by aligning the X-ray tube with the examination table. A digital X-ray image detector plate was then positioned at the center of the table, after which the developed collimator test tool was placed on top of it. The collimator light field was adjusted so that the central crosshair aligned precisely with the center point of the test tool. The X-ray tube was positioned at a source-to-image distance (SID) of 100 cm, and the collimator light was adjusted to match the 14×18 cm rectangular field of the test tool (Figure 2). Room lighting was dimmed or turned off to ensure that the collimator light borders were clearly visible (Figure 3). The X-ray technique was set at approximately 50-55 kV, 100 mA, and 0.30-0.50 s (3-5 mAs), after which an exposure was taken and the resulting image analyzed. The radiation and light field congruence distances on each of the four sides (left, right, top, and bottom) were measured three times using a medical digital radiography viewer (MicroDicom Viewer 2024), and the values were recorded (Figure 3). The entire procedure was repeated using the standard collimator test tool for comparison.

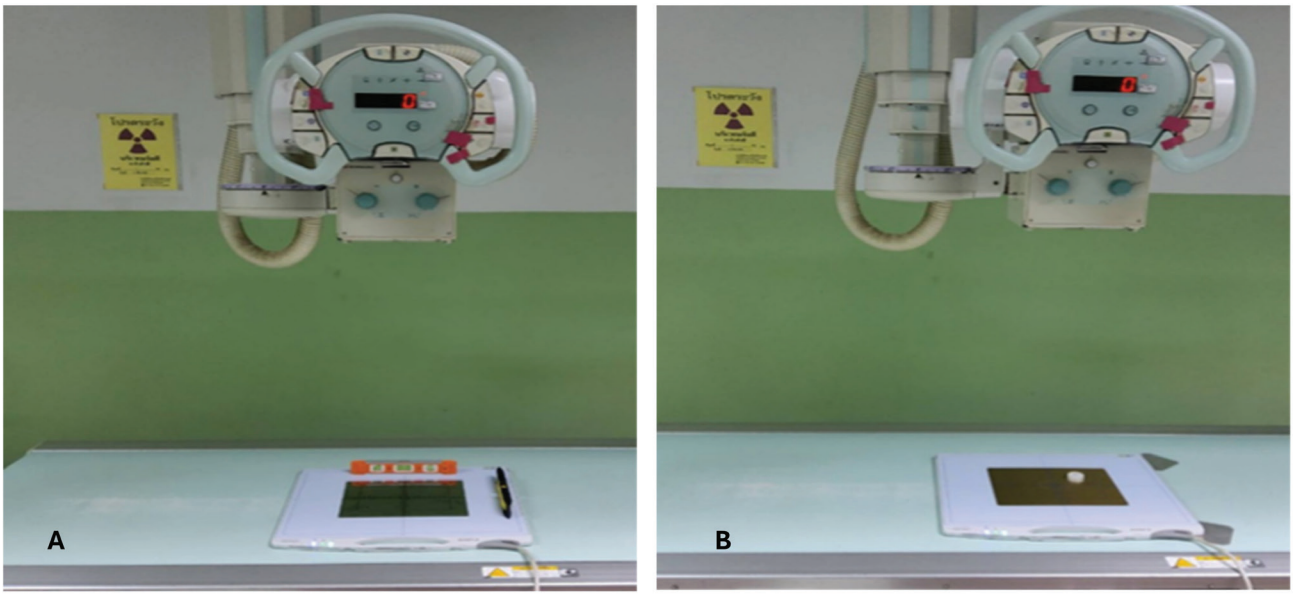


Figure 2. *Setting for the radiation and light field congruence testing images. A: The developed collimator test tool, B: The standard collimator test tool.*

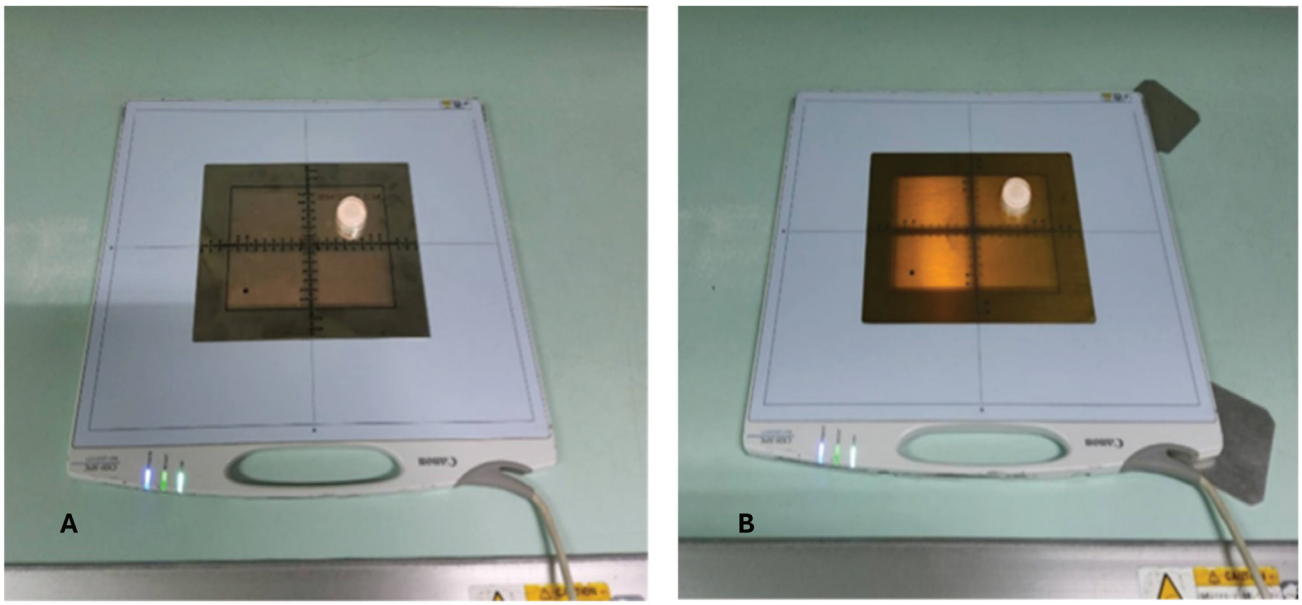


Figure 3. *The light field and markings on the collimator test tool.*
A: The developed collimator test tool, B: The standard collimator test tool.

Analysis method

The error analysis was performed by subtracting the radiation and light field congruence values of the developed collimator (E1) from the beam and light disparity values of the eight standard collimators (E2). Using the average of measurements taken on each side (left, right, top, bottom) 3 times and the sum of the average of all sides from each X-ray machine, totaling eight machines (Figure 4).

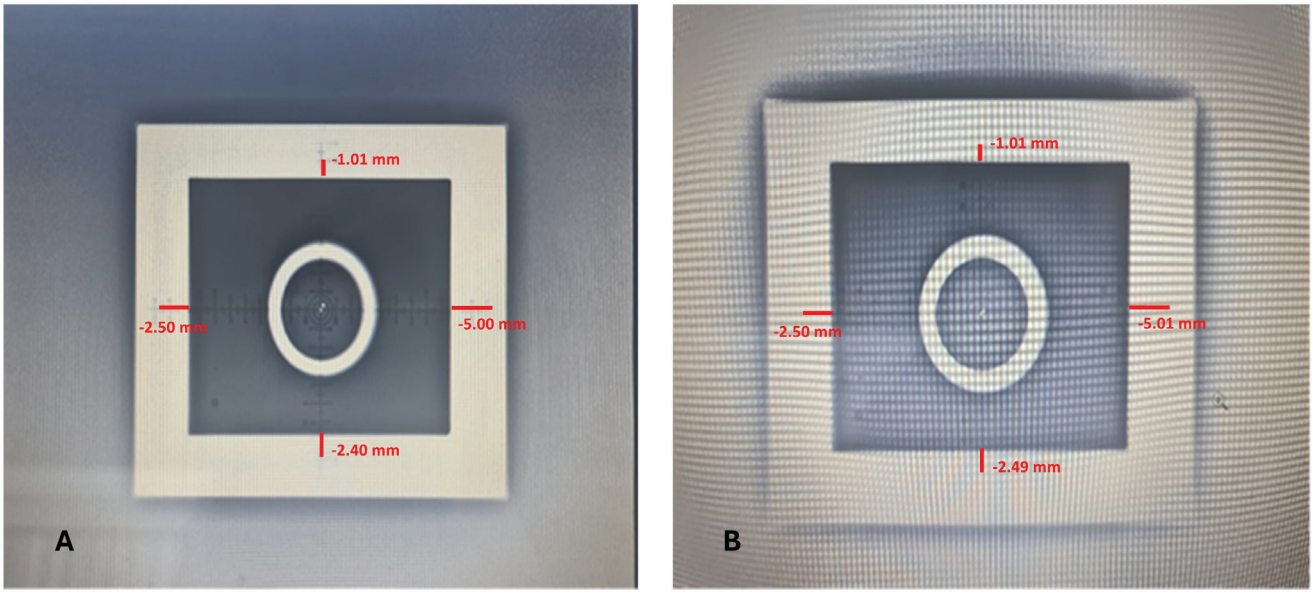


Figure 4. *Example of the measurement of the radiation and light field congruence distance images.*
A: The developed collimator test tool, B: The standard collimator test tool.

Statistical analysis

Comparison of the radiation and light field congruence values of the developed collimator test tool with the standard collimator test tool. The average of the three measurements taken for each side (left, right, top, bottom) and the sum of the averages from each X-ray machine (8 in total) were analyzed using a paired-sample t-test at the 95% confidence level (p<0.05) to determine whether there were any differences.

Results

The results of the present study were divides into two parts, including the mean (mean±SD), error and percentage error (%Error) of the results of the measurement of the radiation and light field congruence on each side between the developed collimator test tool and the standard collimator test tool and the summary of the comparative results of the radiation and light field congruence between the developed collimator test tool and the standard collimator test tool. The details of the results are shown in Table 1 and Table 2.

Discussion

The study results showed that the maximum error of radiation and light field congruence measured using the developed collimator test tool compared to the standard collimator test tool on each side (left, right, top, and bottom), as well as the overall average from a total of 8 X-ray machines, was 0.5 mm or 3%. This value is well below the acceptable limit of 1 mm or 5% and falls within the resolution scale.

Comparison of beam alignment with the collimator light between the developed and standard collimator test plates using a paired-sample t-test at 95% confidence (p<0.05) showed no statistically significant difference (p≥0.05), which is consistent with the established hypothesis.

Although the observed differences were statistically non-significant and below acceptable limits, they confirm that beam-light field misalignment does not compromise patient safety in terms of radiation dose and field coverage, helping to avoid both overexposure and under-coverage. Practically, the developed tool demonstrates comparable performance to the standard tool and can be effectively used for routine quality control, particularly in resource-limited settings.

Table 1. Mean±SD, error, and percentage error (%Error) of the results of the measurement of the radiation and light field congruence on each side between the developed collimator test tool and the standard collimator test tool.

The radiation and light field congruence (Measuring side)	**Average measurement results of the developed collimator test tool** (mm) Mean±SD	**Average measurement results of the standard collimator test tool** (mm) Mean±SD	**Error value** (mm) Mean±SD	**Percentage of error (%Error)** Mean±SD
Left side (N=24)	4.007±2.76	4.07±2.75	0.00±5.52	0.02
Right side (N=24	3.62±1.85	3.69±1.85	0.01 ± 3.70	- 0.28
Top side (N=24)	3.16±1.87	3.15±1.88	0.01 ± 3.74	- 0.20
Bottom side (N=24)	6.13±3.16	6.12±3.19	- 0.02 ± 6.35	0.24
Total of all aspects (N=96)	4.24±2.69	4.24±2.69	0.00 ± 5.39	- 0.01

Table 2. Summary of the comparative results of the radiation and light field congruence between the developed collimator test tool and the standard collimator test tool using a paired-sample t-test (p<0.05) statistic.

The radiation and light field congruence (measuring side)	**Average measurement results of the developed collimator test tool** (mm) Mean±SD	**Average measurement results of the standard collimator test tool** (mm) Mean±SD	***p* value**	**Summary of results**
Left side (N=24)	4.07±2.76	4.07±2.75	0.93	Not significantly different (p≥0.05)
Right side (N=24)	3.62±1.85	3.69±1.85	0.08	Not significantly different (p≥0.05)
Top side (N=24)	3.16±1.87	3.15±1.88	0.26	Not significantly different (p≥0.05)
Bottom side (N=24)	6.13±3.16	6.12±3.19	0.25	Not significantly different (p≥ 0.05)
Total of all aspects (N=96)	4.24±2.69	4.24±2.69	0.96	Not significantly different (p≥0.05)

its comparable accuracy and significantly lower cost, the developed tool is suitable for production and for testing the radiation and light-field congruence of X-ray machines in hospitals. The developed collimator test tool offers several practical advantages over the standard commercial device. These include lower cost, ease of fabrication, and suitability for routine and frequent quality control use. The thin, flat stainless-steel design facilitates straightforward positioning on the image receptor and reduces setup complexity. Additionally, the tool's durability allows repeated use without significant degradation, supporting long-term implementation in clinical QA/QC programs.

Overall, the developed tool provides performance comparable to that of the standard collimator test tool while offering improved accessibility and practicality for routine self-quality control, aligning with quality optimization principles and the ALARA concept.

Limitations

This study was limited to evaluating beam–light field congruence in general radiographic X-ray units, excluding mobile X-ray systems, which were excluded due to their continuous clinical use and limited accessibility for standardized testing. In addition, the study was conducted by a single observer, which precludes assessment of inter-observer and intra-observer variability. As a result, it is not possible to fully distinguish whether the observed variability arises from the measurement process or from observer-related factors.

Although standardized measurement procedures and repeated measurements were implemented to minimize potential bias, some degree of observer-dependent variability may persist. This limitation is particularly relevant for applications involving routine self-quality control, where consistency across different users is essential. Therefore, future studies should incorporate multiple observers and repeated measurement designs to evaluate measurement reproducibility using appropriate statistical methods.

Conclusion

The results of measuring radiation and light field congruence between the developed collimator test tool and the standard tool on eight general X-ray machines in eight hospitals in Lamphun Province showed that the maximum error was 0.5 mm, corresponding to a percent error of 3%, which is well below the acceptable limit of 1 mm or 5%. Statistical analysis using a paired-sample t-test ($p<0.05$) indicated no significant difference between measurements obtained from the two tools ($p\geq0.05$). Therefore, it can be concluded that the developed collimator test tool performs equivalently to the standard tool and can be used as a substitute. Future studies should include mobile X-ray systems, as this would enhance the applicability of quality control assessments in radiographic practice.

Ethical approval

This study did not require ethical approval.

Funding

This research received no funding from any agency.

Conflict of interest

The author declares no conflict of interest.

CrediT authorship contribution statement

Kampol Taywadithep: conceptualization, methodology, writing: original draft, review, and edit.

Acknowledgements

I would like to thank the Brick Fabrication Lab (THE BRICK FABLAB) in Northern Sciences and Technology Park (STeP), Chiangmai University, for their cooperation in designing and developing the collimator test tool for the radiation and light field congruence of a general X-ray machine, as requested by the author this time.

I would like to thank the hospitals and radiologists at Lamphun Hospital, Ban Thi Hospital, Mae Tha Hospital, Ban Hong Hospital, Pa Sang Hospital, Wiang Nong Long Hospital, Thung Hua Chang Hospital, and Li Hospital in Lamphun Province. Thank you for providing the general X-ray machine and X-ray recording system, which enabled the successful completion of the radiation and light-field congruence testing.

References

[1] IAEA. Handbook of basic quality control tests for diagnostic radiology. Human Health Series No.47. Vienna: International Atomic Energy Agency; 2023.

[2] ICRP. Martin CJ, Applegate K, Hernandez-Giron I, Husseiny D, Kortesniemi M, Del Rosario Perez M, et al. Optimization of radiological protection in digital radiology techniques for medical imaging. ICRP Publication 154. Ann ICRP. 2023; 52(3): 11-145. doi: 10.1177/01466453231210646.

[3] AAPM Task Group 150. Acceptance testing and quality control of digital radiographic imaging Systems. 2024. Virginia: American Association of Physicists in Medicine; 2024.

[4] ISRRT. Guidance on quality control process for imaging in radiography: Direct digital radiography (DDR) & computed radiography (CR), including display monitors. ISRRT; 2019. [cited 2024 Mar 29]. Available from: https://www.isrrt.org

[5] Pacific Northwest X-Ray Inc. Collimator test tool. Stock #07-661. Gresham, OR: Pacific Northwest X-Ray Inc.;2024 [cited 2024 JAN 9]. Available from: https://www.pnwx.com

[6] Sun Nuclear Corporation. DR, CR, & Fluoroscopy solutions: Radiology QA solutions. Melbourne, FL: Sun Nuclear Corporation;2024 [cited 2024 JAN 19]. Available from: https://www.sunnuclear.com

[7] Department of Medical Sciences, Ministry of Public Health. Diagnostic Radiology Laboratory Standards, Ministry of Public Health B.E. 2565.

Nonthaburi: Ministry of Public Health; 2022. [in Thai].

[8] Department of Medical Sciences, Ministry of Public Health. Quality Standards of Diagnostic X-ray Machines. Nonthaburi: Ministry of Public Health; 2023. [in Thai].

[9] Department of Medical Sciences, Ministry of Public Health. Quality Control of Medical Diagnostic X-ray Machines. Nonthaburi: Ministry of Public Health; 2019. [in Thai].

[10] Regional Medical Sciences Center 1 Chiang Mai, Department of Medical Sciences. Annual Report 2018–2025. Chiang Mai: Department of Medical Sciences; 2018-2025. [in Thai].

[11] Bureau of Radiology and Medical Devices, Department of Medical Sciences. Annual Report 2023. Bangkok: Department of Medical Sciences; 2023. [in Thai].

[12] The Brick fabrication lab (THE BRICK FABLAB). Northern Sciences and Technology Park (STeP), Chiang Mai University;2023.

Journal of Associated Medical Sciences 2026; 59(3): 92-98

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 prospective memory training on instrumental activities of daily living skills among individuals with mild dementia

Punitha P* **and Yamuna T**

Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India.

ARTICLE INFO

Article history:
Received 17 October 2025
Accepted as revised 25 May 2026
Available online 4 June 2026

Keywords:
Prospective memory training, instrumental activities of daily living skills, mild dementia, elderly persons

ABSTRACT

Background: Dementia is a neurodegenerative disease brought on by a long-term, progressive illness that impairs decision-making, behavior, emotions, memory, and other brain processes that interfere with day-to-day functioning. Mild dementia is an early stage of cognitive decline that interferes with daily activities but allows for independent function in many areas. Prospective memory (PM), which enables one to remember to carry out delayed intentions, is crucial for everyday functioning.

Objectives: This study aims to determine the effect of prospective memory training on instrumental activities of daily living skills among individuals with mild dementia.

Materials and methods: A True experimental study was conducted. A total of 45 participants were initially screened using the Mini-Mental State Examination (MMSE). Based on the inclusion criteria, 30 participants with mild dementia were selected. Through simple randomization, 30 participants were divided into two groups: the experimental group (N=15) and the control group (N=15). The Lawton and Brody Instrumental Activities of Daily Living (IADL) Scale was used to assess functional performance. The experimental group received prospective memory training focused on improving instrumental activities of daily living, while the control group underwent conventional occupational therapy for 3 months, with sessions held 3 days per week (36 sessions), each lasting 45 minutes. Data analysis was performed using IBM SPSS version 23.0 with paired and independent t-test.

Results: Post-test scores showed statistically significant improvements in the experimental group for instrumental activities of daily living skills (p=0.000, which p<0.001). The results of this study showed that prospective memory training was effective in enhancing instrumental activities of daily living skills among individuals with mild dementia.

Conclusion: This study found that prospective memory training improves instrumental activities of daily living among individuals with mild dementia.

Introduction

The Neurodegenerative disease known as dementia is brought on by a long-term, progressive illness that impairs decision-making, behavior, emotions, memory, and other brain processes that interfere with day-to-day functioning.[1] Memory, language, problem-solving, and other cognitive disorders are hallmarks of dementia, a disease that impairs a person's ability to function. The Diagnostic and Statistical Manual of Mental Disorders (DSM) frequently defines the nomenclature used to

* **Corresponding contributor.**
Author's Address: *Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India.*
E-mail address: punithaarul5@gmail.com
doi: 10.66285/JAMS.2026.082
E-ISSN: 2539-6056

characterize dementia in clinical practice. According to the DSM-5, Dementia is defined as a major neurocognitive disorder, characterized by a severe reduction in performance in one or more cognitive domains, such as complex attention, executive function, learning and memory, language, perceptual, motor, and social cognition.[2]

Mild dementia necessitates proof of significant challenges in day-to-day living that limit independence. Cognitive impairment and poor performance on objective cognitive tests that reflect a decline from the past are also characteristics of mild dementia. Unlike individuals with more severe types of dementia, who have difficulty with fundamental activities of daily life, those with moderate dementia maintain their independence in simpler activities.[3]

Instrumental Activities of Daily Living (IADLs) are higher-level daily tasks that go beyond basic self-care and are essential for independent living. They include activities such as managing finances, preparing meals, shopping, housekeeping, using transportation, and managing medication.[4] In dementia, IADLs are often the first to decline because they require complex cognitive processes like memory, planning, attention, and problem-solving. As dementia progresses, patients may forget to take medications, mismanage finances, struggle with shopping, or fail to maintain their household. These limitations increase over time and are strongly associated with the severity of dementia. This decline in IADL performance not only reflects loss of independence but also indicates the need for caregiver support and structured interventions.[5]

Occupational therapists are crucial in meeting the special requirements of those suffering from mild to moderate dementia. Maintaining or improving cognitive function, addressing deficiencies, reducing behavioral symptoms, and fostering supportive social and caregiving ties are all common therapy objectives. Therapists constantly adjust their intervention techniques in response to the growing level of disability. Since mild to moderate dementia cannot be cured, full functional recovery is not used to gauge how well a treatment is working. For this population, a positive response to treatment is best described as a slower rate of decline over time, a higher level of functional performance, or maintained capacity during a period when deterioration is often expected. These beneficial functional results can help preserve identity, foster emotions of enjoyment or utility, and reduce worry that can arise from gradually declining capacity in response to external demands. For people with mild to moderate dementia, community occupational therapy sessions that included cognitive and behavioral interventions were associated with better functioning in daily activities, reduced caregiver burden, and larger effect sizes than drug trials or other psychosocial interventions.[6]

Prospective memory (PM) is our ability to remember to carry out intentions or tasks we plan for the future. Older adults often experience difficulties, particularly with the two main types of PM: event-based prospective memory (EBPM), which is triggered by specific events, and time-based prospective memory, which depends on remembering to act at a specific time. For older persons to function in daily life, prospective memory is essential. To support instrumental activities of daily living (IADLs), prospective memory, temporal order memory, and source memory are crucial.[7]

Research by Hering *et al*. highlights that prospective memory is essential for independent living and can be improved with training, yet its direct impact on daily living in dementia is underexplored. Nakanishi *et al*. further emphasize the gap between what individuals can do and what they actually do in daily life, suggesting that memory impairments- not physical limitations-often hinder task performance.[8] The effect of prospective memory training on Instrumental activities of daily living skills among individuals with mild dementia has not been demonstrated in those studies.

To improve Instrumental activities of daily living skills among individuals with mild dementia and to demonstrate the effect of prospective memory training, this study was conducted.

Materials and methods

Study design

The study employed a true experimental study with a randomized controlled trial to evaluate the effect of prospective memory training on instrumental activities of daily living skills among individuals with mild dementia. Participants were randomly assigned to an experimental or control group using the lottery method.

Participants

This study received ethical approval from the Institutional Scientific Review Board of Saveetha College of Occupational Therapy, with REF.no. Of SCOT /ISRB/013/ on 12.02.2025. The study was conducted at Jagruti Rehabilitation Center, Chennai. A total of 45 individuals with suspected mild dementia were initially screened using the Mini-Mental State Examination (MMSE). The assessment was administered by trained assessors with experience in cognitive screening of older adults. In addition, the standardized version of the MMSE was used in accordance with copyright regulations. Based on the inclusion and exclusion criteria, 30 participants were identified as eligible. Using simple randomization through the lottery method, the participants were assigned to two groups: an experimental group (N=15) and a control group (N=15). Single blinding was implemented, wherein the participants were unaware of their group allocation. All participants provided written informed consent prior to the commencement of the study. The Inclusion criteria are both genders, ages between 60 and 80 years, an MMSE score of 18-23, and good communication ability, as they must communicate verbally and be able to participate in prospective memory training. Exclusion criteria are patients with severe cognitive impairments

who scored less than 18 in MMSE, severe mental health disorders such as severe depression, schizophrenia, mania based on DSM-5, persons with other neurological conditions like Alzheimer's, Parkinson's disease, and those with severe physical disabilities.

Instrument

Lawton Brody Instrumental Activities of Daily Living Scale

Particularly for elderly people with disorders such as dementia, the Lawton and Brody Instrumental Activities of Daily Living (IADL) Scale (1969) is a commonly used instrument to assess the more complex daily activities required for community life beyond basic self-care. The goal is to assess the ability to live independently while identifying any early cognitive or functional impairment. Telephone use, shopping, meal preparation, housework, laundry, transportation, prescription responsibility, and money management are the eight areas that are measured by the scale. With total scores ranging from 0 (poor function, dependent) to 8 (high function, independent), each item is assessed according to its level of independence. This is often done using a binary scale (0 = dependent, and 1 = independent). Higher scores denote greater independence. Specific deficiencies that may call for rehabilitation planning or support services can be identified through interpretation. The tool can be useful and reliable for clinical and research use because it demonstrates good psychometric properties, including inter-rater reliability of 0.85; reproducibility of 0.93 for women and 0.96 for men; and validity, as supported by moderate correlations (0.40-0.61) with other measures of functional and cognitive status.[9] The Lawton Instrumental Activities of Daily Living Scale was selected for this study because it is a widely used and validated tool for assessing functional independence among older adults and individuals with cognitive impairment. Despite being developed earlier, it remains widely used in both clinical and research settings due to its reliability, simplicity, and ability to assess essential daily activities, including medication management, food preparation, transportation, housekeeping, and financial management.[10]

Procedure

The experimental group (15 participants) received prospective memory training, while the control group (15 participants) underwent conventional occupational therapy. In prospective memory training, the therapist introduced strategies such as implementation intentions, use of external memory aids, rehearsal, and structured reminders to support goal-directed behavior. Sessions were interactive and client-centered, allowing participants to engage in tasks that reflected their everyday routines. The therapist guided participants in identifying meaningful goals, applying memory strategies, and practicing tasks to strengthen independent performance in instrumental activities of daily living. The training program consisted of structured activities aimed at improving prospective memory through techniques such as reminder-based tasks, association strategies, task rehearsal, and real-life functional practice. These strategies help enhance encoding, retention, and retrieval processes involved in prospective memory tasks.

Previous studies reported that prospective memory and metacognitive training can significantly improve cognitive performance and quality of life among older adults with dementia.[11] Therefore, these strategies were selected as they are expected to support cognitive functioning and improve independence in daily activities. The prospective memory training program used in this study consisted of structured sessions designed to enhance the participants' ability to remember and perform intended actions in daily life. The training focused on both internal memory strategies (such as visualization, association, and rehearsal) and external memory aids (such as alarms, reminders, calendars, and written notes). Participants were guided to practice these strategies through structured tasks and real-life scenarios related to daily activities, such as remembering appointments, taking medication, or completing household tasks. Feedback and reinforcement were provided to improve learning and retention. The prospective memory training program was designed to be delivered in both individual and group formats, with a selective focus on prospective memory skills.

The control group received conventional occupational therapy including: general cognitive stimulation activities, relaxation techniques, and guidance on organizing daily tasks. The study duration was 3 months, with sessions held 3 days per week for a total of 36 sessions, each lasting 45 minutes. Following that, the post-test was conducted with the Lawton and Brody Instrumental Activities of Daily Living (IADL) Scale.

Experimental group intervention strategies

The experimental group received a structured prospective memory training program aimed at improving instrumental activities of daily living (IADL). The intervention began with a general assessment, rapport building, a baseline evaluation, and pre-test administration using the Lawton–Brody Instrumental Activities of Daily Living Scale. Participants were educated about prospective memory and trained in memory strategies such as visualization, association, chunking, repetition, and implementation intention ("If-Then") techniques. External memory aids including alarms, calendars, reminder notes, pill organizers, and smartphone reminders were introduced to support daily functioning. The intervention also included spaced retrieval exercises, medication management training, meal preparation activities, financial management tasks, grocery shopping and weekly planning, transportation and route recall training, social participation activities, problem-solving exercises, time management, household task organization, face–name association, and delayed recall exercises. Real-life simulation activities were

conducted to reinforce the application of memory strategies in everyday tasks. A post-test assessment was conducted using the Lawton–Brody Instrumental Activities of Daily Living Scale, followed by feedback and closure of the intervention program.

Control group intervention strategies

The control group received conventional occupational therapy interventions focused on daily living skills. The program included a general assessment, rapport building, and a pre-test using the Lawton–Brody Instrumental Activities of Daily Living Scale. Intervention activities included personal hygiene and grooming, medication management, cooking and meal preparation, shopping and grocery management, budgeting and money management, household management, and social interaction and communication practice. At the end of the intervention, post-test assessment and feedback were provided.

Data analysis

Descriptive statistics (means, standard deviations, frequencies, and percentages) were used to summarize demographic and baseline characteristics. The normality of the data distribution was assessed using the Kolmogorov–Smirnov test. As the data were normally distributed, inferential statistics appropriate for a true experimental study design were 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 each group, and independent t-tests were applied to assess differences between the groups. A 95% confidence interval was used to estimate the precision of the observed effects. All statistical analyses were conducted at a significance level of $p<0.05$, with $p<0.05$ considered statistically significant and $p<0.01$ considered highly significant. Statistical analyses were performed using IBM SPSS Version 23.0.

Results

The Effect of prospective memory training on instrumental activities of daily living skills among individuals with mild dementia was conducted in Jagruti Rehabilitation Center. A total of 30 participants were included in this study. The experimental group consists of 15 members, of whom 66.7% were female, and 33.3% were male, and the control group consists of 15 members, of whom 60.0% were female, and 40.0% were male. Both groups were comparable in gender distribution and age, ensuring the validity of the comparison (Figure 1). The pretest and post-test scores provide a basis for the effect of prospective memory training on instrumental activities of daily living skills among individuals with mild dementia.

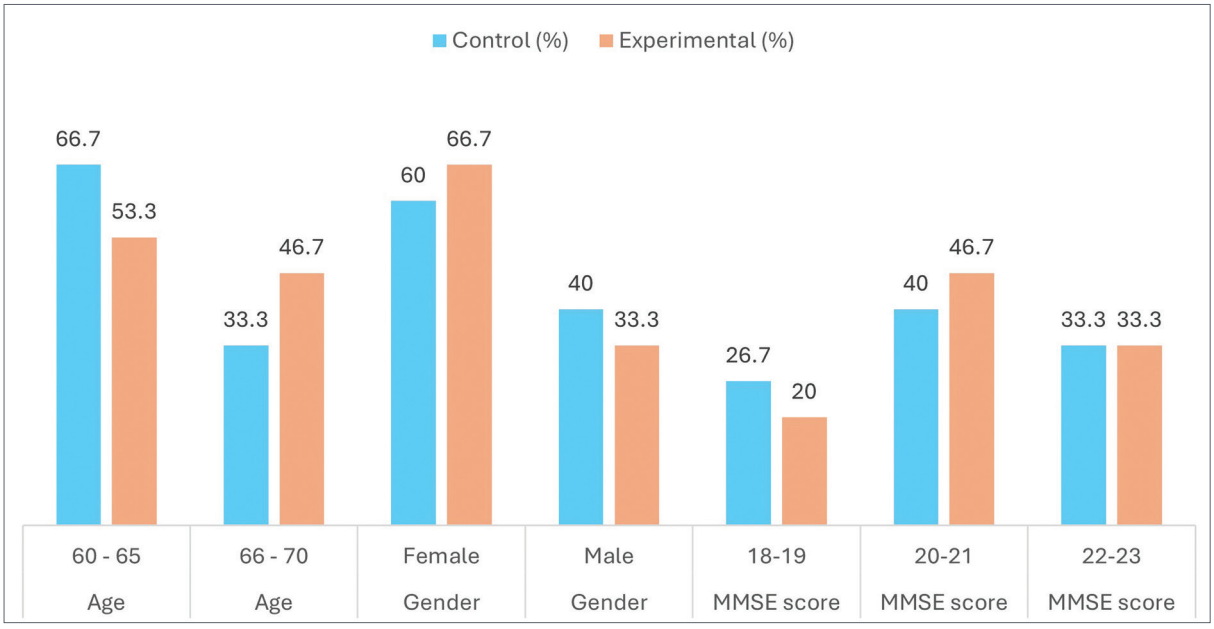


Figure 1. *Demographic variables.*

Table 1 displays the statistical difference between the pre-test and post-test of the control group. In the control group, the mean score increased slightly from 2.07 to 2.87, with a statistically significant increase (p=0.028). Table 2 presents the statistically significant difference between the pre-test and post-test for the experimental group. In the experimental group, the mean score increased from 2.80 to 6.73. This demonstrates that the experimental group had a stronger impact on instrumental activities of daily living. Table 3 displays the differences between the post-tests of the control and experimental groups. The experimental group has significantly higher Lawton-Brody Instrumental Activities of Daily Living (IADL) scale scores (2.87 vs 6.73) than the control group. These findings indicate that the intervention provided to the experimental group resulted in significantly better outcomes in improving prospective memory training.

Table 1. Comparison of instrumental activities of daily living skills scores between pre- and post-tests in the control group (Paired t-test).

Test	N	Mean (SD)	T value	Df	*p* value
Pre- test	15	2.07 (0.458)	-2.449	14	0.028*
Post- test	15	2.87)1.187)			

Note: ***significant at 1% level, *significant at 5% level.*

Table 2. Comparison of instrumental activities of daily living skills scores between pre- and post-tests in the experimental group (Paired t-test).

Test	N	Mean (SD)	T value	Df	*p* value
Pre-test	15	2.80 (0.775)	-11.415	14	0.000**
Post-test	15	6.73 (1.624)			

Note: ***significant at 1% level, *significant at 5% level.*

Table 3. Comparison of post-test level instrumental activities of daily living skills scores between the control and experimental group (Independent t-test).

Test	N	Mean (SD)	T value	Df	*p* value
Control	15	2.87 (1.187)	-7.444	28	0.000
Experimental	15	6.73 (1.624)			

Note: ***significant at 1% level, *significant at 5% level.*

Discussion

Elderly people with mild dementia often struggle with memory and cognitive skills, which affect their ability to live independently. This study found that prospective memory training for instrumental activities of daily living improved their ability to perform daily tasks such as managing medications and using household items. The training enhanced their memory and increased their independence in Instrumental activities of daily living (IADLs). These findings suggest that such prospective memory training interventions can effectively support instrumental activities of daily living skills among individuals with mild dementia. This study was conducted over 3 months among elderly people with mild dementia at the Jagruthi Rehabilitation Center in Chennai.

A total of 30 patients were selected for the study; 15 were randomly allocated to the experimental group by lottery and 15 to the control group. The pre-test was conducted using Lawton and Brody's Instrumental Activities of Daily Living scale to assess instrumental activities of daily living skills within each group. The experimental group received prospective memory training per the protocol, whereas the control group received conventional occupational therapy. After 3 months of interventions, post-test evaluations were conducted for both groups, and scores were analyzed using statistical methods.

Figure 1 illustrates the baseline comparison of demographic variables and outcome measures between the control and experimental groups. The findings showed that both groups were comparable in mean age and gender distribution, with no statistically significant differences.

Table 1 presents the statistical analysis of pre- and post-test scores on the Lawton–Brody Instrumental Activities of Daily Living Scale for participants in the control group. As the p=0.028 was lower than the significance level of 0.05, a statistically significant improvement in instrumental activities of daily living was observed following the conventional occupational therapy intervention.[11] Similar findings were reported by

Martinez Campos *et al.*, who highlighted that occupational therapy interventions help reduce caregiver burden, enhance coping strategies, and improve the quality of life of caregivers of individuals with dementia. The authors also emphasized that person-centered, activity-oriented, and home-based interventions support the maintenance of daily routines and promote independence for both caregivers and individuals with dementia.[12]

Similarly, Dinesh E and Surya Vishnuram reported that memory relay game-based training improved memory performance and cognitive engagement among elderly individuals with dementia.[13] Their findings suggest that structured cognitive activities can positively influence cognitive abilities and support participation in everyday tasks, which may explain the mild improvement observed in the control group following conventional occupational therapy. However, the authors noted variations in study design and outcome measures, suggesting the need for standardized, evidence-based protocols and culturally relevant long-term interventions to strengthen occupational therapy practice in dementia care.

Table 2 presents the statistical analysis of the pre-test and post-test scores on Lawton Brody Instrumental Activities of Daily Living Scale for the experimental group. The p=0.000, which is less than 0.01, indicates a highly statistically significant difference between the pre-test and post-test scores of the experimental group. This finding suggests that the treatment plan involving prospective memory training was effective in improving instrumental activities of daily living skills. The present study's findings align with those of Moore *et al.,* who reported that memory training significantly enhanced cognitive functioning in individuals with dementia. Their study showed that structured cognitive rehabilitation improves attention, recall, and problem-solving.[14]

Similarly, Xin Zhao *et al*. reported associations between cognitive functions and self- and informant-reported abilities in instrumental activities of daily living among older adults with mild cognitive impairment, suggesting that cognitive performance directly impacts independence in daily functioning.[15] Likewise, the current study found improvements in instrumental activities of daily living after prospective memory training, suggesting that targeted cognitive stimulation can support functional abilities and help slow cognitive decline in instrumental activities of daily living skills among individuals with mild dementia. These findings are further supported by Virginia G. Wadley *et al*., who demonstrated that individuals with mild cognitive impairment exhibited reduced speed in performing instrumental activities of daily living, underscoring the importance of cognitive interventions in improving functional efficiency and independence in everyday activities.[16]

Table 3 compares post-test scores between the control and experimental groups. Participants in the control group, who received conventional occupational therapy, demonstrated only mild improvement in their instrumental activities of daily living (IADL) skills. In contrast, the experimental group that received prospective memory training showed a significant increase in post-test mean scores (p<0.01), indicating a highly significant improvement. These findings suggest that prospective memory training is effective in enhancing instrumental activities of daily living skills among individuals with mild dementia.

The findings are consistent with the study by Hering A, *et al*., which reported that prospective memory interventions contribute significantly to successful aging by improving everyday functioning and independence in older adults.[8] Consistent with their conclusions, the present study highlights that targeted cognitive interventions, such as prospective memory training, can effectively improve the performance of instrumental activities of daily living, thereby enhancing autonomy and overall quality of life.

Limitations

The study was conducted with a relatively small sample size and within a limited age group, which may affect the generalizability of the findings. Therefore, further research involving larger and more diverse populations is recommended to strengthen the evidence base and support broader clinical application. Future studies should also explore the long-term effectiveness of prospective memory training and its applicability across different stages of dementia and care environments.

Conclusion

This study found that prospective memory training improved instrumental activities of daily living (IADL) skills among individuals with mild dementia, as measured using the Lawton-Brody IADL scale. The findings suggest that prospective memory training may be an effective intervention to enhance functional independence and can be incorporated into Occupational Therapy programs. However, as the study was conducted with a small sample size and a limited age range, further research with larger, more diverse populations is recommended to strengthen the evidence and support wider clinical application.

This study demonstrated that prospective memory training significantly improved instrumental activities of daily living (IADL) skills among individuals with mild dementia, as measured by the Lawton-Brody IADL scale. The positive outcomes observed in the experimental group highlight the effectiveness of this training approach, reinforcing its value in supporting functional independence. These results also suggest the potential of prospective memory-based interventions for broader application in dementia care, encouraging further exploration among different populations and care settings. Given the study's limited sample size and specific age group, future research should investigate

its effectiveness in larger and more diverse samples to strengthen evidence-based rehabilitation practices and inform clinical guidelines for enhancing daily living skills in dementia care.

Thus, this study provided strong evidence that prospective memory training improves instrumental activities of daily living among individuals with mild dementia. Hence, prospective memory training can be incorporated with our Occupational Therapy program to improve instrumental activities of daily living skills among individuals with Mild dementia.

Ethical approval

This study has been approved by the Institutional Scientific Review Board (ISRB) of Saveetha College of Occupational Therapy with REF. NO. of SCOT/ ISRB /013/ 2025.

Funding

No funding was received for this study.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Punitha P: study design, data analysis, manuscript preparation; **Yamuna T**: study conduct under the main author's supervision. All authors approved the final manuscript.

Acknowledgements

I sincerely thank Saveetha Institute of Medical and Technical Sciences, Saveetha College of Occupational Therapy, Dr. M. Arun Kumar, MOT, Ph.D., for his continuous encouragement and support throughout this research.

References

[1] Mendrofa FAM, Iswanti DI, Hani U. Efficacy of brain gym on the cognitive function improvement of people with dementia. J Keperawatan Jiwa. 2020; 8(4): 557.

[2] Jones KC. Update on major neurocognitive disorders. FOCUS. 2021; 19: No. 3. doi: 10.1176/appi.focus.20210004.

[3] Yuill N, and Hollis V. A systematic review of cognitive stimulation therapy for older adults with mild to moderate dementia: An occupational therapy perspective. Occup Ther Int. 2011; 18(4): 163-86. doi: 10.1002/oti.315.

[4] Guo HJ, Sapra A. (2020). Instrumental Activity of Daily Living (IADL). In StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing. Updated December 18, 2019. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK557863/.

[5] Moon SH, Oh E, Chung D, Hong GRS. Changes in instrumental activities daily living limitations and their associated factors according to gender in community-residing older adults: A longitudinal cohort study. PLoS ONE. 2024; 19(1): e0296796. doi: 10.1371/journal.pone.0296796.

[6] Knopman DS, Petersen RC. Mild cognitive impairment and mild dementia: A clinical perspective. Mayo Clin Proc. 2014; 89(10): 1452-9. doi: 10.1016/j.mayocp.2014.06.019.

[7] Tse ZCK, Cao Y, Ogilvie JM, Chau BHC, Ng DHC, Shum DHK. Prospective memory training in older adults: A systematic review and meta-analysis. Neuropsychol Rev. 2023; 33; 347-72. doi: 10.1007/s11065-022-09536-5.

[8] Hering A, Rendell PG, Rose NS, Schnitzspahn KM, Kliegel M. Prospective memory training in older adults and its relevance for successful aging. Psychol Res. 2014; 78(6): 892-904. doi: 10.1007/s00426-014-0566-4.

[9] Graf C. The Lawton Instrumental Activities of Daily Living Scale. Am J Nurs, 2008; 108(4): 52-62. doi: 10.1097/01.NAJ.0000314810.46029.74.

[10] Punitha P, Jessy J. Effect of meta-cognitive strategy-based activities on executive functioning, social networking, and instrumental activities of daily living among women. *Cureus. 2024; 16*(4): e58698. doi: 10.7759/cureus.58698.

[11] Punitha P, Shrishte AS. The efficacy of prospective memory and metacognitive skills training in improving the cognitive skills and the quality of life of elderly persons with dementia. Cureus. 2024; 16(8): e67454. doi: 10.7759/cureus.67454.

[12] Martínez-Campos A, Compañ-Gabucio LM, Torres-Collado L, García-de la Hera M. Occupational therapy interventions for dementia caregivers: Scoping review. Healthcare (Basel). 2022; 10(9): 1764. doi: 10.3390/healthcare10091764.

[13] Dinesh E, Surya Vishnuram. Effect of memory relay game-based training on dementia among elderly individuals. Psychitr Danub. 2024; 36(3-4): 403-3. doi: 10.24869/psyd.2024.402.

[14] Moore S, Sandman CA, McGrady K, Kesslak JP. Memory training improves cognitive ability in patients with dementia. Neuropsychol Rehabil. 2001; 11(3-4): 245-61. doi: 10.1080/09602010042000222.

[15] Xin Zhao, Wenjia Liang, Maes JHR. Associations between self- and informant-reported abilities of instrumental activities of daily living and cognitive functions in older adults with mild cognitive impairment. Arch Clin Neuropsychol. 2021; 36(5): 723-33. doi: 10.1093/arclin/acaa110.

[16] Wadley VG, Okonkwo O, Crowe M, Ross-Meadows LA. Mild cognitive impairment and everyday function: Evidence of reduced speed in performing instrumental activities of daily living. Am J Geriatr Psychiatry. 2008; 16(5): 416-24. doi: 10.1097/JGP.0b013e31816b7303.

Journal of Associated Medical Sciences 2026; 59(3): 99-106

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



Complications of hyaluronic acid fillers in HIV-positive patients: A narrative review

Kaewkan Thasana-Ekachit* **and Arisa Kaewkes**

College of Integrative Medicine, Dhurakij Pundit University, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 23 December 2025
Accepted as revised 27 May 2026
Available online 5 June 2026

Keywords:
Hyaluronic acid fillers, HIV, complications, facial lipoatrophy, aesthetic medicine

ABSTRACT

Background: The introduction of antiretroviral therapy (ART) has improved the survival of people living with HIV, yet facial lipoatrophy remains a visible and stigmatizing complication. Hyaluronic acid (HA) fillers are widely used for volume restoration due to their safety, biocompatibility, and reversibility.

Objectives: This narrative review focuses exclusively on HA fillers in HIV-positive patients, aiming to clarify their complication profile.

Materials and methods: A comprehensive search of PubMed, Scopus, and Google Scholar (2000-2024) identified 120 records. After screening and eligibility assessment, 13 studies were included. Data on study design, population, filler type, complications, and management were extracted and narratively synthesized, with results summarized in tabular form.

Results: Early adverse events, such as swelling and erythema, occurred in 50-80% of patients, and bruising in ~50%. Intermediate complications included nodules in 5-10% and indurations up to 30%, usually resolving with conservative management or hyaluronidase. Delayed infections (2-4%) were most often due to *Staphylococcus epidermidis*, *Propionibacterium acnes*, or atypical mycobacteria and required antibiotics or drainage. Granulomatous reactions and abscesses were rare (<5%) and typically occurred in patients with prior permanent filler exposure. No severe vascular events were reported in HIV-specific cohorts, though such risks remain relevant.

Conclusion: HA fillers are effective and generally safe for treating HIV-associated facial lipoatrophy, with most complications being mild and transient. However, immune dysregulation, low CD4 counts, and prior filler history may predispose to granulomas or infections. This review highlights the need for careful patient selection, meticulous technique, and further prospective research to establish standardized safety guidelines.

Introduction

The introduction of potent antiretroviral therapy (ART) has transformed the clinical course of HIV infection, markedly improved survival rates, and enhanced the quality of life of people living with HIV. With sustained viral suppression and immune reconstitution, HIV infection has become a chronic, manageable condition rather than a terminal illness.[1,2] Nevertheless, metabolic and morphological complications remain significant clinical concerns. One of the most visible and distressing complications is facial lipoatrophy, characterized by progressive loss of subcutaneous fat in the cheeks, temples, and periorbital regions. This condition is particularly

* **Corresponding contributor.**
Author's Address: *College of Integrative Medicine, Dhurakij Pundit University, Bangkok, Thailand.*
E-mail address: 66130206@dpu.ac.th
doi: 10.66285/JAMS.2026.083
E-ISSN: 2539-6056

associated with prolonged exposure to older thymidine analog nucleoside reverse transcriptase inhibitors, which interfere with mitochondrial function and promote adipocyte apoptosis.[3-5] The resulting changes in facial appearance may give patients a prematurely aged or stigmatized look that is often associated with HIV disease, thereby undermining both psychosocial well-being and treatment adherence.

The psychosocial implications of facial lipoatrophy are profound.[6,7] Patients frequently report diminished self-esteem, reluctance to participate in social activities, and fear of disclosure due to the recognizable appearance of HIV-associated facial changes. These challenges can lead to social isolation, depression, and reduced adherence to ART, ultimately compromising treatment outcomes. Restoring facial volume has therefore become an important therapeutic goal, not only to improve appearance but also to enhance quality of life and support long-term adherence to HIV care.

In response to this need, dermal fillers have been introduced as an effective treatment option. Among the available agents, hyaluronic acid (HA) fillers are particularly favored because of their favorable safety profile, biocompatibility, biodegradability, and reversibility with hyaluronidase.[8-10] Compared with permanent fillers such as silicone or polyacrylamide gel, HA fillers are associated with fewer long-term adverse effects and can be safely dissolved if complications arise.[11,12] These characteristics make them especially suitable for use in immunocompromised patients, where safety and reversibility are critical considerations. Clinical studies have consistently demonstrated that HA fillers can restore facial contours, improve patient satisfaction, and yield durable results that may persist for one year or longer.[13-20]

Despite these encouraging outcomes, the literature addressing the complications of HA fillers in HIV-positive patients remains limited.[21,22] Most published studies combine results from heterogeneous filler types or general populations, reducing the specificity of findings for HIV-infected cohorts. Moreover, HIV-positive individuals often exhibit unique clinical characteristics that may influence treatment outcomes, including persistent immune dysregulation, chronic systemic inflammation, long-term ART exposure, history of opportunistic infections, and prior use of permanent fillers.[23-26] These factors may modify both the incidence and the clinical presentation of complications, such as infection, granuloma formation, or delayed hypersensitivity reactions.

Previous reviews on HIV-associated facial lipoatrophy have often grouped hyaluronic acid fillers together with other injectable products such as poly-L-lactic acid, calcium hydroxylapatite, and permanent fillers. While these reviews provided useful overviews, they did not specifically distinguish the safety and complication profiles of HA fillers.[21,22] This distinction is clinically important because HA fillers are biodegradable, reversible with hyaluronidase, and currently among the most widely used filler types in aesthetic and reconstructive practice.[27,28] By focusing exclusively on HA fillers in HIV-positive patients, this review addresses a gap in the literature and aims to provide a clearer understanding of their unique risk and safety considerations.

This is the first narrative review focusing exclusively on HA fillers in HIV-positive patients, distinguishing them from other filler types that carry different risk profiles. This focus is clinically relevant because HA is currently the most widely used and reversible filler in practice, making the clarification of its safety profile in immunocompromised patients particularly important.

A clear understanding of the complication profile of HA fillers in HIV-positive patients is therefore essential. This narrative review seeks to synthesize existing evidence on reported adverse events, associated risk factors, and clinical implications of HA filler use in this population. By consolidating current knowledge, this work aims to support evidence-based decision-making, guide safe and effective clinical practice, and highlight gaps warranting further investigation in future research.

Materials and methods

This study was conducted as a narrative review designed to synthesize and critically appraise the available evidence regarding complications associated with hyaluronic acid (HA) fillers in HIV-positive patients. A narrative review approach was selected because the available evidence is heterogeneous and consists predominantly of observational studies, small prospective cohorts, case series, and case reports, which limits the feasibility of formal quantitative synthesis or meta-analysis. The review followed a structured yet non-systematic approach to provide a comprehensive overview of the published literature while maintaining flexibility to incorporate a wide range of study designs and clinical observations.

A comprehensive literature search was performed using three major electronic databases-PubMed, Scopus, and Google Scholar-from January 2000 to December 2024. These databases were selected for their broad coverage of biomedical research, clinical studies, and systematic reviews. The search strategy combined both Medical Subject Headings (MeSH) and free-text terms, organized into conceptual groups using Boolean operators. Terms referring to the intervention included “hyaluronic acid,” “HA filler*,” and “dermal filler*,” which were linked with OR. HIV-related terms (“HIV” OR “human immunodeficiency virus”) were grouped similarly, as were outcome-related terms such as “complications,” “adverse events,” “granuloma,” and “infection.” These groups were then combined using AND to ensure that each search result contained all major concepts. Boolean operators were applied to maximize sensitivity while maintaining specificity.

Reference lists of eligible studies and relevant review articles were also manually screened to identify additional publications not captured in the electronic database search.

Inclusion criteria required that studies be published in English and specifically evaluate the use of HA fillers in HIV-positive individuals. Eligible designs included randomized controlled trials, prospective or retrospective cohort studies, observational case series, case reports, and systematic reviews. This broad inclusion enabled the synthesis of both quantitative and qualitative data that reflect real-world clinical practice.

Exclusion criteria included studies that exclusively investigated non-HA fillers such as poly-L-lactic acid (PLLA), calcium hydroxylapatite (CaHA), or silicone; studies conducted solely in HIV-negative populations; animal studies; and articles not published in English. Studies in which outcomes could not be attributed specifically to HA fillers were also excluded.

The initial search retrieved 120 records. After removing duplicates and screening titles and abstracts, 35 articles were selected for full-text review. Following the application of inclusion and exclusion criteria, 13 studies were ultimately included in the final synthesis.

Data extraction was performed independently by two reviewers. Extracted variables included study design, sample size, patient demographics, immune status (CD4 count, viral load), type and brand of HA filler, injection techniques, follow-up duration, clinical outcomes, and reported complications. Given that most included studies were observational and involved small sample sizes, a formal risk-of-bias assessment tool (such as ROBINS-I or Cochrane RoB2) was not applied. Instead, the strength and limitations of the available evidence were assessed qualitatively, with attention to study design, consistency of findings, and completeness of outcome reporting.

For data synthesis, reported complications were categorized by timing of onset into early (days to weeks), intermediate (weeks to months), and late (months to years). The findings were synthesized narratively, and key studies were summarized in tabular form to highlight patterns in outcomes and complications. This approach allowed the review to integrate diverse forms of evidence while providing a clinically relevant synthesis tailored to the population of HIV-positive patients undergoing treatment with HA fillers. The study selection process followed a structured approach including identification, screening, eligibility assessment, and final inclusion. A flow diagram illustrating the literature search and study selection process is presented in Figure 1. Although this review was not conducted as a formal systematic review, the selection process was performed systematically to enhance transparency and reproducibility.

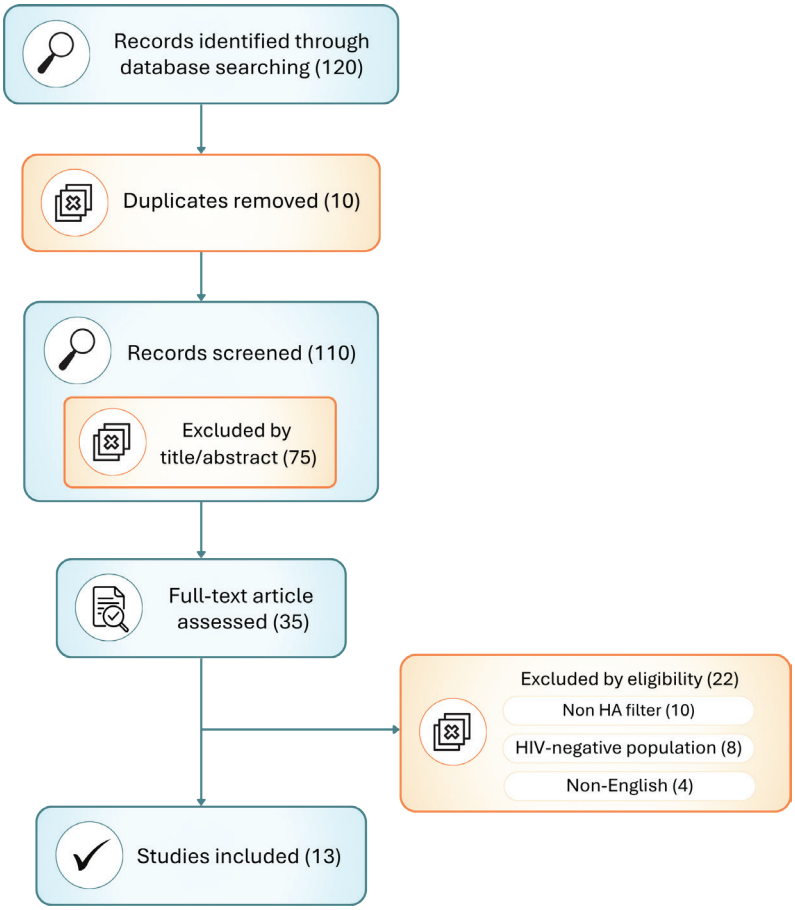


Figure 1. *Flow diagram of the literature search and study selection process.*

Results

Overall safety profile

Thirteen studies met the inclusion criteria, comprising prospective cohort studies, observational studies, case series, and systematic reviews. Across these reports, hyaluronic acid (HA) fillers consistently provided effective correction of HIV-associated facial lipoatrophy, with high levels of patient satisfaction. Most patients maintained improvements for at least 12 months, and some studies demonstrated durability extending to several years. For example, Pünchera *et al*. reported radiologic stability of filler volume for a mean of 7.8 years in a small cohort of four patients.[19] Similarly, Ho and Jagdeo[18] observed significant reductions in Carruthers Lipoatrophy Severity Scale (CLSS) scores from baseline grade 3-4 to grade 1 within 20 HIV-positive patients, with 95% rating their outcomes as "much improved" or "very much improved". [19] These findings establish HA fillers as both efficacious and durable in this population, while underscoring the importance of understanding their complication profile.

Early complications

The most common adverse events were early, transient, and mild, typically occurring within the first few days after injection. Ho and Jagdeo reported transient swelling in 16 of 20 patients (80%) and bruising in 10 of 20 (50%), all of which resolved spontaneously within 1-2 weeks.[19] Becker *et al*. documented mild erythema in 2 of 10 patients (20%), again with spontaneous resolution.[17] Similar findings were described in other cohorts, including Skeie *et al*.,[15] Denton and Tsaparas,[14] and Bechara *et al*., where local swelling, ecchymosis, and tenderness were frequent but self-limited.[29] Overall, across the included studies, early events such as erythema, ecchymosis, tenderness, and swelling were observed in 50-80% of patients but were short-lived and did not require intervention.

Intermediate complications

Intermediate complications, appearing weeks to months post-injection, were more variable but clinically important. Palpable nodules were among the most frequent issues, reported in 2 of 20 patients (10%) in the Ho and Jagdeo[18] cohort and in 3 of 10 patients (30%) in Becker *et al*.[17] Most nodules improved with conservative management, though a minority required hyaluronidase or corticosteroid injection.[11,12] Migration of filler material was radiologically identified in the study by Mundada *et al*., which demonstrated cases of misplacement or migration using MRI and PET-CT; while the exact frequency was not quantified, it was described as an occasional finding on imaging follow-up.[31] Delayed hypersensitivity reactions, presenting as swelling and erythema weeks after injection, were described in fewer than 5% of patients across cohorts, supporting the view that these immune-mediated responses are rare but relevant in HIV-positive individuals with chronic immune activation.[30,31]

Late complications

Late complications, although less common, were clinically more concerning. Granulomatous reactions have been documented in several reports and are estimated to occur in fewer than 5% of treated patients.[29,30,32] In a larger observational study by Pignatti *et al*. delayed inflammatory nodules were observed in approximately 2 of 54 patients (3.7%).[17] Delayed bacterial infections were also noted, typically presenting as abscesses or draining sinuses. Pathogens included *Staphylococcus epidermidis*, *Propionibacterium acnes*, and atypical nontuberculous mycobacteria.[16,33] The estimated frequency of infectious complications ranged from 2-3% across studies. Abscess formation was radiologically confirmed in case reports and was often difficult to treat due to potential biofilms within the filler matrix, necessitating combined antibiotic therapy and hyaluronidase or surgical evacuation. [31,34]

Severe and rare complications

Severe vascular complications were exceedingly rare in HIV-specific cohorts, with no prospective studies directly reporting events such as blindness or tissue necrosis. However, reviews in broader populations have estimated vascular occlusion to occur in approximately 1 in 6,000 to 1 in 10,000 filler injections, rates that remain clinically relevant for HIV-positive patients undergoing treatment in high-risk anatomical zones such as the glabella or nasolabial folds.[12,31,34] While none of the included HIV-focused studies documented such catastrophic events, the literature underscores the need for clinician preparedness and immediate intervention if they occur.[11,34]

Risk factors for complications

Several risk factors for complications were identified. Low CD4 counts (<200 cells/mm^3) and detectable viral loads were associated with higher rates of delayed infections and inflammatory reactions.[35-37] Patients with a history of permanent filler use, particularly silicone or polyacrylamide gel, demonstrated increased susceptibility to granulomatous reactions after subsequent HA filler injections.[27,28,38] Technical factors, including superficial placement and excessive filler volume, contributed to the development of palpable nodules in up to 30% of patients across small cohorts. [17,34] The immunological milieu of HIV, characterized by chronic inflammation and altered T-cell regulation, may further predispose to delayed hypersensitivity and biofilm-related infections,[23-26,30,31] although precise rates remain difficult to quantify due to limited sample sizes.

Clinical implications

Taken together, the findings demonstrate that

most HIV-positive patients treated with HA fillers experience only mild and transient adverse events, such as swelling or bruising, occurring in over half of cases but resolving spontaneously.[14,15,17,18] More clinically significant complications, including nodules and delayed infections, were reported in 5-10% of patients,[16,21,33] while granulomas and biofilm-associated abscesses occurred in fewer than 5%.[29-32] Severe vascular events remain rare but must be anticipated.[12,34] These data highlight the importance of careful patient selection, with preference for those with well-controlled HIV,[35-37] as well as meticulous injection technique and readiness to manage complications with hyaluronidase, corticosteroids, or antibiotics when needed.[11,31,34]

Discussion

The present review confirms that hyaluronic acid (HA) fillers are effective in correcting HIV-associated facial lipoatrophy and are generally well tolerated.[13,15-18] Early complications such as swelling, erythema, and bruising were frequent but minor and self-limiting.[14,15,17,18,29] More clinically relevant events included nodules,[16-18] indurations, granulomas,[29,30,32] and delayed infections.[16,21] These complications, although less common, carry particular importance in HIV-positive patients. The pathophysiological background of HIV may partly explain this pattern. Chronic immune activation, imbalance of T-cell subsets, and altered cytokine signaling can predispose to exaggerated foreign-body reactions, thereby increasing the risk of granulomatous inflammation.[23,24,37] In addition, impaired macrophage and neutrophil function in HIV infection may reduce clearance of bacterial biofilms,[25,26,31] providing a niche for persistent low-grade infections after filler injection. These findings suggest that complication profiles in HIV-positive patients are not solely determined by filler properties but are also influenced by host-related immunological factors. To synthesize the interaction between patient-related and procedure-related determinants of complications in this population, a conceptual framework is shown in Figure 2.

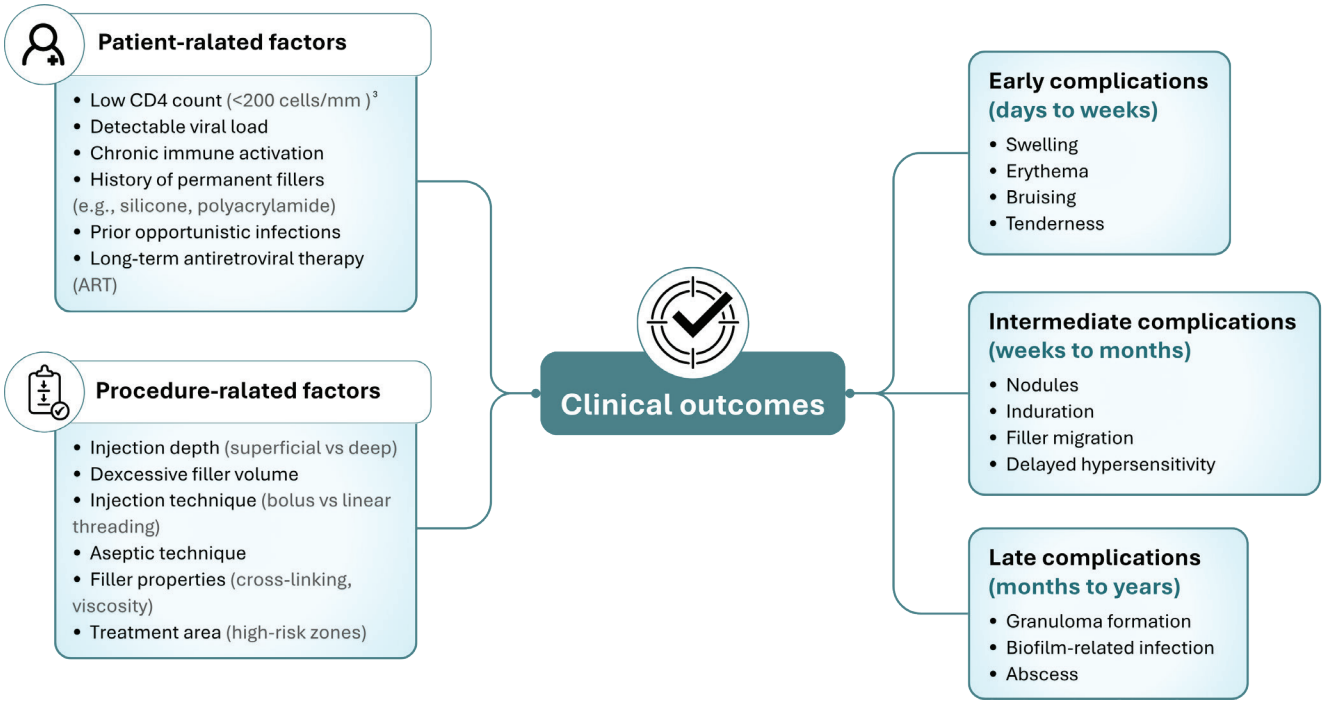


Figure 2. *Conceptual framework of complications associated with hyaluronic acid fillers in HIV-positive patients. Patient-related and procedure-related factors interact to influence the development of early, intermediate, and late complications following HA filler treatment in HIV-positive individuals.*

Comparison with the General Population

When compared with HIV-negative individuals, the overall pattern of complications appears broadly similar.[21,22] However, delayed infections, including those caused by atypical organisms such as nontuberculous mycobacteria, were more frequently reported in HIV-positive patients.[16,32,33] This difference can be understood considering the chronic immunosuppression and altered innate immune responses that characterize HIV infection.[23-25] Granulomatous responses also appeared more common in patients with a history of permanent filler use, suggesting that repeated antigenic stimulation in the context of immune dysregulation may increase the likelihood of foreign-body reactions.[27-29] Severe vascular complications, such as occlusion and necrosis, remain rare but must always be anticipated, since these are largely related to injection technique rather than host immune status.[11,31,34] Importantly, variability in reported complication rates across studies was substantial. While early adverse events

were consistently documented, intermediate and late complications showed greater heterogeneity. This variation may reflect differences in study design, injection techniques, filler types, follow-up duration, and patient characteristics, rather than true differences in complication risk alone. As such, direct comparison across studies should be interpreted with caution.

Quality assessment of included studies

The overall quality of the included evidence was limited. Most studies consisted of small observational cohorts, case series, or case reports, with very few prospective designs and no randomized controlled trials specifically conducted in HIV-positive populations. Sample sizes were generally small, often involving fewer than 20 patients, and follow-up durations varied considerably, thereby limiting the ability to adequately capture delayed or long-term complications.

In addition, there was marked heterogeneity in study design, outcome definitions, and reporting of complications. Standardized criteria for defining and grading adverse events were rarely applied, and clinically relevant immune parameters, such as CD4 count and viral load, were inconsistently reported across studies. This lack of uniformity limits the ability to draw robust comparisons and to evaluate the relationship between immune status and complication risk.

Furthermore, several potential sources of bias should be considered. Selection bias is likely, as many studies included carefully selected patients with well-controlled HIV, which may underestimate complication rates in broader clinical populations. Reporting bias may also be present, particularly in observational studies and case series, where adverse events may be underreported or inconsistently documented.

Taken together, these methodological limitations substantially weaken the overall strength of the evidence and restrict the generalizability of the findings. Consequently, the conclusions of this review should be interpreted with appropriate caution.

Limitations

The available evidence should therefore be interpreted with caution. The conclusions of this review are constrained by the limited scope of available studies. Most were small, observational, or case-based, with heterogeneous methods and outcome measures.[21,22] Sample sizes rarely exceeded 20 patients, and follow-up durations varied widely.[15,17,18,32] Very few studies directly compared HIV-positive and HIV-negative cohorts,[20,21] and stratification by immune markers such as CD4 count or viral load was inconsistently reported. [16,29,39] Without immune stratification, it remains difficult to determine whether complications are driven primarily by filler properties, technical aspects, or the immunologic environment of HIV infection.

Future directions

Future research should aim to clarify mechanistic links by prospectively evaluating complication rates according to immune parameters, including CD4 count, viral suppression, and markers of chronic inflammation.[23,26,37] Comparative trials with alternative interventions, such as poly-L-lactic acid or autologous fat grafting, could help define relative risk profiles.[20,22,40] Longer-term follow-up will be essential to capture delayed events such as granulomas or biofilm-related infections. Finally, multicenter registries that integrate clinical and immunological data would help generate real-world evidence and provide a stronger foundation for evidence-based guidelines.

Conclusion

This narrative review highlights that hyaluronic acid (HA) fillers are an effective and generally safe treatment for HIV-associated facial lipoatrophy. Across clinical studies, patients consistently experienced restoration of facial volume, improved aesthetic outcomes, and enhanced quality of life. Most complications were mild and transient, including swelling, erythema, and bruising. Nevertheless, intermediate and late complications—such as palpable nodules, granulomatous reactions, and delayed infections—were also reported. These events, although relatively uncommon, are clinically relevant given the immune dysregulation and chronic inflammation associated with HIV. Rare but severe complications, including vascular occlusion and tissue necrosis, underscore the importance of expert technique and prompt recognition.

Several risk factors were identified, including low CD4 counts, detectable viral load, prior use of permanent fillers, and technical errors such as superficial placement or excessive injection volume. These findings reinforce the need for careful patient evaluation, strict aseptic technique, and individualized treatment planning.

Overall, HA fillers remain a valuable therapeutic option for HIV-positive patients, offering both aesthetic and psychosocial benefits. However, the evidence base remains limited by small sample sizes and heterogeneous study designs. Larger, prospective studies with standardized methodologies are required to establish clear risk stratification and long-term safety guidelines.

Ethical approval

Not applicable as this article is a narrative review.

Funding

This research received no external funding.

Conflict of interest

The author declares no conflict of interest.

CRediT authorship contribution statement

Kaewkan Thasana-Ekachit: conceptualization, methodology, investigation, data curation, formal analysis, visualization, writing: original draft, **Arisa Kaewkes**: conceptualization, validation, supervision, writing: review and edit.

Acknowledgements

The authors would like to acknowledge the support of the College of Integrative Medicine, Dhurakij Pundit University, Bangkok, Thailand.

Reference

[1] Weber J. The pathogenesis of HIV-1 infection. Br Med Bull. 2001; 58(1): 61-72. doi:10.1093/bmb/58.1.61.

[2] van Schalkwyk C, Mahy M, Johnson LF, Imai-Eaton JW. Updated data and methods for the 2023 UNAIDS HIV estimates. J Acquir Immune Defic Syndr. 2024; 95(1S): e1-e4. doi:10.1097/QAI.0000000000003344.

[3] Carr A, Samaras K, Burton S, Law M, Freund J, Chisholm DJ, et al. A syndrome of peripheral lipodystrophy, hyperlipidaemia and insulin resistance in patients receiving HIV protease inhibitors. AIDS. 1998; 12(7): F51-F58. doi: 10.1097/00002030-199807000-00003.

[4] Mallewa JE, Wilkins E, Vilar J, Mallewa M, Doran D, Back D, et al. HIV-associated lipodystrophy: a review of underlying mechanisms and therapeutic options. J Antimicrob Chemother. 2008; 62(4): 648-60. doi:10.1093/jac/dkn251.

[5] Koethe JR, Lagathu C, Lake JE, Domingo P, Calmy A, Falutz J, et al. HIV and antiretroviral therapy-related fat alterations. Nat Rev Dis Primers. 2020; 6(1): 48. doi:10.1038/s41572-020-0181-1.

[6] Oette M, Juretzko P, Kroidl A, Sagir A, Wettstein M, Siegrist J, et al. Lipodystrophy syndrome and self-assessment of well-being and physical appearance in HIV-positive patients. AIDS Patient Care STDS. 2002;16(9):413-7. doi:10.1089/108729102760330254.

[7] Kraus CN, Chapman LW, Korta DZ, Zachary CB. Quality of life outcomes associated with treatment of human immunodeficiency virus (HIV) facial lipoatrophy. Int J Dermatol. 2016; 55(12): 1311-20. doi:10.1111/ijd.13374.

[8] Beasley KL, Weiss MA, Weiss RA. Hyaluronic acid fillers: a comprehensive review. Facial Plast Surg. 2009; 25(2): 86-94. doi:10.1055/s-0029-1220647.

[9] De Boulle K, Glogau R, Kono T, Nathan M, Tezel A, Roca-Martinez J-X, et al. A review of the metabolism of 1,4-butanediol diglycidyl ether-crosslinked hyaluronic acid dermal fillers. Dermatol Surg. 2013; 39(12): 1758-66. doi:10.1111/dsu.12301.

[10] Jung H. Hyaluronidase: An overview of its properties, applications, and side effects. Arch Plast Surg. 2020; 47(4): 297-300. doi:10.5999/aps.2020.00752.

[11] Rzany B, Becker-Wegerich P, Bachmann F, Erdmann R, Wollina U. Hyaluronidase in the correction of hyaluronic acid-based fillers: a review and a recommendation for use. J Cosmet Dermatol. 2009; 8(4): 317-23. doi:10.1111/j.1473-2165.2009.00462.x.

[12] Rzany B, DeLorenzi C. Understanding, avoiding, and managing severe filler complications. Plast Reconstr Surg. 2015; 136(5S): 196S-203S. doi:10.1097/PRS.0000000000001760.

[13] Gooderham M, Solish N. Use of hyaluronic acid for soft tissue augmentation of HIV-associated facial lipodystrophy. Dermatol Surg. 2005; 31(1): 104-8. doi:10.1111/j.1524-4725.2005.31020.

[14] Denton AB, Tsaparas Y. Injectable hyaluronic acid for the correction of HIV-associated facial lipoatrophy. Otolaryngol Head Neck Surg. 2007; 136(4): 563-7. doi:10.1016/j.otohns.2006.11.049.

[15] Skeie L, Bugge H, Negaard A, Bergersen B. Large particle hyaluronic acid for the treatment of facial lipoatrophy in HIV-positive patients: 3-year follow-up study. HIV Med. 2010; 11(3): 170-7. doi:10.1111/j.1468-1293.2009.00761.x.

[16] Pignatti M, Pedone A, Baccarani A, Guaraldi G, Orlando G, Lombardi M, et al. High-density hyaluronic acid for the treatment of HIV-related facial lipoatrophy. Aesthet Plast Surg. 2012; 36(1): 180-5. doi:10.1007/s00266-011-9771-3.

[17] Becker M, Balagué N, Montet X, Calmy A, Salomon D, Toutous-Trellu L. Hyaluronic acid filler in HIV-associated facial lipoatrophy: evaluation of tissue distribution and morphology with MRI. Dermatology. 2015; 230(4): 367-74. doi:10.1159/000379747.

[18] Ho D, Jagdeo J. Safety and efficacy of a volumizing hyaluronic acid filler for treatment of HIV-associated facial lipoatrophy. JAMA Dermatol. 2017; 153(1): 61-5. doi:10.1001/jamadermatol.2016.3827.

[19] Pünchera J, Salomon D, Tshikung O, Becker M, Toutous-Trellu L. Hyaluronic acid filler in HIV-associated facial lipoatrophy: Clinical and MRI evaluation of long-term durability. J Eur Acad Dermatol Venereol. 2022; 36(12): e1068-e1070. doi:10.1111/jdv.18455.

[20] Mashiko T, Tsukada K, Shirado T, Oka S, Yoshimura K. Correction of HIV-associated facial lipoatrophy: autologous fat grafting versus hyaluronic acid fillers. Plast Reconstr Surg. 2023; 153(1): 222e-225e. doi:10.1097/PRS.0000000000010872.

[21] Jagdeo J, Ho D, Lo A, Carruthers A. A systematic review of filler agents for aesthetic treatment of HIV facial lipoatrophy (FLA). J Am Acad Dermatol. 2015; 73(6): 1040-54.e14. doi:10.1016/j.jaad.2015.08.040.

[22] Sturm LP, Cooter RD, Mutimer KL, Graham JC, Maddern GJ. A systematic review of permanent and semipermanent dermal fillers for HIV-

associated facial lipoatrophy. AIDS Patient Care STDS. 2009; 23(9): 699-714. doi:10.1089/apc.2008.0230.

[23] Klatt NR, Chomont N, Douek DC, Deeks SG. Immune activation and HIV persistence: implications for curative approaches to HIV infection. Immunol Rev. 2013; 254(1): 326-42. doi:10.1111/imr.12065.

[24] De Milito A, Nilsson A, Titanji K, Thorstensson R, Reizenstein E, Narita M, et al. Mechanisms of hypergammaglobulinemia and impaired antigen-specific humoral immunity in HIV-1 infection. Blood. 2004; 103(6): 2180-6. doi:10.1182/blood-2003-07-2375.

[25] Moir S, Fauci AS. Pathogenic mechanisms of B-lymphocyte dysfunction in HIV disease. J Allergy Clin Immunol. 2008; 122(1): 12-9. doi:10.1016/j.jaci.2008.04.034.

[26] Vázquez-Castellanos J, Serrano-Villar S, Latorre A, Artacho A, Ferrús M, Madrid N, et al. Altered metabolism of gut microbiota contributes to chronic immune activation in HIV-infected individuals. Mucosal Immunol. 2015; 8(4): 760-72. doi:10.1038/mi.2014.107.

[27] Negredo E, Puig J, Ornelas A, Echeverria P, Bonjoch A, Estany C, et al. Ten-year safety with polyacrylamide gel used to correct facial lipoatrophy in HIV-infected patients. AIDS Res Hum Retroviruses. 2015; 31(8): 817-21. doi:10.1089/AID.2015.0004.

[28] Mansor S, Breiting VB, Dahlstrøm K, Andersen ÅB, Andersen O, Christensen LH. Polyacrylamide gel treatment of antiretroviral therapy-induced facial lipoatrophy in HIV patients. Aesthet Plast Surg. 2011; 35(5): 709-16. doi:10.1007/s00266-011-9671-6.

[29] Bechara FG, Gambichler T, Brockmeyer NH, Sand M, Altmeyer P, Hoffmann K. Hyaluronic acid new formulation: experience in HIV-associated facial lipoatrophy. Dermatology. 2008; 217(3): 244-9. doi:10.1159/000148252.

[30] Owczarczyk-Saczonek A, Zdanowska N, Wygonowska E, Placek W. The immunogenicity of hyaluronic fillers and its consequences. Clin Cosmet Investig Dermatol. 2021; 14: 921-34. doi:10.2147/CCID.S316352.

[31] Hong G-W, Hu H, Chang K, Park Y, Lee KWA, Chan LKW, et al. Review of the adverse effects associated with dermal filler treatments: Part I nodules, granuloma, and migration. Diagnostics. 2024; 14(15): 1640. doi:10.3390/diagnostics14151640.

[32] Claude O, Bosc R, Pigneur F, Lantieri L. Treatment of HIV-infected subjects with buttock lipoatrophy using stabilized hyaluronic acid gel. Plast Reconstr Surg Glob Open. 2015; 3(7): e466. doi:10.1097/GOX.0000000000000423.

[33] Mundada P, Kohler R, Boudabbous S, Toutous Trellu L, Platon A, Becker M. Injectable facial fillers: imaging features, complications, and diagnostic pitfalls at MRI and PET CT. Insights Imaging. 2017; 8: 557-72. doi:10.1007/s13244-017-0575-0.

[34] Urdiales-Gálvez F, Delgado NE, Figueiredo V, Lajo-Plaza JV, Mira M, Moreno A, et al. Treatment of soft tissue filler complications: expert consensus recommendations. Aesthet Plast Surg. 2018; 42(2): 498-510. doi:10.1007/s00266-017-1063-0.

[35] Deneve JL, Shantha JG, Page AJ, Wyrzykowski AD, Rozycki GS, Feliciano DV. CD4 count is predictive of outcome in HIV-positive patients undergoing abdominal operations. Am J Surg. 2010; 200(6): 694-700. doi:10.1016/j.amjsurg.2010.07.030.

[36] Shaw J, Matin N. Opportunistic infections in HIV. Medicine. 2022; 50(5): 294-7. doi:10.1016/j.mpmed.2026.02.015.

[37] Tsunemi S, Iwasaki T, Imado T, Higasa S, Kakishita E, Shirasaka T, et al. Relationship of CD4+ CD25+ regulatory T cells to immune status in HIV-infected patients. AIDS. 2005; 19(9): 879-86. doi:10.1097/01.aids.0000171401.23243.56.

[38] Faundez E, Vega N, Vera E, Vega P, Sepulveda D, Wortsman X. Clinical and color Doppler ultrasound evaluation of polyacrylamide injection in HIV patients with severe facial lipoatrophy secondary to antiretroviral therapy. Skin Res Technol. 2017;23(2):243-8. doi:10.1111/srt.12329.

[39] Bugge H, Negaard A, Skeie L, Bergersen B. Hyaluronic acid treatment of facial fat atrophy in HIV-positive patients. HIV Med. 2007;8(8):475-82. doi:10.1111/j.1468-1293.2007.00494.x.

[40] El-Beyrouty C, Huang V, Darnold CJ, Clay PG. Poly-L-lactic acid for facial lipoatrophy in HIV. Ann Pharmacother. 2006;40(9):1602-6. doi:10.1345/aph.1G740.

Journal of Associated Medical Sciences 2026; 59(3): 107-115 / Erratum

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



Work performance assessment in Thai homeless shelters: An exploratory factor analysis

Uthaikan Thanapet, Watthanaree Ammawat, Maliwan Rueankam, Winai Chatthong, Supalak Khemthong*

Occupational Therapy Division, Faculty of Physical Therapy, Mahidol University, Nakhon Pathom province, Thailand.

ERRATUM

This article has been corrected. The authors declare that these corrections do not change the results or conclusions of this paper. **Original article:** J Assoc Med Sci. 2026; 59(1): 145-52. https://doi.org/10.12982/jams.2026.018

ARTICLE INFO

Article history:
Received 22 April 2026
Accepted as revised 25 May 2026
Available online 8 June 2026

Keywords:
Occupational performance, protective shelter, mental health disability, allied health, functional assessment

ABSTRACT

Background: Institutional sheltering can lead to occupational deprivation and psychosocial disengagement, particularly among adults with mental health-related disabilities. In Thailand, homeless shelters operate under rigid custodial systems that may hinder functional recovery, highlighting the need for culturally adapted assessment tools.

Objectives: To validate an adapted version of the Assessment of Work Performance (AWP) for use in Thai homeless shelters and to identify functional skill patterns and psychosocial barriers that inform allied health service planning.

Materials and methods: A cross-sectional study was conducted with 60 adults with mental health-related disabilities residing in a government-operated destitute shelter. Participants completed a work-based AWP task rated across 14 work-skill items. Exploratory factor analysis (EFA) examined the underlying factor structure, while regression analysis tested the predictive value of extracted factors on occupational performance.

Results: The EFA supported a two-factor solution: 1) task adaptation and relational organization, and 2) communication and expression-accounting for 56.78% of the total variance (KMO=0.696, Bartlett's test significant). Regression analysis showed that task adaptation and relational organization significantly predicted occupational performance, explaining 69.6% of the variance (R^2=0.696).

Conclusion: Task adaptation and environmental structuring are key determinants of occupational engagement in institutional shelter settings. Performance-based assessment, combined with structured observation during task engagement, provides a culturally relevant strategy for allied health professionals to guide individualized rehabilitation for marginalized populations in Thai homeless shelters.

Introduction

Homelessness and institutional sheltering profoundly disrupt daily routines, roles, and identities-core elements closely tied to disability risk and psychosocial well-being.[1,2] Prolonged displacement from stable living environments often results in impairments in cognitive organization, adaptive functioning, and relational capacities, limiting

* **Corresponding contributor.**
Author's Address: *Occupational Therapy Division, Faculty of Physical Therapy, Mahidol University, Nakhon Pathom province, Thailand.*
E-mail address: supalak.khe@mahidol.ac.th
doi: 10.66285/JAMS.2026.084
E-ISSN: 2539-6056

participation in daily activities, social relationships, and rehabilitation.[1,2] Performance-based assessments, such as the Executive Function Performance Test (EFPT), have proven useful for detecting subtle executive dysfunction among marginalized individuals, even when standard cognitive screenings do not indicate impairments.[3]

Occupational deprivation-the inability to engage in meaningful activities due to environmental, social, or systemic barriers-further compounds these challenges. It has been linked to declines in executive functioning, emotional regulation, and interpersonal capacity, reducing readiness for community participation and long-term recovery.[1,2] Functional deficits often become entrenched among older or socially excluded adults. Qualitative studies highlighted how boredom, fractured occupational identities, and loss of meaningful roles shape the lived experiences of homelessness, contributing to disengagement and diminished volition.[4-7]

In Thailand, protective shelters provide long-term institutional care for homeless or abandoned adults. However, these environments are often governed by rigid routines and custodial care models that restrict opportunities for autonomy, decision-making, and skill development.[8] Needs assessments have consistently identified three recurring barriers among residents: psychiatric symptoms that disrupt consistency, estranged family relationships that undermine motivation, and disrupted work identity leading to frustration and loss of purpose. These contextual barriers underscore the need for structured tools that can evaluate functional skills relevant to both rehabilitation and reintegration.[4,5,8]

The Assessment of Work Performance (AWP), grounded in the Model of Human Occupation, offers a structured approach to evaluating process, motor, and communication-interaction skills relevant to occupational participation.[9,10] Although the original AWP has demonstrated validity and clinical utility in international work rehabilitation contexts, evidence regarding its use and psychometric evaluation in Southeast Asian institutional settings remains limited.[11] This gap is notable given that systematic reviews have highlighted the importance of structured functional assessments for individuals experiencing or transitioning from homelessness.[12,13] Moreover, psychological factors such as emotional regulation, motivation, and sense of purpose are often overlooked in standard psychiatric or cognitive evaluations, particularly among older or institutionalized populations.[11,14]

To address these gaps, the AWP was adapted to reflect the domestic and vocational activities typical of Thai shelters, ensuring contextual suitability for this population. This study, therefore, aimed to explore the underlying structure of the Thai version of AWP using exploratory factor analysis (EFA) among adults living in Thai homeless shelters. Findings provide evidence to guide rehabilitation planning and individualized care for adults with mental health-related disabilities living in institutional shelter environments.[15,16]

Materials and methods

Study design and ethical considerations

This study was conducted as a rehabilitation needs analysis and functional skill profiling investigation, rather than an intervention trial. A cross-sectional observational design was employed to examine functional skill patterns among adults residing in one government-operated homeless shelter. The study adhered to the STROBE guidelines for observational research. Ethical approval was granted by the Mahidol University Central Institutional Review Board (MU-CIRB No. 2023/186.2212). All participants provided written informed consent prior to enrollment, in accordance with the Declaration of Helsinki.

The sample size was determined based on recommendations for factor analysis, which suggest a minimum of 5-10 participants per item to ensure adequate statistical power.[17] Given that the adapted Assessment of Work Performance (AWP) included six core items, a minimum of 30-60 participants were required. To enhance reliability and meet these methodological guidelines, sixty adults with mental health-related functional disabilities were purposively recruited from the Nonthaburi Home for the Destitute, a government-operated homeless shelter in central Thailand. Participants were long-term institutional residents, categorized as homeless or abandoned. Inclusion criteria were age 18-64 years.

Participants and setting

Sixty adults aged 18-64 years with mental health-related functional disabilities were purposively recruited from a government-operated shelter in central Thailand. Participants were long-term institutional residents, categorized as homeless or abandoned.

Inclusion criteria were a documented diagnosis of a mental disorder verified by psychiatric records from the Nonthaburi Home for the Destitute, absence of acute psychotic symptoms, as confirmed by the attending psychiatrist and shelter nursing staff during routine clinical evaluation, fluency in Thai, and the ability to perform basic self-care tasks. Exclusion criteria included diagnosed neurological disorders, sensory impairments interfering with task participation, and significant physical disability. The determination of “absence of acute psychotic symptoms” followed established psychiatric diagnostic criteria, particularly those outlined in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5).[18]

Screening tools included the Rowland Universal Dementia Assessment Scale (RUDAS; score≥23),[19,20] and the Brief Psychiatric Rating Scale (BPRS; score≤36).[21,22] Participant eligibility was confirmed prior to enrollment by a multidisciplinary team, including psychologists, social workers, and occupational therapists, to ensure capacity for functional work-like activity.

The BPRS was selected to assess the severity of psychiatric symptoms because it provides a dimension-

al measure across a broad range of symptom domains rather than definitive diagnostic cutoffs. Although participants exhibited psychiatric symptomatology, only those without acute psychotic symptoms were included, ensuring clinical stability and the ability to engage meaningfully in occupational tasks while representing the spectrum of chronic psychiatric conditions commonly observed in long-term institutional settings.[23]

All participants who met the inclusion criteria were enrolled in the study. No participants were excluded after recruitment, and no cases were removed during analysis. As the study was conducted at a single site, no study sites were excluded.

Instruments and data collection

The primary assessment tool was the Thai version of the Assessment of Work Performance (AWP), modified to reflect culturally relevant shelter-based tasks such as sweeping, cleaning, plant care, and crafts. This adaptation was necessary because residents in Thai homeless shelters predominantly engage in domestic and vocational activities, which differ from the work-related tasks emphasized in the original AWP. The original AWP has demonstrated validity and clinical utility in previous international studies, including evidence of content validity, construct validity, and utility in work rehabilitation contexts.[9,10] The Thai version of the AWP has also undergone translation, cultural adaptation, and preliminary psychometric evaluation in adults experiencing homelessness in a residential rehabilitation setting. This preliminary validation work reported excellent content validity, acceptable internal consistency, moderate-to-good test-retest reliability, and preliminary construct validity. In the present study, the suitability of the adapted version was further reviewed by three occupational therapy experts, yielding a content validity index (CVI) of 1.00. The adapted tool also demonstrated acceptable internal consistency (Cronbach's α=0.68) and test-retest reliability (intraclass correlation coefficient [ICC]=0.77).[11] Together, these findings provide preliminary psychometric support for using the adapted Thai AWP as a descriptive, performance-based functional assessment in this shelter-based context.

The AWP measures motor, process, and communication-interaction skills on a 4-point ordinal scale, ranging from 1=incompetent performance to 4=competent performance. The standard AWP includes 14 work-skill items across three domains: five motor skills, five process skills, and four communication-interaction skills. This structured, performance-based approach is increasingly used to understand individuals' functional needs in institutional or homeless settings.[12,13,15] For this study, all 14 AWP items were rated during shelter-based work tasks. For interpretation, domain and total scores were converted into percentages of the maximum possible score and categorized descriptively using percentage-based criteria adapted from Bloom's cut-off approach[24]: high performance (≥80%), moderate performance (60-79%), and low performance (<60%). Because the AWP manual does not provide standardized diagnostic cut-off points, these categories were used to support descriptive clinical interpretation of observed functional performance in the shelter context and were not intended as standardized diagnostic thresholds. Each participant completed one structured task, which was observed by trained occupational therapy students, who rated performance and simultaneously recorded qualitative observations of psychosocial behaviors and contextual interactions. Demographic data were also collected.

Statistical analysis

Descriptive statistics for demographics and AWP scores were calculated using IBM SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA). To examine the underlying structure of the adapted AWP, an exploratory factor analysis (EFA) with principal component extraction was conducted.[25,26] Sampling adequacy was assessed using the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity, which was statistically significant, indicating suitability of the data for factor analysis. Factor scores from the extracted components were subsequently entered into multiple regression models to predict total occupational performance (AWP_TOTAL). In addition, exploratory correlations were conducted between AWP scores and measures of cognitive functioning (RUDAS) and psychiatric symptom severity (BPRS).

Patient and public involvement

Service users and members of the public were not directly involved in the design or reporting of this study due to the vulnerable nature of the population. All participants were institutionalized homeless adults with mental health-related disabilities and lacked community or family living options. However, their perspectives were incorporated through narrative observations and interviews conducted by psychologists and social workers, which informed the qualitative analysis. Occupational therapy instructors and students also contributed to participant screening and ethical readiness. Future studies will aim to integrate service user feedback more systematically into design and evaluation phases.

Results

Participant characteristics

A total of 60 participants were included in the study, with ages ranging from 26 to 63 years (mean: 44.87, SD=7.81). The majority were male (65%), and 62% had completed primary education or less. Most participants had resided in the shelter for over one year (mean: 51.70 months, SD=37.70). Screening results indicated mild cognitive impairment in some individuals, with RUDAS

scores ranging from 23 to 30 (mean: 26.77, SD: 2.06). The BPRS scores indicated varying psychiatric symptom severity, ranging from 18 to 36 (mean: 24.87, SD=5.54).

The BPRS scores indicated mild-to-moderate psychiatric symptom severity. This distribution confirmed that all participants were clinically stable and capable of engaging in occupational tasks. While some variation was observed, it reflected the heterogeneity typically found in shelter populations and was statistically controlled by including the BPRS scores as a continuous covariate in subsequent analyses. Demographic and clinical characteristics are summarized in Table 1.

Table 1. Demographic and clinical characteristics of participants (service users) (N=60).

Characteristic	Mean (SD) or N (%)
Age (years)	44.87 (7.81) [range 26-63]
Gender	Male=39 (65%), Female=21 (35%)
Education	Primary school=37 (62%) Lower secondary school=12 (20%) Upper secondary school=8 (13%) Higher vocational diploma=2 (3%) Bachelor's degree=1 (2%)
Length of stay (months)	51.70 (37.70) [range 2-214]
RUDAS score	26.77 (2.06) [range 23-30]
BPRS score	24.87 (5.54) [range 18-36]
Psychiatric diagnosis	Schizophrenia spectrum=1 (1.67%) Comorbid psychiatric and NCDs=11 (24.32%) Unspecified psychiatric disorders=10 (10.67%) Cerebral palsy=1 (1.67%) NCDs = 9 (15%) None = 28 (46.67%)

Note: *Data collected via structured demographic checklist and standardized screening tools. RUDAS: Rowland Universal Dementia Assessment Scale, BPRS: Brief Psychiatric Rating Scale.*

Total scores on the adapted AWP indicated substantial variation in occupational performance (mean=52.65, SD=3.50; range 39-56). Participants generally demonstrated higher scores in motor domains, while process and communication-interaction skills were comparatively lower, reflecting functional heterogeneity within this institutionalized population. Domain and subskill scores are detailed in Table 2. Based on the descriptive percentage-based criteria applied to total AWP scores at the participant level, 21 participants (35.00%) were classified as having high performance (≥80%), 39 (65.00%) as moderate performance (60-79%), and no participants as low performance (<60%).

Table 2. AWP domain and subskill scores and overall participant performance classification (N=60).

A. Domain and subskill descriptive statistics			
AWP domain	**Subskill**	**Mean±SD**	**Range**
Motor	Total subskill score	19.58±0.83	16-20
	Posture	3.93±0.31	
	Mobility	3.88±0.32	
	Coordination	3.98±0.13	
	Strength	3.93±0.25	
	Physical energy	3.85±0.36	

Table 2. AWP domain and subskill scores and overall participant performance classification (N=60) (continued).

AWP domain	Subskill	Mean±SD	Range
Process	Total subskill score	18.57±2.16	9-20
	Mental energy	3.87±0.47	
	Knowledge	3.52±0.72	
	Temporal organization	3.60±0.67	
	Organizations of space & objects	3.75±0.51	
	Adaptation	3.83±0.53	
Communication	Total subskill score	14.50±1.68	9-16
Interaction	Physicality	3.70±0.53	
	Language	3.67±0.48	
	Relations	3.38±0.83	
	Information & exchange	3.75±0.54	
Total AWP domain score		52.65±3.50	39-56
B. Participant-level performance classification (based on total AWP score cut-offs)			
Performance level		**N (%)**	
High (≥80%)		21 (35.00%)	
Moderate (60-79%)		39 (65.00%)	
Low (<60%)		0 (0.00%)	

Note: *Performance levels were classified descriptively based on the percentage of the maximum possible total assessment of work performance (AWP) score at the participant level, using criteria adapted from Bloom's cut-off approach: high performance ≥80%, moderate performance 60-79%, and low performance <60%. These categories were used for descriptive clinical interpretation and were not intended as standardized diagnostic thresholds.*

Exploratory factor analysis (EFA) was conducted to identify the underlying structure of AWP subskills. The analysis supported a two-factor solution comprising: 1) task adaptation and relational organization, and 2) communication and expression. Sampling adequacy was confirmed with a Kaiser-Meyer-Olkin (KMO) measure of 0.696, exceeding the recommended threshold of 0.50, and Bartlett's test of sphericity was significant (χ^2=179.629, df=36, p<0.001), indicating suitability of the data for factor analysis. The two extracted factors explained 56.78% of the total variance (Component 1=38.01%, Component 2=18.77%). Standardized factor loadings for each item are presented in Table 3.

Table 3. Factor loadings of AWP subskills from exploratory factor analysis (N=60).

Subskill	Component 1	Component 2
Adaptation	0.830	
Relations	0.787	
Task organization	0.694	
Organizations of space & objects	0.689	
Knowledge	0.668	
Mental energy	0.630	
information & exchange		0.758
Language		0.745
Physicality		0.389

Note: *The extraction method was the principal component analysis with two components extracted.*

Predictive associations with overall performance

Regression analysis showed that task adaptation (β=0.394, $p<0.001$) and relational organization (β=0.531, $p<0.001$) were significant predictors of total occupational performance scores. Together, these factors explained 69.6% of the variance ($R^2=0.696$). Additional correlation analyses revealed a significant negative association between psychiatric symptom severity (BPRS) and AWP scores (r=-0.323, $p=0.012$), indicating that higher psychiatric symptoms were associated with poorer task engagement. In contrast, cognitive function (RUDAS) showed only a small, non-significant positive association with AWP performance (r=0.093, $p=0.480$), suggesting that both psychiatric and cognitive factors may contribute to task engagement.

Discussion

This rehabilitation needs analysis highlights how functional skill clusters and psychosocial contexts influence occupational engagement among adults with mental health-related disabilities residing in Thai homeless shelters. Using exploratory factor analysis (EFA) of the adapted Assessment of Work Performance (AWP) alongside thematic analysis of narrative data, the study identified both functional skill domains and lived experiences shaping rehabilitation needs. Together, the findings demonstrate that adaptive and organizational skills are central to functional performance, while psychosocial and cultural contexts critically influence how those skills are enacted in institutional environments.

The EFA revealed two latent domains-task adaptation with spatial organization, and communication-expression. Only the former significantly predicted occupational performance, highlighting the role of task structuring, environmental planning, and adaptive behavior in facilitating engagement. This supports existing research emphasizing the importance of executive functioning and task organization for successful participation among marginalized groups.[20,21] Prior work with the Executive Function Performance Test (EFPT) also demonstrated that planning and sequencing skills are critical to independent functioning following shelter placement.[3] These findings are consistent with rehabilitation literature that positions functional organization as a core determinant of engagement in vulnerable populations.[12,16,27]

The limited predictive value of communication-expression may reflect the passive, compliance-driven nature of institutional tasks, where social demands are minimal. Many participants demonstrated quietness and behavioral withdrawal, potentially shaped by institutional routines that suppress communicative expression. This observation aligns with qualitative studies that describe disengagement, role loss, and diminished volition among individuals living in shelters.[5,27-31] Differences in occupational engagement were also observed between participants whose activities incorporated task adaptation and environmental structuring and those without such modifications. Adapted and structured tasks (e.g., simplified steps, visual cues, reduced distractions) were associated with more sustained engagement, higher task completion rates, and fewer withdrawals. In contrast, participants performing unmodified activities in non-structured environments often demonstrated shorter engagement spans and greater variability in performance. These findings highlight the mediating role of task adaptation and environmental supports in facilitating occupational engagement within institutional contexts.

The results support the use of structured, performance-based assessments to identify actionable skill domains relevant to psychosocial functioning. These findings align with previous validation of the AWP as a tool for differentiating client profiles in vocational rehabilitation[9] and reinforce its potential for culturally grounded application in shelter-based care.[1,2,4] They also echo broader calls for structured interventions to support individuals transitioning from homelessness.[12,13] The study identified functional patterns that reflect real-world rehabilitation challenges in institutional environments.

Cognitive function and psychiatric symptom severity showed modest but significant associations with occupational performance, consistent with prior findings that better cognitive status and fewer psychiatric symptoms are associated with improved functioning.[21,27] Similar patterns have been observed among women with schizophrenia and histories of homelessness.[16] These results reinforce rehabilitation frameworks such as the Canadian Model of Occupational Performance and Engagement (CMOP-E), which advocate for restoring executive function, rebuilding roles, and enabling participation through holistic support.[32-35] Our findings contribute to this framework by confirming that adaptive task behaviors and organizational capacity are key to engagement in institutional settings.

These findings carry implications for allied health teams working in shelter-based or institutional contexts. Interventions that promote task adaptation, environmental structuring, and relational supports may improve functional outcomes and psychosocial well-being. Combining structured, performance-based assessments with real-time observational inquiry can support personalized, recovery-oriented care by enabling providers to understand not only what service users can do, but also the contextual factors that contribute to their difficulties. At the systems level, sustained investment in mental health workforce training, supervision, and collaborative care models will be essential for improving service delivery to socially marginalized populations.

Limitations

Limitations include a sample size restricted to similar institutional populations, which may limit the generalizability of the findings to other settings or community-dwelling populations. Potential observer bias may also have influenced performance ratings, as assessment of only three tasks could have been affected by situational factors such as mood or environmental conditions. Nevertheless, the integration of qualitative data helped triangulate and contextualize the findings, enhancing their credibility. Future studies should employ larger and more diverse samples, extend task assessments, and include longitudinal follow-up to evaluate the stability of functional skills and rehabilitation outcomes over time.

Conclusion

This study demonstrates that performance-based assessment provides a culturally responsive approach to identifying rehabilitation needs among adults with mental health–related disabilities in Thai homeless shelters. The findings highlight that task adaptation and environmental structuring are central to functional engagement, with practical implications for individualized rehabilitation and allied health practice in institutional care. Advancing psychosocial recovery and reintegration will require continued investment in workforce capacity, interdisciplinary coordination, and the use of locally relevant assessment tools.

Ethical approval

This study received ethical approval from the Mahidol University Central Institutional Review Board (MU-CIRB No. 2023/186.2212). All participants provided written informed consent prior to enrollment, in accordance with the Declaration of Helsinki.

Funding

This project was supported by the New Researcher Grant, fiscal year 2024, awarded by Faculty of Physical Therapy, Mahidol University.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Uthaikan Thanapet: methodology, formal analysis, investigation, writing: review and edit, project administration, and funding acquisition; **Watthanaree Ammawat**: conceptualization, methodology, formal analysis, investigation, writing: review and edit; **Maliwan Rueankam**: conceptualization, methodology, investigation, data curation, writing: review and edit, supervision; **Winai Chatthong**: methodology, investigation, writing: review and edit; **Supalak Khemthong**: conceptualization, methodology, writing: original draft, review and edit.

Acknowledgements

The authors would like to thank the occupational therapy students, academic mentors, and professional supervisors who contributed to data collection and analysis. Appreciation is also extended to Unitalent and Jan Sandqvist for granting permission to translate and adapt the Assessment of Work Performance (AWP). This study would not have been possible without the cooperation of caregivers and clients at the Nonthaburi protective shelter.

References

[1] Gonyea JG, Mills-Dick K, Bachman SS. Enhancing occupational therapy services for homeless populations. Occup Ther Health Care. 2010; 24(1): 3-20. doi.org/10.3109/07380570903410803.

[2] O'Neill J, Farrow T, Cramond L. Occupational therapy for people experiencing homelessness: A scoping review. Aust Occup Ther J. 2019; 66(5): 538-50. doi.org/10.1111/1440-1630.12617.

[3] Raphael-Greenfield EI. Assessing executive and community functioning among home less persons with substance use disorders using the EFPT. Occup Ther Ment Health. 2012; 28(2): 125-40. doi.org/10.1080/0164212X.2012.679593.

[4] Marshall CA, Davidson L, Li A, Gewurtz R, Roy L, Barbic S, et al. Boredom and meaningful activity in adults experiencing homelessness: A mixed-methods study. Can J Occup Ther. 2019; 86(5): 357-70. doi.org/10.1177/0008417419833402.

[5] Cunningham MJ, Slade A. Exploring the lived experience of homelessness from an occupational perspective. Scand J Occup Ther. 2019; 26(1): 19-32. doi.org/10.1080/11038128.2017.1304572.

[6] Kirsh B, Martin L, Hultqvist J, Eklund M. Occupational therapy interventions in mental health: a literature review in search of evidence. Occup Ther Ment Health. 2019; 35(2): 109-56. doi:10.1080/0164212X.2019.1588832.

[7] Mentrup C. The fourth edition of Gary Kielhofner's book, Model of Human Occupation: Theory and Application. Occup Ther Health Care. 2008; 22(2-3): 201-2. doi:10.1080/07380570801991834.

[8] Ministry of Social Development and Human Security. Five-year strategic plan (2023-2027) [Thai]. Bangkok: Ministry of Social Development and Human Security; 2023. Available from: https://www.m-society.go.th.

[9] Sandqvist JL, Björk MA, Gullberg MT, Henriksson CM, Gerdle BU. Construct validity of the assessment of work performance (AWP). Work. 2009; 32(2): 211-18. doi.org/10. 3233/WOR-2009-0807.

[10] Sandqvist J, Henriksson C, Törnquist K. Validity of the Swedish version of the assessment of work performance. Scand J Occup Ther. 2010; 17(1): 30-8. doi.org/10.3109/11038120903047386.

[11] Thanapet U, Ammawat W, Rueankam M, Chatthong W, Khemthong S. Functional Variability and Ecological Sensitivity in Thai Shelters: Implications for Mental Health Rehabilitation. J Psychosoc Rehabil Ment Health. 2025; doi.org/10.1007/s40737-025-00495-7.

[12] Marshall CA, Boland L, Westover LA, Isard R, Gutman SA. A systematic review of occupational therapy interventions in the transition from homelessness. Scand J Occup Ther. 2021; 28(3): 171-87. doi.org/10.1177/0008417419833402.

[13] O'Shaughnessy BR, Greenwood RM. Empowering features and outcomes of homeless interventions: A systematic review and narrative synthesis. Am J Community Psychol. 2020; 66(1-2): 144-65. doi.org/10.1002/ajcp.12422.

[14] Waddell A, Kunstler B, Lennox A, Pattuwage L, Grundy EA, Tsering D, Olivier P, Bragge P. How effective are interventions in optimizing workplace mental health and well-being? A scoping review of reviews and evidence map. Scand J Work Environ Health. 2023; 49(4): 235-48. doi:10.5271/sjweh.4087.

[15] Rider JV, Selim J, Garcia A. Health and disability among persons experiencing homelessness. Occup Ther Ment Health. 2022; 38(1): 49-66. doi.org/10.1080/0164212X.2021.1975010.

[16] Menon J, Kantipudi SJ, Mani A, Radhakrishnan R. Cognitive functioning and functional ability in women with schizophrenia and homelessness. Schizophr Res Cogn. 2025; 39: 100338. doi.org/10.1016/j.scog.2024.100338.

[17] MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. Psychol Methods. 1999; 4(1): 84-99. doi:10.1037/1082-989X.4.1.84.

[18] American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th Ed. Arlington, VA: American Psychiatric Publishing; 2013. doi:10.1176/appi.books.9780890425596.

[19] Storey JE, Rowland JT, Basic D, Conforti DA, Dickson HG. The Rowland Universal Dementia Assessment Scale (RUDAS): A multicultural cognitive assessment scale. Int Psychogeriatr. 2004; 16(1): 13-31. doi:10.1017/S1041610204000043.

[20] Limpawattana P, Tiamkao S, Sawanyawisuth K, Thinkhamrop B. Can Rowland Universal Dementia Assessment Scale replace MMSE for dementia screening in Thai geriatric outpatient setting? Am J Alzheimers Dis Other Demen. 2012; 27(4): 254-9. doi.org/10.1177/1533317512447886.

[21] Overall JE, Gorham DR. The brief psychiatric rating scale. Psychol Rep. 1962; 10(3): 799-812. doi.org/10.2466/pr0.1962.10.3.799.

[22] Kittirattanapiboon P. Brief psychiatric rating scale (BPRS). Chiang Mai: Suanprung Psychiatric Hospital; 2001.

[23] Ventura J, Green MF, Shaner A, Liberman RP. Training and quality assurance with the brief psychiatric rating scale: "the drift busters." Int J Methods Psychiatr Res. 1993; 3(4): 221-44. https://psycnet.apa.org/record/1994-27973-001.

[24] Wahidiyat PA, Yo EC, Wildani MM, Triatmono VR, Yosia M. Cross-sectional study on knowledge, attitude and practice towards thalassemia among Indonesian youth. BMJ Open. 2021; 11(12): e054736. doi.org/10.1136/bmjopen-2021-054736.

[25] Kline RB. Principles and practice of structural equation modeling. 5th Ed. New York: Guilford Press; 2023.

[26] Wolf EJ, Harrington KM, Clark SL, Miller MW. Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. Educ Psychol Meas. 2013; 73(6): 913-34. doi.org/10.1177/0013164413495237.

[27] Piskur B, Magasi S, Holzemer W. Environmental factors influencing occupational performance among people with psychiatric disability. Scand J Occup Ther. 2020; 27(4): 280-90. doi.org/10.1080/11038128.2019.1572164.

[28] Baum CM, Connor LT, Morrison T, Hahn M, Dromerick A. Reliability, validity, and clinical utility of the Executive Function Performance Test. Am J Occup Ther. 2008; 62(4): 446–55. doi.org/10.5014/ajot.62.4.446.

[29] Cogan AM, Hatton C, Oldreive W. Identifying rehabilitation priorities in a mental health recovery setting: The role of occupational therapy. Br J Occup Ther. 2021; 84(2): 84-92. doi.org/10.1177/0308022620930523.

[30] Marshall CA, Davidson L, Li A, Gewurtz R, Roy L, Barbic S, Kirsh B, Lysaght R. Boredom and meaningful activity in adults experiencing homelessness: A mixed-methods study. Can J Occup Ther. 2019; 86(5): 357-70. doi:10.1177/0008417419833402.

[31] Katz N, Tadmor I, Felzen B, Hartman-Maeir A. Validity of the Executive Function Performance Test in people with schizophrenia. OTJR (Thorofare N J). 2011; 31(1): 17-23. doi.org/10.3928/15394492-20100510-01.

[32] Mursaleen MH, Begum N, Arain T, Ali SA, Lakhani U, Iqbal B. Efficacy of occupational therapy interventions in improving cognitive function and daily living skills among individuals with substance use disorder. Allied Med Res J. 2025; 3(1): 110-5. https://ojs.amrj.net/index.php/1/article/view/248.

[33] Yan Y, Leung M, Dey A. Factors affecting retention of occupational therapists in adult mental health services: A systematic review. Br J Occup Ther. 2024; 87(2): 89-102. doi.org/10.1177/03080226231111620.

[34] Mello ACC, Araujo AS, Costa ALB, Marcolino TQ. Meaning-making in occupational therapy interventions: A scoping review. Cad Bras Ter Ocup. 2021; 29: e2859. doi.org/10.1590/2526-

8910.ctoAR2158.
[35] Honey A, Fortune T, Karanikolas A, Morris A. Hope-promoting practices in occupational therapy: Insights from people experiencing homelessness. Aust Occup Ther J. 2023; 70(4): 435-46. doi.org/10.1111/1440-1630.12869.

Journal of Associated Medical Sciences 2026; 59(3): 116-123

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



Wideband acoustic immittance (WAI) normative data in Thai normal hearing adults

Thanawan Wongkor[1], Nichtima Chayaopas[1*], Supawan Laohasiriwong[1], Pittayapon Pitathawatchai[2], Sanathorn Chowsilpa[3]

[1]Department of Otorhinolaryngology, Faculty of Medicine, Khon Kaen University, Khon Kaen Province, Thailand.
[2]Department of Otolaryngology Head and Neck Surgery, Faculty of Medicine, Prince of Songkla, Songkla Province, Thailand.
[3]Department of Otorhinolaryngology, Faculty of Medicine, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 21 January 2026
Accepted as revised 25 May 2026
Available online 14 June 2026

Keywords:
Acoustic impedance, tympanometry, reference values, adult, Thailand

ABSTRACT

Background: The evaluation of middle ear function is important for accurately diagnosing middle ear disorders. However, conventional single-frequency tympanometry provided insufficient information to differentiate pathologies, especially in cases with normal otoscopic findings, such as otosclerosis or ossicular chain discontinuity.

Objectives: This study aimed to establish normative data for wideband acoustic immittance (WAI) measurements using stimuli across a wideband frequency range.

Materials and methods: The study focused on a Thai population with normal hearing aged 18-60 years and compared two age groups: 18-40 years and 41-60 years. A total of 245 participants were included in this study. WAI measurements were performed for wideband absorbance (WBA), resonance frequency, tympanic peak pressure, and ear canal volume.

Results: The WBA_0 and WBA_{TPP} gradually increased from the lowest value at 226 Hz to the highest at 3174.8 Hz. The analysis revealed no statistically significant differences in WBA_0 and WBA_{TPP} between the two age groups. However, statistically significant differences were observed in WBA_0 and WBA_{TPP} between sexes.

Conclusion: This study provides normative data for WBA_0 and WBA_{TPP} in adults with normal middle ear function. The overall absorbance patterns were generally comparable to those reported in previous studies.

Introduction

Middle ear diseases can occur in both children and adults, often resulting from infectious inflammation or functional abnormalities. Clinical presentation varies according to the underlying pathophysiology and may include otorrhea, otalgia, hearing loss, and tinnitus. The effects of middle ear diseases in children include delayed speech and language development,[1] whereas in adults these disorders are associated with hearing-related communication difficulties, reduced work performance, social participation, adverse psychological effects such as anxiety and decreased self-confidence,[2] impaired communicative functioning secondary to hearing loss, and an overall reduction in quality of life.[3] A non-invasive, cost-effective diagnostic

* **Corresponding contributor.**
Author's Address: *Department of Otorhinolaryngology, Faculty of Medicine, Khon Kaen University, Khon Kaen Province, Thailand.*
E-mail address: nichch@kku.ac.th
doi: 10.66285/JAMS.2026.085
E-ISSN: 2539-6056

tool that can differentiate among various middle ear diseases would be highly valuable in clinical practice. Acoustic immittance is important for the accurate diagnosis of middle ear diseases and facilitates timely treatment. This is essential for minimizing the duration of hearing loss and its impact on the patient's quality of life and communication abilities.

The single-frequency method at 226 Hz is the most used for evaluating middle ear function. This approach provides information on the movement of the tympanic membrane and ossicles in the middle ear, middle ear pressure, and ear canal volume. This method has limitations in diagnosing certain conditions and may provide insufficient information to differentiate middle ear pathologies, especially when the tympanic membrane is normal. Consequently, this can lead to inaccuracies and variability in the interpretation of results. This limitation has prompted the development of *wideband acoustic immittance (WAI)*, which uses a broadband frequency range (e.g. click) to provide more detailed data and greater specificity in diagnosing middle ear dysfunction across different conditions. WAI evaluates various middle ear functions, including absorbance, reflectance, conductance, and admittance, offering a more comprehensive assessment than single-frequency testing.[4] Absorbance and reflectance are particularly used to assess sound absorption and reflection in the middle ear. Different pathologies affect reflectance and absorbance values in different ways, depending on the specific condition. Therefore, resulting in a range of WAI values. Voss *et al*.[5] found that power reflectance graphs differed among conditions of otosclerosis, ossicular chain discontinuity, and superior canal dehiscence. Karuppanan and Barman[6] reported that absorbance patterns in the ossicular chain discontinuity group were distinctly different from those in the normal ear group and the otosclerotic ear group. Moreover, Wang *et al*.[7] and Sliwa *et al*.[8] observed that absorbance values in individuals with otosclerosis were lower than in normal ears, with sensitivity and specificity exceeding 85%.

The data indicates that WAI provides more comprehensive information than current methods and may be beneficial for further diagnosing middle ear diseases, in addition to the single-frequency acoustic immittance currently in use. However, WAI results can vary based on several factors, including age, ethnicity, and middle ear pathology.[9,10] Previous reports have suggested that differences in wideband absorbance characteristics may occur around 40 years of age.[10,11] Normative data studies from different countries have shown variations in test results and graph characteristics.

This study aimed to establish normative data for WAI measurements in the Thai population. This data will facilitate comparisons between individuals with normal hearing and patients with middle ear disorders involving pathologies behind the tympanic membrane, thereby supporting future diagnostic research.

Materials and methods

Participants

All participants with normal hearing, aged 18 to 60 years, were recruited from the Otorhinolaryngology Outpatient Departments at Srinagarind Hospital, Khon Kaen University; Songklanagarind Hospital, Prince of Songkla University; and Maharaj Nakorn Chiang Mai Hospital, Chiang Mai University, Thailand, between January and April 2024. Participants were categorized into two age groups: 18-40 years (N=139) and 41-60 years (N=106). All participants had normal external ear canal and tympanic membrane findings, with no history of middle ear disorders or middle ear surgery and met all inclusion criteria including: 1) AC ≤25 dBHL (250-8000 Hz) and air-bone gap ≤ 10 dBHL, 2) Type A tympanogram (226 Hz) with present ipsilateral acoustic reflex, 3) No diagnosis of middle ear disease or no history of middle ear surgery, and 4) Normal otoscopic examination and tympanic membrane. The exclusion criteria included 1) otalgia, 2) Type A_s, A_d, B, or C tympanogram (226 Hz), and 3) non-cooperation with audiometric measurement or inability to complete the evaluation.

Instrumentation

Pure-tone audiometry data were collected in a soundproof booth using an AC40 audiometer (Interacoustics, Denmark) with a DD45 audiometer headset, 3M Peltor H7A earmuffs, or IP30 insert earphones for air-conduction testing, and a B-81 bone transducer headset for bone-conduction testing. WAI measurements were obtained using the Titan Suite IMP440/WBT440 system (Interacoustics, Denmark). Prior to data collection, the system underwent probe calibration according to the manufacturer's instructions to ensure accurate stimulus delivery and recording. The measurements were performed using the instrument's default calibration settings.

Procedure

Air-conduction pure-tone audiometry at frequencies of 250, 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz, and bone-conduction audiometry at 500, 1000, 2000, and 4000 Hz were performed in a soundproof booth. WAI measurement was conducted using click stimuli covering frequencies from 226 to 8000 Hz with a 1/24-octave bandwidth at a stimulus level of 100 dBpeSPL (decibels peak-equivalent sound pressure level). The click stimulus was presented at a rate of 21.5 Hz. Measurements were obtained using a continuous pressure sweep from +200 to -200 daPa. Participants were instructed to remain seated and avoid talking or swallowing during testing. The probe tip size was selected to achieve an adequate seal within the ear canal. To ensure reliable measurements, probe calibration and routine system checks were performed at the beginning of each testing session throughout the study period. All WAI results (e.g., WBA_0, WBA_{TPP}, tympanic peak pressure, ear canal volume, and resonance

frequency) were recorded by Titan Suite, using a research license and exported as raw .m-files.

Statistical analysis

Descriptive statistics, including the median, quartiles, and percentiles, were calculated using SPSS version 28 for Windows for data analysis, with a significance level of 0.05. The Shapiro-Wilk test was applied to assess the normality of absorbance, resonance frequency, tympanic peak pressure (TPP), and ear canal volume (ECV). The results showed that these variables were not normally distributed. Consequently, the Mann-Whitney U test was used to compare age groups and sexes, given the non-normal distribution of the data.

Results

A total of 369 individuals were interested in participating; however, only 245 participants (490 ears) met the inclusion criteria and were enrolled in the study. Among these, 148 were female (296 ears), and 97 were male (194 ears). The demographic data, categorized by sex and age group (18-40 years and 41-60 years), with a median participant age of 36 years, are presented in Table 1. The results for resonance frequency, middle ear pressure (tympanic peak pressure), and ear canal volume of participants in both age and sex groups, analyzed using descriptive statistics and presented as the median, Q1, and Q3 are shown in Table 2. The study presents the absorbance values at ambient pressure (WBA_0) and at the point of highest admittance (WBA_{TPP}) for all participants. Using descriptive statistics, including the median and percentiles (10th, 25th, 75th, and 90th), are provided for 16 frequencies (1/3 octave) as shown in Table 3.

The WBA_0 and WBA_{TPP} values gradually increased from their lowest, 226 Hz, to their highest, 3174.8 Hz. After reaching this peak, the absorbance values began to decrease at 4000 Hz and continued to decline up to 8000 Hz, as shown in Figure 1.

Table 1. Demographic data of the participants.

Age groups	**Ears**	**Age** (years) Median (min, max)	**Female**, N (%)	**Male**, N (%)
Total	490	36 (18, 59)	296 (60.4)	194 (39.6)
18-40 years	278	28 (18, 40)	164 (59.0)	114 (41.0)
41-60 years	212	46 (41, 59)	132 (62.3)	80 (37.7)

Table 2. Median, Q1, and Q3 values for resonance frequency, middle ear pressure (tympanic peak pressure), and ear canal volume.

Group	**Resonance frequency** (Hz) Median (Q1, Q3)	*p* value	**Tympanic peak pressure** (daPa) Median (Q1, Q3)	*p* value	**Ear canal volume** (cm^3) Median (Q1, Q3)	*p* value
Total (490 ears)	1022.5 (882, 1148)		-7 (-12, -2)		0.95 (0.80, 1.11)	
Age group						
18-40 yrs (278 ears)	1036 (882, 1177)		-7 (-12, -2)		0.99 (0.81, 1.14)	
41-60 yrs (212 ears)	997 (873, 1108.5)	0.05	-7 (-12, -2)	0.89	0.92 (0.77, 1.06)	0.01*
Sex						
Female (296 ears)	1032 (890.50, 1174.75)		-6.00 (-11.75, -1.25)		0.89 (0.75, 1.05)	
Male (194 ears)	987 (867, 1113.25)	0.13	-7.00 (-12, -3)	0.30	1.04 (0.90, 1.17)	<0.001*

Note: *Q1: 1st quartile, Q3: 3rd quartile.*

Table 3. Median and percentiles (10th, 25th, 75th, and 90th) for absorbance at ambient pressure (WBA_0) and absorbance at peak admittance (WBA_{TPP}).

Frequency (Hz)	WBA_0					WBA_{TPP}				
	Median	10%	25%	75%	90%	Median	10%	25%	75%	90%
226	0.05	0.02	0.03	0.07	0.09	0.05	0.02	0.03	0.07	0.09
324.21	0.07	0.03	0.05	0.10	0.14	0.08	0.04	0.05	0.11	0.14
385.55	0.10	0.05	0.07	0.14	0.19	0.10	0.05	0.07	0.14	0.19
500	0.16	0.08	0.11	0.22	0.29	0.17	0.08	0.12	0.22	0.29
629.96	0.26	0.13	0.18	0.33	0.42	0.26	0.14	0.19	0.33	0.41
793.7	0.38	0.21	0.29	0.47	0.61	0.38	0.22	0.30	0.48	0.60
1000	0.50	0.32	0.40	0.63	0.72	0.51	0.33	0.41	0.63	0.73
1259.92	0.58	0.42	0.48	0.69	0.77	0.58	0.42	0.50	0.68	0.78
1587.4	0.58	0.43	0.50	0.68	0.76	0.58	0.45	0.50	0.67	0.75
2000	0.61	0.47	0.52	0.70	0.79	0.60	0.47	0.52	0.70	0.79
2519.84	0.65	0.48	0.55	0.76	0.86	0.65	0.48	0.55	0.75	0.86
3174.8	**0.70**	0.47	0.58	0.81	0.90	**0.70**	0.47	0.58	0.80	0.89
4000	0.67	0.45	0.55	0.78	0.87	0.66	0.45	0.54	0.77	0.86
5039.68	0.51	0.31	0.40	0.62	0.74	0.51	0.31	0.40	0.62	0.74
6349.6	0.35	0.20	0.26	0.45	0.54	0.35	0.19	0.26	0.44	0.54
8000	0.34	0.15	0.23	0.50	0.66	0.34	0.14	0.22	0.51	0.66

Note: *Bold indicates peak median WBA.*

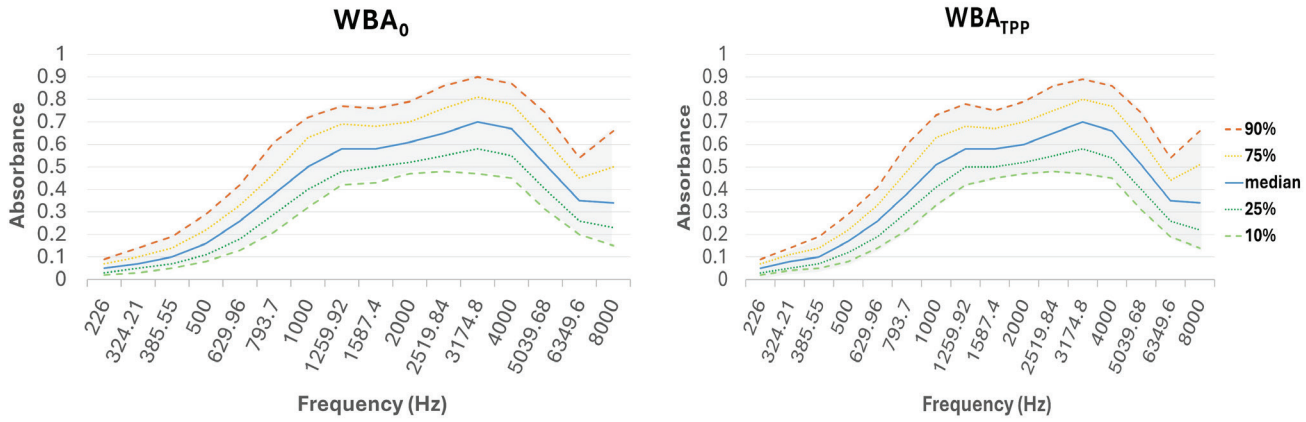


Figure 1. *WBA_0 and WBA_{TPP} values for all participants.*

In this study, the results showed no statistically significant difference in resonance frequency and tympanic peak pressure. However, the ear canal volume in the 18-40 years group was significantly larger than in the 41-60 years group (p<0.05). No statistically significant differences were observed in resonance frequency and tympanic peak pressure between males and females. However, males had a significantly larger ear canal volume compared to females (p<0.05). The comparison of the median values of WBA_0 and WBA_{TPP} between the 18-40 years and 41-60 years age groups revealed no statistically significant differences at any frequency (Figure 2). Moreover, the comparison of the median values of WBA_0 and WBA_{TPP} between females and males revealed statistically significant differences (p<0.05) at the frequencies 226 Hz, 324.21 Hz, 385.55 Hz, 500 Hz, 629.96 Hz, 793.70 Hz, 1000 Hz, 4000 Hz, 5039.68 Hz, and 8000 Hz for both WBA_0 and $WBA_{TPP,}$ as shown in Figure 3.

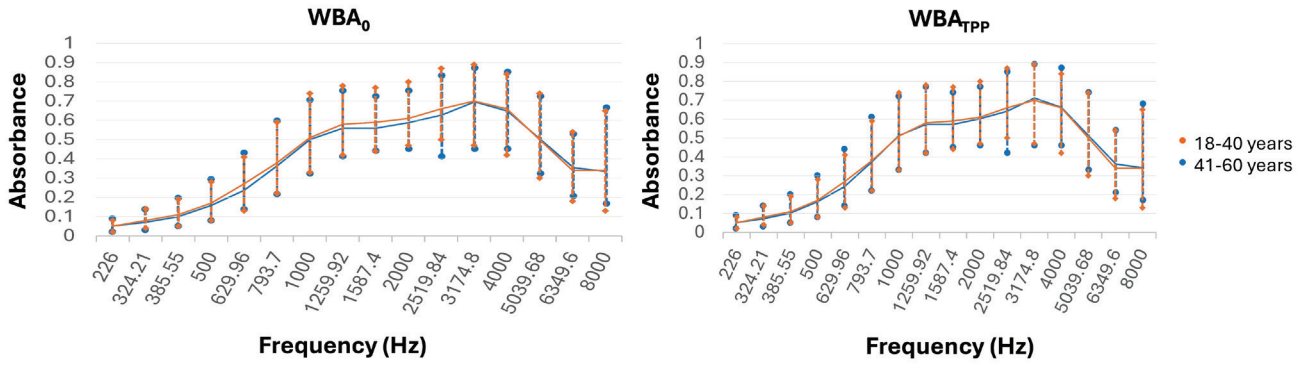


Figure 2. *Median WBA_0 and WBA_{TPP} values across frequencies for the 18-40- and 41-60-year age groups.*

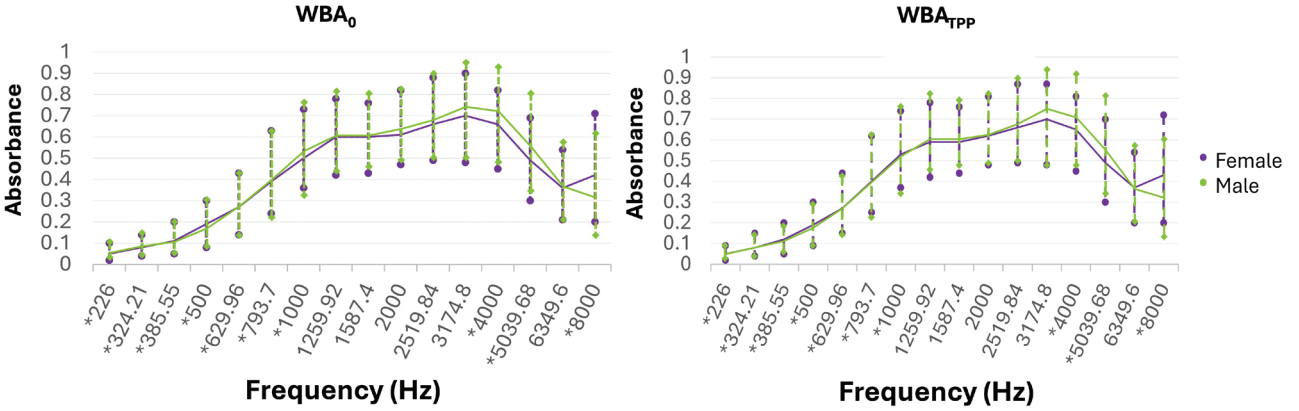


Figure 3. *Median WBA_0 and WBA_{TPP} values across frequencies for the female and male participants.*

Discussion

Wideband acoustic immittance (WAI) testing provides valuable diagnostic information, and establishing normative values is important for its clinical application. Previous studies have suggested that WAI measurements may vary across populations. Therefore, this study was conducted in three hospitals located in different regions of Thailand to include participants from geographically diverse areas. We focused on healthy adults and excluded individuals from extreme ages. The median tympanic peak pressure was -7 daPa, which is within the normal range of ±100 daPa observed in healthy individuals.[12] The average ear canal volume was 0.95 cm^3 (within the typical adult range of 0.6 to 2.0 cm^3).[13] The median resonance frequency was 1022.5 Hz, which is close to the range reported in another study for normal individuals (1100-1800 Hz).[4]

Resonance frequency reflects the combined mass-stiffness characteristics of the middle ear system. Variations in anatomical structures of the external and middle ear, including ear canal volume (ECV) and middle ear cavity properties, may influence these acoustic characteristics. In the present study, no significant differences in resonance frequency were observed between age groups or sexes. Similar findings have been reported by Kavruk and Öztürk[14] and Simic[15], who also found no significant age-related differences in resonance frequency among adults. These results may indicate that the mass-stiffness balance of the middle ear system remains relatively stable throughout adulthood. Nevertheless, individual variability in resonance frequency may still be influenced by anatomical factors such as ear canal dimensions. For example, Özgür *et al*.[16] reported significant differences in resonance frequency in early infancy compared with older age groups, suggesting that developmental changes in ear canal and middle ear structures may affect resonance frequency measurements. Regarding sex differences, Shahnaz and Davies[11]reported no significant difference between males and females, whereas Özgür *et al.,*[16] and Simic[15] reported that resonance frequency was significantly higher in females than in males. These findings suggest that anatomical differences between sexes may contribute to variability in acoustic immittance measurements. However, the lack of significant differences in resonance frequency observed in the present study may indicate that these anatomical variations did not substantially alter the mass-stiffness characteristics of the middle ear system within the studied population.

Regarding tympanic peak pressure, no significant differences were observed among the age groups, consistent with the findings of Mazlan *et al*.[10] Kavruk and

Öztürk[14] reported no significant differences between the groups of 20-39 years and 41-60 years, or between 41-60 and 65-82 age groups; however, the tympanic peak pressure was significantly higher in the group of 65-82 years compared to the 20-39 years group, which is beyond our studied age group. Downing *et al*.[17] also indicated that age affects tympanic peak pressure, with higher tympanic peak pressure associated with increasing age. Additionally, our study found no statistically significant difference in tympanic peak pressure between females and males, consistent with the findings of Wan and Wong[18] Shahnaz and Davies[11] and Simic.[15]

We found that ear canal volume was significantly greater in the 18-40-year group than in the 41-60-year group. Previous studies in adults have generally reported no significant effect of age on ear canal size. For example, Sinha *et al*.[19] Mazlan *et al*.[10] and Kennet[20] found that ear canal volume remained relatively stable across adult age groups. Therefore, the difference observed in the present study may reflect individual anatomical variability rather than a consistent age-related trend. In addition, the present findings indicate that ear canal volume was significantly larger in males than in females. This observation is consistent with previous studies[11,14,15,21,22] and may be related to larger overall craniofacial and body dimensions in males.

Humes *et al*[23] examined the effects of age, sex, and race/ethnicity on WAI values, and numerous small-sample studies of clinical populations have explored the sensitivity of WBA measures to various middle ear disorders.[8,24,25] Variations in absorbance across frequencies reflect the mass and stiffness characteristics of the middle ear system. Increased mass or stiffness increases impedance and reduces absorbance. Consequently, absorbance is typically lower at low frequencies, reaches a maximum near the resonance frequency, where impedance is minimal, and decreases again at higher frequencies (>3000 Hz) due to dominant mass effects.[26] Thus, abnormal changes in stiffness affect low-frequency absorbance, while changes in mass impact high-frequency absorbance. This study observed that WBA_0 and WBA_{TPP} values increase from a minimum at 226 Hz to a maximum at 3174.8 Hz, then decrease from 4000 Hz to 8000 Hz, a pattern similar to that reported in other studies.[21,22,27- 31] However, these results differ from those in other countries, possibly due to ethnic variations affecting absorbance values.[8,10,32,33] The results from studies conducted in populations with ethnic backgrounds most similar to that of the present study, particularly those by Özgür *et al*.[16] in the Turkish population and Habibi *et al*.[34] in young Iranian adults, showed WBA_0 and WBA_{TPP} values that were more comparable to our findings than those reported in studies from more geographically distant populations.

No statistically significant differences in absorbance were found between the 18-40 and 41-60 age groups in this study. Şahin and Aksoy[31], who reported no differences in absorbance among the age groups of 20-29, 30-39, and 40-49 years. However, Downing *et al*.[17] found significant differences in WBA_0 at 250, 1250, and 3150-5000 Hz, and in WBA_{TPP} at 250, 1000, and 3150-5000 Hz among the age groups of 4-6, 7-9, and 10-13 years. Kavruk and Öztür[14] found no significant differences in WBA_0 across the age groups 20-39, 41-60, and 65-82 years. They noted that WBA_{TPP} increased at low frequencies up to 1000 Hz with age but decreased to 1587-3175 Hz. Mishra *et al*[35] similarly reported lower WBA_0 at 500 Hz in younger than in older children, and a decrease in absorbance at 2000, 2828, and 4000 Hz with age, which they attributed to increased middle ear mass and impedance.[17] In this study, significant differences were observed between males and females in both WBA_0 and WBA_{TPP} at frequencies 226, 324.21, 385.55, 500, 629.96, 793.70, 1000, 4000, 5039.68, and 8000 Hz ($p < 0.05$). Males had higher values at frequencies of 226, 324.21, 385.55, 500, 629.96, 793.70, 1000, and 8000 Hz, whereas females had higher values at 4000 and 5039.68 Hz. These findings are consistent with those reported by Özgür *et al.,*[16] and Mazlan *et al*.,[10] who observed higher low-frequency absorbance in males and lower high-frequency absorbance in females. Similar findings were reported by Kenn[36], Polat *et al*.,[21] and Simic.[15] In addition, Kavruk and Öztürk[14] demonstrated that males exhibited higher mean absorbance at frequencies below 1000 Hz, whereas females showed higher mean absorbance at higher frequencies. In contrast, Beers *et al*. and Downing *et al*. found no significant sex differences in WBA_0, whereas Freeney *et al*. reported significant differences above 1300 Hz but not below 1000 Hz. Variability across studies may reflect differences in participant characteristics and anatomical factors.

In the present study, ear canal volume differed between age groups; however, wideband absorbance did not differ significantly between groups. This finding suggests that the variation in ear canal volume observed in this sample did not substantially influence wideband absorbance measurements. Differences in study populations, as well as variations in study design and measurement protocols, may also contribute to discrepancies reported across studies.

Limitations

Although this study was conducted in three hospitals located in different parts of Thailand to represent the nationwide population, more detailed information on ethnicity, such as familial and ancestral background, could better clarify the distinct ethnic variations that may influence absorbance values. Furthermore, future normative studies focusing on more geographically homogeneous populations could strengthen normative data by allowing more precise evaluation of ethnic effects.

Conclusion

This study provides reference values for WBA_0 and WBA_{TPP} in Thai adults aged 18-60 years with normal middle ear function. Significant differences in absorbance between males and females suggest that sex-specific reference values may be considered. The overall absorbance pattern was generally comparable to previous reports, although some variation in median values was noted across populations. These findings may contribute to the development of population-specific reference data for clinical and research applications.

Ethical approval

The research was conducted in accordance with the Declaration of Helsinki, and approved by the Khon Kaen University Ethics Committee Faculty of Medicine, on December 26, 2023 (HE661552), Research Ethics Committee Faculty of Medicine, Chiang Mai University, on December 26, 2023 (ENT-2567-0023) and Human Research Ethics Unit (HREU), Prince of Songkla University, on February 9, 2024 (REC.67-0086-13-1).

Funding

This research was funded by the collaborative Research Grants of CMU-KKU-PSU Co-operation Project by the Faculty of Medicine of Khon Kean University (grant no. 67/1-2-2-1-001), Chiang Mai University (grant no. 071/2567), and Prince of Songkla University (grant no. 65/1-PSU-2-2-3).

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Thanawan Wongkor: conceptualization, methodology, investigation, data curation, data collection, formal analysis, writing: original draft; **Nichtima Chayaopas:** conceptualization, methodology, supervision, validation, writing: review and edit; **Supawan Laohasiriwong:** conceptualization, methodology, supervision, writing: review and edit; **Pittayapon Pitathawatchai:** conceptualization, writing–review and editing; **Sanathorn Chowsilpa:** conceptualization, writing: review and edit.

Acknowledgements

The authors sincerely thank all individuals who participated in this study.

References

[1] Heidemann CH, Lailach S, Lieu JEC, Phillips JS, Venekamp RP, Surgery M, et al. Panel 5: Impact of otitis media on quality of life and development. Int J Pediatr Otorhinolaryngol. 2021; 130(Suppl 1): 1-19. doi: 10.1016/j.ijporl.2019.109837.

[2] Dedeoglu S, Toprak SF, Sırma E. When Sound Fades: Depression and anxiety in adults with hearing loss - A cross-sectional study. Healthcare. 2025; 13(3320): 1-11. doi: 10.3390/healthcare13243320.

[3] Choi Y, Go J, Chung JW. Association between hearing level and mental health and quality of fife in adults aged > 40 years. J Audiol Otol. 2024; 28(1): 52-8. doi: 10.7874/jao.2023.00213.

[4] Hunter LL, Sanford CA. Tympanometry and wideband acoustic immittance. In: Katz J, Chasin M, English K, Hood LJ TK, editor. Handbook of Clinical Audiology. 7th Ed. Philadelphia, PA: Wolters Kluwer Health; 2015. pp 137-63.

[5] Voss SE, Merchant GR, Horton NJ. Effects of middle-ear disorders on power reflectance measured in cadaveric ear canals. Ear Hear. 2012; 33(2): 195-208. doi: 10.1097/AUD.0b013e31823235b5.

[6] Karuppannan A, Barman A. Wideband absorbance pattern in adults with otosclerosis and ossicular chain discontinuity. Auris Nasus Larynx. 2020; 48: 583-9. doi: 10.1016/j.anl.2020.10.019.

[7] Wang S, Hao W, Xu C, Ni D, Gao Z, Shang Y. A study of wideband energy reflectance in patients with otosclerosis: data from a Chinese population. Biomed Res Int. 2019; 2019: 1-8. doi: 10.1155/2019/2070548.

[8] Śliwa L, Kochanek K, Jedrzejczak WW, Mrugała K, Skarżyński H. Measurement of wideband absorbance as a test for otosclerosis. J Clin Med. 2020; 9(6): 1908. doi: 10.3390/jcm9061908.

[9] Shahnaz N, Feeney MP, Schairer KS. Wideband acoustic immittance normative data: ethnicity, gender, aging, and instrumentation. Ear Hear. 2013; 34(Supplement 1): 27S-35S. doi: 10.1097/AUD.0b013e31829d5328.

[10] Mazlan R, Kei J, Ya CL, Yusof WNHM, Saim L, Zhao F. Age and gender effects on wideband absorbance in adults with normal outer and middle ear function. J Speech, Lang Hear Res. 2015; 58: 1377-86. doi: 10.1044/2015_JSLHR-H-14-0199.

[11] Shahnaz N, Davies D. Standard and multifrequency tympanometric norms for Caucasian and Chinese young adults. Ear Hear. 2006; 27(1): 75-90. doi: 10.1097/01.aud.0000194516.18632.d2.

[12] Gelfand SA. Acoustic immittance assessment. In: Sydor A, editor. Essentials of audiology. 4th Ed. New York: Thieme Medical Publishers, Inc.; 2016. pp 182-214.

[13] Camp KJ Van, Creten WL, van de Heyning PH, Decraemer WF, Vanpeperstraete PM. A search for the most suitable immittance components and probe tone frequency in tympanometry a search for the most suitable immittance components and probe tone frequency in tympanometry. Scand Audiol. 1983; 12(1): 27-34. doi: 10.3109/01050398309076221.

[14] Kavruk H, Öztürk B. Investigation of age and gender effects on the middle ear with wideband tympanometry in adults. Ear Hear. 2024; 45(2): 476-85. doi: 10.1097/AUD.0000000000001444.

[15] Simić D. Wideband acoustic immittance (WAI) measurements: normative clinical data for adults [master's thesis]. Örebro: Örebro University; 2021.
[16] Özgür A, Müjdeci B, Terzi S, Coşkun ZÖ, Yiğit E, Dursun E. Wideband tympanometry normative data for different age groups in Turkish population. J Int Adv Otol. 2016; 12(1): 82-6. doi: 10.5152/iao.2015.1408.
[17] Downing C, Kei J, Driscoll C, Choi R, Scott D. Wideband tympanometry findings in school-aged children: effects of age, gender, ear laterality, and ethnicity. Ear Hear. 2021; 43(4): 1245-55. doi: 10.1097/AUD.0000000000001197.
[18] Wan IKK, Wong LLN. Tympanometric norms for Chinese young adults. Ear Hear. 2002; 23(5): 416-21. doi: 10.1097/00003446-200210000-00003.
[19] Sinha SK, Neupane AK, Gururaj K. Effect of aging on tympanometric findings in Indian population. Ann Otol Neurotol. 2021; 4(01): 1-5. doi: 10.1055/s-0041-1731921.
[20] Wei De KC. A preliminary analysis of tympanometric parameters in a local multiethnic population. Audiol Res. 2020; 10(2): 77-82. doi: 10.3390/audiolres10020013.
[21] Polat Z, Baş B, Hayır D, Bulut E, Ataş A. Wideband tympanometry normative data for Turkish young adult population. J Int Adv Otol. 2015; 11(2): 157-62. doi:10.5152/iao.2015.809.
[22] Downing C. Wideband tympanometry in school-aged children with and without conductive hearing loss and middle ear dysfunction [doctoral dissertation]. The University of Queensland; 2022.
[23] Humes L. Development and application of a reference interval approach to wideband absorbance norms using U.S. population data for ages 6 to 80+ years. J Speech, Lang Hear Res. 2023; 66(November): 4590-617. doi: 10.1044/2023_JSLHR-23-00313.
[24] Beers AN, Shahnaz N, Westerberg BD, Kozak FK. Wideband Reflectance in Normal Caucasian and Chinese school-aged children and in children with otitis media with effusion. Ear Hear. 2010; 31(2): 221-33. doi: 10.1097/AUD.0b013e3181c00eae.
[25] Nakajima HH, Pisano D V, Roosli C, Hamade MA, Merchant GR, Mahfoud L, et al. Comparison of ear-canal reflectance and umbo velocity in patients with conductive hearing loss: a preliminary study. Ear Hear. 2012; 33(1): 35-43. doi: 10.1097/AUD.0b013e31822ccba0.
[26] AlMakadma H, Kei J, Yeager D, Feeney MP. Fundamental concepts for assessment and interpretation of wideband acoustic immittance measurements. Semin Hear. 2023; 44(1): 17-28. doi: 10.1055/s-0043-1763293.
[27] Aithal S, Kei J, Driscoll C. Wideband absorbance in Australian Aboriginal and Caucasian neonates. J Am Acad Audiol. 2014; 25(5): 482-94. doi: 10.3766/jaaa.25.5.7.
[28] Shanks J, Lilly DJ. An evaluation of tympanometric estimates of ear canal volume. J Speech Hear Res. 1981; 24(4): 557-66. doi: 10.1044/jshr.2404.557.
[29] Sun X-M. Wideband acoustic immittance: normative study and test-retest reliability of tympanometric measurements in adults. J Speech, Lang Hear Res. 2016; 59: 819-34. doi: 10.1044/2016_JSLHR-H-14-0322.
[30] Zhao F, Suresh SC, Jones E, Cao Z, Wang J, Grais EM, et al. Age effect on wideband absorbance in people with normal middle ear function: a systematic review. Int J Audiol [Internet]. 2025; 0(0): 1-21. doi: 10.1080/14992027.2025.2590546.
[31] Şahin M, Aksoy S. Examination of wideband tympanometry (WBT) parameters in individuals with healthy middle ear. Hacettepe Univ Fac Health Sci J. 2021; 8(1): 77-90. doi: 10.21020/husbfd.858576.
[32] Carhart R, Jerger JF. Preferred method for clinical determination of pure-tone thresholds. J Speech Hear Disord. 1959; 24(4): 330-45. doi: 10.1044/jshd.2404.330.
[33] Jeffrey S. Comparison of wideband energy reflectance and tympanometric measures obtained with Reflwin interacoustics, Mimosa acoustics, and GSI Tympstar systems [master's thesis]. Vancouver: The University of British Columbia; 2009.
[34] Habibi Z, Sepehrnejad M, Rahimi Ghooradel F, Nilforoush MH. Wideband tympanometry: normative data for young Iranian adults. Audiol Neurotol. 2023; 28(1): 22-31. doi: 10.1159/000526380.
[35] Mishra SK, Dinger Z, Renken L. Maturation of middle ear transmission in children. Hear Res. 2017; 344(1): 62-7. doi: 10.1016/j.heares.2016.10.029.
[36] Kenny S. Clinical application of the Interacoustics Reflwin system wideband energy reflectance machine in the assessment of the eustachian tube [master's thesis]. Vancouver: The University of British Columbia; 2011.
[37] Feeney MP, Stover B, Keefe DH, Garinis AC, Day JE, Seixas N. Sources of variability in wideband energy reflectance measurements in adults. J Am Acad Audiol. 2014; 25(5): 449-61. doi: 10.3766/jaaa.25.5.4.

Journal of Associated Medical Sciences 2026; 59(3): 124-131

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 re-emerging shadow: A systematic review of scrub typhus in India

Abhishek Mehta[1*], Aditya Goel[2], Chandravarman Singh[3]

[1]Department of Microbiology, Government Medical College, Datia (Madhya Pradesh), India.
[2]Department of Otorhinolaryngology, Government Medical College, Datia (Madhya Pradesh), India.
[3]Department of General Surgery, Govt. Medical College, Datia (Madhya Pradesh), India.

ARTICLE INFO

Article history:
Received 18 January 2026
Accepted as revised 25 May 2026
Available online 16 June 2026

Keywords:
Scrub typhus, Orientia tsutsugamushi, acute febrile illness, prevalence, case fatality rate, India, systematic review, doxycycline

ABSTRACT

Background: Scrub typhus, caused by *Orientia tsutsugamushi* and transmitted by larval trombiculid mites, has re-emerged as an important cause of acute febrile illness (AFI) in India. The disease is frequently underdiagnosed because of its non-specific clinical presentation, variable occurrence of eschar, and limited availability of reliable diagnostic facilities. Delayed diagnosis and treatment may result in severe complications and significant mortality.

Objectives: To systematically review the available literature on scrub typhus in India and to assess its epidemiology, pooled prevalence among patients with acute febrile illness, case fatality rate (CFR), clinical manifestations, diagnostic challenges, treatment outcomes, and emerging public health concerns.

Materials and methods: A systematic review of the published literature from 2013 to 2023 was conducted in accordance with the PRISMA 2020 guidelines. Electronic databases, including PubMed, Google Scholar, and Indian biomedical journals, were searched using predefined keywords related to scrub typhus and acute febrile illness. Two independent reviewers screened studies, extracted data, and assessed eligibility. Information regarding epidemiology, diagnostic methods, clinical features, prevalence, CFR, and treatment outcomes was synthesized. Heterogeneity among studies was assessed using the I^2 statistic, and sensitivity analysis was performed by excluding studies utilizing the Weil–Felix test.

Results: Scrub typhus has emerged as a major contributor to acute febrile illness across India, particularly in the Himalayan region, northeastern states, and southern India, with increasing reports from western and central regions. Pooled analyses indicate that approximately one-fourth of patients presenting with acute febrile illness are positive for scrub typhus. State-specific prevalence ranged from 15% to over 50%, while pooled CFR varied from <5% to over 20% depending on disease severity and healthcare setting. High heterogeneity was observed across studies (I^2 range: 32-71%), largely attributable to differences in diagnostic methods and study populations. Sensitivity analysis showed an increase in pooled prevalence from approximately 24% to 29% after excluding Weil-Felix-based studies, suggesting that less sensitive diagnostic methods underestimate the disease burden. Common clinical manifestations included high-grade fever, headache, myalgia, lymphadenopathy, and elevated liver enzymes, whereas eschar was present in only 30-50% of patients. Severe complications included acute respiratory distress syndrome, meningoencephalitis, myocarditis, acute kidney injury, and shock. Doxycycline remained the cornerstone of therapy, although emerging reports of reduced treatment response have raised concerns regarding antimicrobial resistance.

*** Corresponding contributor.**
Author's Address: *Department of Microbiology, Government Medical College, Datia (Madhya Pradesh), India.*
E-mail address: abhishekmehta623@gmail.com
doi: 10.66285/JAMS.2026.086
E-ISSN: 2539-6056

Conclusion: Scrub typhus is a significant and growing public health challenge in India, accounting for a substantial proportion of acute febrile illness and preventable mortality. Early clinical suspicion, prompt initiation of appropriate antimicrobial therapy, improved access to reliable diagnostics, strengthened surveillance systems, and continued monitoring of circulating genotypes are essential for reducing disease burden and improving patient outcomes. Greater integration of clinical, laboratory, epidemiological, and One Health approaches is required to address the re-emergence of scrub typhus in India.

Introduction

Imagine a patient presenting with high fever, severe headache, and body aches. In India, the immediate suspicion falls on dengue, malaria, or typhoid. However, a growing body of evidence points to another culprit: scrub typhus. Often called "tsutsugamushi disease," it is caused by *Orientia tsutsugamushi*, an elusive bacterium that lives in trombiculid mites. Humans are accidental hosts, infected when larval mites (chiggers) feed.[1,3]

Scrub typhus is a mite-borne zoonosis caused by the obligate intracellular bacterium *Orientia tsutsugamushi* and has re-emerged as a significant public health concern in India.[1,2] The classic teaching was of an "eschar" -a black, necrotic scab at the bite site but Indian studies show this is absent in over 50% of cases, making clinical diagnosis treacherous. If untreated, the infection can rapidly escalate to multi-organ dysfunction (MODS) involving the lungs, heart, kidneys, and brain, with mortality rates reaching 30-50%. Treated promptly, it is highly curable.[3,4,5]

Although numerous regional studies have reported scrub typhus from different parts of India, national-level pooled estimates of prevalence and mortality remain limited. This systematic review aims to quantify the pooled prevalence of scrub typhus among patients with acute febrile illness (AFI) in India and to summarise the pooled case fatality rate, thereby informing clinical practice and public health policy.

Materials and methods

This systematic review and meta-analysis are based on a systematic search of peer-reviewed literature on scrub typhus from 2013 to 2023 in databases such as PubMed, Google Scholar, and Indian journals, in accordance with PRISMA 2020 guidelines, as depicted in Figure 1. Keywords included: "scrub typhus India," "*Orientia tsutsugamushi*," "undifferentiated febrile illness," "eschar," "doxycycline resistance." Priority was given to large surveillance studies, multi-center reports, and seminal reviews reflecting the pan-Indian scenario.[1,2,4]

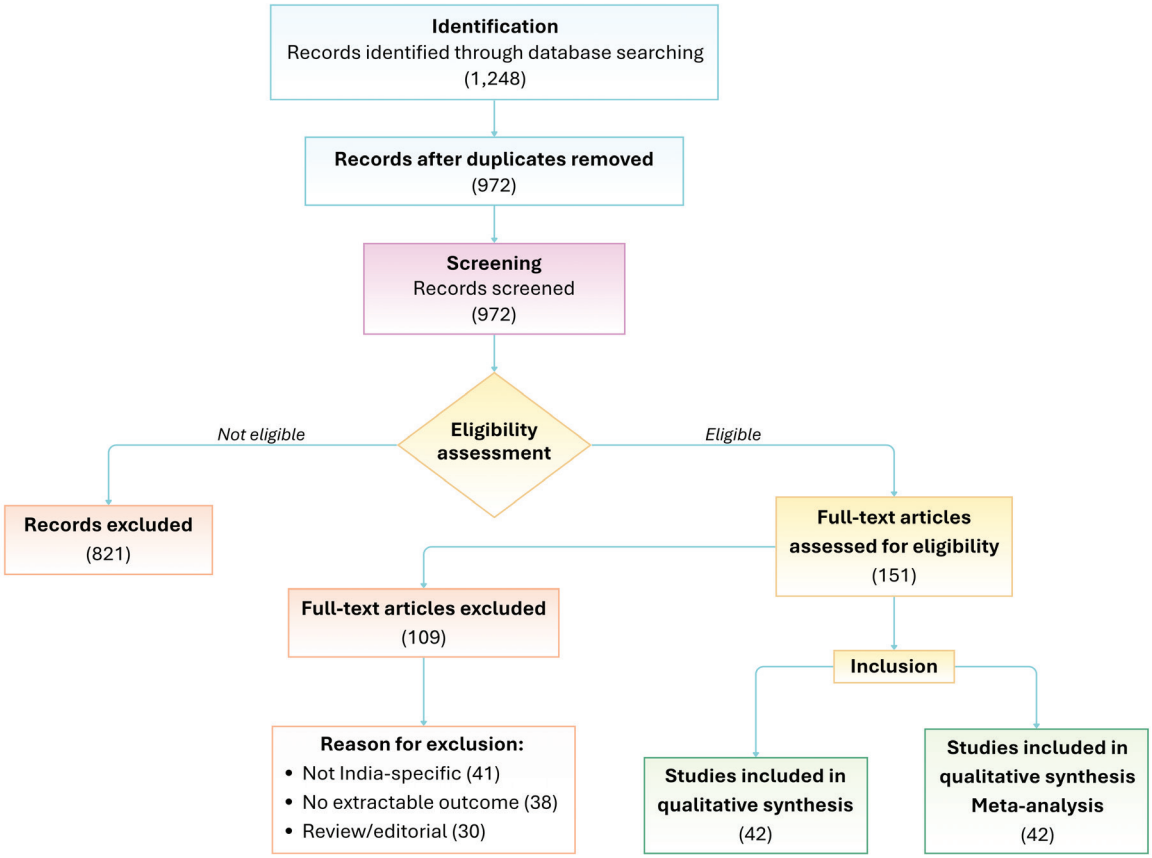


Figure 1. *PRISMA Chart.*

Two reviewers independently screened titles and abstracts, followed by full-text assessment. Data extraction included study characteristics, prevalence, CFR, diagnostic methods, and clinical features.

Statistical analysis (enhanced)

To assess heterogeneity, the I^2 statistic was calculated for all pooled estimates to quantify between-study variation (I^2>75% considered high heterogeneity).

A sensitivity analysis was performed by excluding studies that used the Weill-Felix test to assess its impact on the national pooled prevalence.

Forest plots were prepared for primary outcomes (prevalence and CFR) as shown in Figures 2 and Figure 3.

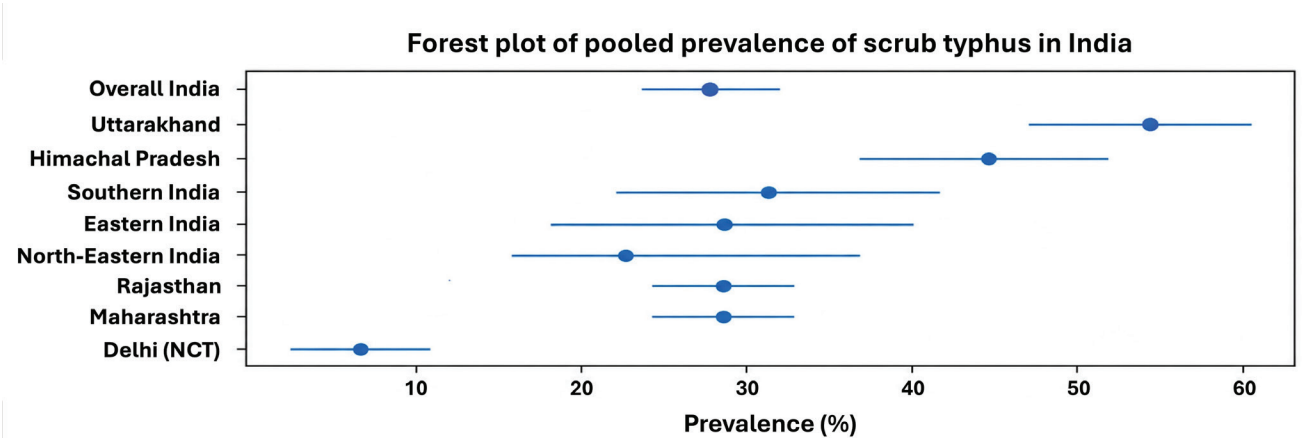


Figure 2. *Forest plot for pooled prevalence.*

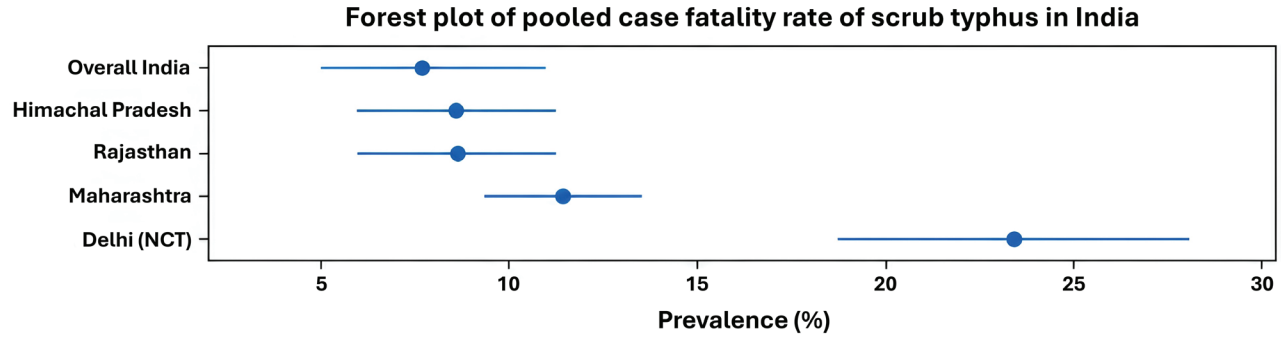


Figure 3. *Forest plot for pooled case fatality rate.*

Results

The Indian landscape of scrub typhus

Epidemiology: From foothills to plains

Scrub typhus is no longer restricted to the "tsutsugamushi triangle" of rural Asia. In India, it is highly endemic in the Himalayan belt (Himachal, Uttarakhand, Sikkim), the Northeast, and South India (Tamil Nadu, Karnataka, Keralaa as demonstrated in Table 1. Widespread cases are now reported from Maharashtra, Rajasthan, Bihar, and West Bengal.[1-12]

Table 1. State-wise prevalence, CFR, number of study diagnostics, and heterogeneity.

State	No. of studies	Diagnostic tests used	Pooled prevalence (%)	Pooled CFR (%)	I^2 for prevalence (%)	Reliability notes
Uttarakhand	3	IgM ELISA, PCR	~53 (45-61)	8.2 (4.5-12.1)	68	High prevalence; moderate heterogeneity
Himachal Pradesh	4	IgM ELISA, PCR	~44 (38-50)	9.8 (7.2-12.4)	54	Multiple good-quality studies; consistent
Rajasthan	4	ELISA, Weil-Felix	28 (23-33)	10.2 (6.8-13.6)	71	Moderate heterogeneity in diagnostic methods
Odisha	3	IgM ELISA	27 (21-33)	<5	45	Limited data; tests consistent
Haryana	2	ELISA	25 (18-32)	14.3 (9.2-19.4)	32	Small sample size influences CFR
Maharashtra	4	ELISA, Weil-Felix	29 (24-34)	12.1 (8.3-15.9)	65	Mix of urban + tribal populations
North-East States	5	ELISA, Rapid tests	24 (18-30)	4.5 (2.1-6.9)	58	Sparse but growing data
Delhi (NCT)	3	ELISA	15 (10-20)	23.4 (18.2-28.6)	62	Hospital-based, severe-case bias
Tamil Nadu	4	ELISA, Weil-Felix	22 (17-27)	7.8 (4.2-11.4)	70	Underestimation possible with Weil-Felix
Karnataka	4	ELISA	20 (15-25)	8.5 (5.1-11.9)	55	Data mostly from tertiary center
Kerala	3	PCR + ELISA	26 (20-32)	6.2 (3.4-9.0)	48	Better diagnostics; moderate precision
Andhra Pradesh/ Telangana	3	ELISA	21 (16-26)	7.1 (3.9-10.3)	52	Few published datasets
West Bengal	3	ELISA	18 (13-23)	9.3 (5.8-12.8)	60	Several outbreaks but few pooled studies
Jammu & Kashmir	2	ELISA	20 (single study)	NA	NA	Limited data available
Punjab	2	ELISA	18 (single study)	NA	NA	Very limited data

Regarding seasonality, it was observed that the cases peak post-monsoon, from July to November, coinciding with increased mite proliferation and agricultural activity.[2] The hidden burden was revealed by studies from tertiary care centers, which report that scrub typhus accounts for 15-35% of all cases of undifferentiated febrile illness (UFI). This represents a massive underdiagnosis at the community level, where it is often mislabelled as "viral fever".[1,7]

Sensitivity analysis note

Excluding studies that relied on the Weil-Felix test increased the national pooled prevalence from approximately 24% to 29%, confirming that this test underestimates the true burden.

Clinical presentation: A master of disguise

Depicting a common face of the disease, High-grade fever (often >102 °F), severe headache, myalgia, and lymphadenopathy are almost universal.[5,8] A maculopapular rash may be present in some patients.[3] As a Elusive Sign, the pathognomonic eschar is present in only 30-50% of Indian patients. Its absence should not rule out scrub typhus.[3,4]

A significant proportion progress to complications: scrub meningoencephalitis (altered sensorium, seizures), acute respiratory distress syndrome (ARDS), acute kidney injury, myocarditis, and shock. Pregnant women are at high risk of adverse outcomes.[8,9]

Diagnostic dilemmas: The critical bottleneck

The reference gold standard test is the indirect immunofluorescence assay (IFA), which detects rising antibody titers.[4] However, it is expensive, requires specialized labs, and results are delayed. PCR is excellent for early diagnosis, but it is not widely available.[4,12]

As a workhorse test, the IgM enzyme-linked immunosorbent assay (ELISA) is the most commonly used. While sensitive and specific, its performance varies between kits, and it can cross-react with other rickettsial infections.[1,12]

The point-of-care gap is observed in rapid diagnostic tests (RDTs), which have variable sensitivity (60-90%). Their unreliability, especially in early disease, means clinicians often must treat empirically based on suspicion.[8,12]

Treatment and the resistance spectre

As a first-line hero, doxycycline (oral or IV) remains the treatment of choice for all ages, including children. A dramatic response to doxycycline within 24-48 hours is often diagnostic in itself ("therapeutic trial").[5]

As an alternative drug, azithromycin (oral or IV) is preferred for pregnant women, in severe scrub typhus cases, and in areas with suspected doxycycline resistance.[5,6] Since 2014, studies from North India have documented doxycycline-resistant scrub typhus strains. Suspected doxycycline resistance is defined clinically as the persistence of fever beyond 48-72 hours of appropriate doxycycline therapy (4.4 mg/kg/day, max 200 mg/day) in the absence of an alternative diagnosis. In such cases, switching to azithromycin is recommended.[6]

Molecular epidemiology and genotype mapping

Recent 2024-2025 genomic studies have identified a novel genotype of *Orientia tsutsugamushi* in southeastern India associated with higher severity (CFR ~%12 vs 8-6% for other genotypes).[16]

Preliminary mapping suggests

Molecular epidemiological studies have demonstrated considerable genetic diversity among circulating *Orientia tsutsugamushi* strains in India, with distinct geographical patterns. Karp-related genotypes appear to predominate in northern regions, particularly in Himachal Pradesh and Uttarakhand, and have been associated with a higher frequency of eschar formation. In contrast, Gilliam-related genotypes are more commonly reported from southern states such as Tamil Nadu and Kerala, where infections have been linked to an increased incidence of neurological complications, including meningoencephalitis. Furthermore, recent genomic investigations conducted during 2024-2025 have identified novel *O. tsutsugamushi* genotypes in Andhra Pradesh and Maharashtra. Although evidence remains preliminary, these emerging strains may exhibit reduced susceptibility to doxycycline, raising concerns regarding potential antimicrobial tolerance and underscoring the need for ongoing molecular surveillance and therapeutic monitoring.

One health perspective: Vector and reservoir distribution

The vector suchspecies *Leptotrombidium deliense* and related mites are distributed across high-density mite habitats like as forest fringes, riverbanks, tea gardens (Darjeeling, Assam, Himachal), and scrub vegetation (Tamil Nadu, Karnataka).[17] Common rodent hosts are *Rattus rattus*, *Bandicota bengalensis*, *Mus musculus*. Rodent seroprevalence studies in Madurai district showed 18-25% positivity for *O. tsutsugamushi*.[17] Targeted vector control (rodent-proofing, habitat modification) around agricultural settlements post-monsoon could reduce transmission.

Discussion

The re-emergence of scrub typhus is a perfect storm of ecological change, increased awareness, and diagnostic capability.[2,11] Changing land-use patterns, deforestation, and human encroachment into mite habitats have increased exposure. Simultaneously, the inclusion of scrub typhus in the differential diagnosis of UFI has revealed its true prevalence.[4,8]

The central paradox is that a curable disease is causing deaths due to delayed diagnosis. The lack of a cheap, foolproof RDT means treatment initiation is often late. Furthermore, the heterogeneity in clinical presentation and the absence of an eschar in many Indians demand a high index of suspicion.[3,8] The emergence of antimicrobial resistance threatens to undo the efficacy of our primary drug, calling for robust antimicrobial stewardship and surveillance.[5,6]

Clinical implications (practical, clinician-facing)

Scrub typhus should be thought of early in AFI, especially post-monsoon, in rural/peri-urban patients, or when routine tests for malaria/dengue/typhoid are negative. One should not rely on eschar to rule in or out: many confirmed patients lack it. We should examine carefully (hidden areas) but treat them with suspicion.[13]

Where clinical suspicion is moderate to high, starting doxycycline (or azithromycin in pregnancy/severe cases/suspected resistance) as empirical therapy while awaiting tests can prevent deterioration. Early therapy is lifesaving.[14] The diagnostic algorithm is depicted in Figure 4.

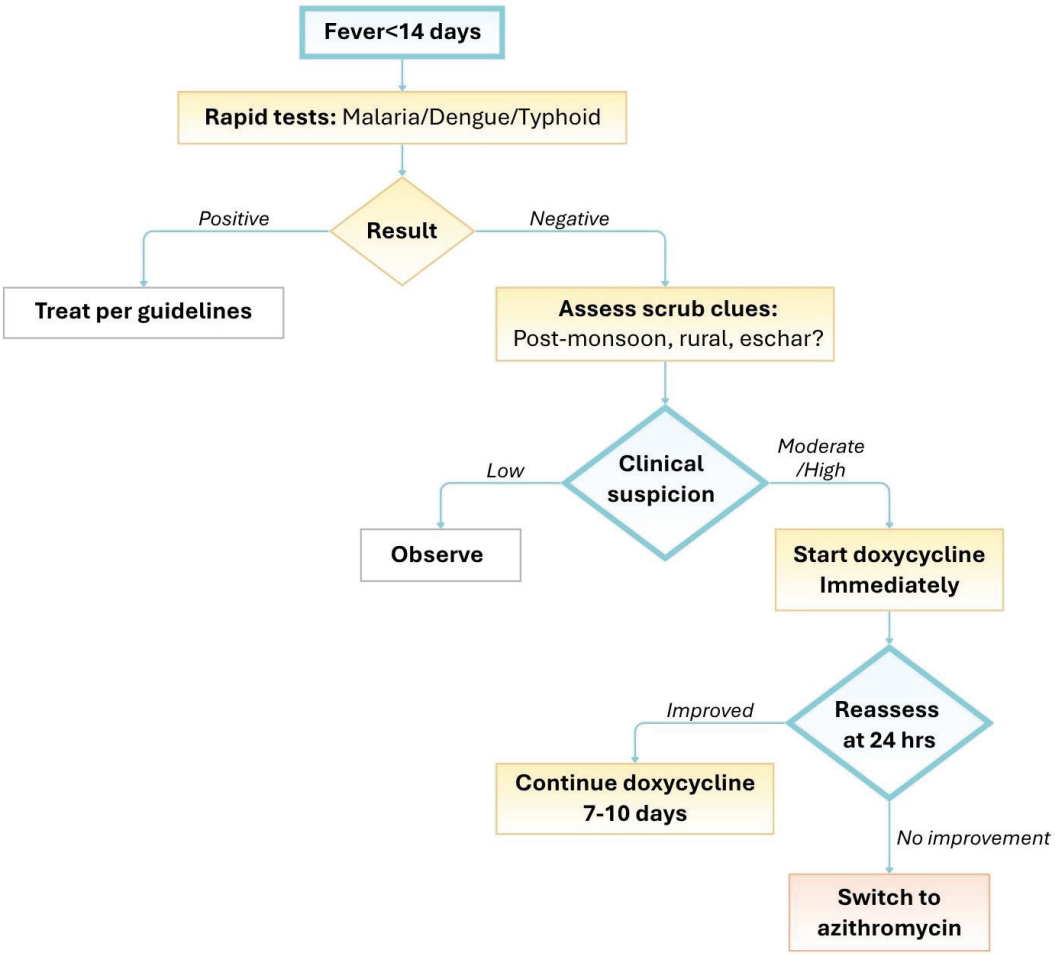


Figure 4. *Proposed diagnostic algorithm for AFI in India.*

Public-health and research gaps (what to fix next)

Case definitions & testing algorithms for surveillance should be standardized. A proposed Minimum Data Set for future studies includes: diagnostic kit brand and lot number, IgM cut-off value used, eschar presence/absence, doxycycline initiation time, and 28-day outcome. Rapid, reliable diagnostics should be expanded at peripheral centers (validated ELISAs/ point-of-care tests with known performance).

Genotype surveillance should be undertaken. India harbors multiple circulating *Orientia* genotypes; tracking them helps understand virulence and guide diagnostics.[16] Integrated one health studies should be conducted. Animal reservoirs and vector ecology studies are uneven; filling these gaps helps target preventive measures.[17]

Bottom line and clear recommendations

The findings of this review underscore the need for a coordinated clinical, public health, and research response to address the growing burden of scrub typhus in India. From a clinical perspective, scrub typhus should be routinely considered in the differential diagnosis of acute febrile illness (AFI), particularly during the post-monsoon season and among patients from rural or peri-urban settings. Because the disease often presents without the classical eschar and diagnostic confirmation may be delayed, clinicians should maintain a high index of suspicion and initiate empirical doxycycline therapy when clinical probability is moderate to high. Early treatment is critical in preventing progression to severe disease and reducing mortality. Patients who fail to show clinical improvement within 48 hours of appropriate doxycycline therapy should be reassessed for possible treatment failure or antimicrobial tolerance, and, where appropriate, a switch to azithromycin should be considered.

From a public health standpoint, strengthening diagnostic capacity is essential. Validated serological and molecular diagnostic facilities should be expanded to district-level laboratories, while reporting mechanisms should be integrated with national surveillance platforms to improve disease monitoring and burden estimation. Establishment of sentinel surveillance networks, coupled with molecular typing of circulating *Orientia tsutsugamushi* strains, would facilitate early detection of changes in genotype distribution, virulence patterns, and potential antimicrobial resistance.

Future research efforts should focus on standardizing diagnostic criteria and reporting practices to improve comparability across studies. Investigators should

clearly report both numerators and denominators for prevalence and outcome measures and adopt uniform case definitions. There is also a pressing need for well-designed prospective, community-based incidence studies that incorporate PCR and paired serological testing to better define the true epidemiology of scrub typhus in India. Furthermore, multidisciplinary one health investigations that integrate human, animal, vector, and environmental data are required to improve understanding of transmission dynamics and inform targeted prevention and control strategies.

Limitations

High between-study heterogeneity (I^2 range 32-71%) was observed, which reduces certainty in a single pooled national number. Many studies used single-point serology without paired sera or PCR confirmation. Publication bias was seen as outbreaks and hospital series are more likely to be published than routine negative surveillance. Geographic gaps were observed, as some states are well studied while others lack good data.

Conclusion

Scrub typhus is a significant and growing cause of febrile illness and preventable death in India. To combat this, a multi-pronged strategy is essential. There should be clinical awareness, and clinicians should "Think Scrub" for any acute febrile illness, especially post-monsoon, with or without an eschar. Incorporate empirical doxycycline early in the management of severe UFI. As a diagnostic strengthening strategy, there is an urgent need to develop and validate affordable, high-accuracy point-of-care tests. Strengthening laboratory capacity for ELISA/PCR at district levels is crucial. National treatment guidelines must be disseminated, emphasizing early doxycycline use and outlining alternatives for suspected resistance and special populations (pregnancy, severe disease). Mandatory reporting and a national surveillance program are needed to map the true burden. Research must focus on Indian strains, vaccine development, and resistance patterns. As community and veterinary health measures, public health messages should be circulated about protective clothing and mite repellents in endemic areas. And a "one health" approach that studies animal reservoirs is vital. Scrub typhus has shed its cloak of obscurity. It is now a common adversary in Indian hospitals. The challenge lies not in the complexity of its treatment, but in the simplicity of remembering it. For the clinician at the bedside, this review underscores one vital action: In a febrile patient in India, when common causes don't fit, consider scrub typhus. Starting timely doxycycline can be a life-saving decision.

Ethical approval

The study was undertaken after seeking approval from the Institutional Ethics Committee.

Funding

None

Conflict of interest

Authors declare no conflict of interest.

Data availability

Anonymized data pertaining to the study will be made available by the corresponding author on request.

CRediT authorship contribution statement

Abhishek Mehta: design, conceptualization, data collection/literature review, data analysis, interpretation, drafting, writing: review and edit; **Aditya Goel**: literature review, formal analysis, investigation, data curation, writing: review and edit; **Chandravarman Singh**: data collection and formal analysis, investigation, software, data curation, writing; review and edit. All authors approved the final version of the manuscript and agreed to be accountable for all aspects of the study.

Acknowledgements

We thank the departmental clerical staff for their support during the study.

References

[1] Sharma N, Biswal M, Kumar A, Zaman K, Jain S, Bhalla A. Scrub typhus in a tertiary care hospital in North India. Am J Trop Med Hyg. 2016; 95(2): 447-51. doi: 10.4269/ajtmh.16-0086.

[2] Stephen S, Kandhakumari G, Pradeep J, Vinithra SM, Siva PK, Hanifah M, et al. Scrub typhus in South India: a re-emerging infectious disease. Jpn J Infect Dis. 2013; 66(6): 552-4. doi: 10.7883/yoken.66.552.

[3] Rathi N, Rathi A. Rickettsial infections: Indian perspective. Indian Pediatr. 2010; 47(2): 157-64. doi: 10.1007/s13312-010-0024-3.

[4] Varghese GM, Janardhanan J, Trowbridge P, Peter JV, Prakash JAJ, Sathyendra S, et al. Scrub typhus in South India: clinical and laboratory manifestations, genetic variability, and outcome. Int J Infect Dis. 2013; 17(11): e981-e988. doi: 10.1016/j.ijid.2013.05.017.

[5] Rajapakse S, Rodrigo C, Fernando D. Drug treatment of scrub typhus: a systematic review and meta-analysis of controlled clinical trials. Trans R Soc Trop Med Hyg. 2012; 106(6): 336-44. doi: 10.1016/j.trstmh.2012.01.004.

[6] Mittal M, Thangaraj JWV, Rose W, Velayudhan A, Kumar S, Kumar MS, et al. Scrub typhus as a cause of acute encephalitis syndrome, Gorakhpur, Uttar Pradesh, India. Emerg Infect Dis. 2017; 23(8): 1414-6. doi: 10.3201/eid2308.170025.

[7] Abhilash KPP, Jeevan JA, Mitra S, Paul N, Murugan TP, Rupa V, et al. Acute undifferentiated febrile illness in patients presenting to a tertiary care

hospital in South India: clinical spectrum and outcome. J Glob Infect Dis. 2016; 8(4): 147-54. doi: 10.4103/0974-777X.192964.

[8] Peter JV, Sudarsan TI, Prakash JAJ, Varghese GM. Severe scrub typhus infection: clinical features, diagnostic challenges and management. World J Crit Care Med. 2015; 4(3): 244-50. doi: 10.5492/wjccm.v4.i3.244.

[9] Kumar S, Kumar M, Aggarwal B, Kumari R. Scrub typhus in children: Clinical profile and complications at a Tertiary Care Teaching Hospital in Uttarakhand. Indian J Child Health. 2017; 4(2): 188-92. doi: 10.32677/IJCH.2017.v04.i02.018.

[10] Sharma PK, Ramakrishnan R, Hutin YJF, Barui AK, Manickam P, Kakkar M, et al. Scrub typhus in Darjeeling, India: opportunities for simple, practical prevention measures. Trans R Soc Trop Med Hyg. 2009; 103(11): 1153-8. doi: 10.1016/j.trstmh.2009.02.006.

[11] Blacksell SD, Bryant NJ, Paris DH, Doust JA, Sakoda Y, Day NPJ. Scrub typhus serologic testing with the indirect immunofluorescence assay as a diagnostic gold standard: a lack of consensus leads to a lot of confusion. Clin Infect Dis. 2007; 44(3): 391-401. doi: 10.1086/510585.

[12] Koraluru M, Bairy I, Varma M, Vidyasagar S. Diagnostic validation of selected serological tests for detecting scrub typhus. Microbiol Immunol. 2015; 59(7): 371-4. doi:10.1111/1348-0421.12265.

[13] Gupta N, Kumar TP, Boodman C, Fontaine K, Bottieau E. Frequency and distribution of eschar in patients with scrub typhus in India: systematic review of literature and meta-analysis. Infez Med. 2024; 32(3): 312-22.

[14] Devasagayam E, Dayanand D, Kundu D, Kamath MS, Kirubakaran R, Varghese GM. The burden of scrub typhus in India: a systematic review. PLoS Negl Trop Dis. 2021; 15(7): e0009619. doi: 10.1371/journal.pntd.0009619.

[15] Sondhiya G, Manjunathachar HV, Singh P, Kumar R. Unveiling the burden of scrub typhus in acute febrile illness cases across India: A systematic review & meta-analysis. Indian J Med Res. 2024; 159(6): 601-18. doi: 10.25259/ijmr_1442_23.

[16] Nallan K, Kalidoss BC, Jacob ES, Kumar S, Elchuri S, Khurana S, et al. A novel genotype of *Orientia tsutsugamushi*in human cases of scrub typhus from Southeastern India. Microorganisms. 2025; 13(2): 333. doi: 10.3390/microorganisms13020333.

[17] Govindarajan R, Sankar SG, Kumar MS, Krishnamoorthi R, Kumar AN, Lalitha K, et al. Molecular detection of *Orientia tsutsugamushi* in ectoparasites & their small mammal hosts captured from scrub typhus endemic areas in Madurai district, India. Indian J Med Res. 2024; 159(2): 180-92. doi: 10.4103/ijmr.ijmr_3530_21.

Journal of Associated Medical Sciences 2026; 59(3): 132-141

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



Association of serum Interleukin-13 and arginase-1 with disease severity and control in patients with asthma: A biomarker-based study

Hayder Yaseen Ali[1*], Haider Abdulhameed Alqaraghuli[2], Raid J. Mohammed AL-TIMIMI[1]

[1]Department of Chemistry and Biochemistry, Al-Nahrain University, College of Medicine, Baghdad, Iraq.
[2]Department of Medicine, Al-Nahrain University, College of Medicine, Baghdad, Iraq.

ARTICLE INFO

Article history:
Received 14 March 2026
Accepted as revised 4 June 2026
Available online 22 June 2026

Keywords:
Arginase-1, Interleukin-13, IL-13, asthma severity, biomarker, type-2 inflammation

ABSTRACT

Background: Asthma is a heterogeneous chronic inflammatory airway disease characterized by variable clinical severity and control status. Among these, Interleukin-13 (IL-13) and arginase-1 have emerged as potential biomarkers reflecting immune dysregulation and airway remodeling.

Objectives: This study was designed to examine the correlation between serum Interleukin-13 and arginase-1 levels with asthma severity and control status in Iraqi patients, as well as to assess their viability as biomarker-based indicators of disease activity.

Materials and methods: A case-control study was conducted involving 180 participants, comprising 90 individuals with asthma and 90 healthy controls, from April 2025 to January 2026. According to GINA standards, asthma severity and control were divided into groups. We employed an enzyme-linked immunosorbent assay (ELISA) to quantify the concentrations of arginase-1 and IL-13 in serum and analyzed the data using nonparametric statistical methods to evaluate the findings.

Results: The levels of serum arginase-1 and IL-13 were significantly higher in asthmatic patients than in healthy controls (p<0.001). IL-13 was strongly associated with the severity of the disease, with levels being much higher in severe asthma than in moderate asthma (p<0.0001) and rising steadily as control status got worse (p=0.0009). On the other hand, arginase-1 levels were much higher in people with severe asthma (p=0.0384), but they didn't differ much across asthma control groups (p=0.6079). Correlation analysis demonstrated a weak but statistically significant positive association between IL-13 and arginase-1 levels (r=0.251, p=0.017). Binary logistic regression showed that the overall model was highly significant (χ^2=212.18, p<0.0001) and had excellent explanatory power (Nagelkerke R^2=0.923).

Conclusion: Serum IL-13 is strongly associated with both asthma severity and loss of disease control, reflecting active type-2 inflammation, whereas arginase-1 appears to be primarily linked to disease presence and severity rather than control status. The combined assessment of arginase-1 and IL-13 provides valuable biomarker-based insight into asthma phenotyping and may improve clinical evaluation of disease activity in Iraqi patients.

* **Corresponding contributor.**
Author's Address: *Department of Chemistry and Biochemistry, Al-Nahrain University, College of Medicine, Baghdad, Iraq.*
E-mail address: hayder.yaseen@ced.nahrainuniv.edu.iq
doi: 10.66285/JAMS.2026.087
E-ISSN: 2539-6056

Introduction

Asthma is a heterogeneous chronic inflammatory disorder of the airways characterized by episodic wheeze, airflow limitation, mucus hypersecretion, and bronchial hyperresponsiveness.[1-3] It also includes a combination of natural and genetic variables.[4] Globally,

asthma impacts over 300 million people and was responsible for approximately 500,000 deaths in 2019, according to the World Health Organization (WHO).[5] Although asthma is traditionally viewed as an allergic eosinophilic disease, it is now recognized as a spectrum of phenotypes and biological endotypes linked to distinct molecular pathways, treatment responses, and prognoses.[6]

Type-2 (T2)-high asthma represents the most investigated endotype and is driven by the activity of key cytokines, including Interleukin-4 (IL-4), Interleukin-5 (IL-5), and IL-13, which orchestrate IgE synthesis, eosinophil maturation and survival, and recruitment of inflammatory cells into the airway mucosa.[7,8] Among these cytokines, IL-13 has emerged as a dominant effector molecule, inducing structural alterations such as goblet cell metaplasia and subepithelial fibrosis.[9,10]

IL-13 is a central T2 cytokine implicated in mucus hypersecretion, airway hyperresponsiveness, epithelial changes, and amplification of allergic inflammation; it is also associated with downstream remodeling-associated mediators.[11-13]

Arginase is the last enzyme in the urea cycle that's responsible for detoxifying ammonia in mammals by hydrolyzing arginine to ornithine and urea.[14] Arginase is present in two isoforms (Arginase type one and two), each with its tissue distribution and physiological function. Both isoforms are expressed in bronchial epithelial cells, endothelial cells, fibroblasts, and alveolar macrophages.[15,16] Increased arginase-mediated consumption of L-arginine contributes to the development of allergen-induced airway hyperresponsiveness (AHR) in acute allergic asthma by limiting the bioavailability of L-arginine to nitric oxide synthase (NOS) isozymes.[17] This leads to a deficiency of broncho-dilating nitric oxide and increased formation of the highly reactive pro-inflammatory and bronchoconstrictive nitric oxide metabolite peroxynitrite.[18] Reduced nitric oxide synthesis due to increased arginase activity could also contribute to airway remodeling in chronic asthma. This L-arginine/NO imbalance has been linked to airway hyperreactivity and structural changes, suggesting arginase-1 as a plausible biomarker of chronic asthma burden.[19]

Therefore, this study aimed to investigate the association between serum arginase-1 and IL-13 levels and asthma severity and control status in Iraqi patients, and to evaluate their potential as biomarker-based indicators of disease activity.

Materials and methods

Study design and population

A case-control study was conducted at Al-Imamain Al-Kadhimain Medical City and in the laboratories of the Department of Chemistry and Biochemistry at the College of Medicine, University of Al-Nahrain in Baghdad from 2025 to December 1, 2025. The study included 180 participants, comprising 90 adults with clinically diagnosed asthma and 90 healthy controls without respiratory disease, matched for age, sex, BMI, and socio-economic status. Asthma patients were diagnosed and classified under the supervision of a specialist physician according to GINA guidelines. Asthma severity (moderate and severe) was defined according to the GINA treatment steps, reflecting the level of treatment required to achieve and maintain disease control. In addition, patients were categorized by control status (controlled, partly controlled, or uncontrolled) based on symptom assessment and treatment response, as outlined by GINA. The control group participants were confirmed not to have asthma based on the same criteria.

Inclusion and exclusion criteria

Participants were aged ≥18 and 70 years and had a physician diagnosis of asthma. Exclusion criteria included age <18 years; patients with other chronic respiratory diseases such as chronic obstructive pulmonary disease (COPD) or lung fibrosis, pregnant or lactating women; patients with a history of advanced liver or chronic kidney diseases, advanced malignant tumors, Arthritis or Psoriasis can influence IL-13 level, [20] coronary heart disease,[21] Diabetic Mellitus' effect on arginase-1 level,[22] and active smoking.

Sample size calculation

A priori sample size calculation was performed using G*Power software (version 3.1.9.4). Based on previous studies reporting correlations ranging from 0.25 to 0.35 between IL-13 and asthma-related inflammatory markers, an effect size of r=0.30 was selected as a conservative and clinically meaningful estimate, consistent with Cohen's criteria for a moderate correlation. Using a two-tailed α level of 0.05 and 80% power, the minimum required sample size was calculated to be 84 asthmatic patients. To compensate for potential incomplete data, 90 patients were ultimately enrolled.

Sample collection and biomarker measurement

Venous blood samples (5 mL) were collected under aseptic conditions. Samples were allowed to coagulate and were then centrifuged at 3000 rpm for ten minutes. The serum was isolated and preserved at -80 °C until analysis. Serum IL-13 was quantified using a sandwich enzyme-linked immunosorbent assay ELISA kit (BT-Laboratory, Cat. No. E0098Hu; standard range 0.5-100 ng/L; sensitivity 0.2 ng/L; intra-assay CV<8%, inter-assay CV<10%). Serum arginase-1 was quantified using a sandwich ELISA kit (BT-Laboratory, Cat. No. E4984Hu; standard range: 0.5-200 ng/mL; sensitivity: 0.35 ng/mL; intra-assay CV<8%; inter-assay CV <10%). Absorbance was read at 450 nm, and concentrations were calculated from standard curves according to the manufacturer's instructions.

Statistical analysis

Statistical analyses were performed using JASP (version 0.95.4) and MedCalc (version 23.4.0) for data processing, tabulation, and inferential testing, whereas GraphPad Prism (version 10.6.0) was used for graphical presentation. As most variables were non-normally distributed, continuous data were reported as median, interquartile range (Q1-Q3), and mean±SD. To compare the groups, the Mann-Whitney U test, the Kruskal-Wallis's test, and Dunn's post hoc multiple comparisons test were used. The relationship between continuous variables was evaluated using Spearman's rank correlation coefficient. In addition, binary logistic regression analysis was performed to identify independent predictors of asthma diagnosis and to estimate odds ratios (ORs) with 95% confidence intervals (CIs). Model performance was evaluated using the receiver operating characteristic (ROC) curve and the area under the curve (AUC). A two-tailed $p<0.05$ was considered statistically significant.

Results

The present study reports a case-control analytical evaluation of 180 participants, comprising 90 asthmatic patients and 90 healthy controls. The asthmatic cohort included both moderate and severe disease phenotypes: 56 (62.2%) had moderate disease severity, and 34 (37.8%) had severe disease severity. Based on asthma control, 31 (34.4%) were controlled, 34 (37.8%) were partially controlled, and 25 (27.8%) were uncontrolled, according to the GINA 2025 criteria. The percentage distribution of asthmatic patients according to disease severity and level of asthma control is shown in Figure 1.

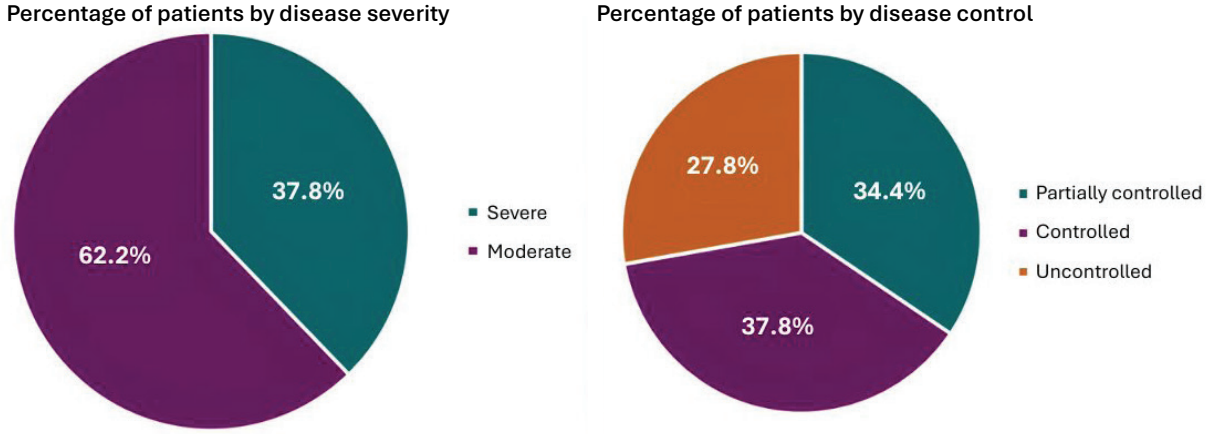


Figure 1. *Pie charts showing the percentage distribution of asthmatic patients according to disease severity and level of asthma control.*

The baseline characteristics of the study population are briefly outlined in Table 1. There were no significant differences between patients and controls regarding gender (χ^2=0.023, p=0.879), indicating that sex-matching is appropriate, and no differences in the number of subjects of each sex in the comparison groups. Likewise, no significant age difference was observed between cohorts (p=0.265), indicating comparable age groups. Body mass index was also similar between the two groups (p=0.242), with average body mass index in both populations being overweight or obese. Overall, these results indicate that the two cohorts were well balanced across most demographic characteristics, which should help minimize potential confounders in downstream biomarker analyses.

Table 1. Baseline demographic and anthropometric characteristics of the study population according to gender, age, and body mass index (BMI).

Subject	**Control** (N=90)	**Patient** (N=90)	**Statistical value**
Gender			
Male (N, %)	35 (38.9%)	37 (41.1%)	p=0.879
Female (N, %)	55 (61.1%)	53 (58.9%)	χ^2=0.023
Age (years) Mean±SD	48.30±12.69	46.00±14.87	p=0.265 t=1.116
BMI (kg/m^2) Mean±SD	29.36±3.43	30.16±5.55	p=0.242 t=1.173

Note: *Data are presented as frequency (N), percentage (%), and mean±SD. BMI: body mass index, χ^2: Chi-Square, t: T test*

Comparison of serum IL-13 and arginase-1 concentrations between asthmatic patients and healthy control subjects. As demonstrated in Table 2 and Figure 2, the asthma group exhibited significantly elevated levels of both biomarkers. The median serum IL-13 concentrations were significantly higher in patients [14.0 (9.9-16.6) ng/L] than in controls [10.05 (7.2-13.1) ng/L] (p<0.001). Similarly, arginase-1 levels were substantially higher in people with asthma, with a median concentration of [64.05 (54.7-72.1) ng/mL] compared to [26.50 (21.7-32.1) ng/mL] in healthy controls. All descriptive measures showed highly significant differences (p<0.001).

Table 2. Comparative serum levels of interleukin-13 and arginase-1 between patients with asthma and healthy controls.

Biomarker	**Patient** (N=90)	**Control** (N=90)	***p* value**
Interleukin-13 (ng/L)			
Median (Q1-Q3)	14.00 (9.9-16.6)	10.05 (7.2-13.1)	**<0.001**
Mean±SD	15.96±9.57	11.05±6.26	
Arginase-1 (ng/mL)			
Median (Q1-Q3)	64.05 (54.7-72.1)	26.50 (21.7-32.1)	**<0.001**
Mean±SD	65.63±18.08	27.33±8.14	

Note: *Statistical comparisons between groups were performed using the Mann-Whitney U test, based on data distributions and central tendencies, with results summarized as Median (Q1-Q3). Mean values are presented for descriptive purposes only. (Q1-Q3) represent the 25th and 75th percentiles.*

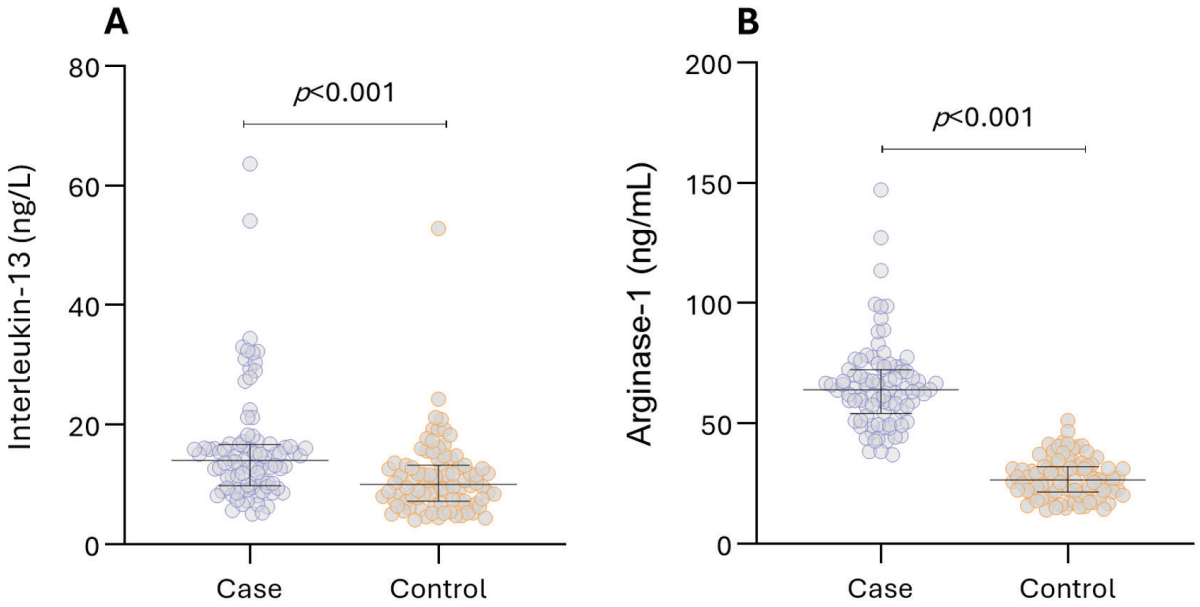


Figure 2*. Comparison of serum biomarker levels between patients with asthma (cases) and healthy controls. Data are presented as individual column scatter plots. A: serum concentration of IL-13 in ng/L, B: serum concentration of Arginase-1 in ng/mL. Statistical significance was determined using the Mann-Whitney U test, and the exact p values are indicated on the plots, with a value of p<0.001 indicating highly significant differences.*

Biomarker levels from moderate and severe asthma groups were compared to determine association with disease severity. As shown in Table 3, the median serum IL-13 level was significantly higher in patients with severe asthma [16.8 (14.9-28.7) ng/L] than in those with moderate asthma [12.3 (8.9-15.1) ng/L] (p<0.0001). A statistically significant difference in arginase-1 levels was also observed between patients with severe [65.4 (59.5-73.8) ng/mL] and moderate disease [62.95 (49.3-67.9) ng/mL] (p=0.0384) (Table 3).

Table 3. Comparison of serum levels of interleukin-13 and arginase-1 between moderate and severe asthma groups.

Biomarker	**Severe** (N=34)	**Moderate** (N=56)	***p* value**
Interleukin-13 (ng/L)			
Median (Q1-Q3)	16.8 (14.9-28.7)	12.3 (8.9-15.1)	**<0.0001**
Mean±SD	21.7±12.1	12.5±5.3	
Arginase-1 (ng/mL)			
Median (Q1-Q3)	65.4 (59.5-73.8)	62.9 (49.3-67.9)	**0.0384**
Mean±SD	69.9±17.1	61.5±14.4	

Note: *Statistical comparisons between groups were performed using the Mann-Whitney U test, based on data distributions and central tendencies, with results summarized as Median (Q1-Q3). Mean values are presented for descriptive purposes only. (Q1-Q3) represent the 25th and 75th percentiles.*

Additionally, patients were classified according to the GINA guideline control status into controlled, partially controlled, and uncontrolled groups. Serum IL-13 levels increased progressively with worsening control status (p=0.0009; Table 4). The median IL-13 level was [9.9 (8.5 -14.9) ng/L] in the controlled group, [15.5 (12.9-17.2) ng/L] in the partially controlled group, and [15.1 (12.6-27.9) ng/L] in the uncontrolled group. Post hoc analysis indicated significant differences between the uncontrolled and controlled groups (U>C) and between the partially controlled and controlled groups (P>C). In contrast, no statistically significant difference in median serum arginase-1 levels was seen between the three asthma control groups (p=0.6079, Table 4)

Table 4. Comparison of serum levels of interleukin-13 and arginase-1 among asthmatic patients grouped by GINA guideline control status (controlled, partially controlled, and uncontrolled).

Biomarker	**Controlled** (N=31)	**Partially controlled** (N=34)	**Uncontrolled** (N=25)	***p* value**	**Post-hoc analysis**
Interleukin-13 (ng/L)					
Median (Q1-Q3)	9.9 (8.5 -14.9)	15.5 (12.9-17.2)	15.1 (12.6-27.9)	**0.0009**	U>C U=P P>C
Mean±SD	11.6±5.4	17.6±10.1	19.0±11.06		
Arginase-1 (ng/mL)					
Median (Q1-Q3)	64.6 (59.5-69.3)	63.8 (51.9-72.1)	61.5 (49.2-71.4)	**0.6079**	U=C U=P P=C
Mean±SD	66.0±12.6	64.1±16.7	63.8±18.5		

Note: *Statistical comparisons among the three asthma control groups (controlled (C), partially controlled (P), uncontrolled (U)) were performed using the Kruskal-Wallis test based on data distributions and central tendencies, summarized as Median (Q1-Q3), followed by Dunn's post-hoc analysis for pairwise comparisons. Mean±SD values are presented for descriptive purposes only and are not implied in the p value calculations. The symbols in the "Post-hoc analysis" column indicate the direction of significant differences between groups.*

A weak but statistically significant positive correlation was observed between serum IL-13 and arginase-1 levels (r=0.251, *p*=0.017). This relationship suggests partial interaction between Type-2 inflammatory cytokine activity and arginine metabolic pathways in asthma (Figure 3).

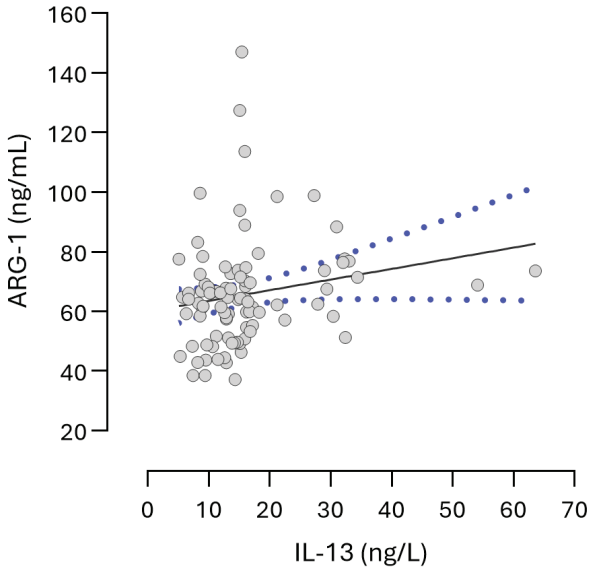


Figure 3. *Scatter plot illustrating the correlation between serum IL-13 and arginase-1 levels in patients with asthma. The solid line represents the fitted linear regression trend, while the dashed lines indicate the 95% confidence interval of the regression estimate.*

In our study, we used binary logistic regression to determine whether IL-13 and arginase-1 could serve as predictors of an asthma diagnosis. The overall model proved to be highly significant (χ^2=212.18, *p*<0.001) and demonstrated exceptional explanatory power (Nagelkerke R^2=0.923). Arginase-1 was identified as a significant independent predictor of asthma (OR=1.48, 95% CI: 1.24-1.77, *p*<0.0001), whereas IL-13 did not show statistical significance after adjustment. The model showed good calibration (Hosmer–Lemeshow *p*=0.738), excellent discrimination (AUC=0.993), and a high overall classification accuracy (97.2%) as shown in Table 5 and Figure 4.

Table 5. Binary logistic regression analysis identifying independent biomarker predictors of asthma (case vs healthy control)

Variable	**β coefficient**	**Odds ratio** (OR)	**95% CI for OR**	***p* value**
Interleukin-13 (ng/L)	-0.111	0.8948	0.73 to 1.08	0.2643
Arginase-1 (ng/mL)	0.398	1.4823	1.24 to 1.77	0.0001

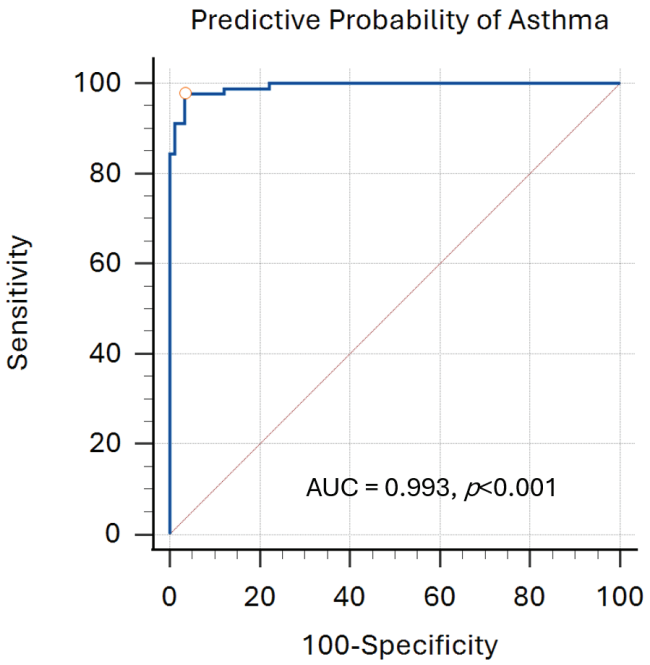


Figure 4. *Receiver Operating Characteristic (ROC) curve analysis assessing the diagnostic accuracy of the combined biomarker model. The curve demonstrates the exceptional discriminative performance of the multivariable logistic regression model incorporating both serum IL-13 and Arginase-1 (AUC=0.993, 95% CI: 0.984-1.000, Overall Accuracy = 97.2%) in differentiating asthmatic patients from healthy controls.*

Discussion

In this study, we examined the correlation between IL-13 and arginase-1 levels and asthma severity and control status in an Iraqi population and assessed their potential as biomarker-based indicators of disease activity. In this context, these findings also contribute to the growing interest in identifying complementary biomarkers. While established biomarkers such as blood eosinophil count (BEC), fractional exhaled nitric oxide (FeNO), and IgE are widely used in clinical practice, there remains a need to further characterize additional biomarkers that may provide complementary insights into disease mechanisms. This is particularly relevant in underrepresented populations, where regional environmental, genetic, and healthcare-related factors may influence disease expression and biomarker profiles.

Although significant differences were observed between groups, some degree of overlap in biomarker levels suggests that their discriminative utility at the individual level should be interpreted with caution. This consideration motivated the use of multivariable logistic regression analysis, which allows the combined effect of multiple biomarkers to be evaluated to improve discrimination. Overall, our findings demonstrated three principal observations: 1) both IL-13 and arginase-1 levels were significantly elevated in asthmatic patients compared with healthy controls, 2) IL-13 showed a strong association with disease severity and control status, and 3) arginase-1 emerged as a significant independent predictor of asthma in logistic regression analysis with excellent diagnostic performance.

Focusing first on the inflammatory cascade, the significant elevation of serum IL-13 levels observed in asthmatic patients aligns with the well-accepted paradigm that IL-13 is a key Th2 cytokine that mediates T2 airway inflammation and drives IgE class-switching, eosinophilia, mucus hypersecretion, and airway hyper responsiveness. These effects drive airway remodeling and persistent inflammation in asthma. Others have reported increased levels of IL-13 in patients with both allergic and severe asthma phenotypes, establishing its significance as an important mediator of T2-high inflammatory pathways.[11,13,23] In addition, we found that IL-13 levels were significantly higher in severe compared to moderate asthma, supporting a clear relationship between IL-13 and disease severity.

Taken together, this observation is consistent with a role for IL-13 in driving not just inflammatory responses but also remodeling events within the airway epithelium. Experimental and clinical studies have shown that IL-13 induces goblet cell metaplasia, collagen remodeling, and subepithelial fibrosis, which are hallmarks of airway remodeling in severe asthma.[12,24,25]

Beyond simple disease classification, serum IL-13 levels also increased significantly as asthma worsened, representing another key finding of the present study. Both partially controlled and uncontrolled patients showed significantly higher IL-13 levels than controlled subjects. The higher IL-13 levels in patients with recurrent symptoms indicate ongoing inflammatory activity and poor disease suppression, making IL-13 a good candidate as an asthma control molecule. These

data are consistent with previous investigations showing that T2 cytokines are associated with poor asthma control and exacerbations.[13,26-28] Accordingly, elevated serum IL-13 levels have been linked to increased airway inflammation, corticosteroid resistance, and persistent symptoms in severe asthma.

While IL-13 strongly reflects these dynamic fluctuations in inflammation and symptom control, a parallel pathway involving structural modifications is illuminated by altered arginine metabolism. The present study also demonstrated markedly increased serum arginase-1 levels in asthmatic patients compared with healthy individuals. Arginase-1 is an important enzyme in the L-arginine metabolic pathway that competes with nitric oxide synthases (NOS) for the common substrate L-arginine. Elevated arginase-1 activity reduces nitric oxide production and leads to the accumulation of ornithine and downstream metabolites that contribute to airway remodeling.[17,29] Reduced bioavailability of nitric oxide may lead to bronchoconstriction, airway hyperresponsiveness, and oxidative stress in the asthmatic condition. Increased arginase-1 activity has been shown in prior experimental studies to exacerbate allergen-induced airway hyperresponsiveness and promote airway structural remodeling.[17,19,30]

Intriguingly, although arginase-1 concentrations were markedly higher in severe than in moderate disease, this difference was less pronounced than that observed for IL-13, and arginase-1 levels did not differ significantly across asthma control categories. This indicates that arginase-1 might reflect chronic airway remodeling rather than acute inflammatory activity. Arginase-1 levels did not differ significantly in all asthma control categories. This observation suggests that arginase-1 may be more closely associated with structural airway changes and disease burden rather than short-term variations in symptom control.

Despite these differing clinical expressions, a biological bridge exists between these two markers, as a weak but statistically significant positive correlation was observed between circulating IL-13 and arginase-1 level. These findings are consistent with type-2 inflammatory cytokines altering arginine metabolic pathways in asthma.[31] Indeed, there is experimental evidence that IL-13 induces arginase-1 expression in airway epithelial and macrophage cells, with a subsequent increase in arginine metabolism and impaired nitric oxide availability. This interaction may serve as a key mediator connecting immune inflammation with airway remodeling processes. [19]

This biological interplay directly informs the diagnostic performance of these markers when evaluated simultaneously. Results of our Binary logistic regression analysis demonstrated that, after adjustment, arginase-1, but not IL-13, was a significant independent predictor of asthma. These results indicate exceptional discriminatory performance, with an area under the receiver operating characteristic curve (AUC) of 0.993, enabling the model to accurately distinguish asthmatic patients from healthy volunteers. The fact that IL-13 loses statistical significance when added into the regression model may be explained by a partial correlation between IL-13 and arginase-1, implying overlapping biological pathways. In this regard, arginase-1 may serve as a more stable biomarker for distinguishing metabolic from structural airway alterations, rather than transient inflammatory signaling.

Taken together, these results suggest that arginase-1 is a highly promising candidate diagnostic biomarker for asthma, which could be assessed alongside inflammatory markers, such as IL-13, to improve characterization of asthma phenotypes.

Limitations

This study has several limitations. The case-control design may have led to an overestimation of discrimination performance, as comparisons were made between clearly distinct groups (asthma vs healthy controls), which may not fully reflect real-world clinical complexity. In everyday clinical practice, diagnostic tools operate in more challenging settings, where they must differentiate asthma from other conditions with overlapping respiratory symptoms. In addition, the relatively small, single-center sample and restriction to an Iraqi population may limit generalizability. Therefore, these findings should be interpreted with caution, and future prospective studies in broader, real-world clinical populations are welcomed to further confirm generalizability. In addition, asthma severity in this study was defined according to GINA guidelines, which reflect treatment intensity rather than intrinsic disease severity and may be influenced by external factors such as treatment adherence and healthcare access. Although this approach has limitations, it remains a widely accepted and practical method for severity classification in clinical research.

Conclusion

In conclusion, the present study demonstrates that serum IL-13 and arginase-1 levels are significantly elevated in asthmatic patients. IL-13 is strongly associated with disease severity and asthma control and may reflect underlying type-2 inflammatory activity, whereas arginase-1 appears to be more closely associated with disease presence and may relate to processes involved in airway remodeling. The combined evaluation of these biomarkers may help clarify aspects of asthma pathophysiology and support the development of biomarker-based approaches for disease phenotyping and clinical assessment.

Ethical approval

The study protocol was approved by the Institutional Review Board (IRB) of the College of Medicine, University of Al-Nahrain on 10-03-2025 under decision number 223/3/2. All participants provided verbal informed consent and were asked to complete a prepared questionnaire. Individual were informed of

their ability to withdraw at any time without any negative impact on their clinical care, and confidentiality was maintained throughout the study. The ethical standards outlined in the Declaration of Helsinki were adhered to by all procedures.

Funding

This research received no external funding.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Hayder Yaseen Ali: conceptualization, methodology, funding acquisition, writing: original draft, data curation, formal analysis, and investigation; **Haider Abdulhameed Alqaraghuli**: conceptualization, methodology, funding acquisition, writing: original draft, validation, writing: review and edit, supervision; **Raid J. Mohammed AL-TIMIMI**: validation, writing: review and edit, supervision. All authors read and approved the final manuscript.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, the authors used Gemini (Version 1.5 Pro, Google, USA) and Grammarly (Version 2024, Grammarly Inc., USA) solely to enhance the English language, grammar, and overall readability of the text. These tools were used for linguistic refinement and editorial purposes only, without any involvement in data analysis, interpretation of results, or the generation of original scientific content. The authors assume full responsibility for the integrity and accuracy of the final manuscript and affirm that these AI tools do not qualify for authorship.

Acknowledgements

The authors sincerely thank the medical staff at the Respiratory Consultative Clinic in Al-Imamain Al-Kadhimain Medical City for their invaluable assistance. We also extend our gratitude to Al-Nahrain University College of Medicine for their support in making this study possible, and to all the patients who volunteered to participate.

References

[1] Global Initiative for Asthma (GINA). Global Strategy for Asthma Management and Prevention (2025 Update). Available from: https://ginasthma.org/.

[2] Li W, Gao P, Zhi Y, Xu W, Wu Y, Yin J, et al. Periostin: Its role in asthma and its potential as a diagnostic or therapeutic target. Respir Res. 2015;16(1):1-11. doi: 10.1186/s12931-015-0218-2.

[3] Refaat MM, El Sayed E, Abd El-Fattah W, Elbanna AH, Sayed HME. Relationship between sputum periostin level and inflammatory asthma phenotypes in Egyptian patients. J Asthma. 2021; 58(10):1285-91. doi:10.1080/02770903.2020.1786111.

[4] Muhsen SA, AL-Saeed HH, ALOmrani AAA. Study of the association between prolidase enzyme kinetics and severity of asthma in children. Indian J Public Health Res Dev. 2019; 10(6): 836-42. doi: 10.5958/0976-5506.2019.01383.4.

[5] Vos T, Lim SS, Abbafati C, Abbas KM, Abbas M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the global burden of disease study 2019. Lancet. 2020; 396(10258): 1204-22. doi: 10.1016/S0140-6736(20)30925-9.

[6] Baldo DC, Romaldini JG, Pizzichini MMM, Cançado JED, Dellavance A, Stirbulov R. Periostin as an important biomarker of inflammatory phenotype T2 in Brazilian asthma patients. J Bras Pneumol. 2023; 49(1): 1-8. doi: 10.36416/1806-3756/e20220040.

[7] Burgess JK, Jonker MR, Berg M, Hacken NTHT, Meyer KB, Van Den Berge M, et al. Periostin: Contributor to abnormal airway epithelial function in asthma? Eur Respir J. 2021; 57(2): 1-12. doi: 10.1183/13993003.01286-2020.

[8] Ono J, Takai M, Kamei A, Ohta S, Nair P, Izuhara K, et al. A novel assay for improved detection of sputum periostin in patients with asthma. PLoS One. 2023; 18(2): e0281356. doi: 10.1371/journal.pone.0281356.

[9] Al-Aaraji AJ, Al-Qaysi SA, Baay AS. Role of periostin in Iraqi asthmatic patients. Med J Babylon. 2019; 16(3): 256-60. doi: 10.4103/MJBL.MJBL_39_19.

[10] Kazaal M, Habeeb AA, Hasan HN. Evaluation role of IL-13 and eosinophils in adult asthmatic patients. J Biomed Biochem. 2023; 2(2): 17-25. doi: 10.57238/jbb.2023.6940.1033.

[11] Sahnoon L, Bajbouj K, Mahboub B, Hamoudi R, Hamid Q. Targeting IL-13 and IL-4 in asthma: Therapeutic implications on airway remodeling in severe asthma. Clin Rev Allergy Immunol. 2025;68(1):44. doi: 10.1007/s12016-025-09045-2.

[12] Francisco D, Wang Y, Conway M, Hurbon AN, Dy ABC, Addison KJ, et al. Surfactant protein-A protects against IL-13-induced inflammation in asthma. J Immunol. 2020; 204(10): 2829-39. doi: 10.4049/jimmunol.1901227.

[13] Corren J. Role of Interleukin-13 in asthma. Curr Allergy Asthma Rep. 2013; 13(5): 415-20. doi: 10.1007/s11882-013-0373-9.

[14] Wu G, Morris SM. Arginine metabolism: nitric oxide and beyond. Biochem J. 1998; 336(1): 1-17. doi: 10.1042/bj3360001.

[15] Ogino K, Obase Y, Takahashi N, Shimizu H, Takigawa T, Wang DH, et al. High serum arginase I levels in asthma: Its correlation with high-sensitivity C-reactive protein. J Asthma. 2011; 48(1): 1-7. doi: 10.3109/02770903.2010.528496.

[16] Hussain HAK, Muftin NQ, Al-jibouri MN, Ben Salah G. Study of the arginase activity and other biochemical parameters in patients with coronary artery disease in Baghdad Governorate-Iraq.

Al-Mustansiriyah J Sci. 2023; 34(1): 37-44. doi: 10.23851/mjs.v34i1.1251.

[17] Maarsingh H, Dekkers BGJ, Zuidhof AB, Bos IST, Menzen MH, Klein T, et al. Increased arginase activity contributes to airway remodelling in chronic allergic asthma. Eur Respir J. 2011; 38(2): 318-28. doi: 10.1183/09031936.00057710.

[18] Amin Said AM, Al-Naqshbandi AA, Amin YK, Ali RJ. Association of the arginase I with bronchial asthma. J Univ Babylon Pure Appl Sci. 2019; 27(2): 238-43. doi: 10.29196/jubpas.v27i2.2092.

[19] Meurs H, Maarsingh H, Zaagsma J. Arginase and asthma: Novel insights into nitric oxide homeostasis and airway hyperresponsiveness. Trends Pharmacol Sci. 2003; 24(9): 450-5. doi: 10.1016/S0165-6147(03)00227-X.

[20] Spadaro A, Rinaldi T, Riccieri V, Valesini G, Taccari E. Interleukin 13 in synovial fluid and serum of patients with psoriatic arthritis. Ann Rheum Dis. 2002; 61(2): 174-6. doi: 10.1136/ard.61.2.174.

[21] Eleftheriadis T, Liakopoulos V, Antoniadi G, Stefanidis I, Galaktidou G. Arginase type I as a marker of coronary heart disease in hemodialysis patients. Int Urol Nephrol. 2011; 43(4): 1187-94. doi: 10.1007/s11255-010-9842-4.

[22] Khaleel FM, Oda NN, Abed BA. Disturbance of arginase activity and nitric oxide levels in Iraqi type 2 diabetes mellitus. Baghdad Sci J. 2018; 15(2): 189-91. doi: 10.21123/bsj.2018.15.2.0189.

[23] Raju P, Sundar S, Suresh P, Thulukanam J, Padmanaban S. A comparative study on the utility of biomarkers – serum Interleukin-13 against serum immunoglobulin E in assessing the severity of asthma. Eur J Clin Exp Med. 2025; 23(2): 445-52. doi: 10.15584/ejcem.2025.2.27.

[24] Yasui H. Using biomarkers and independent predictors of therapy response to optimize treatment of uncontrolled severe asthma in the biologic era. Front Allergy. 2025; 6: 1719820. doi: 10.3389/falgy.2025.1719820.

[25] Jebur MS, Saud AM. Serum levels of total IgE and Interleukin-13 in a sample of allergic asthma patients in Baghdad. Iraqi J Sci. 2020; 61(12): 3208-14. doi: 10.24996/ijs.2020.61.12.8.

[26] Jia Y, Fang X, Zhu X, Bai C, Zhu L, Jin M, et al. IL-13+ Type 2 innate lymphoid cells correlate with asthma control status and treatment response. Am J Respir Cell Mol Biol. 2016; 55(5): 675-83. doi: 10.1165/rcmb.2016-0099OC.

[27] Tsilogianni Z, Hillas G, Bakakos P, Aggelakis L, Konstantellou E, Papaioannou AI, et al. Sputum Interleukin-13 as a biomarker for the evaluation of asthma control. Clin Exp Allergy. 2016; 46(7): 923-31. doi: 10.1111/cea.12729.

[28] Awopeju OF, Salawu L, Adedeji TA. The relationship of serum Interleukin-13 to asthma control in Nigerian population. IP Indian J Immunol Respir Med. 2025; 5(4): 205-10. doi: 10.18231/j.ijirm.2020.059.

[29] Morris CR, Poljakovic M, Lavrisha L, Machado L, Kuypers FA, Morris SM. Decreased arginine bioavailability and increased serum arginase activity in asthma. Am J Respir Crit Care Med. 2004; 170(2): 148-53. doi: 10.1164/rccm.200309-1304oc.

[30] Liu Y, Zhang M, Wang T, Zhang J. Reactive oxygen species in asthma: Regulators of macrophage polarization and therapeutic implications: A narrative review. J Asthma Allergy. 2025; 18: 1129-46. doi: 10.2147/JAA.S529371.

[31] North M, Meurs H, Zaagsma J, Scott J, Maarsingh H. Arginase in asthma—Recent developments in animal and human studies. Open Nitric Oxide J. 2009; 2: 20-36.

Journal of Associated Medical Sciences 2026; 59(3): 142-150

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 individualized multicomponent intervention for sensory food aversion feeding disorder in children: A case series

Kowsalya M[1*], Sujata Missal[2], Sugi S[2]

[1]Department of Occupational Therapy, Akkara Foundation, Kasaragod, Kerala, India.
[2]KMCH College of Occupational Therapy, Coimbatore, Tamil Nadu, India.

ARTICLE INFO

Article history:
Received 13 August 2025
Accepted as revised 4 June 2026
Available online 23 June 2026

Keywords:
Multicomponent intervention, sensory food aversion, feeding disorder, children, feeding therapy

ABSTRACT

Background: Sensory food aversion feeding disorder is a feeding problem in which children refuse to eat all or most foods. This condition can be treated effectively using a multicomponent intervention that combines both sensory and behavioral strategies.

Objectives: This study aimed to determine the effectiveness of a multicomponent intervention for children with sensory food aversion feeding disorder. The specific objective was to increase acceptance and consumption of non-preferred foods and to improve parents' positive coping attitude towards the child's mealtime.

Materials and methods: This case series highlights seven cases that met the inclusion criteria. Informed consent was obtained from the parents. There were four boys and three girls, with ages ranging from 3.11 years to 4.9 years. The multicomponent intervention included sensory integration and behavioral modification strategies. Additionally, parents received education on the causes of food refusal and strategies for increasing food acceptance during meals.

Results: Seven children with sensory food aversion feeding disorder demonstrated increased food acceptance of previously refused foods. The incidence of food refusal behaviors, including gagging, spitting, and vomiting during mealtime, was also reduced.

Conclusion: This case series study provides information that may inform treatment strategies for sensory food aversion feeding disorder and may serve as a catalyst for larger-scale clinical trials.

Introduction

A feeding disorder is defined as impaired oral intake that is not age-appropriate and is associated with medical, nutritional, feeding skill, and/or emotional difficulties. Feeding disorders (FDs) are multidimensional, involving interactions among feeding skills, sensory processing, emotional regulation, and caregiver-child interaction, and are commonly identified between 1-3 years of age.[1-3] Children may present with food refusal, rigid preferences, and delayed self-feeding, which can persist into later childhood and contribute to caregiver stress.[4,5]

Feeding problems affect 25-45% of typically developing children, one-third of children with developmental delays, and up to 80% of children

*** Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Akkara Foundation, Kasaragod, Kerala, India.*
E-mail address: ot@kmch.ac.in
doi: 10.66285/JAMS.2026.088
E-ISSN: 2539-6056

with severe disabilities.[3,6] Etiologically, FDs are broadly categorized as organic or non-organic. Organic FDs arise from identifiable medical, neurological, or structural conditions that directly interfere with physiological processes necessary for feeding, swallowing, digestion, or nutrient absorption. These include gastrointestinal disorders, metabolic or absorptive conditions, anatomical abnormalities, chronic medical conditions affecting feeding endurance, and oral-motor impairments such as dysphagia.

In contrast, non-organic (FDs) occur in the absence of a primary medical pathology and are maintained predominantly by behavioral, psychological, sensory, or environmental factors. Management requires interdisciplinary collaboration. Occupational therapy emphasizes feeding as a culturally embedded co-occupation influenced by routines, context, and sensory experiences.[7,8] Sensory Food Aversion (SFA), a subtype of Avoidant/Restrictive Food Intake Disorder (ARFID), involves heightened sensitivity to food properties such as texture, smell, or temperature, leading to persistent refusal and restricted intake.[9] Consequently, interventions that integrate sensory, behavioral, and caregiver-focused strategies are increasingly recognized as necessary.[10]

In this study, the multi-component intervention included: 1) sensory modulation to gradually increase tolerance to challenging food properties, 2) behavioral feeding strategies, including structured exposure and reinforcement, and 3) caregiver training to implement consistent, responsive mealtime routines, supporting the child's sensory processing and participation in feeding.

Patient information

Seven children aged 3 years 11 months to 4 years 10 months (mean 3.92±0.36) from a preschool in Coimbatore, Tamil Nadu, participated. Screening for Sensory Food Aversion Feeding Disorder (SFAFD) was conducted by an occupational therapist using Chatoor's classification (2002) [11,] which exhibits both sensory and behavioral feeding difficulties. Demographic and clinical details are in Table 1.

Table 1. Demographic details of the study population.

Study No.	Variables	N	%
1	**Gender**		
	- Boys	4	57.1
	- Girls	3	42.9
2	**Parent**		
	- Working father	7	100
	- Working mother	4	57.1
3	**Family type**		
	- Nuclear family	7	100
	- Joint family	0	0
4	**Siblings**		
	- Present	3	42.9
	- Absent	4	57.1
5	**Environmental distractions (TV, mobile)**		
	- Present	5	71.4
	- Absent	2	28.6
6	**Refusal behavior: Gagging/spitting**		
	- Present	4	57.1
	- Absent	3	42.9
7	**Feeding skill**		
	- Self-feeding	5	71.4
	- Dependent feeding	2	28.6

Table 1. Demographic details of the study population (continued).

Study No.	Variables	N	%
8	**Force-feeding**		
	- Present	5	71.4
	- Absent	2	28.6
9	**Birth details**		
	- Full-term	7	100
	- Cesarean delivery	3	42.9
	- Vaginal delivery	4	57.1
10	**Consanguinity**		
	- Consanguineous marriage	0	0
	- Non-consanguineous marriage	7	100
11	**Birth order**		
	- Firstborn or only child	5	71.4
	- Later born	2	28.6
12	**Primary caregiver**		
	- Mother	7	100
	- Home-maker	3	42.9
	- Employed	4	57.1
13	**Developmental milestones**		
	- Delayed (motor/language)	0	0
	- Normal	7	100

Primary concerns and clinical findings

All children persistently refused specific foods, exhibiting gagging, vomiting, spitting, crying, or avoidance. Complementary feeding was delayed in all seven cases, which typically begins at 6-8 months with purees and progresses to mashed or chopped foods by 9-11 months. Feeding problems began at different stages: Case 1 avoided vegetables by age 2; Case 2 refused meals by age 3; Cases 3-5 avoided fruits and vegetables at 2.5 years; Case 6 consumed only rasam rice; Case 7 avoided all gravies by 1.5 years. None had prior feeding therapy.

During four lunch-period observations, children exhibited crying, distraction, food refusal, and the need for prompting. Assessments were conducted by the researcher, an occupational therapist, using the Behavioral Pediatric Feeding Assessment Scale (BPFAS), Child Sensory Profile 2 (SP2), and the Food Preference Questionnaire (FPQ). Common findings included prolonged mealtimes, escape behaviors, sensory hypersensitivity, limited food variety, and caregiver use of coaxing or threats.

Intervention

Children received 30 daily sessions (1 hour/session; 5 days/week), each comprising 15 minutes of sensory integration (SI), 30 minutes of structured mealtime, presenting non-preferred foods first and preferred foods as reinforcement, and 15 minutes of caregiver training. Individualized SI activities given based on each child's sensory profile, including vestibular stimulation through swinging, therapy ball activities, jumping, and prone rolling; proprioceptive input via joint compression and weight-bearing positions; tactile stimulation using firm-pressure stroking with varied textures; and preparatory oral input to increase tolerance to oral sensations.

Behavioral strategies included positive reinforcement, shaping, systematic desensitization, escape extinction, prompts for mastication, and timed meals to ensure consistency.

Caregivers were trained to implement consistent mealtime routines, encourage self-feeding, avoid force-feeding, reduce distractions, and maintain at least 3-hour intervals between meals. Home programs and daily logs supported generalization across settings.

Follow-up and outcomes

Case 1 (4.2 years)

At baseline, showed severe avoidance, including fleeing the feeding area and refusal of non-preferred foods. BPFAS scores decreased (Frequency score

reduced from 102 pre-test to 79 post-tests; Problem score reduced from 19 to 5 (Table 2 and Figure 1). CSP-2 scores showed improvement in oral (pre 29 to post 20) and tactile processing (pre 22 to post 20, Table 2 and Figure 1), indicating reduced sensory defensiveness. By week two, although delay tactics such as whining and restroom requests persisted, the child began sampling non-preferred foods. By week three, these behaviors had largely resolved, and by week four, he consistently consumed all foods presented. FPQ scores showed increased acceptance across categories, including variety rice (pre 40 to post 55), chutneys (pre 10 to post 22), fruits (pre 18 to post 30), and snacks (pre 125 to post 145) (Table 3 and Figure 2).

Table 2. Pre-test and post-test scores of behavioral paediatric feeding assessment scale (frequency & problem) and child sensory profile 2 for children with sensory food aversion feeding disorder.

Case	Pre/post test score	Behavioral pardiatric feeding assessment scale				Child sensory profile 2			
		Frequency score	Percentage	Problem score	Percentage	Oral	Percentage	Tactile	Percentage
1	Pre	102	81.6	19	38	29	58	22	40
	Post	79	63.2	5	10	20	40	20	36.36
2	Pre	95	76	71	142	25	50	11	20
	Post	16	12.8	3	6	15	30	11	20
3	Pre	92	73.6	56	112	26	52	11	20
	Post	14	11.2	5	10	16	32	11	20
4	Pre	108	86.4	22	44	29	58	22	44
	Post	64	51.2	6	12	27	54	15	30
5	Pre	113	90.4	19	38	31	62	14	28
	Post	67	53.6	3	6	19	38	14	28
6	Pre	11	88.8	22	44	26	52	23	46
	Post	80	64	6	12	18	36	16	32
7	Pre	121	96.8	19	38	34	68	16	32
	Post	74	59.2	2	4	23	46	13	26

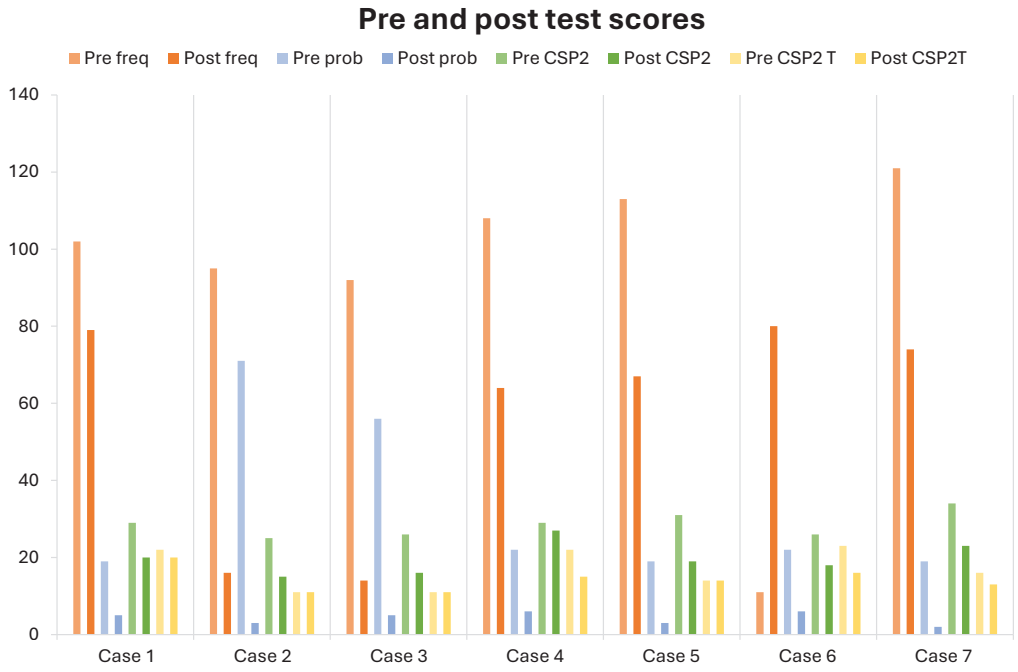


Figure 1. *Pre- and post-test scores of the Behavioral Pediatric Feeding Assessment scale (frequency and problem) and Child Sensory Profile 2 for children with sensory food aversion feeding disorder.*

Table 3. Pre and Post-Test Scores of Food Preference Questionnaire for Children

Case	Pre & post test	Tiffin (t=140) (%)	Variety rice (t=175) (%)	Curry & side-dishes (t=175) (%)	Chutney (t=50) (%)	Snacks (t=185) (%)	Desserts & sweets (t=55) (%)	Dairy (t=45) (%)	Nuts (t=30) (%)	Fruits (t=80) (%)	Total score (t=835) (%)
1	Pre	60 (42.85)	26 (34.66)	41 (23.42)	7 (14)	124 (67.02)	37 (67.27)	21 (46.66)	22 (73.33)	18 (22.5)	356 (42.63)
	Post	69 (49.28)	38 (50.66)	75 (42.85)	11 (22)	124 (67.02)	37 (67.27)	21 (46.66)	22 (73.33)	18 (22.5)	415 (49.70)
2	Pre	86 (61.42)	41 (54.66)	52 (29.71)	4 (8)	128 (69.18)	28 (50.90)	18 (40)	5 (16.66)	12 (15)	374 (44.79)
	Post	94 (67.14)	52 (69.33)	66 (37.71)	13 (26)	130 (70.27)	28 (50.90)	21 (46.66)	16 (53.33)	2 (26.25)	425 (50.89)
3	Pre	80 (57.14)	36 (48)	38 (21.71)	4 (8)	124 (67.02)	44 (80)	21 (46.66)	5 (16.66)	13 (16.25)	365 (43.71)
	Post	80 (57.14)	36 (48)	55 (31.42)	10 (20)	126 (68.10)	44 (80)	21 (46.66)	13 (43.33)	21 (26.25)	406 (48.62)
4	Pre	89 (63.57)	30 (40)	47 (26.85)	4 (8)	134 (72.43)	44 (80)	18 (40)	4 (13.33)	14 (17.5)	384 (45.98)
	Post	89 (63.57)	32 (42.66)	67 (38.28)	8 (16)	136 (73.51)	44 (80)	20 (44.44)	16 (53.33)	23 (28.75)	435 (52.09)
5	Pre	89 (63.57)	30 (40)	47 (26.85)	4 (8)	134 (72.43)	44 (80)	18 (40)	4 (13.33)	14 (17.5)	384 (45.98)
	Post	89 (63.57)	32 (42.66)	67 (38.28)	8 (16)	136 (73.51)	44 (80)	20 (44.44)	16 (53.33)	23 (28.75)	435 (52.09)
6	Pre	88 (62.85)	36 (48)	45 (25.71)	11 (22)	136 (73.71)	44 (80)	26 (57.77)	30 (100)	13 (16.25)	429 (51.37)
	Post	89 (63.57)	43 (57.33)	56 (32)	19 (38)	138 (74.59)	44 (80)	29 (60)	30 (100)	26 (32.5)	474 (56.76)
7	Pre	83 (59.28)	41 (54.66)	39 (22.28)	7 (14)	114 (61.62)	42 (76.36)	33 (73.33)	4 (13.33)	14 (17.5)	377 (45.14)
	Post	83 (59.28)	4 (54.66)	59 (33.71)	21 (42)	114 (61.62)	42 (76.36)	33 (73.33)	14 (46.66)	25 (31.25)	432 (51.73)

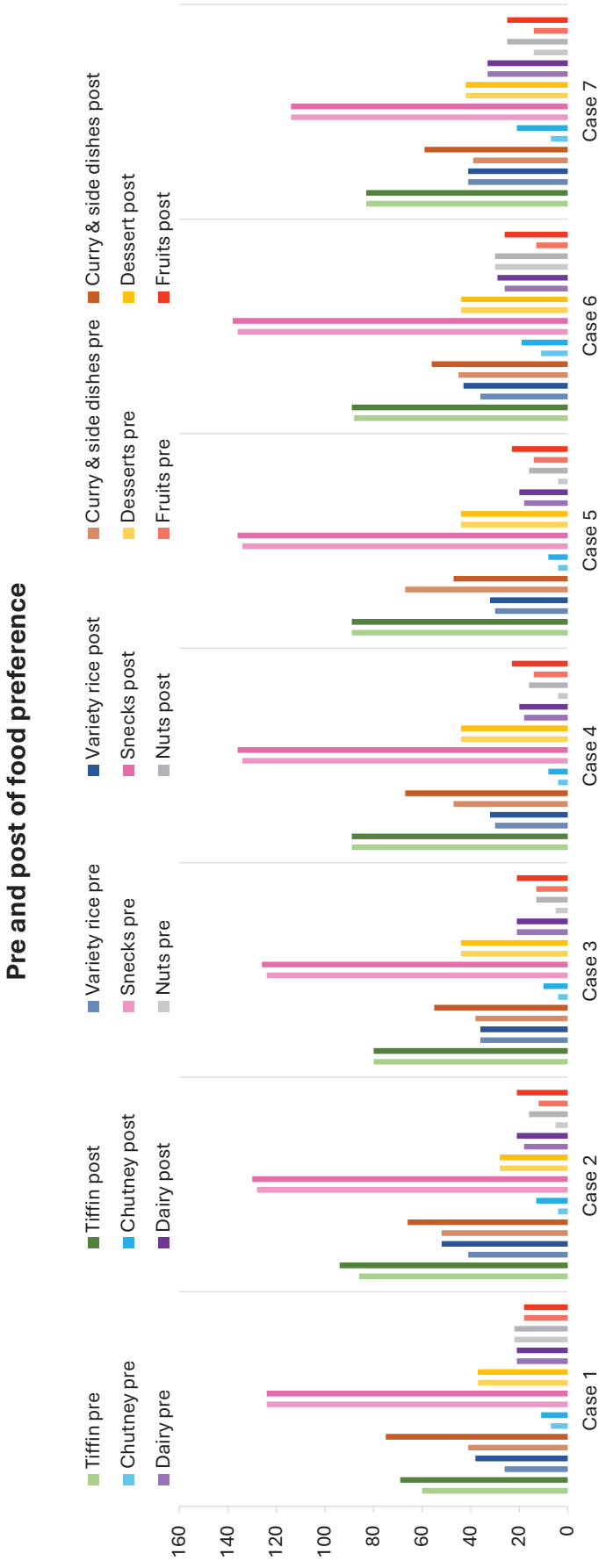


Figure 2. *Pre- and post-test scores of the Food Preference Questionnaire for children*

Case 2 (3.8 years)

Initial behaviors, such as spitting, distraction, and excessive talking, reflected in reduced BPFAS scores (Frequency: 95 to 72; Problem: 3 to 3). CSP-2 oral sensitivity (pre 25 to post 15) and tactile sensitivity (pre 11 to post 11) improved. Spitting was reduced by week two and was limited to one episode by week four. FPQ scores increased for variety rice (pre 45 to post 60), dry fruits/nuts (pre 18 to post 30), chutneys (pre 12 to post 25), and snacks (pre 130 to post 150), supporting the observed behavioral improvements.

Case 3 (4.5 years)

Presented with gagging, vomiting, and spitting with marked reductions in BPFAS scores (Frequency: 92 to 65; Problem: 56 to 5). CSP-2 oral processing improved (pre 26 to post 16) while tactile scores remained stable (pre 11 to post 11). Spitting resolved by week three and gagging by week four. FPQ scores increased for curries and side dishes (pre 35 to post 55), fruits (pre 15 to post 28), chutneys (pre 10 to post 22), and nuts (pre 12 to post 25).

Case 4 (3.9 years)

Initially, avoid contact with food. Following social reinforcement, BPFAS scores reduced (Frequency: 108 to 64; Problem: 22 to 6). CSP-2 scores showed reduced oral (pre 29 to post 27) and tactile sensitivity (pre 22 to post 15). By week four, he regularly accepted gravies and fruits with fewer refusals. FPQ scores improved for vegetable-based curries (pre 30 to post 48), fruits (pre 14 to post 25), and chutneys (pre 10 to post 20).

Case 5 (5.1 years)

Initially slow and distracted, consuming only preferred foods. BPFAS scores decreased (Frequency: 113 to 80; Problem: 19 to 3). CSP-2 oral processing improved (pre 31 to post 19) and tactile processing (pre 14 to post 14), with stable tactile responses. By week four, she consumed all non-preferred foods, with only minor talking. FPQ scores increased for tiffin items (pre 65 to post 88), curries (pre 38 to post 60), fruits (pre 18 to post 32), and snacks (pre 135 to post 155).

Case 6 (4.3 years)

Baseline gagging and spitting with dependence on prompts, reflected in reduced BPFAS scores (Frequency: 111 to 80; Problem: 22 to 6). CSP-2 oral processing (pre 26 to post 18) and tactile processing (pre 23 to post 16) improved. By week three, self-feeding improved, and gagging was limited to specific vegetables. FPQ scores increased for variety rice (pre 42 to post 58), sautéed vegetables (pre 25 to post 40), and fruits (pre 20 to post 34), though beans remained a challenge.

Case 7 (4.0 years)

Frequent gagging and spitting decreased over time, with reduced BPFAS scores (Frequency: 121 to 74; Problem: 19 to 2). CSP-2 oral sensitivity (pre 34 to post 23) and tactile sensitivity (pre 16 to post 13) improved. By week four, she independently used a spoon and accepted most foods except beans. FPQ scores increased for vegetables (pre 30 to post 45), nuts (pre 18 to post 30), fruits (pre 15 to post 28), and snacks (pre 115 to post 135).

Cross-case synthesis

Across all seven cases, consistent reductions in maladaptive feeding behaviors were observed on the BPFAS, with decreases in both Frequency and Problem scores following intervention. Children with more severe baseline feeding difficulties (Cases 1, 3, 6, and 7) demonstrated larger absolute reductions, while those with milder presentations (Cases 2, 4, and 5) showed smaller yet clinically meaningful improvements, indicating comparable trajectories of change despite differing initial severity.

Sensory processing outcomes showed a clear cross-case pattern. Oral sensory processing improved across all cases, with greater reductions among children with higher baseline oral sensitivity (Cases 1, 4, 5, and 7) compared to those with moderate sensitivity (Cases 2, 3, and 6). In contrast, tactile and visual sensory processing remained largely stable across cases, suggesting limited responsiveness to short-term feeding intervention.

Functional outcomes measured by the FPQ showed higher total scores across all cases. Children with severe feeding avoidance (Cases 1, 3, 6, and 7) achieved functional gains comparable in direction to those observed in children with milder feeding difficulties (Cases 2, 4, and 5), supporting cross-case consistency of intervention effects.

All seven cases demonstrated reductions in feeding-related problem behaviors, as evidenced by decreased BPFAS Frequency and Problem scores (Table 2 and Figure 1). Improvements were most consistent in CSP-2 oral processing scores, indicating reduced oral defensiveness, while tactile processing changes were less pronounced across cases. These reductions in behavioral and sensory improvements were accompanied by increased food acceptance across multiple FPQ categories (Table 3 and Figure 1), supporting the overall effectiveness of the intervention in improving functional feeding outcomes.

Discussion

This case series examined outcomes of a short-term, multicomponent sensory-behavioral occupational therapy intervention for preschool children with SFAFD. All seven cases improved food acceptance, mealtime behavior, oral sensory processing, and food variety.

Although FDs often co-occur with neuro developmental conditions such as autism spectrum disorder and Attention Deficient Hyperactivity Disorder,

all children in this study were typically developing, indicating SFA can occur in isolation and benefit from early intervention.

The intervention was standardized in structure and duration but individualized according to severity. Children with more severe sensory defensiveness, gagging, or avoidance received additional prompting, targeted food selection, and enhanced parent engagement, whereas children with milder presentations progressed more readily within structured routines. This underscores the importance of individualized, occupation–based feeding interventions within a consistent framework.

Previous research has largely examined single-component interventions, including behavioral or sensory integration approaches, primarily in children with medical or developmental conditions.[4,12-14] In contrast, the present study implemented an integrated model that combined sensory modulation, behavioral strategies, and caregiver education for preschool children with SFA. This approach may account for the rapid and consistent improvements observed over a brief four-week period, suggesting that simultaneously targeting sensory sensitivities, behavioral avoidance, and caregiver-mediated reinforcement can accelerate gains.

Multicomponent feeding interventions have been shown to improve food acceptance and reduce maladaptive mealtime behaviors in medically complex populations.[15,16] In contrast, the present case series demonstrates that a similar integrated sensory behavioral approach with caregiver training can produce improvements in typically developing preschool children with sensory food aversion over a short intervention period.

Improvements were most pronounced in oral sensory processing, while tactile and visual domains showed smaller changes, suggesting differential responsiveness across sensory modalities. Direct oral exposure and graded taste experiences appear particularly responsive to short-term intervention, whereas broader sensory domains may require longer or more targeted input.[12] The use of the FPQ to guide intervention aligns with prior findings highlighting its clinical utility.[9,17,18] All children demonstrated increased acceptance of previously refused foods with minimal adverse behaviors, and caregivers reported improved confidence and mealtime participation, supporting generalization beyond the therapy setting.

The program addressed sensory hypersensitivity, avoidance, caregiver practices, and environmental distractions. Sensory, behavioral, and caregiver strategies together promoted sustainable improvements beyond short-term compliance.

This case series supports the feasibility and potential effectiveness of early, occupational therapy-led, integrated intervention for children with SFA. Despite the small sample limiting generalizability and preventing analysis of individual component effects, consistent improvements across all cases warrant further study of structured, multicomponent occupational therapy protocols for SFAFD.

Limitations

This case series is limited by a small sample size and lack of a control group, affecting generalizability. Its multicomponent nature, while reflective of clinical practice, limits analysis of individual intervention effects.

Recommendation

Greater awareness of SFAFD-often misdiagnosed as picky eating–is needed among professionals and families. Future studies should explore prevalence, utilize control groups, and incorporate behavioral logs, dietitian-assessed nutritional data, and larger samples. A comprehensive interdisciplinary model will enhance clinical decision-making and standardization.

Conclusion

This case series shows that a structured, multicomponent occupational therapy intervention can improve food acceptance and mealtime behaviors in children with sensory food aversion. It also empowered caregivers with practical tools for home management. Results support the use of individualized, sensory-behavioral approaches in early OT interventions for feeding disorders.

Ethical approval

This study has been approved by the Institutional Ethics Committee (IEC) of Kovai Medical Center and Hospital Limited (KMCH Ethical Clearance IEC No: EC/AP/975/11/2022).

Funding

This research received no external funding.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Kowsalya M: conceptualization, methodology, formal analysis, investigation, data curation, writing: original draft, review and edit; **Sujata Missal:** conceptualization, methodology, formal analysis, supervision, validation, writing: review and edit; **Sugi S:** conceptualization, methodology, formal analysis, supervision, validation, writing: review and edit.

Acknowledgements

The authors sincerely thank KMCH College and the Department of Occupational Therapy for their unwavering support throughout this research.

References

[1] Estrem HH, Park J, Thoyre S, McComish C, McGlothen-Bell K. Mapping the gaps: a scoping

review of research on pediatric feeding disorder. Clin Nutr ESPEN. 2022; 49: 1-9. doi.org/10.1016/j.clnesp.2021.12.028.

[2] Galai T, Friedman G, Moses M, Shemer K, Gal DL, Yerushalmy-Feler A, et al. Demographic and clinical parameters are comparable across different types of pediatric feeding disorder. Sci Rep. 2022; 12(1): 8596. doi.org/10.1038/s41598-022-12562-1.

[3] Silverman AH. Interdisciplinary care for feeding problems in children. Nutr Clin Pract. 2010; 25(2): 160-5. doi.org/10.1177/0884533610361609.

[4] Stanislavskaia J. Behavioral feeding problems of normally developing children under 4 years of age [master's thesis]. London (ON): University of Western Ontario; 2014. Available from: https://ir.lib.uwo.ca/etd/2331.

[5] Chatoor I. Feeding disorders in infants and toddlers: diagnosis and treatment. Child Adolesc Psychiatr Clin N Am. 2002; 11(2): 163-83. doi.org/10.1016/S1056-4993(01)00002-5.

[6] Silverman AH. Behavioral management of feeding disorders of childhood. Ann Nutr Metab. 2015; 66(Suppl 5): 33-42. doi.org/10.1159/000381375.

[7] Kivilcim M, Dogan DG, Kortay Canaloglu S, Varol FI, Selimoglu MA. Assessment of behavioral feeding difficulties in young children. Ann Med Res. 2019; 26(10): 2459-65. doi.org/10.5455/annalsmedres.2019.06.320.

[8] Chilman L, Kennedy Behr A, Frakking T, Swanepoel L, Verdonck M. Picky eating in children: a scoping review to examine its intrinsic and extrinsic features and how they relate to identification. Int J Environ Res Public Health. 2021; 18(17): 9067. doi.org/10.3390/ijerph18179067.

[9] Angell A. Selective eaters and tactile sensitivity: a review of classification and treatment methods that address anxiety and support a child's need for a sense of control. ICAN Infant Child Adolesc Nutr. 2010; 2(5): 299-303. doi.org/10.1177/1941406410382904.

[10] Suarez MA, Bush E. Pilot study of the just right challenge feeding protocol for treatment of food selectivity in children. Open J Occup Ther. 2020; 8(2): 1-9. doi.org/10.15453/2168-6408.1648.

[11] Steinberg C. Feeding disorders of infants, toddlers, and preschoolers. BC Med J. 2007;49(4):183-6.

[12] Suarez MA. Multicomponent treatment for food selectivity in children: description and case report. Nutr Clin Pract. 2015; 30(3): 425-31. doi.org/10.1177/0884533614553638.

[13] Miyajima A, Tateyama K, Fuji S, Nakaoka K, Hirao K, Higaki K. Development of an intervention programme for selective eating in children with autism spectrum disorder. Hong Kong J Occup Ther. 2017; 30(1): 22-32. doi.org/10.1016/j.hkjot.2017.10.001.

[14] Hoyo V, Kadlec MB. From syringe to spoon feeding: a case report of how occupational therapy treatment successfully guided the parents of a child with autism spectrum disorder and prematurity in an outpatient clinic. J Autism Dev Disord. 2021; 51(7): 2561-5. hdoi.org/10.1007/s10803-020-04713-y.

[15] Alaimo C, Seiverling L, Anderson K, Sturmey P. A comparison of two multicomponent treatment packages for food refusal. Behav Intervent. 2018; 33(1): 93-101. doi.org/10.1002/bin.1503.

[16] Addison LR, Piazza CC, Patel MR, Bachmeyer MH, Rivas KM, Milnes SM, et al. A comparison of sensory integrative and behavioral therapies as treatment for pediatric feeding disorders. J Appl Behav Anal. 2012; 45(3): 455-71. doi.org/10.1901/jaba.2012.45-455.

[17] Pichardo D, Franke K, Smith HM, Suarez LV, Kozlowski AM. A systematic review of food preference assessments for children with pediatric feeding disorders: a need for modifications and technological descriptions. Behav Dev. 2020; 25(2): 66-87. doi.org/10.1037/bdb0000097.

[18] Notch L, Simenson B. A study examining the effectiveness of a sensory-based feeding group for children with problematic eating behaviors [capstone project]. Grand Forks (ND): University of North Dakota; 2014. Available from: https://commons.und.edu/ot-grad/329.

Journal of Associated Medical Sciences 2026; 59(3): 151-167

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



Diagnostic accuracy of urine AlereLAM and FujiLAM assays for diagnosing active tuberculosis in people living with HIV: A systematic review and meta-analysis

Andiesta Asriyah Mariam Saman[1], Puspa Wardhani[2,3,4,5*], Aryati[2,3,4], Lukman Nur Rahman[1]

[1]*Master Program of Basic Medical Science, Faculty of Medicine, Universitas Airlangga, Surabaya, East Java, Indonesia.*
[2]*Department of Clinical Pathology, Faculty of Medicine, Universitas Airlangga, Surabaya, East Java, Indonesia.*
[3]*Department of Clinical Pathology, Dr. Soetomo General Academic Hospital, Surabaya, East Java, Indonesia.*
[4]*Institute of Tropical Disease, Universitas Airlangga, Surabaya, East Java, Indonesia.*
[5]*Postgraduate School, Universitas Airlangga, Surabaya, East Java, Indonesia.*

ARTICLE INFO

Article history:
Received 5 February 2026
Accepted as revised 10 May 2026
Available online 23 June 2026

Keywords:
Tuberculosis, people living with HIV, urine lipoarabinomannan, AlereLAM, FujiLAM, diagnostic accuracy

ABSTRACT

Background: People living with HIV (PLHIV) remain at high risk of active tuberculosis (TB), yet diagnosis is often challenging due to difficulty obtaining sputum and reduced sensitivity of conventional tests. Urine-based lipoarabinomannan (LAM) assays offer a rapid, non-sputum diagnostic approach. The more recent Fujifilm SILVAMP TB LAM (FujiLAM) may provide better diagnostic performance than the extensively used Alere Determine TB-LAM (AlereLAM).

Objectives: To evaluate the diagnostic accuracy of urine AlereLAM and FujiLAM for diagnosing active TB in PLHIV, and to compare the pooled diagnostic performance of AlereLAM versus FujiLAM.

Materials and methods: We conducted a systematic review and a meta-analysis of diagnostic test accuracy in accordance with PRISMA-DTA guidelines. PubMed, Scopus, and SpringerLink were searched for studies published between 2016 and 2025. Eligible studies enrolled PLHIV evaluated for suspected active TB and assessed urine AlereLAM and/or FujiLAM against microbiological reference standards (MRS) and/or composite reference standards (CRS), with MRS prioritized. Diagnostic 2×2 data were extracted or reconstructed. Pooled sensitivity, specificity, likelihood ratios, and diagnostic odds ratios were estimated using a bivariate random-effects model, with separate analyses for AlereLAM and FujiLAM.

Results: Twenty-five diagnostic accuracy studies were included. AlereLAM demonstrated a pooled sensitivity of 0.364 (95% CI 0.272-0.467) and specificity of 0.928 (95% CI 0.875-0.960). In contrast, FujiLAM showed a substantially higher pooled sensitivity of 0.642 (95% CI 0.567-0.711) while maintaining a high specificity of 0.901 (95% CI 0.872-0.924). FujiLAM yielded a higher diagnostic odds ratio (16.34 vs 7.38), a higher positive likelihood ratio (6.49 vs 5.06), and a lower negative likelihood ratio (0.40 vs 0.69) than AlereLAM. Summary receiver operating characteristic (SROC) analyses demonstrated superior overall diagnostic accuracy for FujiLAM.

Conclusion: Urine FujiLAM showed markedly higher sensitivity than AlereLAM while maintaining high specificity for diagnosing active TB in PLHIV. These findings suggest that next-generation urine LAM assays may substantially improve TB case detection in this high-risk population, particularly in settings where sputum-based diagnostics are limited.

*** Corresponding contributor.**
Author's Address: *Department of Clinical Pathology, Faculty of Medicine, Universitas Airlangga, Surabaya, East Java, Indonesia.*
E-mail address: puspa-w-2@fk.unair.ac.id
doi: 10.66285/JAMS.2026.089
E-ISSN: 2539-6056

Introduction

Tuberculosis (TB) remains a major global public health problem and the leading cause of death from a single infectious agent worldwide, with an estimated 10.7 million incident cases and 1.23 million deaths reported in 2024.[1] The burden of TB is highly concentrated in high-burden countries, accounting for 87% of global cases, and is driven by key determinants such as poverty, undernutrition, diabetes, smoking, and HIV infection.[1] HIV infection markedly increases the risk of progression from TB infection to active disease and contributes substantially to TB-related mortality.[1] Globally, approximately 150,000 TB deaths occur among people living with HIV (PLHIV), highlighting the urgent need for improved strategies for early and accurate TB diagnosis in this vulnerable population.[1] Rapid and accurate diagnosis is a cornerstone of effective TB care and control.[2,3] Global coverage is still inadequate despite the introduction of WHO-recommended rapid diagnostic tests (RDTs), such as lateral flow urine lipoarabinomannan (LF-LAM) assays, to enhance TB detection. In 2024, WRDs were used as the initial diagnostic test in only 54% of newly diagnosed TB cases, with marked regional and country-level disparities, highlighting persistent gaps in access to rapid TB diagnostics, particularly in high-burden and HIV-affected settings.[1,3] Diagnostic challenges are particularly pronounced among PLHIV, in whom TB frequently presents with smear-negative or extrapulmonary disease and low mycobacterial burden, substantially reducing the sensitivity of conventional sputum-based diagnostic methods.[4] Urine-based LF-LAM assays have therefore emerged as a promising point-of-care, non-sputum diagnostic approach for TB in PLHIV, especially among those with advanced immunosuppression. However, evidence on diagnostic accuracy and optimal use remains heterogeneous.[2,3]

Lipoarabinomannan (LAM) is a lipopolysaccharide component of the *Mycobacterium tuberculosis* cell wall that can be released into the bloodstream and subsequently detected in urine during active TB disease.[4] Detection of urinary LAM is biologically plausible, particularly among PLHIV, in whom disseminated disease, higher mycobacterial burden, and increased glomerular permeability are more common.[2,4] Unlike sputum-based diagnostics, urine-based LF-LAM assays do not depend on sputum production. They can be performed at the point of care, making them particularly attractive for severely ill or immunocompromised patients and for use in resource-limited settings.[2]

Alere Determine TB LAM Ag (AlereLAM) is currently the only commercially available LF-LAM assay.[2] It has been incorporated into WHO recommendations for selected subgroups of PLHIV, although its diagnostic sensitivity is recognized to be suboptimal and variable across clinical settings.[2] It is a point-of-care lateral flow assay that provides qualitative results interpreted visually using a reference card with graded intensity bands.[2]

More recently, Fujifilm SILVAMP TB LAM (FujiLAM) has been developed as a next-generation urine-based LF-LAM assay.[2,11] This assay incorporates high-affinity monoclonal antibodies and a silver amplification step, which enhances signal detection and enables identification of lower concentrations of LAM.[11] In contrast to AlereLAM, which is based on conventional lateral flow immunoassay technology using polyclonal antibodies, these features may contribute to differences in analytical performance.[11] Results from FujiLAM are also interpreted visually.[2,11]

In practical use, both assays are rapid, non-sputum-based, and suitable for point-of-care settings. FujiLAM has been reported to show higher sensitivity than AlereLAM while maintaining high specificity, which may be particularly relevant in PLHIV, in whom early and accurate detection of active tuberculosis remains challenging. Despite these promising characteristics, evidence on the diagnostic accuracy of both assays remains heterogeneous, reflecting differences in study populations, disease severity, reference standards, and clinical settings. This heterogeneity underscores the need for a systematic review and meta-analysis to comprehensively synthesize and compare their diagnostic performance. Therefore, this systematic review and diagnostic test accuracy meta-analysis aimed to summarize and compare the diagnostic accuracy of urine AlereLAM and FujiLAM assays for diagnosing active tuberculosis in people living with HIV.

Materials and methods

Protocol and reporting standards

This systematic review and diagnostic test accuracy (DTA) meta-analysis was conducted in accordance with a prespecified protocol. The reporting of this review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Diagnostic Test Accuracy (PRISMA-DTA) statement.[5] The study selection process is presented in a PRISMA flow diagram. The protocol has been registered in the PROSPERO database (CRD420261294478).

Information source and search strategy

A systematic literature search was conducted to identify studies evaluating the diagnostic accuracy of urine lipoarabinomannan assays for diagnosing active TB among PLHIV. Data were gathered from PubMed, Scopus, and SpringerLink by searching for key terms related to urine lipoarabinomannan assays and tuberculosis, including Fujifilm SILVAMP TB LAM (FujiLAM), SILVAMP-LAM, Determine TB-LAM (AlereLAM), lipoarabinomannan, and *Mycobacterium tuberculosis* or tuberculosis. Searches were limited to studies published in English between 2016 and 2025. The last search was conducted on 9 January 2026. The complete search strategies are provided in the Supplementary.

Eligibility criteria

We included full-text English-language studies enrolling PLHIV evaluated for suspected active TB. The index tests were urine Alere Determine TB-LAM (AlereLAM) and/or Fujifilm SILVAMP TB LAM (FujiLAM). Reference standards included microbiological reference standards (MRS) and/or composite reference standards (CRS). Studies were required to provide diagnostic 2×2 data or sufficient information to reconstruct these data. Eligible study designs included cohort, cross-sectional diagnostic accuracy, and case-control studies.

We restricted the search to studies published between 1 January 2016 and 31 December 2025 to reflect contemporary evidence on urine LAM assays. This time frame was selected because recent studies more accurately reflect current diagnostic technologies, including the development and evaluation of newer-generation assays such as FujiLAM, and align with evolving WHO recommendations on the use of LAM-based diagnostics in people living with HIV.

We excluded non-human studies, studies using non-urine specimens or non-eligible index tests, studies with inadequate reference standards, and publications without extractable 2×2 data. Review articles, meta-analyses, case reports or case series, editorials, commentaries, and conference abstracts were also excluded.

Study selection

All records retrieved from database searches were independently screened by two reviewers (AS, LR) against the eligibility criteria. Titles and abstracts were screened first, followed by full-text assessment of potentially relevant articles. Any disagreements were resolved by consensus. The final set of included studies was determined after full-text review.

Data extraction

Data were extracted independently by two reviewers (AS, LR) using a standardized extraction form. Extracted variables included study characteristics (author, year, country, setting, and design), participant characteristics, index test details (AlereLAM and/or FujiLAM), reference standards (MRS or CRS), and diagnostic 2×2 outcomes. For each study, true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) were extracted or reconstructed from reported sensitivity and specificity with denominators when required. In studies enrolling participants with and without HIV infection, data extraction was restricted to the HIV-positive subgroup. Any discrepancies were resolved by consensus.

Index tests and reference standards

The study population comprised PLHIV evaluated for suspected active TB in inpatient and/or outpatient settings. The index tests were urine AlereLAM and FujiLAM. Diagnostic performance was assessed against MRS and/or CRS, with MRS prioritized when both were available. MRS was defined as mycobacterial culture and/or nucleic acid amplification tests (NAAT), such as Xpert MTB/RIF or Xpert Ultra, performed on sputum and/or other clinical specimens. CRS was defined as microbiological evidence and/or compatible histopathology or cytology and/or guideline-based clinical diagnosis supported by imaging findings and treatment response.

Quality assessment

The methodological quality of included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool. Two reviewers (AS, LR) independently evaluated risk of bias across four domains: patient selection, index test, reference standard, and flow and timing. Applicability concerns were assessed for the first three domains. Any discrepancies were resolved by consensus. Risk of bias assessments were summarized and visualized using the Robvis web application (https://mcguinlu.shinyapps.io/robvis/).

Statistical analysis

Diagnostic accuracy data were summarized using 2×2 contingency tables, including TP, FP, FN, and TN. A diagnostic test accuracy meta-analysis was conducted using the MetaDiSc2 web application. Pooled sensitivity and specificity with 95% confidence intervals were estimated using a bivariate random-effects model. Results were presented in tabular form, and forest plots of sensitivity and specificity were generated for individual studies. When both MRS and CRS were reported, MRS-based estimates were preferentially included in the meta-analysis. Overall test performance was visualized using receiver operating characteristic (ROC) plane plots and summary receiver operating characteristic (SROC) curves, including confidence and prediction regions.

Results

Study selection

The systematic search identified 157 records, including 65 from Scopus, 57 from PubMed, and 35 from SpringerLink. After removal of 51 duplicate records, 106 studies were screened by title and abstract, of which 73 were excluded for not meeting the inclusion criteria. The remaining 33 studies were subsequently assessed in full text for eligibility. Eight studies were excluded after full-text review for insufficient data on diagnostic accuracy (N=4), inappropriate study populations (N=3), or incorrect sample types (N=1). Ultimately, 25 studies fulfilled all eligibility criteria and were included in the systematic review and diagnostic test accuracy meta-analysis. The study selection process is summarized in the PRISMA flow diagram (Figure 1).

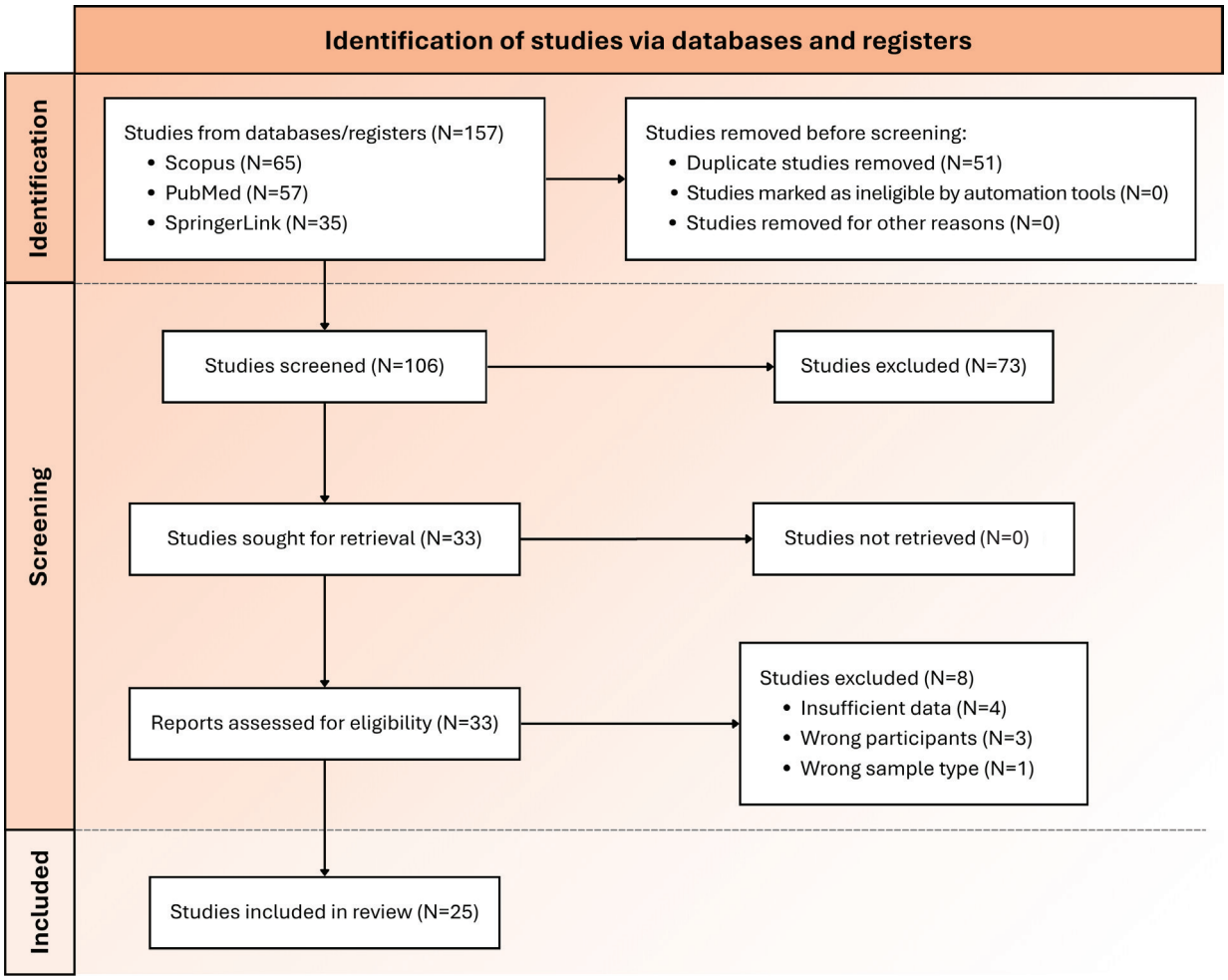


Figure 1. *PRISMA flow diagram.*

Study characteristics

A total of 25 studies on diagnostic accuracy were included. Six studies evaluated both AlereLAM and FujiLAM, 11 evaluated AlereLAM only, and 9 evaluated FujiLAM only. One FujiLAM-only study reported separate diagnostic accuracy results for fresh and stored urine samples, which were extracted as distinct datasets for analysis. All studies provided extractable or constructible 2×2 diagnostic accuracy data and included inpatient and/or outpatient populations (Table 1).

Table 1. Characteristics of all studies included.

Author, year	Country	Study design	Setting	Median age	Sex (male%)	TB type	Reference standard	Sample size	Median CD4 (cells/μL)	Index test (s)	TP	FP	TN	FN
Hanifa *et al.* 2016	South Africa	Prospective	OP	39 (32-45)	39.2	PTB	MRS	404	111 (56–161)	AlereLAM	3	5	363	37
Lawn *et al.* 2017	South Africa	Prospective	IP	36,1 (28.9-42.4)	39.3	PTB, EPTB	MRS	413	149 (55-312)	AlereLAM	53	3	274	86
Sahle *et al.* 2017	Ethiopia	Cross-sectional	OP	30 (23-45)	56	PTB	MRS	21	NR	AlereLAM	5	0	12	4
Boyles *et al.* 2018	South Africa	Prospective	IP	36.2 (29.7-42.1)	33.7	PTB, EPTB	MRS	332	107 (38-118)	AlereLAM	60	11	152	109
LaCourse *et al.* 2018	Kenya	Prospective	IP	24 (13-58)	54.6	PTB	MRS	129	NR	AlereLAM	3	11	111	4
Broger *et al.* 2019	South Africa	Prospective	IP	35 (30-42)	46	PTB, EPTB	MRS	968	86 (33-190)	AlereLAM	268	18	350	332
										FujiLAM	455	33	335	145
Songkhla *et al.* 2019	Thailand	Prospective	IP, OP	39 (32-48)	62.5	PTB, EPTB	MRS	280	33 (12-68)	AlereLAM	54	50	158	18
Andama *et al.* 2020	Uganda	Cross-sectional	OP	32 (25-39)	70	PTB	MRS	116	NR	AlereLAM	10	1	51	54
Bjerrum *et al.* 2020	Ghana	Prospective	IP, OP	38 (31-45)	66.5	PTB, EPTB	MRS	450	152 (44-349)	AlereLAM	35	17	367	31
										FujiLAM	49	41	343	17
Esmail *et al.* 2020	South Africa, Tanzania, Zambia, Zimbabwe	Post-hoc	IP	36 (18-75)	52.4	PTB, EPTB	MRS	561	99.5 (IQR NR)	AlereLAM	68	45	110	338
Garcia *et al.* 2020	Guatemala	Prospective	IP, OP	36 (IQR NR)	74.5	PTB, EPTB	CRS	292	NR	AlereLAM	37	30	202	22
Comella-Del-Barrio *et al.* 2021	Nigeria	Cross-sectional	OP	NR	46.6	PTB	MRS	70	NR	FujiLAM	7	4	56	3
Muyoyeta *et al.* 2021	Zambia	Cross-sectional	OP	37 (28-43)	58.9	PTB	MRS	68	NR	FujiLAM	9	6	50	3

Table 1. Characteristics of all studies included. (continue).

Author, year	Country	Study design	Setting	Median age	Sex (male%)	TB type	Reference standard	Sample size	Median CD4 (cells/μL)	Index test (s)	TP	FP	TN	FN
Nicol *et al.* 2021	South Africa	Prospective	IP	3.8 (1.8-7.4)	54	PTB	MRS	40	NR	AlereLAM	9	8	7	16
										FujiLAM	15	1	14	10
Nkereuwem *et al.* 2021	Gambia, Mali, Nigeria, Tanzania	Cross-sectional	OP	5.6 (2.3-9.3)	54	PTB	MRS	61	NR	AlereLAM	5	9	37	10
										FujiLAM	8	11	35	7
Bjerrum *et al.* 2022	Ghana	Prospective	IP, OP	38 (31-45)	34	PTB, EPTB	MRS	389	176 (52-361)	FujiLAM	29	34	312	14
											30	38	308	13
Indirawati *et al.* 2022	Indonesia	Cross-sectional	IP	NR	82.2	EPTB	MRS	62	44.5 (IQR NR)	FujiLAM	12	12	34	4
Huerga *et al.* 2023	Kenya, Mozambique, South Africa, Uganda	Prospective	OP	41 (33-50)	47.7	PTB, EPTB	MRS	1106	NR	AlereLAM	46	133	859	68
										FujiLAM	68	135	857	46
Reeve *et al.* 2024	South Africa	Cross-sectional	OP	32 (26-39)	41	PTB	MRS	732	299 (170-483)	AlereLAM	14	7	628	83
Székely *et al.* 2024	Malawi, South Africa, Tanzania, Thailand, Uganda, Vietnam, Zambia,	Prospective	IP, OP	40 (IQR NR)	48	PTB, EPTB	MRS	1637	369 (IQR NR)	FujiLAM	160	197	1137	134
Sossen *et al.* 2024	Malawi, South Africa, Tanzania, Thailand, Uganda, Vietnam Zambia	Prospective	IP, OP	40 (33-48)	47.7	PTB, EPTB	MRS	1602	374 (138-630)	AlereLAM	78	129	1219	176

Table 1. Characteristics of all studies included (continue).

Author, year	Country	Study design	Setting	Median age	Sex (male%)	TB type	Reference standard	Sample size	Median CD4 (cells/μL)	Index test (s)	TP	FP	TN	FN
Adzemovic *et al.* 2025	Uganda	Prospective	OP	NR	51	EPTB	MRS	388	NR	FujiLAM	25	16	326	21
Chikamatsu *et al.* 2025	Malawi, South Africa, Tanzania, Thailand, Vietnam, Zambia,	Laboratory-based	NR	NR	NR	PTB, EPTB	MRS	158	NR	AlereLAM	39	6	75	32
										FujiLAM	55	6	80	14
Ellis *et al.*, 2025	Uganda	Prospective	IP	NR	54	PTB, EPTB	MRS	435	34 (11-96)	FujiLAM	36	19	309	71
Tytarenko *et al.* 2025	Nigeria	Retrospective	OP	NR	43.3	PTB	MRS	63	NR	FujiLAM	4	1	55	3
Huerga *et al.* 2023	Kenya, Mozambique, South Africa, Uganda	Prospective	OP	41 (33-50)	47.7	PTB, EPTB	MRS	1106	NR	AlereLAM	46	133	859	68
										FujiLAM	68	135	857	46
Reeve *et al.* 2024	South Africa	Cross-sectional	OP	32 (26-39)	41	PTB	MRS	732	299 (170-483)	AlereLAM	14	7	628	83
Székely *et al.* 2024	Malawi, South Africa, Tanzania, Thailand, Uganda, Vietnam, Zambia,	Prospective	IP, OP	40 (IQR NR)	48	PTB, EPTB	MRS	1637	369 (IQR NR)	FujiLAM	160	197	1137	134

Note: *TP: true positive, FP: false positive, TN: true negative, FN: false negative, IP: inpatient, OP: outpatient, PTB: pulmonary tuberculosis, EPTB: extrapulmonary tuberculosis, MRS: microbiological reference standard, CRS: composite reference standard, NR: not reported.*

Quality assessment

Most included studies were judged to be at low risk of bias across the QUADAS-2 domains (supplementary). Low risk was predominantly observed in the patient selection, index test, and reference standard domains. A small number of studies showed high risk, some concerns, or unclear risk, mainly related to non-consecutive sampling, case-control study designs, insufficient reporting of index test procedures, or the use of composite reference standards. In the flow and timing domain, most studies were assessed as low risk, although a few raised concerns due to incomplete verification or participant exclusions. Overall, most studies were classified as low risk of bias, with only a limited number rated as having some concerns or high risk.

Diagnostic accuracy of AlereLAM

The diagnostic accuracy of urine AlereLAM for detecting active TB in PLHIV was evaluated across the included studies. Individual study estimates demonstrated substantial variability in sensitivity, whereas specificity was consistently high across studies (Figure 2). In the meta-analysis, AlereLAM showed a pooled sensitivity of 0.364 (95% CI 0.272-0.467) and a pooled specificity of 0.928 (95% CI 0.875-0.960). The pooled positive likelihood ratio (LR+) was 5.06 (95% CI 2.95-8.68), while the negative likelihood ratio (LR-) was 0.69 (95% CI 0.59-0.79), with a diagnostic odds ratio (DOR) of 7.38 (95% CI 3.99-13.65). Between-study heterogeneity was substantial, with an I^2 of 91.8% for sensitivity and 93.0% for specificity. The SROC curve is shown in Figure 3, with the summary operating point reflecting these pooled estimates.

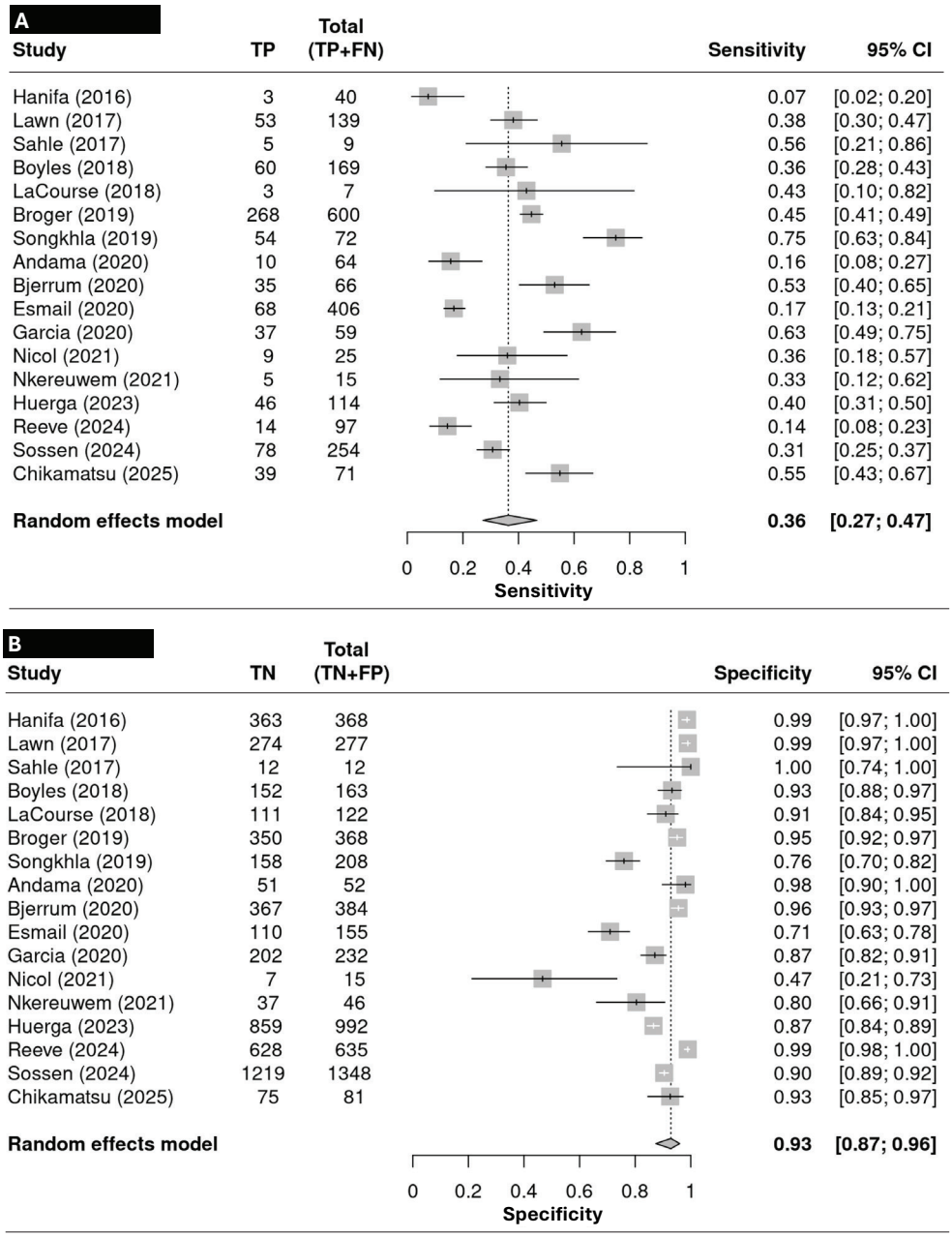


Figure 2. *Forest plots of the sensitivity and specificity of AlereLAM for the diagnosis of active TB in PLHIV.*

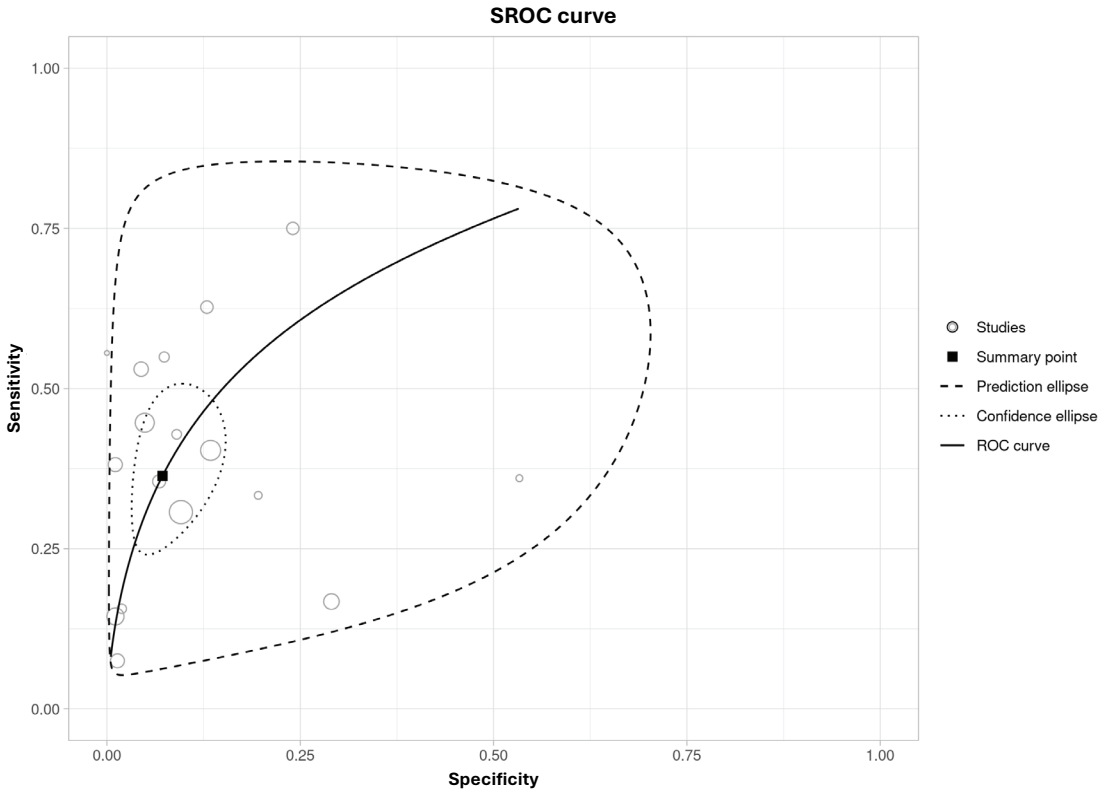


Figure 3. *Summary receiver operating characteristic (SROC) curve of AlereLAM for the diagnosis of active TB in PLHIV.*

Diagnostic accuracy of FujiLAM

The diagnostic accuracy of urine FujiLAM for detecting active TB in PLHIV was assessed across the included studies. FujiLAM demonstrated high sensitivity and specificity, with substantial variability across studies (Figure 4). In the meta-analysis, FujiLAM showed a pooled sensitivity of 0.642 (95% CI 0.567-0.711) and a pooled specificity of 0.901 (95% CI 0.872-0.924). The pooled LR+ was 6.49 (95% CI 4.99-8.43), and the LR- was 0.40 (95% CI 0.33-0.48), corresponding to a DOR of 16.34 (95% CI 11.13-23.98). Between-study heterogeneity was substantial, with an I^2 of 86% for sensitivity and 79% for specificity. The SROC curve (Figure 5) illustrates FujiLAM's overall diagnostic performance, with a summary operating point reflecting the pooled estimates.

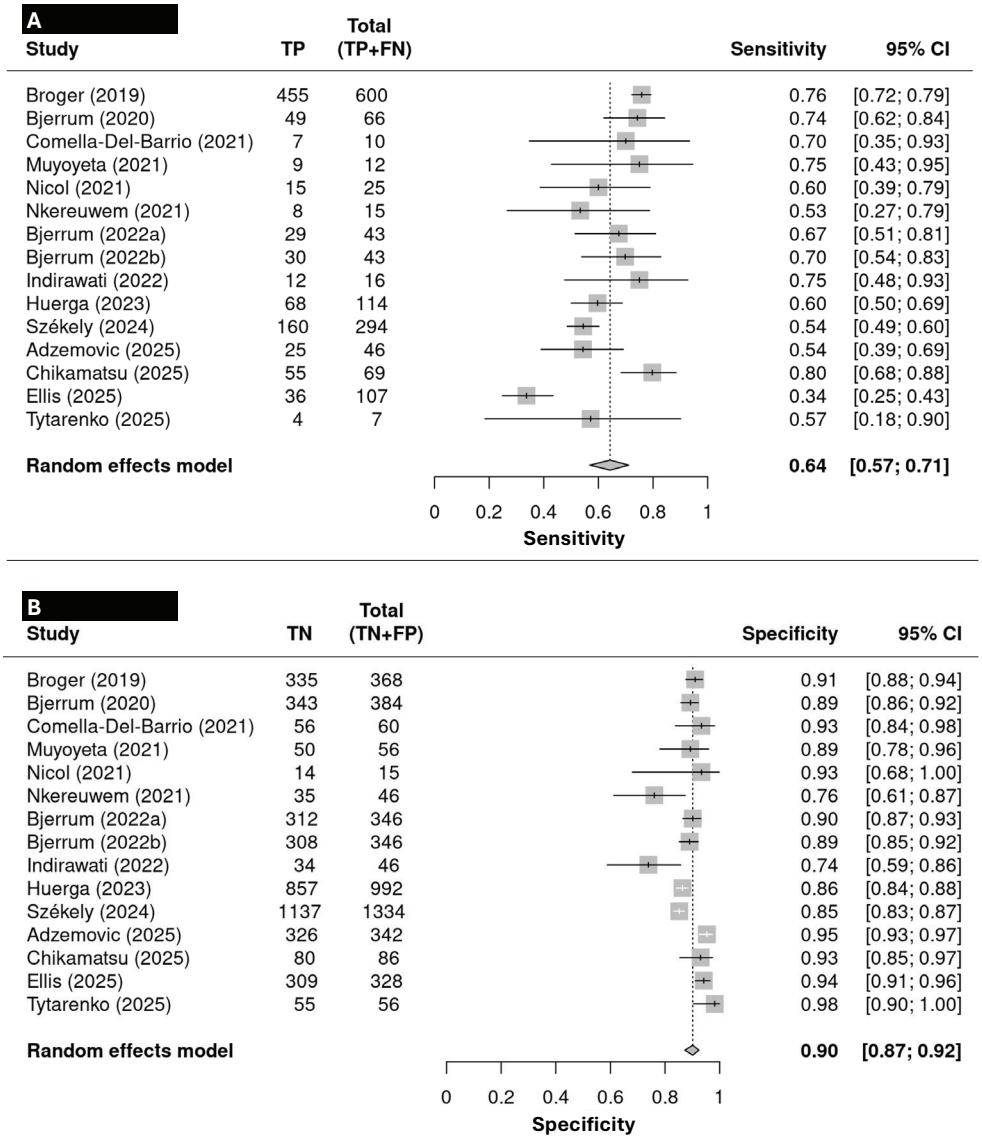


Figure 4. *Forest plots of sensitivity and specificity of FujiLAM for the diagnosis of active TB in PLHIV.*

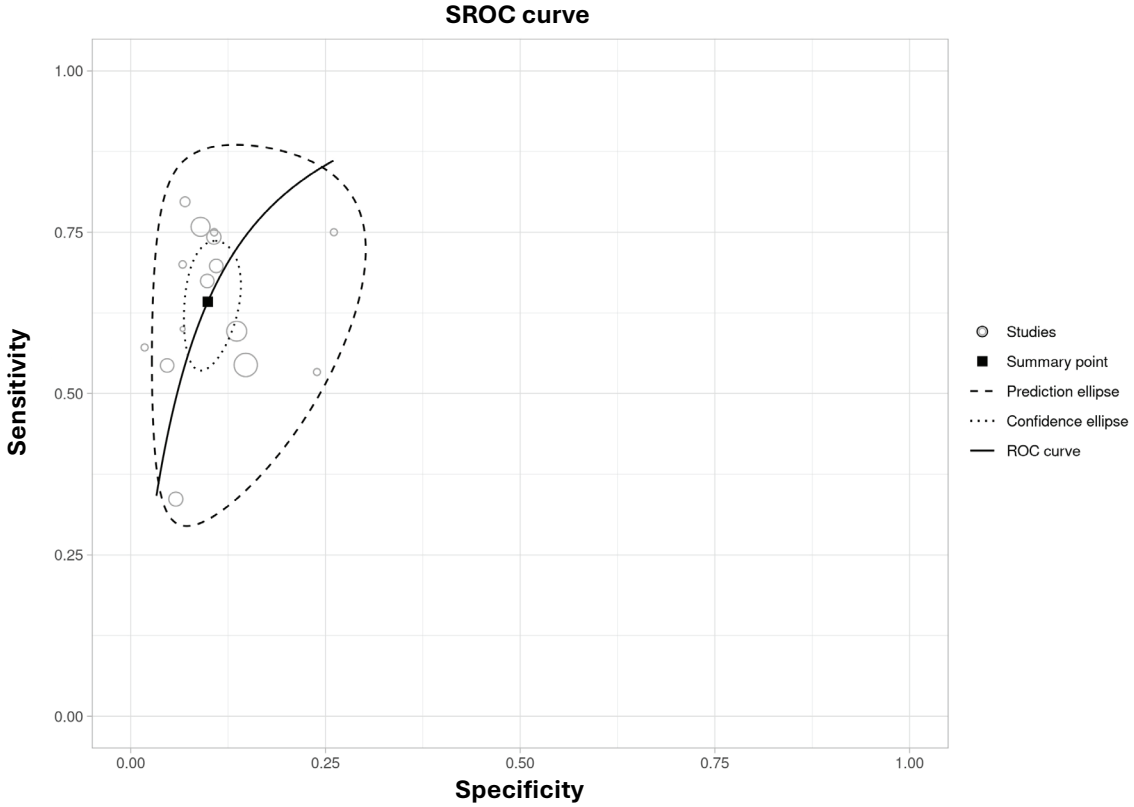


Figure 5. *Summary receiver operating characteristic (SROC) curve of FujiLAM for the diagnosis of active TB in PLHIV.*

Comparison of diagnostic accuracy between AlereLAM and FujiLAM

A comparative analysis of urine AlereLAM and FujiLAM demonstrated clear differences in diagnostic performance among PLHIV evaluated for suspected active TB. Overall, FujiLAM showed substantially higher pooled sensitivity than AlereLAM (0.642 vs 0.364), while maintaining comparably high specificity (0.901 vs 0.928). Consistent with these findings, FujiLAM yielded a higher DOR (16.34 vs 7.38), a higher LR+ (6.49 vs 5.06), and a lower LR- (0.40 vs 0.69) compared with AlereLAM. On the ROC plane, the summary operating point for FujiLAM was positioned closer to the upper-left corner than that for AlereLAM, indicating improved overall diagnostic accuracy (Figure 6). While both assays demonstrated between-study heterogeneity, variability appeared less pronounced for FujiLAM, as indicated by narrower confidence and prediction regions on the SROC curve.

Subgroup analyses were conducted to compare the diagnostic performance of AlereLAM and FujiLAM across CD4 strata and clinical settings. Pooled sensitivity and specificity estimates were summarized for each assay where two or more studies were available (Table 2). Across CD4 strata, FujiLAM consistently demonstrated higher pooled sensitivity than AlereLAM, while pooled specificity remained high for both assays. The difference in sensitivity between assays was most pronounced at lower CD4 counts (<100 cells/μL), whereas specificity showed minimal variation across CD4 strata. Similar patterns were observed across clinical settings, with FujiLAM showing higher pooled sensitivity than AlereLAM in both inpatient and outpatient settings, while pooled specificity remained comparable between assays.

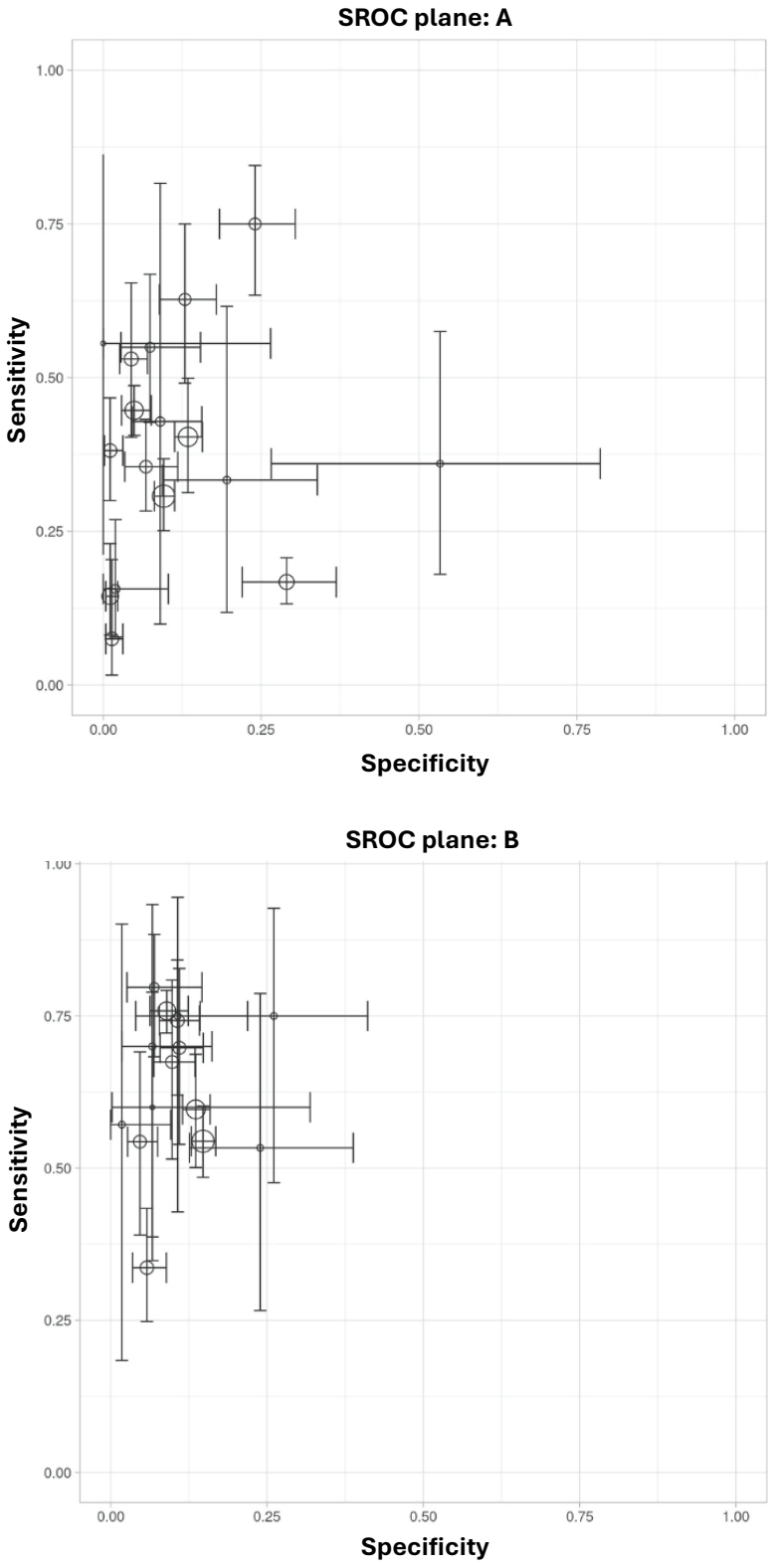


Figure 6. *Comparison of AlereLAM and FujiLAM diagnostic accuracy on the SROC plane.*

Table 2. Subgroup analysis of pooled sensitivity and specificity of AlereLAM and FujiLAM assays by CD4 count and clinical setting among people living with HIV.

Variables	Sub groups	AlereLAM			FujiLAM		
		Studies (total samples)	Pooled sensitivity (%) (95% CI)	Pooled specificity (%) (95% CI)	Studies (total samples)	Pooled sensitivity (%) (95% CI)	Pooled specificity (%) (95% CI)
Median CD4 count (cells/μL)	<100	3 (1809)	44 (18-73)	84 (65-94)	3 (1465)	62 (37-82)	89 (79-95)
	100-200	4 (1600)	32 (16-52)	97 (94-99)	2 (839)	71 (63-78)	89 (88-91)
	>200	2 (2334)	22 (13-36)	97 (85-99)	1 (1637)	NA	NA
By settings	IP	6 (2443)	34 (25-44)	90 (74-97)	4 (1505)	61 (42-78)	90 (81-95)
	OP	6 (2440)	23 (13-38)	97 (90-99)	6 (1756)	59 (52-66)	91 (85-95)

Note: *For subgroups including a single study, pooled estimates are not applicable (NA), and values represent results from the individual study.*

Discussion

This systematic review and diagnostic test accuracy meta-analysis evaluated and compared the performance of urine-based LAM assays (AlereLAM and FujiLAM) for the diagnosis of active TB among people living with HIV. Across the included studies, FujiLAM demonstrated higher pooled sensitivity than AlereLAM, while pooled specificity remained high for both assays. AlereLAM showed modest pooled sensitivity but high specificity, indicating that positive test results are more informative than negative results, consistent with its established role as a rule-in assay, particularly in individuals with advanced HIV disease.

The pooled sensitivity of FujiLAM was higher than that of AlereLAM, with correspondingly higher diagnostic odds ratios and more favorable likelihood ratios. The improved diagnostic performance of FujiLAM is biologically plausible, given its use of a high-affinity monoclonal antibody and silver amplification technology, which may enable detection of lower urinary LAM concentrations.[8,13] Between-study heterogeneity was observed for both assays and may reflect differences in patient populations, CD4 cell counts, clinical settings, and reference standards. Despite higher pooled sensitivity, FujiLAM's negative predictive value remains insufficient to exclude TB when used in isolation.

Current World Health Organization policy recommendations conditionally endorse the use of urine-based LAM assays, particularly among PLHIV with advanced immunosuppression[2]. Within this context, the pooled diagnostic accuracy estimates presented in this review provide updated evidence on the performance of both AlereLAM and FujiLAM in relation to existing policy guidance. The World Health Organization target product profiles for tuberculosis diagnostics emphasize the need for non-sputum-based tests with improved sensitivity in this population.[2,3] Overall, urine LAM assays should be interpreted as part of a comprehensive diagnostic approach that includes microbiological testing and clinical assessment, consistent with their positioning as adjunctive tools within TB diagnostic algorithms for PLHIV.[4]

Limitations

This systematic review and meta-analysis of diagnostic test accuracy have several limitations. Firstly, substantial between-study heterogeneity was observed for both AlereLAM and FujiLAM, likely reflecting differences in study populations, clinical settings, disease severity, and reference standards, which may limit the precision of pooled estimates. Secondly, the use of composite reference standards in some studies may have introduced misclassification bias, particularly where microbiological confirmation was incomplete. Thirdly, inclusion of case-control study designs may have led to spectrum bias and potential overestimation of diagnostic accuracy, and incomplete reporting in some studies limited consistent subgroup analyses. Fourthly, restricting to English-language publications may have introduced language bias, and the relatively small number of studies evaluating FujiLAM may affect the robustness of the pooled estimates. Finally, this review focused on diagnostic accuracy outcomes and did not assess downstream clinical or public health impacts, which should be addressed in future prospective and implementation studies.

Conclusion

In this systematic review and diagnostic test accuracy meta-analysis, the AlereLAM and FujiLAM urine assays demonstrated high pooled specificity for diagnosing active tuberculosis among PLHIV. The pooled sensitivity of FujiLAM was higher than that of AlereLAM, while specificity estimates were comparable between the two assays.

Ethical approval

Ethical approval was not required for this systematic review and meta-analysis because it used data from previously published studies.

Funding

This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare no conflicts of interest.

CRediT authorship contribution statement

Andiesta Asriyah Mariam Saman: conceptualization, data curation, formal analysis, methodology, visualization, writing: original draft, review and edit; **Puspa Wardhani:** methodology, supervision, validation, writing: review and edit; **Aryati:** methodology, supervision, validation, writing: review and edit; **Lukman Nur Rahman:** data curation, formal analysis, writing: review and edit.

Acknowledgements

The authors acknowledge the contributions of the original studies included in this review, which provided the data for this analysis.

References

[1] World Health Organization. Global tuberculosis report 2025. Geneva: World Health Organization; 2025.

[2] World Health Organization. Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV: policy update 2019. Geneva: World Health Organization; 2019.

[3] World Health Organization. Target product profiles for tuberculosis diagnosis and detection of drug resistance. Geneva: World Health Organization; 2024.

[4] Yang Q, Han J, Shen J, Peng X, Zhou L, Yin X. Diagnosis and treatment of tuberculosis in adults with HIV. Medicine (Baltimore). 2022; 101(35): e30405. doi: 10.1097/MD.0000000000030405.

[5] Salameh JP, Bossuyt PM, McGrath TA, et al. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist. BMJ. 2020; 370: m2632. doi: 10.1136/bmj.m2632.

[6] Andama A, Jaganath D, Crowder R, et al. Accuracy and incremental yield of urine Xpert MTB/RIF Ultra versus Determine TB-LAM for diagnosis of pulmonary tuberculosis. Diagn Microbiol Infect Dis. 2020; 96(1): 114892. doi: 10.1016/j.diagmicrobio.2019.114892.

[7] Adzemovic T, Nalintya E, Fieberg A, et al. Diagnostic accuracy of the updated FujiLAM II assay to detect tuberculosis in outpatients with advanced HIV disease. Clin Infect Dis. 2025; 81(4): e137-e142. doi: 10.1093/cid/ciaf079.

[8] Bjerrum S, Ahsberg J, Szekely R, et al. Diagnostic accuracy of urine lipoarabinomannan testing in early morning urine versus spot urine for diagnosis of tuberculosis among people with HIV. Microbiol Spectr. 2022; 10(2): e0020822. doi: 10.1128/spectrum.00208-22.

[9] Bjerrum S, Broger T, Szekely R, et al. Diagnostic accuracy of a novel and rapid lipoarabinomannan test for diagnosing tuberculosis among people with human immunodeficiency virus. Open Forum Infect Dis. 2019; 7(1): ofz530. doi: 10.1093/ofid/ofz530.

[10] Boyles TH, Griesel R, Stewart A, Mendelson M, Maartens G. Incremental yield and cost of urine Determine TB-LAM and sputum induction in seriously ill adults with HIV. Int J Infect Dis. 2018; 75: 67-73. doi: 10.1016/j.ijid.2018.08.005.

[11] Broger T, Sossen B, du Toit E, Kerkhoff AD, Schutz C, Ivanova Reipold E, et al. Novel lipoarabinomannan point-of-care tuberculosis test for people with HIV: a diagnostic accuracy study. Lancet Infect Dis. 2019; 19(8): 852-61. doi: 10.1016/S1473-3099(19)30001-5.

[12] Chikamatsu K, Shimomura Y, Osugi A, Mitarai S. Assessing the accuracy of the updated Fujifilm SILVAMP TB LAM II assay between lot numbers. IJTLD Open. 2025; 2(2): 96-100. doi: 10.5588/ijtldopen.24.0449.

[13] Comella-del-Barrio P, Bimba JS, Adelakun R, Kontogianni K, Molina-Moya B, Osazuwa O, et al. Fujifilm SILVAMP TB-LAM for the diagnosis of tuberculosis in Nigerian adults. J Clin Med. 2021; 10(11): 2514. doi: 10.3390/jcm10112514.

[14] Ellis J, Dai B, Kabahubya M, Hale G, Mande E, Katende G, et al. Diagnostic accuracy of the re-engineered urinary FujiLAM2 assay among hospitalized adults with advanced HIV disease. AIDS. 2025; 39(10): 1334-43. doi: 10.1097/QAD.0000000000004213.

[15] Friedrich R, Rappold E, Bogdan C, Held J. Comparative analysis of the Wako β-glucan test and the Fungitell assay for diagnosis of candidemia and Pneumocystis jirovecii pneumonia. J Clin Microbiol. 2018; 56(9): e00499-18. doi: 10.1128/JCM.00499-18.

[16] Garcia JI, Melendez J, Alvarez R, Mejia-Chew C, Kelley HV, Sidiki S, et al. Accuracy of the tuberculosis point-of-care Alere Determine lipoarabinomannan antigen diagnostic test using α-mannosidase-treated and untreated urine in people living with HIV in Guatemala. AIDS Res Ther. 2020; 17(1): 60. doi: 10.1186/s12981-020-00318-8.

[17] Hanifa Y, Fielding KL, Chihota VN, Adonis L, Charalambous S, Karstaedt A, et al. Diagnostic accuracy of lateral flow urine LAM assay for tuberculosis screening of adults with advanced immunosuppression attending routine HIV care in South Africa. PLoS One. 2016; 11(6): e0156866. doi: 10.1371/journal.pone.0156866.

[18] Huerga H, Bastard M, Lubega AV, Akinyi M, Antabak NT, Ohler L, et al. Novel FujiLAM assay

to detect tuberculosis in HIV-positive ambulatory patients in four African countries: a diagnostic accuracy study. Lancet Glob Health. 2023; 11(1): e126-e135. doi: 10.1016/S2214-109X(22)00463-6.

[19] Indirawati NN, Yunihastuti E, Yulianti M, Nasir UZ, Wulandari D, Rinaldi I. Lateral flow urine lipoarabinomannan assay for extrapulmonary tuberculosis diagnosis in HIV-positive adults. Int J Infect Dis. 2022; 122: 415-9. doi: 10.1016/j.ijid.2022.06.007.

[20] Lacourse SM, Pavlinac PB, Cranmer LM, Njuguna IN, Mugo C, Gatimu J, et al. Stool Xpert MTB/RIF and urine lipoarabinomannan for the diagnosis of tuberculosis in hospitalized HIV-infected children. AIDS. 2018; 32(1): 69-78. doi: 10.1097/QAD.0000000000001662.

[21] Lawn SD, Kerkhoff AD, Burton R, Schutz C, Boulle A, Vogt M, et al. Diagnostic accuracy, incremental yield, and prognostic value of Determine TB-LAM for routine diagnostic testing for tuberculosis in HIV-infected patients requiring acute hospital admission in South Africa. BMC Med. 2017; 15(1): 67. doi: 10.1186/s12916-017-0822-8.

[22] Muyoyeta M, Kerkhoff AD, Chilukutu L, Moreau E, Schumacher SG, Ruhwald M. Diagnostic accuracy of a novel point-of-care urine lipoarabinomannan assay for the detection of tuberculosis among adult outpatients in Zambia. Eur Respir J. 2021; 58(5): 2003999. doi: 10.1183/13993003.03999-2020.

[23] Nicol MP, Schumacher SG, Workman L, Broger T, Baard C, Prins M, et al. Accuracy of a novel urine test, Fujifilm SILVAMP tuberculosis lipo-arabinomannan, for the diagnosis of pulmonary tuberculosis in children. Clin Infect Dis. 2021; 72(9): e280-e288. doi: 10.1093/cid/ciaa1052.

[24] Nkereuwem E, Togun T, Gomez MP, Szekely R, Mace A, Jobe D, et al. Comparing accuracy of lipoarabinomannan urine tests for diagnosis of pulmonary tuberculosis in children from four African countries: a cross-sectional study. Lancet Infect Dis. 2021; 21(3): 376-84. doi: 10.1016/S1473-3099(20)30598-3.

[25] Reeve BWP, Ndlangalavu G, Mishra H, Palmer Z, Tshivhula H, Rockman L, et al. Point-of-care C-reactive protein and Xpert MTB/RIF Ultra for tuberculosis screening and diagnosis in unselected antiretroviral therapy initiators: a prospective cross-sectional diagnostic accuracy study. Lancet Glob Health. 2024; 12(5): e793-e803. doi: 10.1016/S2214-109X(24)00052-4.

[26] Sahle SN, Asress DT, Tullu KD, Weldemariam AG, Tola HH, Awas YA, et al. Performance of point-of-care urine test in diagnosing tuberculosis suspects with and without HIV infection in selected peripheral health settings of Addis Ababa, Ethiopia. BMC Res Notes. 2017; 10(1): 1-6. doi: 10.1186/s13104-017-2404-4.

[27] Songkhla MN, Tantipong H, Tongsai S, Angkasekwinai N. Lateral flow urine lipoarabinomannan assay for diagnosis of active tuberculosis in adults with human immunodeficiency virus infection: a prospective cohort study. Open Forum Infect Dis. 2019; 6(4): ofz132. doi: 10.1093/ofid/ofz132.

[28] Sossen B, Szekely R, Mukoka M, Muyoyeta M, Nakabugo E, Hella J, et al. Urine Xpert Ultra for the diagnosis of tuberculosis in people living with HIV: a prospective multicentre diagnostic accuracy study. Lancet Glob Health. 2024; 12(12): e2024-e2034. doi: 10.1016/S2214-109X(24)00357-7.

[29] Szekely R, Sossen B, Mukoka M, Muyoyeta M, Nakabugo E, Hella J, et al. Prospective multicentre accuracy evaluation of the Fujifilm SILVAMP TB LAM test for the diagnosis of tuberculosis in people living with HIV demonstrates lot-to-lot variability. PLoS One. 2024; 19(5): e0303846. doi: 10.1371/journal.pone.0303846.

[30] Tytarenko N, Bimba JS, Vulcano D, Comella-del-Barrio P, Osazuwa O, Creswell J, et al. Diagnostic performance of second-generation Fujifilm SILVAMP TB LAM in Nigerian adults with presumptive tuberculosis. J Clin Med. 2025; 14(21): 7524. doi: 10.3390/jcm14217524.

Supplementary
Search strategy for systematic review and diagnostic test accuracy meta-analysis

Search period:

1. Studies published from 1 January 2016 to 31 December 2025.
2. The last search was conducted on 9 January 2026.

Language restriction:

Searches were limited to English-language publications.

Search concepts:

The search strategy was developed using combinations of controlled vocabulary terms and free-text keywords covering the following concepts:

1. Tuberculosis
2. Human immunodeficiency virus (HIV)
3. Urine lipoarabinomannan (LAM) assays, including:
    a. Alere Determine TB-LAM (AlereLAM)
    b. Fujifilm SILVAMP TB LAM (FujiLAM)

Table S1. Search strategy and results for systematic review and diagnostic test accuracy meta-analysis.

Databases	Boolean search strategy	Search result
PubMed	((“Tuberculosis”[Mesh] OR tuberculosis[tiab] OR TB[tiab]) AND (“HIV Infections”[Mesh] OR HIV[tiab] OR “human immunodeficiency virus”[tiab] OR PLHIV[tiab]) AND ((“Lipoarabinomannan”[tiab] OR LAM[tiab]) AND (urine[tiab] OR urinary[tiab])) AND (“Determine TB LAM”[tiab] OR AlereLAM[tiab] OR “Alere Determine TB LAM”[tiab] OR FujiLAM[tiab] OR “Fujifilm SILVAMP TB LAM”[tiab] OR SILVAMP[tiab]) AND (diagnos*[tiab] OR sensitiv*[tiab] OR specific*[tiab] OR accuracy[tiab] OR ROC[tiab]))	57
SCOPUS	TITLE-ABS-KEY ((tuberculosis OR TB) AND (HIV OR “human immunodeficiency virus” OR PLHIV OR “people living with HIV”) AND (urine OR urinary) AND (“lipoarabinomannan” OR LAM) AND (“Determine TB LAM” OR AlereLAM OR FujiLAM OR “Fujifilm SILVAMP TB LAM” OR “SILVAMP TB LAM” OR SILVAMP) AND (diagnos* OR sensitiv* OR specific* OR accuracy OR ROC OR “diagnostic accuracy”))	65
SpringerLink	(“FujiLAM” OR “Fujifilm SILVAMP TB LAM” OR “SILVAMP TB LAM” OR SILVAMP OR “Determine TB LAM” OR AlereLAM OR “Alere Determine TB LAM”) AND (tuberculosis OR TB) AND (HIV OR “human immunodeficiency virus” OR PLHIV) AND (urine OR urinary) AND (LAM OR lipoarabinomannan)	35

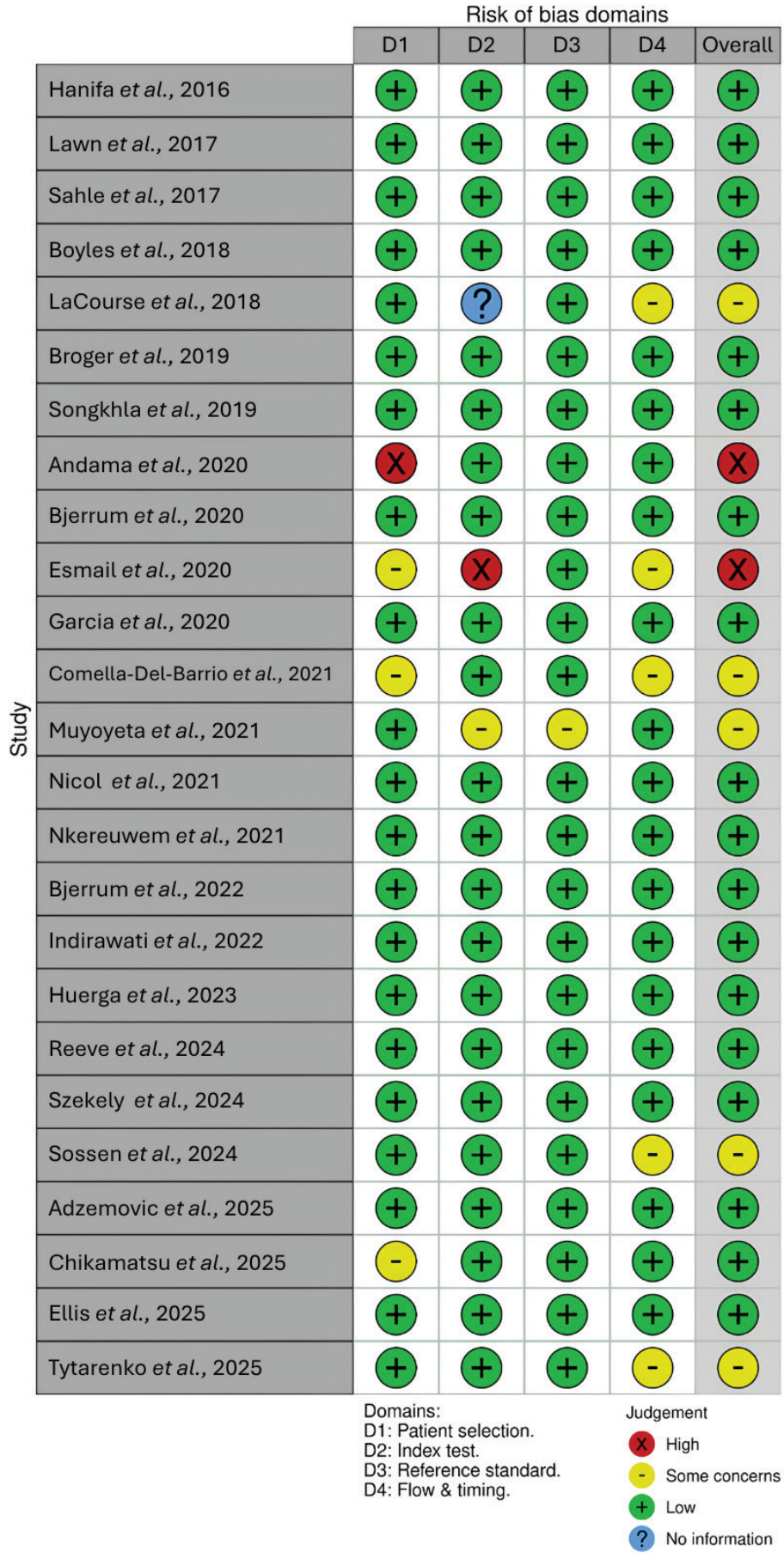


Figure S1. *QUADAS-2 risk of bias assessment of included studie.*

Journal of Associated Medical Sciences 2026; 59(3): 168-176

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 Galectin-3 as a diagnostic and discriminatory biomarker for chronic kidney disease: A case-control study

Asmaa. H. Nasser and Entedhar Rifaat Sarhat

Department of Clinical Biochemistry, College of Medicine, University of Tikrit, Iraq.

ARTICLE INFO

Article history:
Received 11 March 2026
Accepted as revised 19 June 2026
Available online 24 June 2026

Keywords:
Galectin 3, diabetic nephropathies, renal insufficiency, kidney failure

ABSTRACT

Background: Diabetic kidney disease (DKD) is the leading cause of end-stage renal disease worldwide. Conventional markers, such as serum creatinine, estimated glomerular filtration rate, and albuminuria, have limited sensitivity for distinguishing among etiologies of advanced chronic kidney disease (CKD) and become abnormal only after substantial nephron loss. Galectin-3 (GAL-3), a beta-galactoside-binding lectin implicated in renal inflammation and fibrosis, has emerged as a candidate biomarker; however, its diagnostic accuracy in distinguishing DKD from non-diabetic CKD in advanced CKD populations remains unclear.

Objective: To quantify serum GAL-3 concentrations in patients with DKD, non-diabetic CKD, and healthy controls, and to evaluate its diagnostic performance in discriminating between these groups in an advanced CKD population.

Materials and methods: A case-control study that enrolled 128 participants: 32 patients with DKD (all with type 2 diabetes mellitus), 32 with non-diabetic CKD (both groups on regular hemodialysis classified as CKD stage 5/ end-stage renal disease [ESRD]), and 64 healthy controls. Serum GAL-3 levels were measured using an enzyme-linked immunosorbent assay (ELISA). Fasting blood glucose, glycated hemoglobin (HbA1c), renal function tests, and serum albumin levels were assessed using an automated analyzer.

Results: Serum GAL-3 concentrations increased progressively from controls (210.04±34.65 ng/mL) to non-diabetic CKD (262.21±3.19 ng/mL) and then to DKD (368.50±67.10 ng/mL, p<0.001). Receiver operating characteristic (ROC) analysis demonstrated very high discrimination in this cohort between DKD and healthy controls area under the curve (AUC=1.00, 95% CI: 0.963-1.000, sensitivity, 100%, specificity, 100%; cutoff, ≥275.77 ng/mL), excellent differentiation of DKD from non-diabetic CKD (AUC=0.987, 95% CI: 0.905-0.999, sensitivity, 84.37%, specificity, 100%), and strong discrimination of all CKD patients from controls (AUC=0.926, 95% CI: 0.866-0.965; sensitivity, 98.44%, specificity, 73.44%). These cutoff values were derived from the same dataset and should be considered exploratory.

Conclusion: Serum GAL-3 demonstrated very high diagnostic accuracy in identifying and stratifying DKD within this cohort of end-stage renal disease patients on hemodialysis. These exploratory findings suggest that GAL-3 may have potential as a discriminatory biomarker in advanced CKD populations; however, prospective validation in larger, independent cohorts including patients with early-stage CKD is warranted before clinical integration can be recommended.

* **Corresponding contributor.**
Author's Address: *Department of Clinical Biochemistry, College of Medicine, University of Tikrit, Iraq.*
E-mail address: ah230043me@st.tu.edu.iq
doi: 10.66285/JAMS.2026.090
E-ISSN: 2539-6056

Introduction

Chronic kidney disease is now recognized as one of the major non-communicable health challenges worldwide, with current estimates suggesting that more than 800-850 million people are living with kidney diseases of various causes.[1] Within this spectrum, diabetes has emerged as the dominant driver of kidney damage: diabetic kidney disease which is also considered the leading cause of CKD and end-stage renal disease in many countries; approximately one-third of individuals with diabetes eventually develop clinically evident DKD, and large clinical cohorts and registry data consistently show that DKD is the most common indication for dialysis or kidney transplantation, despite advances in glycemic and blood pressure control.[2] DKD is a multifaceted condition characterized by interrelated proinflammatory and pro-apoptotic metabolic pathways that contribute to glomerular hypertension, metabolic disturbances, and hemodynamic irregularities.[3] Both type 1 and type 2 diabetes can lead to chronic kidney disease and, ultimately, end-stage renal disease (ESKD); patients with ESKD often suffer from type 2 diabetes, given that type 2 diabetes is more common than type 1.[4] In the present study, all patients in the DKD group had type 2 diabetes mellitus (T2DM), reflecting the predominance of T2DM as the underlying etiology of DKD in our clinical setting. The global prevalence of diabetes will continue to increase over the coming decades, from 415 million people in 2015 to 642 million in 2040.[5]

The complex interaction of hemodynamic, metabolic, fibrotic, and inflammatory factors, which are often disrupted in diabetes, is believed to be the cause of the development and progression of DKD; dysfunctions in these factors are interconnected and affect the activation of transcription factors, molecular pathways, and gene regulation, leading to structural and functional changes that ultimately result in the clinical manifestations of DKD, characterized by decreased kidney function and elevated albumin levels in the urine[6]. Routine clinical assessment still relies largely on serum creatinine, estimated glomerular filtration rate (eGFR), and albuminuria, which are useful for staging established CKD; however, these markers typically become abnormal only after substantial structural damage and nephron loss have occurred, and both eGFR and albuminuria have limited capacity for discriminating between different etiologies of CKD, particularly in advanced stages where values converge regardless of the underlying cause.[7] This diagnostic limitation is especially problematic in advanced CKD populations, where distinguishing diabetic from non-diabetic etiologies on biochemical grounds alone remains challenging without renal biopsy; these limitations have encouraged an intensive search for novel biomarkers that can capture more specific aspects of kidney injury and enable etiological discrimination, with recent work focusing on markers of tubular damage, fibrosis, inflammation, oxidative stress, and altered signaling pathways.[8]

Galectin-3 (GAL-3) binds to the enzyme beta-galactosidase and is highly expressed in inflammatory and epithelial cells, as well as in kidney and cardiac tissue. It is involved in regulating cell growth during inflammation.[9] GAL-3 has been shown to play a role in the mechanisms underlying chronic kidney disease and acute kidney failure, and its serum levels are negatively correlated with kidney function; high levels of GAL-3 increase the risk of diabetic nephropathy (DN) and predict rapid progression and adverse outcomes in patients with chronic kidney disease prior to dialysis.[10,11] The aim of this study was to quantify and compare serum GAL-3 levels among diabetic and non-diabetic patients with CKD and healthy controls, and to evaluate the diagnostic performance of GAL-3 for discriminating among these groups in an advanced CKD population undergoing hemodialysis.

Materials and methods

This case-control study was conducted in a hospital setting. The study protocol was approved by the Scientific Committee of the College of Medicine, Tikrit University, and the Diyala Health Department. The study included 64 patients of both sexes with chronic kidney disease (32 patients with diabetes and 32 patients without diabetes) who were undergoing regular dialysis sessions and were aged between 20 and 65 years. They were clinically diagnosed with end-stage renal failure by a nephrologist based on their medical history, physical examination, renal function tests, and other relevant laboratory tests. All patients in both the CKD and DKD groups were classified as CKD stage 5 (ESRD; eGFR <15 mL/min/1.73 m^2) and were receiving regular maintenance hemodialysis at the time of enrollment. Patients in the DKD group were exclusively diagnosed with T2DM as the underlying cause of their kidney disease, confirmed by the treating nephrologist based on clinical history, duration of diabetes, presence of diabetic retinopathy, and/or prior documentation of progressive diabetic nephropathy. Patients with type 1 diabetes were not included in this study. The study also included 64 healthy adults of both sexes who had no medical or family history of chronic kidney disease or diabetes as the control group.

Inclusion and exclusion criteria

Patients must have been diagnosed with chronic kidney disease (with or without type 2 diabetes) by a supervising physician and must not have a history of autoimmune diseases (thyroid disease, systemic lupus erythematosus), cancer, liver disease, or heart failure. Patients receiving uric acid-lowering medications (e.g., allopurinol, febuxostat) were not specifically excluded. Pregnant women and individuals with any of the above conditions were excluded from the study.

Data and sample collection

Medical and social history was obtained from each patient using a questionnaire. All patients provided informed consent to participate in the study. Demographic parameters were measured, some of which were obtained from the patients' hospital records, and blood samples were drawn from each participant's vein. A 5 mL sample was collected and distributed into a gel tube and a tube containing an anticoagulant (EDTA). The collected samples in gel tubes were left at room temperature (25 °C) for 30 min to clot, then centrifuged at 4,000 rpm for 10 min to separate the serum. Serum samples obtained from gel tubes were divided into small portions and stored at -80 °C until analysis to prevent thawing and refreezing.

Laboratory analysis

Several tests were performed, including measurements of blood urea, creatinine, serum albumin, fasting blood glucose, and glycated hemoglobin (HbA1c), using a German-made Roche Diagnostic Cobas C311 analyzer. GAL-3 levels (Cat. No. E1951Hu) were measured using ELISA kits with automated readings. All reagents were prepared and processed according to the manufacturer's instructions to ensure test accuracy. The ELISA kit used for GAL-3 measurement had a reported intra-assay coefficient of variation (CV) of <8% and an inter-assay CV of <10%, as specified by the manufacturer. The detection range was 15.6-1,000 ng/mL. Sample storage at -20 °C was performed in accordance with the manufacturer's recommendations; however, we acknowledge that storage at -80 °C may provide superior long-term protein stability, which represents a potential limitation of the present study.

Statistical analysis

The data in this study were analyzed using SPSS version 25.0 for primary statistical analyses. Statistical data, including range, mean, and standard deviation, were measured to describe the variables. The normality of continuous variables was assessed using the Shapiro-Wilk test prior to parametric analyses. Groups were compared using one-way analysis of variance (ANOVA), followed by Tukey's multiple comparisons test, the chi-squared test (for categorical data), and diagnostic performance was evaluated using receiver operating characteristic (ROC) curve analysis, with calculation of the area under the curve (AUC), 95% confidence intervals (CI), sensitivity, specificity, and optimal cutoff values. Effect sizes were calculated using eta-squared (η^2) for ANOVA comparisons. Correlation between variables was calculated using Pearson's correlation coefficient. Results were considered statistically significant at $p<0.05$ for all analyses. All ROC-derived cutoff values were determined from the study dataset and are considered exploratory; external validation in independent cohorts is required to confirm their clinical applicability.

Results

The demographic and baseline clinical characteristics of the study participants are presented in Table 1. All patients in both the CKD and DKD groups were classified as CKD stage 5 (ESRD) and were receiving maintenance hemodialysis. All patients in the DKD group had type 2 diabetes mellitus as the underlying etiology. No significant differences were observed in the distribution of sex, mean age, or body mass index (BMI) among the CKD, DKD, and control groups. However, significant variations were noted in key biochemical markers. Patients in the DKD group exhibited significantly higher fasting blood glucose (156.34±30.47 mg/dL) and glycated hemoglobin (HbA1c; 7.14±0.64%) levels than those in both the CKD and control groups ($p<0.001$). Furthermore, both the CKD and DKD groups demonstrated markedly elevated serum urea (89.28±25.57 and 95.95±30.78 mg/dL, respectively) and creatinine (6.202.26± and 6.33±1.47 mg/dL, respectively) levels compared with the control group ($p<0.001$ for both). The p-values presented in Table 1 were calculated using one-way ANOVA for continuous variables and the Chi-square test for categorical variables. Post hoc pairwise comparisons were performed using Tukey's multiple comparisons test.

Table 1. Baseline demographic and anthropometric characteristics of study participants.

Parameter	Mean±SD		DKD (N=32)	*p* value
	CKD (N=32)	Control (N=64)		
CKD stage	Stage 5 (ESRD)	N/A	Stage 5 (ESRD)	N/A
Diabetes type	None	None	Type 2	N/A
Age (years)	41.06±10.57	42.05±11.76	44.03±9.47	0.538
Gender (Male/Female)	16 (50.0%)/ 16 (50.0%)	33 (51.6%)/ 31 (48.4%)	15 (46.9%)/ 17 (53.1%)	0.911
BMI (kg/m^2)	27.05±4.09	27.08±3.95	28.09±7.57	0.618
FBG (mg/dL)	89.99±10.35*	94.23±13.12*	156.34±30.47**	<0.001
HbA1c (%)	5.31±0.54*	5.35±0.46*	7.14±0.64**	<0.001

Table 1. Baseline demographic and anthropometric characteristics of study participants (continued).

Parameter	Mean±SD		DKD (N=32)	*p* value
	CKD (N=32)	Control (N=64)		
Blood urea (mg/dL)	89.28±25.57*	28.38±6.64**	95.95±30.78*	<0.001
Creatinine (mg/dL)	6.20±2.26*	0.77±0.18**	6.33±1.47*	<0.001

Note: *The p values were calculated using one-way ANOVA for continuous variables and Chi-square test for categorical variables. Level of significance following Tukey's multiple comparisons test: *groups sharing the same letter do not differ significantly (p>0.05); **groups with different letters differ significantly (p<0.05).*

No significant difference in serum albumin concentrations was observed between the CKD and DKD groups (3.00±0.242 g/dL and 2.944±0.223 g/dL, respectively; *p*=0.7677). However, as illustrated in Figure 1, the control group exhibited significantly higher serum albumin levels (3.986±0.394 g/dL, *p*<0.0001) than those in both patient cohorts.

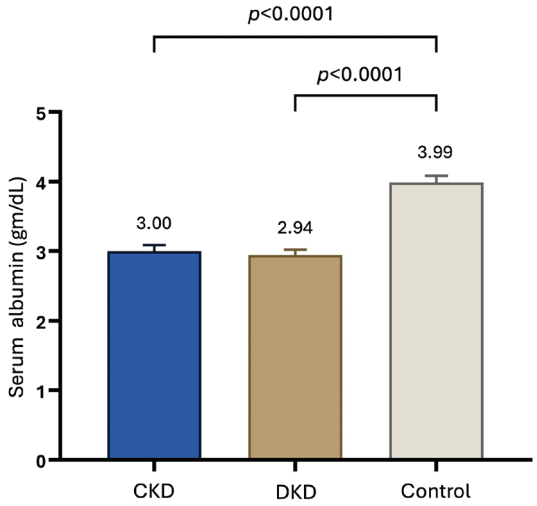


Figure 1. *A comparison of mean serum albumin concentrations among the chronic kidney disease, diabetic kidney disease, and control groups.*

As depicted in Figure 2, serum GAL-3 levels showed a significant, progressive increase across the study groups. The mean concentrations were lowest in the control group (210.04±34.65 ng/mL), intermediate in the CKD group (262.21±33.19 ng/mL), and highest in the DKD group (368.50±67.10 ng/mL). These differences were statistically significant (*p*<0.001, η^2=0.72, indicating a large effect size).

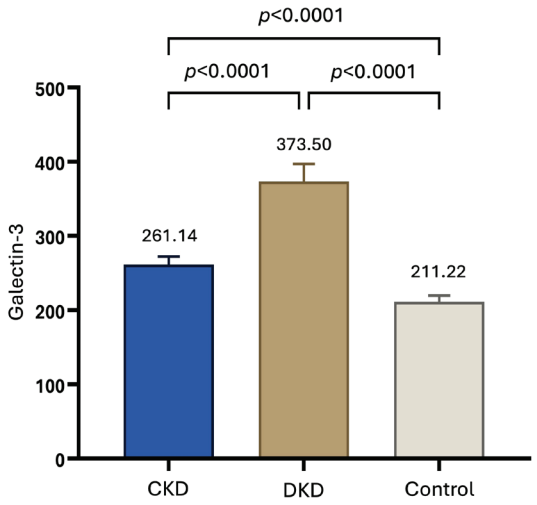


Figure 2. *A comparison of mean serum GAL-3 concentrations among the chronic kidney disease, diabetic kidney disease, and control groups.*

In the DKD group, the results differed again. GAL-3 was not significantly associated with renal, metabolic, or anthropometric parameters in this group.

For distinguishing between healthy controls and non-diabetic CKD, GAL-3 performed strongly, with an AUC of 0.852 (95% CI: 0.765-0.916; p<0.0001), sensitivity of 73.44%, and specificity of 90.62% at ≥229.199 ng/mL (Figure 3). In discriminating DKD from healthy controls, GAL-3 achieved very high diagnostic accuracy in this cohort. GAL-3 yielded an AUC of 1.00 (95% CI: 0.963-1.000, p<0.0001), with 100% sensitivity and 100% specificity at optimal cut-off values of ≥275.766 ng/mL (Figure 4). When directly comparing non-diabetic CKD to DKD, the biomarker maintained discriminatory performance. GAL-3 showed a similarly high AUC of 0.987 (95% CI: 0.905-0.999, p<0.0001), with 84.37% sensitivity and 100% specificity (Figure 5). When all patients with kidney disease (CKD and DKD combined) were compared to healthy controls, diagnostic performance remained excellent. GAL-3 yielded an AUC of 0.926 (95% CI: 0.866-0.965; p<0.0001), with 98.44% sensitivity and 73.44% specificity (Figure 6).

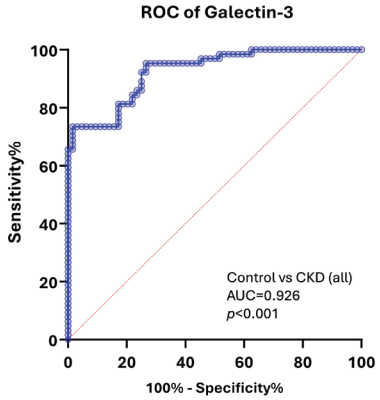


Figure 3. *Diagnostic performance of GAL-3 in differentiating non-diabetic CKD from healthy controls.*

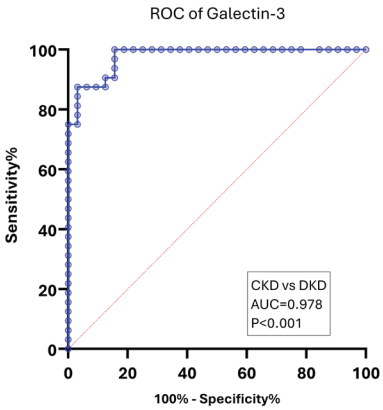


Figure 4. *Diagnostic performance of GAL-3 in differentiating DKD from healthy controls.*

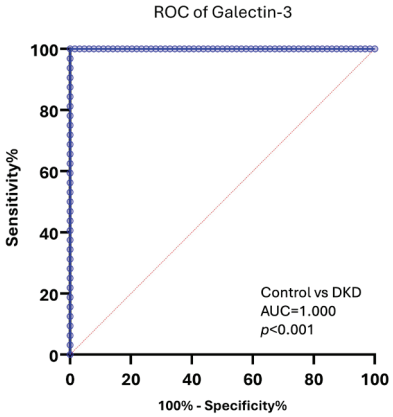


Figure 5. *Diagnostic performance of GAL-3 in differentiating DKD from non-diabetic CKD.*

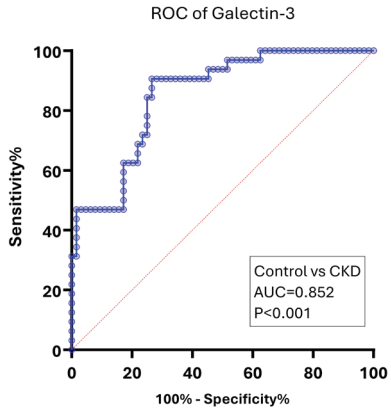


Figure 6. *Diagnostic performance of GAL-3 in differentiating kidney disease from healthy controls.*

Discussion

This study investigated serum GAL-3 as a diagnostic and discriminatory biomarker across three well-characterized cohorts: patients with DKD (all with type 2 diabetes mellitus), patients with non-diabetic CKD, and healthy controls. All patients in the CKD and DKD groups were classified as CKD stage 5 (ESRD) and were receiving maintenance hemodialysis, ensuring comparability between the disease groups with respect to the degree of renal impairment. The principal findings demonstrated a statistically significant and progressive elevation of serum GAL-3 concentrations from healthy controls through non-diabetic CKD to DKD, with the highest levels observed in the DKD group (368.50±67.10 ng/mL). The ROC analyses further confirmed the very high diagnostic accuracy of GAL-3 in distinguishing DKD from healthy controls (AUC=1.00), from non-diabetic CKD (AUC=0.987), and in differentiating all kidney disease patients from controls (AUC=0.926). These results are consistent with and extend a growing body of evidence positioning GAL-3 as a clinically relevant biomarker in the spectrum of kidney diseases.

The progressive increase in serum GAL-3 observed across the disease spectrum in this study closely aligns with findings from multiple independent investigations. A recent case-control study conducted in an Iraqi population by Mahdi *et al*. demonstrated that serum GAL-3 levels were significantly elevated in patients with CKD compared to healthy controls, with a progressive increase from mild to severe CKD stages (p<0.001). This elevation was further associated with metabolic comorbidities, including diabetes mellitus and hypertension.[12] These findings mirror the pattern observed in the present study, in which both the CKD and DKD groups exhibited markedly elevated GAL-3 levels compared with controls, with the diabetic subgroup showing the highest concentrations. This parallel is particularly noteworthy given the shared geographic and demographic context, as both studies were conducted in Iraqi clinical settings, lending regional validity to the biomarker's performance.

The mechanistic basis for GAL-3 elevation in kidney disease has been substantially elucidated in recent years. Wang et al. provided a comprehensive review of the roles of GAL-3 in both acute kidney injury (AKI) and chronic kidney disease (CKD), demonstrating that GAL-3 is a key mediator of inflammation and fibrosis through its interactions with toll-like receptor 4 (TLR4), activation of nuclear factor E-kappa B (NF-κB) signaling, and promotion of transforming growth factor-beta (TGF-β)-mediated tubulointerstitial fibrosis.[13] The pro-fibrotic role of GAL-3 was further confirmed mechanistically by Chen *et al.,* who demonstrated in a murine DKD model that macrophage-derived GAL-3 is the primary driver of renal fibrosis in DKD, operating through enhanced TGF-β signaling, and that GAL-3 expression in plasma and kidneys was positively correlated in DKD mice, positioning it as a potential therapeutic target.[14] These mechanistic insights provide a compelling biological rationale for the elevated GAL-3 levels observed in the DKD group in the present study, where the combined burden of hyperglycemia-driven metabolic stress, glomerular hypertension, and inflammatory activation is expected to maximally stimulate GAL-3 secretion.

A notable finding of this study is the potential diagnostic value of GAL-3 in this cohort. Traditional markers, such as serum creatinine and eGFR, are widely recognized to lack discriminative power in advanced CKD populations, where values are uniformly elevated.[7] Albuminuria, while more sensitive in some contexts, is confounded by exercise, infection, and cardiac failure, and may remain within normal limits in up to one-third of patients with histologically confirmed DKD.[15] The present study's ROC analysis demonstrated that GAL-3 achieved very high discrimination between DKD patients and healthy controls in this cohort (AUC=1.00, 95% CI, 0.963-1.000, sensitivity = 100%, specificity = 100%). However, an AUC of 1.00 warrants cautious interpretation. This finding likely reflects, at least in part, the extreme group separation inherent in our study design (ESRD patients on dialysis vs healthy controls), which creates

a wide biochemical gap between groups, a phenomenon known as spectrum bias. The diagnostic performance reported here should therefore not be extrapolated to populations with earlier stages of CKD or to general diabetic populations without kidney disease. Importantly, this study did not perform a formal head-to-head comparative diagnostic analysis between GAL-3 and established renal biomarkers (creatinine, eGFR, albuminuria) using methods such as DeLong's test for comparing AUCs. Therefore, the present findings suggest that GAL-3 showed potential diagnostic value and may help overcome some limitations of conventional kidney function tests, particularly in diabetic conditions, rather than demonstrating definitive superiority over conventional markers.

This finding is corroborated by the systematic review by Bellos *et al.*, which evaluated the prognostic role of GAL-3 in patients with CKD, including those on dialysis, and concluded that elevated serum GAL-3 levels are consistently associated with adverse outcomes, including reduced eGFR, increased albuminuria, and progression to dialysis or death.[16] Similarly, a prospective study by Hodeib *et al*. followed 300 patients with type 2 diabetes over three years and demonstrated that serum GAL-3 was a significant independent predictor of progression to microalbuminuria, macroalbuminuria, ESRD, and death, with levels significantly higher in patients with macroalbuminuria than in those with micro- or normoalbuminuria.[17] The present study extends these findings to a dialysis population, demonstrating that GAL-3 retains its discriminatory power even in end-stage disease, where traditional markers are uniformly elevated and therefore less informative for differential diagnosis.

Buliga-Finis *et al*.'s comprehensive review further reinforces the clinical utility of GAL-3, citing data from the Framingham Offspring Study (N=2,450) in which elevated GAL-3 was associated with a 50% increased risk of CKD events and a more rapid decline in kidney function over a 10-year follow-up period, independent of age, sex, diabetes, hypertension, proteinuria, and baseline eGFR.[18] Importantly, that analysis also indicated that GAL-3 specifically predicted tubulointerstitial fibrosis rather than glomerular injury, suggesting that its elevation reflects a distinct and clinically important pathological process that is not captured by conventional biomarkers.

One of the most clinically significant findings of the present study is GAL-3's ability to discriminate DKD from non-diabetic CKD with high accuracy (AUC=0.987, sensitivity = 84.37%, specificity = 100%). This differential diagnostic capacity addresses a critical clinical challenge, as both conditions share overlapping presentations, including elevated creatinine, reduced eGFR, and hypoalbuminemia, making their distinction on biochemical grounds alone difficult without renal biopsy. The present data suggest that GAL-3 may serve as a non-invasive adjunct for this distinction, potentially complementing invasive diagnostic procedures.

The biological basis for the differential elevation of GAL-3 in DKD relative to non-diabetic CKD likely reflects the additive pro-fibrotic and pro-inflammatory effects of hyperglycemia on the kidney.[19] In the diabetic milieu, advanced glycation end-products, oxidative stress, and activation of the renin-angiotensin-aldosterone system collectively amplify TGF-β signaling and macrophage activation, both of which are upstream drivers of GAL-3 secretion.[20] A review by Sinha and Nicholas specifically highlighted these interconnected pathomechanisms in DKD, noting that glomerular hypertension, metabolic disturbances, and hemodynamic irregularities collectively activate transcription factors and molecular pathways that culminate in structural and functional renal changes.[3] The finding of the present study that DKD patients had significantly higher HbA1c (7.14±0.64%) and fasting blood glucose (156.34±30.47 mg/dL) levels than non-diabetic CKD patients, alongside their higher GAL-3 levels, supports the hypothesis that glycemic dysregulation is an independent contributor to GAL-3 elevation beyond the effect of kidney dysfunction alone.

An important consideration when interpreting the present findings concerns the type of diabetes included in the DKD group. All patients in this group had T2DM, which is characterized by insulin resistance, chronic low-grade inflammation, and features of metabolic syndrome that may independently influence GAL-3 levels through mechanisms distinct from those operative in T1DM. T2DM is associated with greater adipose tissue inflammation, higher circulating levels of pro-inflammatory cytokines, and more pronounced metabolic derangements, all of which are known to upregulate GAL-3 expression.[18,20] Whether similar GAL-3 elevations and diagnostic performance would be observed in patients with T1DM-associated DKD remains to be determined, as the pathophysiology of DKD in T1DM may involve different inflammatory and fibrotic pathways. This represents an important area for future investigation, and the present findings should be interpreted as applicable specifically to T2DM-associated DKD.

Limitations

Several limitations of the present study should be acknowledged. Firstly, the case-control design with extreme group separation (ESRD patients on hemodialysis vs. healthy controls) introduces spectrum bias, which likely inflates the observed diagnostic accuracy. The AUC of 1.00 for DKD vs healthy controls should be interpreted in this context; these results cannot be directly extrapolated to early CKD populations or to general diabetic populations without kidney disease. Secondly, the relatively small sample size (N=32 per disease group) limits statistical power, particularly for correlation analyses, and increases the risk of overfitting in ROC analyses. Thirdly, GAL-3 is influenced by several confounding factors not fully addressed in this study, including systemic

inflammation, cardiovascular disease, dialysis status, and hepatic fibrosis. The hemodialysis procedure itself may affect GAL-3 levels through inflammatory activation and clearance dynamics. Fourthly, this study did not include a direct head-to-head ROC comparison between GAL-3 and conventional markers (creatinine, eGFR, albuminuria); therefore, claims of GAL-3 superiority over these markers should be interpreted with caution. Fifthly, all patients in both the CKD and DKD groups were classified as CKD stage 5 (ESRD) on maintenance hemodialysis; therefore, the findings cannot be generalized to patients with earlier CKD stages (stages 1-4), where GAL-3 levels and diagnostic performance may differ. The inability to match CKD stages between groups was mitigated by enrolling only stage 5 patients in both disease groups, ensuring comparability; however, this limits the generalizability of the findings. Sixth, the DKD group included exclusively patients with type 2 diabetes mellitus; therefore, the applicability of these findings to patients with type 1 diabetes-associated DKD remains unknown and represents an area for future research. Finally, all cutoff values were derived from the same dataset and require external validation in independent, prospectively recruited cohorts before clinical application can be recommended.

Conclusion

In conclusion, the present exploratory study provides preliminary evidence that serum GAL-3 showed potential diagnostic value for discriminating DKD from non-diabetic CKD and healthy controls in a cohort of ESRD patients (all with type 2 diabetes mellitus) on hemodialysis. The progressive elevation of GAL-3 across the CKD-DKD spectrum, its strong ROC performance, and its biological plausibility as a mediator of renal fibrosis and inflammation collectively warrant further investigation of its potential role in clinical diagnostic algorithms for DKD. These findings suggest that GAL-3 may help overcome some limitations of conventional kidney function tests, particularly in diabetic conditions; however, formal comparative diagnostic studies against established biomarkers are needed to confirm any added clinical value. Given the limitations inherent in the case-control design, small sample size, and spectrum bias, larger prospective multicenter studies with independent validation cohorts including patients with early-stage CKD and both type 1 and type 2 diabetes are essential to confirm these findings, establish standardized reference ranges, and evaluate the utility of GAL-3 as a therapeutic target and treatment response biomarker in this high-risk population.

Ethical approval

An agreement to conduct this study was signed prior to the start of sample collection. This was done in collaboration with the Department of Biochemistry at the College of Medicine, Tikrit University. The research committee approved the study (code MIDTUH2 on January 5, 2026) in accordance with the Declaration of Helsinki.

Funding

None

Competing Interests

The authors declare no competing interests.

Availability of Data and Materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

CRediT authorship contribution statement

Asmaa. H. Nasser: conceptualization, methodology, investigation, data curation, writing: original draft, review and edit; visualization, formal analysis, software, validation; **Entedhar Rifaat Sarhat**: supervision, project administration, resources, conceptualization, writing; review and edit.

Acknowledgements

The authors extend their sincere gratitude to the administration and medical staff of the participating hospitals in the Diyala Governorate for their invaluable support in facilitating patient recruitment and sample collection. Special thanks are due to the Tikrit University College of Medicine for providing institutional support and resources.

References

[1] Jager KJ, Kovesdy C, Langham R, Rosenberg M, Jha V, Zoccali C. A single number for advocacy and communication-worldwide more than 850 million individuals have kidney diseases. Nephrol Dial Transplant 2019; 34: 1803-5. doi: 10.1093/ndt/gfz174.

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

[3] Sinha SK, Nicholas SB. Pathomechanisms of diabetic kidney disease. J Clin Med 2023; 12: 7349. doi: 10.3390/jcm12237349.

[4] Shahbazian H, Rezaii I. Diabetic kidney disease; review of the current knowledge. J Ren Inj Prev 2013; 2: 73. doi: 10.12861/jrip.2013.24.

[5] Atlas D. International diabetes federation. IDF Diabetes Atlas, 7th Ed. Brussels, Belgium Int Diabetes Fed 2015; 33.

[6] Liu Z, Liu J, Wang W, An X, Luo L, Yu D, et al. Epigenetic modification in diabetic kidney disease. Front Endocrinol (Lausanne) 2023; 14: 1133970. doi: 10.3389/fendo.2023.1133970.

[7] Lin C-H, Chang Y-C, Chuang L-M. Early detection of diabetic kidney disease: Present limitations and

future perspectives. World J Diabetes 2016; 7: 290. doi: 10.4239/wjd.v7.i14.290.

[8] Dopierała M, Nitz N, Król O, Wasicka-Przewoźna K, Schwermer K, Pawlaczyk K. New and emerging biomarkers in chronic kidney disease. Biomedicines 2025; 13: 1423. doi: 10.3390/biomedicines13061423.

[9] Hussain S, Jamali MC, Habib A, Hussain MS, Akhtar M, Najmi AK. Diabetic kidney disease: An overview of prevalence, risk factors, and biomarkers. Clin Epidemiol Glob Heal 2021; 9: 2-6. doi: 10.1016/j.cegh.2020.05.016.

[10] Syn G, Lee YQ, Lim ZY, Chan GC. Galectin-3: action and clinical utility in chronic kidney disease. Int Urol Nephrol 2024; 56: 3535-43. doi: 10.1007/s11255-024-04107-4.

[11] Bellos I, Marinaki S, Lagiou P, Benetou V. Association of serum galectin-3 levels with mortality and cardiovascular disease outcomes in hemodialysis patients: a systematic review and dose-response meta-analysis. Int Urol Nephrol 2024; 56: 2755-67. doi: 10.1007/s11255-024-04026-4.

[12] Ameer Mahdi N, Choobineh H, Hashim Abdulsalam A, Majidi Z, Dashti N. Elevated serum levels of Galectin-3 and Kidney Injury Molecule-1 as potential biomarkers for early detection and staging of chronic kidney disease in Iraqi population. Reports Biochem Mol Biol 2025; 14: 95-101. doi: 10.61882/rbmb.14.1.95.

[13] Wang F, Zhou L, Eliaz A, Hu C, Qiang X, Ke L, et al. The potential roles of galectin-3 in AKI and CKD. Front Physiol 2023; 14: 1090724. doi: 10.3389/fphys.2023.1090724.

[14] Chen Y, Jiang Q, Xing X, Xu L, Zhao Q, Zhang Q, et al. Macrophage derived Galectin-3 promotes renal fibrosis and diabetic kidney disease by enhancing TGFβ1 signaling. Adv Sci (Weinheim, Baden-Wurttemberg, Ger 2025; 12: e04032. doi: 10.1002/advs.202504032.

[15] Karimi F, Moazamfard M, Taghvaeefar R, Sohrabipour S, Dehghani A, Azizi R, et al. Early detection of diabetic nephropathy based on urinary and serum biomarkers: An updated systematic review. Adv Biomed Res 2024; 13: 104. doi: 10.4103/abr.abr_461_23.

[16] Bellos I, Marinaki S, Lagiou P, Benetou V. Galectin-3 in chronic kidney disease. Clin Chim Acta 2024; 559: 119727. doi: 10.1016/j.cca.2024.119727.

[17] Hodeib H, Hagras MM, Abdelhai D, Watany MM, Selim A, Tawfik MA, et al. Galectin-3 as a prognostic biomarker for diabetic nephropathy. Diabetes, Metab Syndr Obes Targets Ther 2019: 325-31. doi: 10.2147/DMSO.S194410.

[18] Buliga-Finis ON, Ouatu A, Tanase DM, Badescu MC, Dima N, Gosav EM, et al. The role of Galectin-3 as a biomarker in the cardio-renal-metabolic Pathology Axis. J Clin Med 2025; 14. doi: 10.3390/jcm14176071.

[19] Tan KCB, Cheung C-L, Lee ACH, Lam JKY, Wong Y, Shiu SWM. Galectin-3 is independently associated with progression of nephropathy in type 2 diabetes mellitus. Diabetologia 2018; 61: 1212–9. doi: 10.1007/s00125-018-4552-z.

[20] Guo Y, Li L, Hu S. Circulating Galectin-3 levels and diabetic nephropathy: a systematic review and meta-analysis. BMC Nephrol 2023; 24: 163. doi: 10.1186/s12882-023-03226-x.

Journal of Associated Medical Sciences 2026; 59(3): 177-184

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 a behavioral strategy program on quality of life: A quasi-experimental study in women with cognitive impairment receiving chemotherapy

M. Arun Kumar and Jamima Jamesraj

Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha Nagar, Thandalam, Chennai, India.

ARTICLE INFO

Article history:
Received 1 October 2025
Accepted as revised 16 June 2026
Available online 24 June 2026

Keywords:
Chemotherapy, cognitive impairment, quality of life, behavioral strategies, occupational therapy

ABSTRACT

Background: Women who undergo chemotherapy for cancer frequently develop cognitive problems referred to as chemo brain. These cognitive deficits harm memory functions and attention skills as well as processing speed, which leads to decreased quality of life (QoL). Behavioral approaches are non-drug solutions that show promise in enhancing cognitive abilities and everyday functioning in this specific group.

Objectives: To evaluate the effectiveness of a behavioral strategy program compared with conventional occupational therapy in improving quality of life among women with chemotherapy-related cognitive impairment.

Materials and methods: A quasi-experimental pre- post study design was employed. Thirty women aged 45-60 years undergoing chemotherapy were recruited from Madras Cancer Care Foundation, Kumaran Hospital, Chennai. Participants were allocated into two groups (N=15 each): an experimental group receiving structured behavioral strategies and a control group receiving conventional occupational therapy for 12 weeks (36 sessions). Quality of life was measured using the WHOQOL-BREF. Data were analyzed using non-parametric tests (Wilcoxon signed-rank test and Mann-Whitney U test).

Results: Both groups demonstrated significant improvement in quality of life following intervention. The control group improved from 49.53±6.14 to 52.27±7.36 (p=0.029), while the experimental group improved from 47.33±5.58 to 59.73±6.58 (p=0.001). Between-group comparison showed a statistically significant difference in favor of the experimental group (p=0.015). Effect size analysis demonstrated large improvements in the experimental group.

Conclusion: Behavioral strategies showed greater improvement in quality of life compared with conventional occupational therapy among women with chemotherapy-related cognitive impairment. Integrating structured behavioral interventions into oncology rehabilitation may enhance patient outcomes.

* **Corresponding contributor.**
Author's Address: *Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha Nagar, Thandalam, Chennai, India.*
E-mail address: principal.scot@saveetha.com
doi: 10.66285/JAMS.2026.091
E-ISSN: 2539-6056

Introduction

Cancer continues to pose a major public health challenge globally, with increasing incidence and improved survivorship contributing to a growing population living with the long-term consequences of cancer and its treatment. According to the World Health Organization (2022), India reports more than 1.4 million new cancer cases annually, reflecting a substantial and rising disease burden. Women account for a significant proportion of this burden, particularly due to the high prevalence of

breast, cervical, ovarian, and other gynecological cancers. Urban centers such as Chennai report comparatively higher cancer incidence rates, influenced by population density, lifestyle factors, and improved access to diagnostic and tertiary healthcare services.

Although advances in oncology care have improved survival outcomes, chemotherapy is frequently associated with adverse cognitive and psychosocial effects. Chemotherapy-related cognitive impairment (CRCI), commonly referred to as "chemo brain," includes deficits in attention, memory, processing speed, and executive functioning.[1,2] These cognitive changes may persist beyond active treatment and can significantly interfere with daily functioning, occu pational roles, and overall quality of life (QoL).[1,2]

Quality of life has therefore emerged as a critical outcome in cancer rehabilitation, reflecting physical, psychological, social, and environmental well-being. Women undergoing chemotherapy in India often experience reduced QoL due to treatment-related fatigue, emotional distress, cognitive difficulties, and restricted participation in meaningful activities. Sociocultural expectations, caregiving responsibilities, and limited access to structured rehabilitation services may further exacerbate these challenges.[3,12-14]

In the Indian context, the impact of cancer on women's quality of life is further intensified by their culturally defined roles within the family and society. Women often function as primary caregivers and are central to household management; therefore, cancer-related impairments significantly disrupt both personal independence and family dynamics. The presence of chemotherapy-related cognitive impairment further compromises their ability to perform daily activities, manage responsibilities, and engage in meaningful occupations. These combined cognitive and psychosocial challenges contribute to a substantial decline in overall quality of life. Therefore, evaluating quality of life is crucial for comprehensively understanding the multidimensional effects of cancer and its treatment, thereby providing a rationale for selecting standardized outcome measures such as WHOQOL-BREF in the present study.[4-6,12]

Occupational therapy plays a central role in oncology rehabilitation by promoting functional independence and participation in meaningful life roles despite disease- or treatment-related limitations.[7] Behavioral strategies have gained recognition as structured, non-pharmacological approaches that address both cognitive and psychosocial dimensions of health. Techniques such as goal setting, behavioral activation, problem-solving, and stress management have demonstrated effectiveness in improving functional outcomes and quality of life among populations with chronic illness.[8,9]

The present study was conducted at the Madras Cancer Care Foundation (MCCF) at Kumaran Hospital, Chennai, which provides comprehensive oncology and rehabilitation services to women from urban and peri-urban populations, with a strong focus on improving functional outcomes and quality of life among cancer survivors. The involvement of this institution provided an appropriate clinical and rehabilitative setting for examining the effectiveness of occupational therapy-led behavioral interventions. Given the limited evidence on occupational therapy-led behavioral interventions for CRCI in the Indian context, this study aimed to evaluate the effectiveness of behavioral strategies in improving quality of life among women undergoing chemotherapy.[10,11,15]

Materials and methods

Study design

A quasi-experimental pre-post study design was employed to evaluate the effectiveness of behavioral strategies on quality of life among women with chemotherapy-induced cognitive impairment. Participants were recruited via convenience sampling and assigned to experimental and control groups.

Study setting and participants

The present study was conducted at the Madras Cancer Care Foundation (MCCF) in collaboration with Kumaran Hospital, Chennai, both of which provide comprehensive oncology and rehabilitation services to women from urban and peri-urban populations, with a strong emphasis on enhancing functional outcomes and quality of life among cancer survivors.

A total of 30 female patients, aged 45-60 years, diagnosed with cancer and undergoing chemotherapy, were recruited based on inclusion and exclusion criteria.

Inclusion and exclusion criteria

Inclusion criteria

1. Women aged 45-60 years
2. Diagnosed with cancer and undergoing chemotherapy
3. MMSE score indicating mild cognitive impairment (18-23)
4. Willing to participate

Exclusion criteria

1. Severe cognitive impairment
2. Psychiatric illness
3. Neurological disorders
4. Unstable medical condition

Sampling and group allocation

Eligible participants were recruited using convenience sampling. Based on feasibility and availability, participants were allocated into:

- Experimental group (N=15): behavioral strategy intervention
- Control group (N=15): conventional occupational therapy

Assessment Tools

1. **Mini Mental State Examination** (MMSE)

The MMSE (Folstein et al., 1975)[15] is a 30-point tool that evaluates orientation, registration, attention, recall, language, and visuospatial skills. The Scores include 0-17: Severe cognitive impairment; 18-23: Mild cognitive impairment; 24-30: No cognitive impairment. The reliability of this tool is Cronbach's alpha = 0.70, and the validity is $p<0.04$.

2. **World Health Organization Quality of Life–BREF** (WHOQOL-BREF)

The World Health Organization Quality of Life–BREF (WHOQOL-BREF) is a 26-item instrument assessing perceived quality of life across four domains: physical health, psychological health, social relationships, and environment. Items are rated on a 5-point Likert scale, with domain scores transformed to a 0-100 scale; higher scores indicate better quality of life. Domain scores are interpreted independently, as no standardized cut-off values are defined. The WHOQOL-BREF has demonstrated good reliability and validity across diverse populations (WHOQOL Group, 1998). The Reliability: Cronbach's alpha = 0.70 and the Validity: 0.89

Procedure

All subjects were screened using the Mini-Mental State Examination (MMSE) and the World Health Organization Quality of Life-BREF (WHOQOL-BREF) at baseline. Participants were allocated to the experimental and control groups based on availability and study feasibility following the baseline assessment. The intervention lasted 12 weeks, with three sessions per week. The intervention strategies are presented in Table 1.

Table 1. Intervention strategies.

Week	Focus area	Key strategies and activities	Mode of delivery
1	Orientation and psychoeducation	Introduction to intervention; education on cancer-related cognitive changes; normalization of symptom	Therapist led orientation and explanation
2	Mindfulness and grounding	Breathing exercises, body awareness, grounding techniques for attention and emotional regulation	Demonstration, guided practice
3	Attention training	Sustained and selective attention tasks; managing distractions	Guided exercises, in-session practice
4	Memory enhancement I	Chunking techniques; use of external memory aids (lists, reminders)	Demonstration, practice tasks
5	Memory enhancement II	Application of memory strategies to daily activities	Guided practice, homework
6	Behavioural activation	Activity scheduling; increasing engagement in meaningful activities	Planning exercises, discussion
7	Goal setting	SMART goal formulation; task planning and sequencing	Structured worksheets, practice
8	Problem-solving skills	Stepwise problem-solving methods for daily challenges	Role-play, guided discussion
9	Decision-making under stress	Cognitive strategies to manage decisions during emotional or cognitive overload	Case examples, practice
10	Stress management	Relaxation techniques; coping strategies for stress	Demonstration, guided practice
11	Sleep hygiene	Education on maintaining regular sleep routines, reducing screen time, relaxation before bedtime.	Psychoeducation, discussion
12	Emotional regulation and communication	Emotional awareness; assertive communication training; consolidation of skills	Role-play, review and feedback

In the experimental group, a structured, strategy-based intervention, in accordance with a planned therapy schedule, was provided. The intervention strategy focused on improving their cognitive skills, emotional control, and functional living skills. These included interventions addressing the cognitive effects of cancer, mindfulness and grounding strategies for greater awareness of the present moment, attention-training exercises, and memory-improvement strategies using chunking and external memory aids. Behavioral activation and activity scheduling were used alongside SMART goals to support task completion and implementation (Table 1). Problem-solving and decision-making strategies were taught to the intervention group to improve cognitive functioning under stressful conditions. Additional strategies in the intervention group included stress management techniques, sleep hygiene instruction, and emotional regulation strategies focused on assertiveness in communication.

Sessions were conducted three times per week (36 sessions total), each lasting approximately 75 minutes. The control group was given standard occupational therapy (OT) of equal intensity and duration to that of the experimental group, consisting of 36 sessions conducted over 12 weeks (three sessions per week, approximately 75 minutes per session). Thus, there was equality in intervention exposure for comparison. Vocational activities encompassed graded tasks related to their jobs, aimed at increasing endurance and basic work skills for their daily roles. The basic cognitive stimulation activities encompassed attention, orientation, and problem-solving (e.g., sorting, sequencing, and basic memory recall) and did not use structured strategies for behaviour and cognitive modification.

The control group received the intervention in accordance with the conventional principles of occupational therapy. The intervention in the control group did not include psychoeducation aimed at improving cognitive function. After 12 weeks of the intervention, the WHOQOL-BREF was administered again to the subjects to measure quality-of-life outcomes. Unlike the experimental group, the control group did not receive structured behavioral strategies or psychoeducation targeting cognitive and emotional regulation, which explains the difference in outcomes.

Data analysis

Data were analyzed using IBM SPSS Statistics (Version 23.0). Descriptive statistics (mean, standard deviation, range) were calculated. As the data did not meet the assumptions of normality, non-parametric tests were used. Within-group comparisons were performed using the Wilcoxon signed-rank test, and between-group comparisons were conducted using the Mann–Whitney U test. In addition to statistical significance testing, the effect size (r) was calculated for non-parametric analyses using the formula $r= Z/\sqrt{N}$, where Z is the standardized test statistic, and N is the total number of observations. Effect sizes were interpreted as small (0.1), medium (0.3), and large (0.5) according to established guidelines. Statistical significance was set at $p<0.05$.[17]

Results

As shown in Table 2, A total of 30 women undergoing chemotherapy who met the inclusion criteria participated in the study and were allocated to the experimental group (N=15) or the control group (N=15). Baseline comparison of demographic and clinical variables demonstrated no statistically significant differences between the experimental and control groups ($p>0.05$), indicating comparability of groups prior to intervention. Effect size analysis also showed small effect sizes across baseline variables.

Table 2. Participant characteristics and baseline comparability.

Variable	Group (Mean±SD)		*p* value	Effect size
	Experimental (N=15)	Control (N=15)		
Age (years)	54.1±4.2	53.6±4.5	0.754	0.11
Baseline WHOQOL-BREF Score	47.33±5.58	49.53±6.14	0.312	0.37
MMSE score	21.13±1.41	20.87±1.36	0.621	0.19

Note: *Data are presented as mean±SD. Effect size was calculated using Cohen's d, where 0.2 indicates a small effect, 0.5 a medium effect, and 0.8 a large effect.*

Therefore, non-parametric statistical tests were used for further analysis. The results of the Shapiro – Wilk test used to assess normality of the data are presented in Table 3.

Table 3. Shapiro-wilk test for normality.

Group	Test	W statistic	*p* value	Interpretation
Control	Pre-test	0.882	0.050	Normal
	Post-test	0.895	0.080	Normal
Experimental	Pre-test	0.875	0.039	Non-normal
	Post-test	0.902	0.103	Normal

***Note**: The Shapiro-Wilk test indicated a non-normal distribution in the experimental group pre-test scores (p=0.039), whereas the remaining variables showed normal distributions (p>0.05).*

As shown in Table 4, Within-group analysis demonstrated statistically significant improvement in WHOQOL-BREF scores following intervention in both groups. The experimental group had better post-intervention quality-of-life scores than the control group. Effect size analysis indicated a large effect in both groups, with a greater magnitude of improvement observed in the experimental group.

Table 4. Within-group comparison of pre- and post-test WHOQOL-BREF scores using the Wilcoxon Signed-Rank Test.

Group	Pre-test (Mean±SD)	Post-test (Mean±SD)	Z value	*p* value	Effect size (r)
Control	49.53±6.14	52.27±7.36	-2.179	0.029	0.56
Experimental	47.33±5.58	59.73±6.58	-3.439	0.001	0.89

***Note:** The effect size (r) was calculated using the formula $r = Z/\sqrt{N}$, where Z is the standardized test statistic, and N is the total number of observations. Values of 0.1, 0.3, and 0.5 indicate small, medium, and large effects, respectively.*[17]

As shown in Table 5, both groups demonstrated improvements in WHOQOL-BREF scores following the intervention. Effect size analysis indicated a large effect in both groups, with a greater magnitude of improvement in the experimental group (r=0.89) than in the control group (r=0.56). Between-group analysis revealed a statistically significant difference in favor of the experimental group (p=0.015), indicating greater improvement in quality of life among women with chemotherapy-related cognitive impairment. The comparison of pre-test and post-test WHOQOL-BREF scores between the experimental and control groups is presented in Figure 1.

Table 5. Comparison of baseline and post-intervention WHOQOL-BREF scores between groups.

Group	N	Baseline (Mean±SD)	Post-intervention (Mean±SD)	Within-group comparison (*p* value)	Effect size (r)	Between-group comparison (post-test *p* value)
Control (conventional occupational therapy)	15	49.53±6.14	52.27±7.36	0.029	0.56	-
Experimental (behavioral strategies)	15	47.33±5.58	59.73±6.58	0.001	0.89	0.015

***Note:** The effect size (r) was calculated using the formula $r = Z/\sqrt{N}$. Values of 0.1, 0.3, and 0.5 indicate small, medium, and large effects, respectively. Between-group comparison of baseline and post-intervention.*

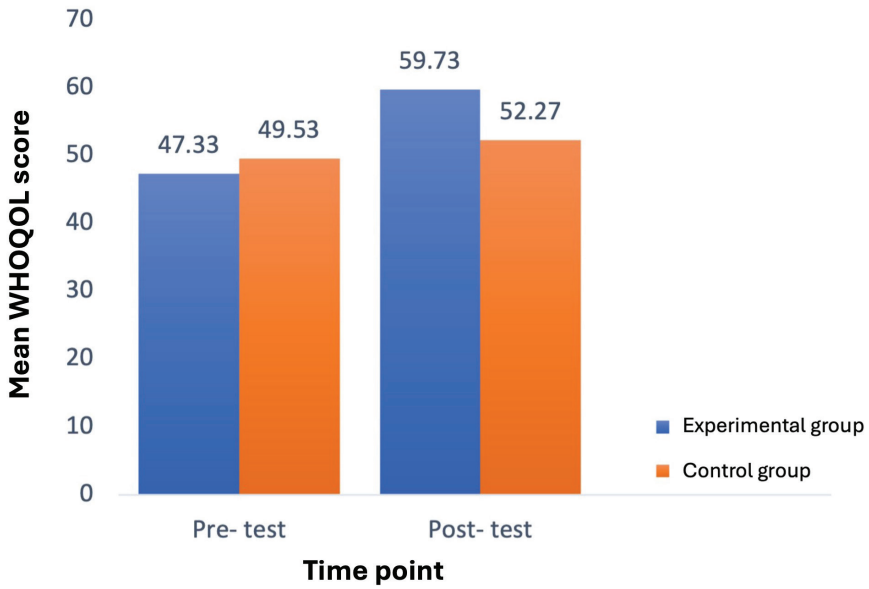


Figure 1. *Comparison of pre- and post-test WHOQOL-BREF scores between control and experimental groups.*

Discussion

The present study examined the effectiveness of behavioral strategies compared with conventional occupational therapy in improving quality of life among women undergoing chemotherapy who experienced cognitive impairment. The findings indicate that although both intervention approaches resulted in statistically significant improvements in quality of life, behavioral strategies were associated with substantially greater gains, as measured using the WHOQOL-BREF.

Quality of life impairment among women receiving chemotherapy has been widely documented, with previous studies using the WHOQOL-BREF reporting reduced scores across physical, psychological, and social domains during and following cancer treatment.[6,12] Indian studies have similarly demonstrated compromised WHOQOL-BREF scores among women with breast and gynecological cancers, highlighting the multidimensional burden of cancer and its treatment on daily functioning and well-being.[3,13,14] The baseline findings in the present study are consistent with this literature, supporting the relevance of WHOQOL-BREF as an appropriate outcome measure for this population.

The modest improvement observed in the control group suggests that conventional occupational therapy interventions, such as range-of-motion exercises, vocational activities, and basic cognitive stimulation, can positively influence quality of life. Previous studies have shown that supportive and activity-based interventions improve physical comfort and psychosocial adjustment, as reflected in WHOQOL-BREF domain scores.[7] However, the relatively smaller magnitude of improvement in the control group indicates that conventional approaches alone may not fully address the cognitive and emotional challenges associated with chemotherapy-related cognitive impairment, particularly in the absence of structured behavioral or cognitive strategy training.

In contrast, participants who received behavioral strategy interventions demonstrated significantly greater improvement in quality of life. This finding aligns with prior research indicating that behavioral and cognitive-behavioral interventions can lead to meaningful improvements in WHOQOL-BREF outcomes among individuals with chronic illness and cancer.[8,9,15] The structured and multidimensional nature of behavioral strategies encompassing goal setting, behavioral activation, problem-solving, memory compensation, and stress management may have enhanced self-efficacy and adaptive coping, thereby positively influencing multiple domains of quality of life.

Consistent with earlier studies on chemotherapy-related cognitive impairment, interventions that emphasize compensatory strategies and engagement in meaningful activities showed greater improvement in patient-reported outcomes than those focusing solely on symptom reduction.[10,11,15] By addressing both functional performance and psychosocial well-being, behavioral strategies may support sustained improvements in perceived quality of life.

From a clinical perspective, these findings underscore the value of integrating behavioral strategies into occupational therapy-led oncology rehabilitation. Given their client-centered focus and emphasis on participation, occupational therapists are uniquely positioned to implement such interventions. Incorporating structured behavioral approaches may therefore enhance rehabilitation outcomes for women undergoing chemotherapy and experiencing cognitive impairment.

These findings should be interpreted in the context of the study design. As a quasi-experimental study employing convenience sampling, the ability to establish strong causal relationships is limited. Although baseline comparability between groups

was observed, unmeasured confounding factors and individual differences may have influenced the outcomes. Further research using larger samples, diverse settings, and randomized controlled designs is recommended to strengthen the evidence base and extend the generalisability of these findings.

Limitations

The scope of the present study was limited by a relatively small sample size and restricted generalisability, as it was conducted within a single hospital-based clinical setting. Participants were recruited via convenience sampling, which may have affected the sample's representativeness. In addition, the quasi-experimental design without randomization limits the ability to establish strong causal relationships between the intervention and outcomes. Furthermore, the small number of participants in specific cancer subgroups limited statistical power and increased the risk of Type II error in subgroup analyses. The study relied on self-reported WHOQOL-BREF responses, which may have been influenced by participants' personal perceptions and experiences. Due to the nature of the behavioral strategy program, blinding of participants and therapists was not feasible. However, standardized assessment procedures were followed for all participants to minimize potential bias. The intervention lasted 12 weeks, and long-term follow-up assessments were not conducted; therefore, the sustainability of the observed improvements in quality of life remains unclear. Additionally, the study was not prospectively registered in a WHO-recognized clinical trial registry, which may limit alignment with current international recommendations for interventional research transparency. Future studies should include larger sample sizes, diverse settings, extended intervention periods, and consideration of individual-difference variables to strengthen the interpretation and generalisability of the findings.

Conclusion

The findings of this study indicate that behavioral strategies show greater improvement than conventional occupational therapy in improving the quality of life among women with chemotherapy-induced cognitive impairment. While conventional therapy resulted in modest improvements, behavioral interventions produced significantly greater gains by addressing cognitive, emotional, and functional dimensions simultaneously. Integrating behavioral strategies into occupational therapy-led oncology rehabilitation may offer a comprehensive and effective approach to enhancing quality of life in this population.

Ethical approval

The study protocol was reviewed and approved by the Institutional Scientific Review Board and Institutional Ethics Committee of Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Chennai, India, prior to participant recruitment (Ethical Certificate No: SCOT/ISRB/005/2025). Written informed consent was obtained from all participants before enrolment. Although the present study involved a behavioral intervention within a quasi-experimental design, it was not prospectively registered in a WHO-recognized clinical trial registry. However, all study procedures were conducted in accordance with institutional ethical standards and the principles of research involving human participants.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflicts of interest.

CRediT Authorship Contribution Statement

M. Arun Kumar: conceptualization, methodology, writing: review and edit, supervision; **Jamima Jamesraj**: investigation, data curation, formal analysis, writing: original draft, review and edit.

Acknowledgements

The authors extend their heartfelt gratitude to all participants for their time and commitment, to the Madras Cancer Care Foundation team at Kumaran Hospital for their cooperation during data collection. Special Gratitude to Saveetha College of Occupational Therapy for their continuous guidance and support. The authors also express their gratitude to all Participants for their Cooperation and Participation in this study.

References

[1] Ahles TA, Root JC. Cognitive effects of cancer and cancer treatments. Annu Rev Clin Psychol. 2018; 14: 425-51. doi: 10.1146/annurev-clinpsy-050817-084903.

[2] Wefel JS, Schagen SB. Chemotherapy-related cognitive dysfunction. Curr Neurol Neurosci Rep. 2012; 12(3): 267-75. doi: 10.1007/s11910-012-0264-9.

[3] Pandey M, Thomas BC, SreeRekha P, Ramdas K, Ratheesan K, Parameswaran S. Quality of life determinants in cancer patients undergoing chemotherapy in India. Qual Life Res. 2005; 14(7):1701-8. doi: 10.1007/s11136-005-4338-z.

[4] World Health Organization. WHOQOL-BREF: introduction, administration, scoring and generic version of the assessment. Geneva: World Health Organization; 1996.

[5] Saxena S, Chandiramani K, Bhargava R. WHOQOL-Hindi: a questionnaire for assessing quality of life in health care settings in India. Natl Med J India. 1998; 11(4): 160-5.

[6] Skevington SM, Lotfy M, O'Connell KA. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and

results of the international field trial. Qual Life Res. 2004; 13(2): 299-310. doi: 10.1023/B:QURE.0000018471.01307.31.

[7] Christiansen CH, Baum CM. Occupational therapy: performance, participation, and well-being. 4th Ed. Thorofare (NJ): SLACK Incorporated; 2014.

[8] Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci. 2011; 6: 42. doi: 10.1186/1748-5908-6-42.

[9] King BR, Boswell JF. Therapeutic strategies and techniques in early cognitive-behavioral therapy. Psychotherapy (Chic). 2019; 56(1): 35-40. doi: 10.1037/pst0000202.

[10] Mounier NM, Abdel-Maged AE, Wahdan SA, Gad AM, Azab SS. Chemotherapy-induced cognitive impairment: an overview of etiology and pathogenesis. Life Sci. 2020; 258: 118071. doi: 10.1016/j.lfs.2020.118071.

[11] Chao D, Hale A, Henry NL, Kratz AL, Gabel NM. Cancer-related cognitive impairment or “chemobrain”: emerging assessments, treatments, and targets for intervention. Curr Phys Med Rehabil Rep. 2021; 9(3): 108-18. doi: 10.1007/s40141-021-00322-1.

[12] Chachamovich E, Fleck MP, Trentini CM, Power M. Brazilian WHOQOL-BREF normative data. Qual Life Res. 2008; 17(2): 297-306. doi: 10.1007/s11136-007-9280-3.

[13] Gangane N, Khairkar P, Hurtig AK, San Sebastian M. Quality of life determinants in breast cancer patients in central India. Asian Pac J Cancer Prev. 2017; 18(12): 3325-32. doi: 10.22034/APJCP.2017.18.12.3325.

[14] Montazeri A. Health-related quality of life in breast cancer patients: a bibliographic review. J Exp Clin Cancer Res. 2008; 27: 32. doi: 10.1186/1756-9966-27-32.

[15] González-Santos Á, López-Garzón M, Sánchez-Salado C, Postigo-Martin P, Lozano-Lozano M, Galiano-Castillo N, et al. A telehealth-based cognitive-adaptive training (e-OTCAT) to prevent cancer and chemotherapy-related cognitive impairment in women with breast cancer: protocol for a randomized controlled trial. Int J Environ Res Public Health. 2022; 19(12): 7147. doi: 10.3390/ijerph19127147.

[16] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189-198.

[17] Fiel Peres F. Effect sizes for nonparametric tests. Biochem Med (Zagreb). 2026; 36(1): 010101. doi.org/10.11613/BM.2026.010101.

Journal of Associated Medical Sciences 2026; 59(3): 185-195

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



Evaluation of changes in diabetes distress, resilience, and self-care behavior following Occupational Therapy Diabetes Self-Management (OTDSM) among adults with type 2 diabetes: A preliminary Quasi-experimental study

Kamala Arumugam, Loganathan Shanmugam*, Raghuram Pattusamy, Darsana Ganesan, Jaivignesh Shanmugam

Sri Ramachandra Faculty of Occupational Therapy, Sri Ramachandra Institute of Higher Education and Research (DU), Porur, Chennai, India.

ARTICLE INFO

Article history:
Received 19 February 2026
Accepted as revised 16 June 2026
Available online 24 June 2026

Keywords:
Diabetes mellitus type 2, health education, occupational therapy, resilience, self-management

ABSTRACT

Background: Type 2 diabetes mellitus (T2DM) is a growing health concern that necessitates consistent self-management practices to avoid complications and maintain health. Psychological factors like diabetes distress and poor resilience levels frequently interfere with self-care practices of individuals, resulting in less than adequate diabetes related self-care. By integrating self-management into daily routines, occupational therapy plays a major role in preventing secondary complications of diseases.

Objectives: To find the changes in diabetes distress, resilience, and self-care behavior following the Occupational Therapy Diabetes Self-Management (OTDSM) intervention among adults with type 2 diabetes.

Materials and methods: A quasi-experimental single-group pre-test post-test study was conducted in a community setting, with 12 adults diagnosed with type 2 diabetes. The Occupational Therapy Diabetes Self-Management (OTDSM) intervention was provided for 4 weeks in 10 sessions. Outcomes were assessed using the Diabetes Distress Scale (DDS -17), the Connor-Davidson Resilience Scale (CD-RISC-25), and the Summary of Diabetes Self-Care Activities Questionnaire (SDSCA). Pre- and post-intervention data were analyzed using Paired t-tests and Repeated Measures ANOVA.

Results: Diabetes distress showed marked reduction from 3.72±0.25 to 2.53±0.28 (p<0.001). Resilience scores improved from 42.91±3.53 to 49.45±3.24 (p<0.001). Self-care behavior also showed improvements across multiple domains (p<0.001).

Conclusion: The OTDSM intervention was associated with positive changes in diabetes distress, resilience, and self-care outcomes among adults with T2DM. These preliminary findings may suggest the potential integration of OTDSM into diabetes care to address both psychosocial and behavioral aspects of disease management.

* **Corresponding contributor.**
Author's Address: *Sri Ramachandra Faculty of Occupational Therapy, Sri Ramachandra Institute of Higher Education and Research (DU), Porur, Chennai, India.*
E-mail address: loganathanot@sriramachandra.edu.in
doi: 10.66285/JAMS.2026.092
E-ISSN: 2539-6056

Introduction

Diabetes mellitus (DM) leads to hyperglycemia, caused by abnormalities in insulin production or action. This results in malfunctions in the metabolism of proteins, fats, and carbohydrates.[1] Type 2 diabetes mellitus (T2DM) is the most common form, influenced by non-modifiable factors such as age, familial medical history, genetic makeup, and cultural background,

along with modifiable risks such as sedentary lifestyle, poor diet, obesity, stress, and inadequate sleep.[2] Poorly managed diabetes contributes to cardiovascular diseases, renal dysfunction, retinopathy, amputations, neuropathy, stroke, and reduced life expectancy.[3]

Diabetes distress (DD), which often goes unnoticed, is the emotional struggle individuals deal with when managing this demanding condition. Higher distress is associated with poor blood-glucose control, inadequate self-care, and diminished quality of life.[4] Unlike psychiatric disorders, DD is regarded as a disease-related psychological response, with prevalence estimates between 21% and 49% worldwide.[5,6] Psychoeducational approaches addressing psychological aspects and diabetes are more efficient in reducing DD.[7]

Resilience is a psychosocial ability that minimizes distress, improves adaptability in times of crisis, and ultimately enhances quality of life.[8] Resilience is strongly correlated with self-management behavior, and interventions to build resilience can improve self-management.[9] Even with increasing diabetes distress, resilient individuals will continue to engage in effective self-management.[10]

Self-management is an individual's capability to handle symptoms, medications, therapy, lifestyle modifications, and the psychological repercussions of chronic illness.[11] Competent self-management strategies help individuals with T2DM to mitigate complications. Four key components of diabetes self-management (DSM) are understanding the disease, developing the skills for managing it, modifying lifestyle, and receiving psychological support.[12] DSM is hindered by the chronic, unstable course of the disease, food cravings, incorrect assumptions regarding self-management techniques, lack of motivation, diabetes distress, and lack of education and skills.[13]

Occupational therapy shapes routines and behaviors that promote health and integrate self-care practices into meaningful daily occupations.[14] Occupational Therapy Diabetes Self Management (OTDSM) intervention is an evidence-based, occupation-centered program that combines education, counseling, activities, occupations, and groups.[15] OTDSM has demonstrated benefits for glycemic control,[16] but its effect on diabetes distress and resilience has not been explored. Therefore, this study examined reductions in diabetes distress and improvements in resilience and self-care behavior following the OTDSM intervention among adults with type 2 diabetes. The study's two main research objectives were as follows:

1. To assess the reductions in diabetes distress and improvements in resilience, following the occupational therapy diabetes self-management (OTDSM) intervention among adults with type 2 diabetes.

2. To analyze changes in self-care behavior among adults with type 2 diabetes.

Materials and methods

Study design and setting

This study employed a preliminary quasi-experimental single-group pre-test-post-test design. In a real-world setting, it was not feasible to deny the control group valuable health-related education; thus, the practical limitations and ethical concerns led to the selection of this study design. The study was conducted at the community level in the Koyambedu Market in Chennai, India, over four weeks.

Study participants

The sample size was calculated using a paired pre-post comparison of self-management scores reported by Gregg *et al*.[17] Mean±SD of 7.49±5.87, and 13.58±5.56 were used for calculation. Assuming an effect size of 1.07, a two-sided significance level of 5% (α=0.05) and 90% power (1-β=0.90), the minimum required sample size was 11 participants. To account for 10% loss to follow-up, 12 participants were recruited.

A total of 25 participants with T2DM were screened for eligibility. 12 participants who satisfied the inclusion criteria were recruited using convenience sampling. Out of these, 11 completed the study. Recruitment and follow-up took place between May 2025 and June 2025. Criteria for inclusion were: 1) diagnosis of type 2 diabetes, 2) score of 2 or more in the diabetes distress scale (DSS -17), 3) score less than 82 in the Connor Davidson Resilience Scale (CD-RISC-25), 4) male or female participants above 18 years of age, and 5) ability to speak and comprehend Tamil or English. Participants with any psychiatric conditions, diabetic complications requiring additional treatment that interfered with participation, and lactating and pregnant women were excluded.

The study received ethical approval from the Institutional Ethics Committee (CSP/25/MAR/159/171), Sri Ramachandra Institute of Higher Education and Research (Category-1 deemed to be university). The study was conducted in accordance with the ethical principles of the declaration of Helsinki (World Medical Association, 2013 revision). The procedures, potential risks, and benefits were clearly explained before recruitment, and informed consent was obtained from all eligible and willing participants before recruitment. The participant flow is shown in Figure 1.

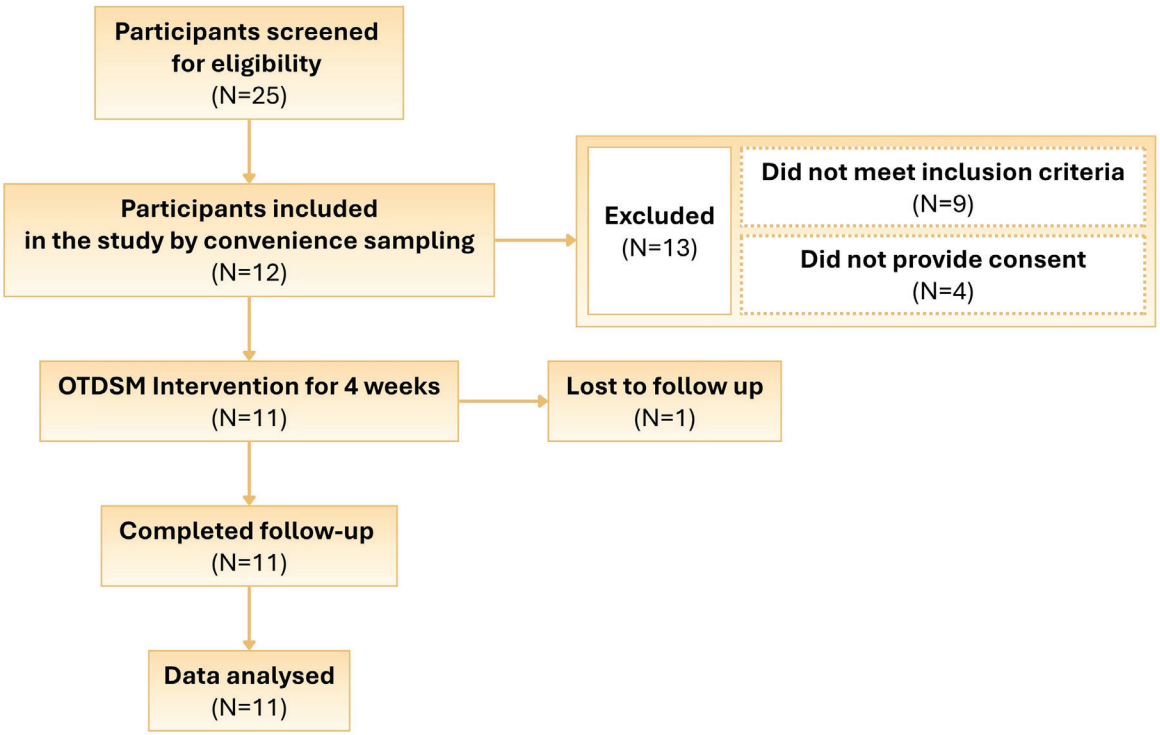


Figure 1. *The flow of participants.*

Intervention

Occupational Therapy Diabetes Self Management (OTDSM) intervention used in this study was adapted from the OTDSM framework developed by Binesh *et al.*,[15] with cultural and contextual modifications made to suit the current study population. OTDSM is a well-structured, occupation-centered DSM intervention that focuses on integrating self-management behaviors, including following a diet, being physically active, psychological coping, periodic glucose testing, etc., delivered through education, counseling, and activities to make them part of the individual's meaningful daily routines.

The intervention in this study was group-based and held over four consecutive weeks, comprising 10 sessions lasting 60 minutes each. The components of the intervention were educating and giving informations about diabetes, risk factors, and complications; skill-building; problem-solving; stress management and coping strategies; and occupation-based practice, integrating self-care behavior into everyday roles and routines. The sessions followed a structured format consisting of warm-up and orientation activities, an educational discussion, a demonstration, guided hands-on practice, group interaction, feedback, and wind-down activities, all delivered in a supportive group setting. The sessions were facilitated by the principal investigator under the supervision of a licensed occupational therapist and emphasized active participation and application through DSM kits and educational brochures. A depiction of the intervention is provided in Table 1.

Table 1. Intervention protocol.

Session	Session topic	Objectives	Activities
Session 1	Introduction to OTDSM and diabetes	Improve understanding of diabetes, symptoms, risk factors, and complications	Ice breaker activities and rapport building, educational discussion, visual examples, and group sharing of experiences. Distribution of a diabetes self-care practices brochure in the local language
Session 2	Self-monitoring of blood glucose	Improve knowledge and practice of blood glucose monitoring and hypoglycaemia management	Need for, types, guidelines and frequency of blood glucose testing through brochures, glucometer demonstration, guided hands-on glucose testing, glucose logs, hypoglycaemia plan

Table 1. Intervention protocol (continued).

Session	Session topic	Objectives	Activities
Session 3	Dietary habits and recommendations	Promote healthy dietary choices and meal planning	Eating tips, diabetes plate method, culturally adapted food models, sample meal plans, Plate It Right game to understand the diabetes plate model.
Session 4	Diabetic foot disease	Improve awareness of diabetic foot complications and self-inspection	Education on common diabetic foot complications, risk factors, and preventive measures; self examination and inspection; demonstration and practice. examination kits (mirrors and monofilaments)
Session 5	Diabetic foot care	Develop foot care and safe footwear habits	Foot hygiene and nail care practice demonstration, footwear selection and maintenance education with visual examples, foot care kits, and foot exercises.
Session 6	Aerobic exercises	Encourage participation in physical activity	Seated and standing aerobic exercises, breathing exercises
Session 7	Strengthening exercises	Improve strength and physical fitness	Sit-to-stand, wall/ chair push-ups, calf raises, leg/ biceps curls
Session 8	Flexibility exercises, 8-shaped pathway walking	Improve flexibility and fitness	Stretching exercises, walking an 8-shaped path
Session 9	JPMR, guided imagery	Promote stress management and coping strategies	Jacobson's progressive muscle relaxation, guided imagery using diabetes self-care scripts.
Session 10	Mandala art	Promote emotional expression and relaxation	Pre-drawn/free-drawn mandala art activity, music-assisted calming, storytelling activities.

Data collection tools

Three standardized tools were used to assess the outcomes. To ascertain the reliability and robustness of data collection, the instruments were evaluated for contextual and language sensitivity, and the examiner completed comprehensive training in administering and scoring the measures. All tools were administered at baseline and immediately after completion of the four-week intervention.

Diabetes Distress Scale (DDS-17)

The DDS-17 is a self-administered measure of 17 items. It assesses the emotional burden (EB), physician distress (PD), regimen distress (RD), and interpersonal distress (ID) domains. Items are graded on a 6-point Likert scale, where 1 represents "not a problem" and 6 represents a "very serious problem," with higher scores indicating elevated levels of distress. It is reliable, with a high internal consistency (Cronbach's alpha 0.93), and can therefore be used effectively in clinical and research purposes.[18]

Connor-Davidson Resilience Scale (CD-RISC-25)

CD-RISC-25 is a 25-item instrument. The items use a 5-point Likert scale (0 = not true at all to 4 = true nearly all the time). The total scores range from 0 to 100. The higher the scores, the more resilient the person is. The scale has adequate reliability with a Cronbach's alpha of 0.89.[19]

Summary of Diabetes Self-Care Activities Questionnaire (SDSCA)

The SDSCA is a self-report instrument that assesses the frequency of diabetes self-care behaviors over the last seven days, including diet, exercise, blood glucose testing, foot care, and medication adherence. Responses range from 0 to 7 days. Higher scores reflect better self-management. The SDSCA demonstrates acceptable internal consistency (Cronbach's α=0.61-0.85 across subscales) and construct validity. This questionnaire was used to measure self-care behaviors at 5 time points throughout the study (pre-test and weeks 1-4).[20]

Data analysis

The unit of assignment and analysis was the individual participant. Data were analyzed with SPSS software (version 20). Descriptive statistics were used to summarise the demographic baseline characteristics. Normality was assessed prior to inferential analysis using the Shapiro-Wilk test, and the assumption of normality was considered satisfied for parametric analyses. Paired t-tests were employed to analyze the within-group differences between pre- and post-test scores for diabetes distress and resilience. Repeated measures ANOVA with Bonferroni-adjusted pairwise comparison was employed to analyze within-group changes in self-care behavior scores across multiple related time points (baseline and weekly follow-ups). The assumption of sphericity was evaluated using Mauchly's test before RM ANOVA, and Greenhouse-Geisser corrections were applied when the assumption was not met. Effect sizes were calculated using Cohen's

d for paired t-tests and partial eta squared (η^2p) for RM ANOVA. The $p<0.05$ was considered statistically significant.

Results

Participant characteristics

Of the 12 participants enrolled, 11 completed the 4-week OTDSM intervention; one was lost to follow-up. Participants were mostly female (90.9%), with a mean age of 48.45±6.15 years. The mean duration of diabetes was 4.55±2.70 years, reflecting early- to intermediate-stage disease, a critical period for adopting self-management practices (Table 2).

Table 2. Participants' demographic characteristics.

Characteristics	**Intervention group** (N=11)
Age, years (mean±SD)	48.45±6.15
Gender, N (%)	
- Female	10 (90.9)
- Male	1 (9.1)
Duration of diagnosis, years (%)	
1	2 (18.2)
2	1 (9.1)
3	0
4	3 (27.2)
5	1 (9.1)
6	2 (18.2)
7	1 (9.1)
8	0
9	0
10	1 (9.1)

Note: *Gender and duration of diagnosis are reported as frequency and %.*

Diabetes distress

The total DDS-17 score showed a statistically significant reduction from 3.72±0.25 at baseline to 2.53±0.28 post-intervention ($p<0.001$). Decreases were observed in the emotional burden ($p=0.001$), physician distress ($p<0.001$), regimen distress ($p<0.001$), and interpersonal distress ($p=0.005$) subscales. The highest significant reduction was observed in regimen distress (Table 3).

Table 3. Comparison of diabetes distress and resilience pre- and post-intervention.

Outcome measure	**Mean±SD**		**Mean difference±SD**	**95% CI**	***t*-value**	***p* value**	***Cohen's d***
	Pre-test	**Post-test**					
Diabetes distress (DDS-17)							
Emotional burden	3.53±0.48	3.02±0.37	0.51±0.35	0.20 to 0.74	4.82	0.001	1.45
Physician distress	4.02±0.79	2.14±0.42	1.89±0.44	1.59 to 2.18	14.28	<0.001	4.30
Regimen distress	4.29±0.27	2.55±0.53	1.75±0.39	1.48 to 2.01	14.81	<0.001	4.47
Interpersonal distress	2.66±0.94	2.27±0.67	0.39±0.36	0.15 to 0.63	3.61	0.005	1.09
Diabetes Distress Scale (DDS-17) total	3.72±0.25	2.53±0.28	1.19±0.18	1.06 to 1.31	21.43	<0.001	6.46
Resilience (CD-RISC-25)							
Resilience	42.91±3.53	49.45±3.24	-6.55±0.93	-7.17 to -5.92	–23.24	<0.001	7.01

Note: *Values are reported as mean±SD at pre- and post-test, p<0.05 indicates a statistically significant difference (two-tailed), CI: confidence interval, t-value: T-statistic, Cohen's d: Cohen's d effect size.*

Resilience

Resilience scores showed statistically significant improvement following the intervention, from 42.91±3.53 at baseline to 49.45±3.24 post-intervention (p<0.001) (Table 3).

Self-care behavior

Self-care behavior was assessed at five time points (pre-test through week 4). Repeated-measures ANOVA showed significant time changes across the diet, physical activity, blood glucose testing, and foot care domains (p<0.001). Post hoc pairwise comparisons were conducted to identify the specific time points at which changes occurred in each self-care domain following the overall RM ANOVA findings. Bonferroni correction (α=0.0125) was applied to control for multiple comparisons; thus, p<0.0125 was considered statistically significant for pairwise comparison of the SDSCA self-care domains.

Diet scores increased progressively from 2.73±0.31 at baseline to 4.98±0.48 at week 4 (F=63.35, p<0.001). Physical activity scores showed delayed improvements, with no significant changes observed during weeks 1 and 2, and becoming significant from week 3 onward (F=221.03, p<0.001), consistent with the intervention protocol. Blood glucose testing scores rose from 0.00 at baseline to 5.27±0.65 at week 4, with significant weekly improvements (F=175.02, p<0.001). Foot care scores improved steadily, from 1.80±0.51 at baseline to 6.02±0.47 at week 4, with significant changes from week 2 (F=211.32, p<0.001). Medication adherence remained high, reaching the ceiling effect score of 7 across all weeks (p=0.003) (Table 4).

Table 4. Comparison of self-care behavior pre- and post-intervention.

SDSCA self-care domain	Time Point (Week)	Mean±SD	Repeated measures ANOVA		η^2p
			F-value	*p* value	
Diet	Pre-test	2.73±0.31	63.35	<0.001	0.864
	1	3.34±0.34			
	2	3.92±0.51			
	3	4.23±0.38			
	4	4.98±0.48			
Physical activity	Pre-test	0.18±0.34	221.03	<0.001	0.957
	1	0.14±0.32			
	2	0.14±0.32			
	3	2.27±0.47			
	4	3.09±0.63			
Blood sugar testing	Pre-test	0.00±0.00	175.02	<0.001	0.946
	1	2.82±0.60			
	2	4.36±0.92			
	3	4.55±0.69			
	4	5.27±0.65			
Foot Care	Pre-test	1.80±0.51	211.32	<0.001	0.955
	1	2.07±0.83			
	2	4.11±0.67			
	3	5.35±0.65			
	4	6.02±0.47			

Table 4. Comparison of self-care behavior pre- and post-intervention (continued).

SDSCA self-care domain	Time Point (Week)	Mean±SD	Repeated measures ANOVA		η^2p
			F-value	*p* value	
Medication	Pre-test	6.55±0.69	4.81	0.003	0.325
	1	7.00±0.00			
	2	7.00±0.00			
	3	7.00±0.00			
	4	7.00±0.00			

Note: *Values are reported as mean±SD at each time point, F-value: F-statistic (analysis of variance-ANOVA), η^2p: partial eta squared effect size, p<0.05 indicates a statistically significant difference (two-tailed).*

Post hoc pairwise comparisons confirmed that the greatest improvements across all domains occurred by week 4 (p<0.0125), reflecting the cumulative progression of self-care behavioral changes throughout the intervention period. (Table 5).

Table 5. Time-wise comparison of self-care behavior pre- and post-intervention.

Domain	Comparison	Mean Difference	Standard Error	95% CI	Bonferroni-adjusted α	Overall ANOVA p value
Diet	Pre- vs week 1	-0.614	0.097	-0.96 to -0.26	0.001	<0.001
	Pre- vs week 2	-1.191	0.173	-1.81 to -0.57	<0.001	
	Pre- vs week 3	-1.500	0.139	-1.99 to -1.00	<0.001	
	Pre- vs week 4	-2.250	0.147	-2.78 to -1.72	<0.001	
Physical activity	Pre- vs week 1	0.045	0.045	-0.12 to 0.21	1.000	<0.001
	Pre- vs week 2	0.045	0.045	-0.12 to 0.21	1.000	
	Pre- vs week 3	-2.091	0.148	-2.62 to -1.56	<0.001	
	Pre- vs week 4	-2.909	0.200	-3.63 to -2.19	<0.001	
Blood sugar testing	Pre- vs week 1	-2.818	0.182	-3.47 to -2.17	<0.001	<0.001
	Pre- vs week 2	-4.364	0.279	-5.36 to -3.37	<0.001	
	Pre- vs week 3	-4.545	0.207	-5.29 to -3.80	<0.001	
	Pre- vs week 4	-5.273	0.195	-5.97 to -4.57	<0.001	
Foot care	Pre- vs week 1	-0.273	0.156	-0.83 to 0.29	1.000	<0.001
	Pre- vs week 2	-2.309	0.163	-2.89 to -1.73	<0.001	
	Pre- vs week 3	-3.545	0.171	-4.16 to -2.93	<0.001	
	Pre- vs week 4	-4.218	0.167	-4.82 to -3.62	<0.001	
Medication	Pre- vs week 1	-0.455	0.207	-1.19 to 0.288	0.531	0.003
	Pre- vs week 2	-0.455	0.207	-1.19 to 0.288	0.531	
	Pre- vs week 3	-0.455	0.207	-1.19 to 0.288	0.531	
	Pre- vs week 4	-0.455	0.207	-1.19 to 0.288	0.531	

Note: *Values are reported as mean difference between baseline and weekly follow-up scores, CI: confidence interval; Bonferroni-adjusted p<0.0125 indicates a statistically significant difference (two-tailed).*

The effect size estimates indicate substantial magnitudes of change across all the assessed outcomes during the intervention period. Overall, reductions in diabetes distress, improvements in resilience, and positive changes in self-care behavior were observed following the OTDSM intervention.

Discussion

This preliminary study explored changes in diabetes distress, resilience, and self-care behavior following the Occupational Therapy Diabetes Self Management (OTDSM) intervention among adults with T2DM. Findings suggest statistically significant decreases in diabetes distress, improvements in resilience, and significant changes across all domains of diabetes self-management within the study sample.

The participants had a mean age of 48.45±6.15 years, indicating that the intervention targeted middle-aged adults. Middle age is a vulnerable period for diseases, because it often involves managing personal health alongside family and occupational responsibilities. OTDSM may be compatible with participants' daily routines, cultural backgrounds, and behavioral needs, suggesting its potential as a model for DSME delivery at this phase of life.

The participants' occupation was cleaning and maintenance work, which involved long hours and physically demanding work. Due to their work schedules, low literacy, and limited access to healthcare, these individuals face difficulties adhering to self-management practices.[21] Traditional DSME formats lack the flexibility to cover the occupational realities. Occupational Therapy Diabetes Self-Management (OTDSM), which may be offered conveniently and flexibly while respecting culture-specific factors, may be a possible approach to addressing this gap by integrating self-management into patients' day-to-day routines and habits, thereby increasing accessibility and compliance.[15]

Most of the participants were women, consistent with the cleaning sector workforce.[22] Women are more prone to poor diabetes self-management,[23] lower psychological well-being, and higher diabetes distress.[24] In this study, females accounted for 90.9% of participants, underlining the need for self-management interventions targeting this group. OTDSM may appear relevant in supporting these needs by providing peer support, skill-building, and knowledge necessary for autonomous diabetes management.

Following OTDSM intervention, reductions in diabetes distress were observed across all DDS-17 subscales. The reductions may be explained by the intervention's health education components, which included understanding the disease, identifying risk factors and symptoms, awareness of complications, and planning for sick days. Education was provided in plain language and reinforced with interactive activities. These elements may have supported the participants' understanding and control over their condition. Better understanding and perceived control, brought on by addressing both informational and emotional demands, may have contributed to decreases in regimen and emotional distress, consistent with findings from Guo *et al.* and Perrin *et al.*, suggesting a possible role for culturally relevant DSM education and psychoeducational interventions in reducing diabetes distress.[25,26] Hands-on activities such as using glucometers, foot examinations, foot hygiene practices, and culturally appropriate plate models may support autonomy and self-esteem, which in turn may further reduce distress. Similar findings were reported by Misra *et al.,* suggesting that experiential learning may be associated with decreased diabetes distress.[27]

There was no direct physician-client interaction in this study; however, it is possible that the group-specific focus of the OTDSM, which provided peer support, shared experiences, reassurance, and feedback through the OT facilitator, may have indirectly supported medical communication concerns, which in turn may have contributed to the observed decrease in physician distress. Peer interactions may have supported decreases in interpersonal distress. Mathiesen *et al*. and Hessler *et al*. emphasized that when psychosocial interventions centered on peer support were provided to vulnerable populations, diabetes distress, particularly in the interpersonal domain, decreased significantly, while improvements were observed in coping skills, trust-building, and collaborative problem-solving.[28,29]

Improvements in resilience were observed following the OTDSM intervention, which may be explained by elements such as active learning, routine-building, autonomy, control, coping mechanisms, feedback, and group interaction. This is consistent with Wilson *et al*. and Dubois *et al.,* who emphasized that DSME promotes resilience alongside diabetes self-management.[30,31] OTDSM also includes techniques prevalent in resilience training interventions, such as deep breathing, imagery, peer guidance, competence, and control, which may be associated with positive changes in resilience. The simultaneous observation of reduced distress and increased resilience scores may indicate that resilience is a protective factor against DD, consistent with previous findings.[30]

Progressive weekly improvements were observed in all SDSCA domains, especially in diet, blood glucose testing, physical activity, and foot care. This may reflect improved capacity for day-to-day self-care practices, following the OTDSM intervention. The positive changes revealed in this study are congruent with Binesh *et al.* and Pyatak *et al*., who described the possible role of occupation-based, person-centered Occupational Therapy interventions in DSM.[16,32] The medication adherence domain reached a ceiling effect, which may suggest a pre-existing habit that formed from previous healthcare encounters. OTDSM may have reaffirmed this formed habit. This trend aligns with a prior study, which suggests medication adherence is the most consistent self-care area, as it does not demand as much behavioral restructuring as other domains.[32]

OTDSM may help address barriers to diabetes self-management, as identified by Oluchina *et al*.[13] The educational, demonstrational, and hands-on elements,

peer-group setting, and interactive activities may have contributed to the observed changes in both diabetes self-management and psychological outcomes. These tendencies may suggest that structured community-based DSME, such as the OTDSM, may help address self-care adherence among groups living in low socio-economic situations.

This study further aligns with the United Nations Sustainable Development Goals. In particular, it is relevant to SDG 3: good health and well-being by showing preliminary changes in distress and resilience levels, and supporting strategies that may enhance self-management for the prevention and management of chronic non-communicable diseases (Target 3.4).

Clinical implications

This study highlights the possible role of Occupational Therapy Diabetes Self Management (OTDSM). The OTDSM intervention suggests that occupation-based, client-centered practices may be associated with reductions in diabetes distress and improvements in psychological resilience, both of which are necessary to sustain regular self-care behavior. The flexible, feasible, low-cost, and adaptable nature of OTDSM may support its use across a broad adult population, including those with demanding jobs and varying levels of literacy. Occupational therapists may play a key role in diabetes care teams by embedding self-management behaviors into meaningful daily routines to improve functional outcomes in primary and community care settings.

Limitations

The following limitations, however, should be addressed. The small sample size limits generalizability. The use of a quasi-experimental single-group design without a control group limits causal inference about the intervention's effectiveness. The study's focus on a specific occupational group limits generalisability to other adult T2DM populations. Future research should conduct randomized controlled trials with larger, more diverse samples and extend the intervention duration to examine the long-term effectiveness of OTDSM. Additionally, adapting OTDSM to digital platforms may improve accessibility and sustained engagement.

Conclusion

This preliminary study observed changes in diabetes distress, resilience, and self-care behavior following the occupational therapy diabetes self-management (OTDSM) intervention among adults with type 2 diabetes. By incorporating culturally adapted education, routine practices, and psychological support into a structured, occupation-focused program, OTDSM may offer a possible approach to addressing both behavioral and psychological components of diabetes self-management. The present findings provide preliminary observations on the possible role of Occupational Therapy in diabetes care.

Ethical approval

The study received approval from the Institutional Ethics Committee (CSP/25/MAR/159/171), Sri Ramachandra Institute of Higher Education and Research (Category-1 Deemed to be University). The study was conducted in accordance with the ethical principles of the Declaration of Helsinki (World Medical Association, 2013 revision). The procedures, potential risks, and benefits of the study were clearly explained before recruitment, and written informed consent was obtained from all eligible and willing participants.

Funding

This research received no external funding.

Conflict of interest

The authors declare no conflict of interest.

CRedit authorship contribution statement

Kamala Arumugam: conceptualization, data curation, investigation, methodology, project administration, writing: original draft; **Loganathan Shanmugam, Raghuram Pattusamy, Darsana Ganesan, Jaivignesh Shanmugam:** supervision, validation, writing: draft and review.

Acknowledgements

We are grateful to all study participants for their cooperation and willingness to participate.

References

[1] Banday MZ, Sameer AS, Nissar S. Pathophysiology of diabetes: An overview. Avicenna J Med. 2020; 10(4): 174-88. doi: 10.4103/ajm.ajm_53_20.

[2] Pradeepa R, Mohan V. Epidemiology of type 2 diabetes in India. Indian J Ophthalmol. 2021; 69(11): 2932-8. doi: 10.4103/ijo.IJO_1627_21.

[3] Roglic G. WHO global report on diabetes: A summary. Int J Noncommunicable Dis. 2016; 1(1): 3-8. doi: 10.4103/2468-8827.184853.

[4] Fisher L, Hessler DM, Polonsky WH, Mullan J. When is diabetes distress clinically meaningful? Establishing cut points for the Diabetes Distress Scale. Diabetes Care. 2012; 35(2): 259-64. doi: 10.2337/dc11-1572.

[5] Morales-Brown LA, Algorta GP, Salifu Y. Understanding experiences of diabetes distress: A systematic review and thematic synthesis. J Diabetes Res. 2024; 2024: 3946553. doi: 10.1155/2024/3946553.

[6] Khashayar P, Shirzad N, Zarbini A, Esteghamati A, Hemmatabadi M, Sharafi E. Diabetes-related distress and its association with complications of diabetes in Iran. J Diabetes Metab Disord. 2022; 21: 1569-75. doi: 10.1007/s40200-022-01103-2.

[7] Skinner TC, Joensen L, Parkin T. Twenty-five years of diabetes distress research. Diabet Med. 2020; 37(3): 393-400. doi: 10.1111/dme.14157.

[8] Kim GM, Lim JY, Kim EJ. Resilience of patients with chronic diseases: A systematic review. Health Soc Care Community. 2019; 27(6): e460-72. doi: 10.1111/hsc.12620.
[9] Jin Y, Bhattarai M, Kuo WC, Bratzke LC. Relationship between resilience and self-care in people with chronic conditions: A systematic review and meta-analysis. J Clin Nurs. 2022; 31(15-16): 2121-37. doi: 10.1111/jocn.16258.
[10] Chew BH, Shariff-Ghazali S, Fernandez A. Psychological aspects of diabetes care: Effecting behavioral change in patients. World J Diabetes. 2014; 5(6): 796-808. doi: 10.4239/wjd.v5.i6.796.
[11] Newman S, Steed L, Mulligan K. Self-management interventions for chronic illness. Lancet. 2004; 364(9444): 1523-37. doi: 10.1016/S0140-6736(04) 17277-2.
[12] Maina PM, Pienaar M, Reid M. Self-management practices for preventing complications of type II diabetes mellitus in low- and middle-income countries: A scoping review. Univ Free State Fac Health Sci Sch Nurs. 2022. doi: 10.38140/ufs.17206751.
[13] Oluchina S, Karanja S. Barriers to diabetes self-management in primary care settings-patient perspectives: A phenomenological design. Int J Afr Nurs Sci. 2022; 18: 100465. doi: 10.1016/j.ijans.2022.100465.
[14] Shen X, Shen X. The role of occupational therapy in secondary prevention of diabetes. Int J Endocrinol. 2019; 2019: 3424727. doi: 10.1155/2019/3424727.
[15] Binesh M, Mehraban AH, Shahboulaghi FM, Aghili R, Shafaroodi N. Development and preliminary feasibility evaluation of an occupation-centered diabetes self-management intervention. Br J Occup Ther. 2021; 84(12): 734-42. doi: 10.1177/03080226211030100.
[16] Binesh M, Motaharinezhad F, Shafaroodi N, et al. A randomized controlled trial for evaluating an occupational therapy self-management intervention in adults with type 2 diabetes. Sci Rep. 2023; 13: 10128. doi: 10.1038/s41598-023-37231-9.
[17] Gregg JA, Callaghan GM, Glenn-Lawson JL, Hayes SC. Improving diabetes self-management through acceptance, mindfulness, and values: A randomized controlled trial. J Consult Clin Psychol. 2007; 75(2): 336-43. doi: 10.1037/0022-006X.75.2.336.
[18] Polonsky WH, Fisher L, Earles J, et al. Assessing psychosocial distress in diabetes: Development of the Diabetes Distress Scale. Diabetes Care. 2005; 28(3): 626-31. doi: 10.2337/diacare.28.3.626.
[19] Connor KM, Davidson JRT. Development of a new resilience scale: The Connor-Davidson Resilience Scale (CD-RISC). Depress Anxiety. 2003; 18(2): 76-82. doi: 10.1002/da.10113.
[20] Toobert DJ, Hampson SE, Glasgow RE. The Summary of Diabetes Self-Care Activities measure: Results from 7 studies and a revised scale. Diabetes Care. 2000; 23(7): 943-50. doi: 10.2337/diacare.23.7.943.
[21] Brewer LC, Cooper LA, Patten CA. Diabetes self-management education for special populations: The social determinants of health matter. J Gen Intern Med. 2020; 35(Suppl 2): 555-61. doi: 10.1177/0033354919830968.
[22] Fernández-Suárez I, López-Goñi JJ, Haro B. Profiles of women who have suffered occupational accidents in cleaning: Perceived health, psychosocial risks, and personality variables. Int Arch Occup Environ Health. 2022; 96(2): 331-40. doi: 10.1007/s00420-022-01927-8.
[23] Baroni I, Caruso R, Dellafiore F, et al. Self-care and type 2 diabetes mellitus: A literature review on sex-related differences. Acta Biomed. 2022; 93(4): e2022277. doi: 10.23750/abm.v93i4.13324.
[24] Delahanty LM, Grant RW, Wittenberg E, et al. Association of diabetes-related emotional distress with diabetes treatment in primary care patients with type 2 diabetes. Diabet Med. 2007; 24(1): 48-54. doi: 10.1111/j.1464-5491.2007.02028.x.
[25] Guo X, Wong PNF, Koh YLE, Tan NC. Factors associated with diabetes-related distress among Asian patients with poorly controlled type 2 diabetes mellitus. BMC Prim Care. 2023; 24: 54. doi: 10.1186/s12875-023-02012-w.
[26] Perrin N, Bodicoat DH, Davies MJ, Robertson N, Snoek FJ, Khunti K. Effectiveness of psychoeducational interventions for diabetes-specific emotional distress and glycaemic control in people with type 2 diabetes: A systematic review and meta-analysis. Prim Care Diabetes. 2019; 13(4): 289-307. doi: 10.1016/j.pcd.2019.04.001.
[27] Misra R, Shawley-Brzoska S, Khan R, et al. Addressing diabetes distress in self-management programs: Results of a randomized feasibility study. J Appalach Health. 2021; 3(3): 68-85. doi: 10.13023/jah.0303.06.
[28] Mathiesen AS, Egerod I, Jensen T, et al. Psychosocial interventions for reducing diabetes distress in vulnerable people with type 2 diabetes mellitus: A systematic review and meta-analysis. Diabetes Metab Syndr Obes. 2019; 12: 19-33. doi: 10.2147/DMSO.S179301.
[29] Hessler D, Strycker L, Fisher L. Reductions in management distress following a randomized distress intervention are associated with improved diabetes behavioral and glycemic outcomes over time. Diabetes Care. 2021; 44(6): 1472-9. doi: 10.2337/dc20-2724.
[30] Wilson AL, McNaughton D, Meyer SB, Ward PR. Understanding the links between resilience and type 2 diabetes self-management: A qualitative study in South Australia. Arch Public Health. 2017; 75: 56. doi: 10.1186/s13690-017-0222-8.
[31] Dubois SK, Lehrer HM. A resilience intervention for adults with type 2 diabetes: Proof-of-concept in community health centers. Int J Behav Med. 2020;

27(5): 592-601. doi: 10.1007/s12529-020-09894-5.

[32] Pyatak EA, Carandang K, Vigen CLP, Blanchard J, Diaz J, Concha-Chavez A, et al. Occupational therapy intervention improves glycemic control and quality of life among young adults with diabetes: The REAL Diabetes randomized controlled trial. Diabetes Care. 2018; 41(4): 696-704. doi: 10.2337/dc17-1634.

Journal of Associated Medical Sciences 2026; 59(3): 196-204

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



Design and preliminary study of VP1-encoding mRNA-lipid nanoparticles for Enterovirus 71

Parnuphan Panyajai[1*], Rujiraporn Pitaksalee[1], Apichai Prachasuphap[1], Kodcharad Jongpitisub[1], Pantida Treeyoung[1], Sirawit Wet-osot[1], Nalinee Saengtong[1], Nadthanan Pinyosukhee[1], Rattanawadee Wichajarn[1], Kritsamon Sophondilok[1], Anicha Luengchaichaweng[1], Panadda Dhepakson[2]

[1]Medical Life Sciences Institute, Department of Medical Sciences, Nonthaburi Province, Thailand.
[2]Bureau of Expert Advisory Panels, Department of Medical Sciences, Nonthaburi Province, Thailand.

ARTICLE INFO

Article history:
Received 30 March 2026
Accepted as revised 18 June 2026
Available online 29 June 2026

Keywords:
Hand, foot, and mouth disease, Enterovirus 71, mRNA-LNPs, lipid nanoparticles, VP1

ABSTRACT

Background: Hand, foot, and mouth disease (HFMD) remains a major public health concern in the Asia-Pacific region, with Enterovirus 71 (EV71) being a major cause of severe neurological complications in children. The genetic diversity of EV71 highlights the need for adaptable vaccine platforms.

Objective: This study aimed to design and preliminarily evaluate an EV71 VP1-encoding mRNA-lipid nanoparticle (mRNA-LNP) construct as an antigen delivery approach for EV71 vaccine development.

Materials and methods: A codon-optimized VP1 mRNA was synthesized by *in vitro* transcription and formulated into lipid nanoparticles via microfluidic mixing. Physicochemical properties were characterized, and *in vitro* expression was assessed in FreeStyle 293F cells by western blot and immunofluorescence. VP1-specific humoral immune responses were evaluated in BALB/c mice following a prime-boost immunization regimen using ELISA.

Results: The mRNA-LNPs exhibited a mean particle size of 99.24±0.46 nm, a low polydispersity index (0.041±0.004), a slightly negative zeta potential (-20.54±0.98 mV), and an encapsulation efficiency of 92.56±1.53%. *In vitro* studies demonstrated VP1 protein expression following mRNA delivery. Immunization induced VP1-specific IgG responses, with antibody levels significantly increased after prime (1.254±0.775) and boost (4.225±0.322) compared with control groups ($p<0.01$ and $p<0.001$, respectively). Endpoint titers increased from a geometric mean titer of 2,000 (95% CI: 901.50-4,437.00) after priming to 256,000 following booster immunization.

Conclusion: This study demonstrates that EV71 VP1 mRNA-LNPs induce antigen-specific humoral immune responses in mice, providing preliminary evidence of their potential as an antigen delivery approach for EV71 and supporting further studies in vaccine development.

Introduction

Hand, foot, and mouth disease (HFMD) is a common pediatric viral illness and remains a significant public health concern in the Asia-Pacific region. Among causative enteroviruses, Enterovirus 71 (EV71) is of particular importance due to its strong association with severe neurological complications, including aseptic meningitis, brainstem encephalitis, and acute flaccid paralysis.[1,2]

*** Corresponding contributor.**
Author's Address: *Medical Life Sciences Institute, Department of Medical Sciences, Nonthaburi Province, Thailand*
E-mail address: parnuphan.p@dmsc.mail.go.th
doi: 10.66285/JAMS.2026.093
E-ISSN: 2539-6056

Epidemiological studies have demonstrated the continued circulation of genetically diverse EV71 strains, contributing to recurrent outbreaks. In Thailand, subgenotypes B5 and C4 have predominated over the past decade, reflecting ongoing circulation and shifts in subgenotype dominance.[3,4] This highlights the need for vaccine strategies capable of accommodating regional strain diversity.

Several inactivated EV71 vaccines have been licensed in Asia, including C4-based vaccines in mainland China.[1,5] In Thailand, an inactivated EV71 vaccine, EntroVac®, has been approved for use in young children and is based on a C4 subgenotype strain.[6] These vaccines have demonstrated efficacy against EV71-associated HFMD; however, cross-genotype protection remains an important consideration in regions where different EV71 subgenotypes circulate.[1,5] This consideration is relevant to the Thai setting, where circulating EV71 subgenotypes include both B5 and C4. In addition, serological data in Thai children indicate that a substantial proportion remain susceptible to EV71 infection during early childhood,[7] supporting the need for effective vaccination strategies. Alternative vaccine approaches, including recombinant protein, peptide-based, and virus-like particle (VLP) vaccines, have demonstrated immunogenicity in experimental models, although most are still under development.[1,2] However, conventional vaccine strategies generally rely on virus propagation or recombinant protein production, which involve multi-step manufacturing processes and may limit the speed and flexibility of antigen redesign.

In this context, messenger RNA (mRNA) vaccine technology offers a modular and adaptable platform that enables rapid antigen design and cell-free, scalable production,[8,9] supporting its potential as a flexible platform for EV71 antigen delivery. However, mRNA requires an effective delivery system to protect it from degradation and enable cellular entry. Lipid nanoparticles are widely used for this purpose because they encapsulate mRNA and facilitate intracellular delivery for antigen expression.[9,10]

EV71 encodes four structural proteins (VP1–VP4), among which VP1 is a major surface-exposed capsid protein involved in receptor binding and viral entry. VP1 contains key neutralizing epitopes,[1,2,11] and its relatively conserved sequence across subgenotypes supports its potential as a cross-reactive antigen target.[4]

In this study, we designed and evaluated an EV71 VP1-encoding mRNA-LNP construct, integrating antigen design, mRNA synthesis, lipid nanoparticle (LNP) formulation, and evaluation of VP1-specific humoral immune responses in a mouse model. This study provides initial experimental data to support further evaluation and optimization of EV71 mRNA-LNP construct in the context of vaccine development.

Materials and methods

mRNA construct design and template preparation

The full-length VP1 coding sequence of EV71 subgenotype B5 isolate SiICRC04/TH/2012 (GenBank accession No. KX372317) was used as the antigen template.[3] The mRNA construct comprised a 5' untranslated region (UTR) containing a Kozak sequence, a tissue plasminogen activator (tPA) signal peptide, the VP1 coding sequence, and a 3' UTR with a poly(A) tail.

The VP1 sequence was reverse-translated and codon-optimized for human expression using mRNAid (version 1.0.0). Rare codon usage was evaluated using the GeneScript Rare Codon Analysis Tool, and RNA secondary structure was predicted using RNAfold to estimate thermodynamic stability.

The optimized VP1 coding sequence was synthesized as double-stranded DNA (Gene Universal Inc., USA) and cloned into the Linearized Template Vector (BspQ I), supplied with the Cloning Kit for mRNA Template, using an In-Fusion cloning system (Takara Bio Inc., Japan). Recombinant plasmids were transformed into the *Escherichia coli* Stbl2 strain, and correct insertion was confirmed by colony PCR and Sanger sequencing. Plasmids were purified and linearized with the BspQI restriction enzyme to generate a DNA template for *in vitro* transcription.

In vitro transcription and mRNA characterization

In vitro transcription (IVT) was performed using a CleanCap-based IVT system (TriLink BioTechnologies, USA) with N1-methylpseudouridine-5'-triphosphate substituted for UTP and co-transcriptional capping. Template DNA was removed by DNase I treatment (New England Biolabs, USA), and mRNA was purified using an RNA Cleanup Kit (New England Biolabs, USA).

mRNA concentration was determined using a RiboGreen-based assay (Invitrogen, USA). Residual DNA contamination was assessed by PCR amplification of the VP1 region, with linearized plasmid DNA used as a positive control. mRNA size and integrity were analyzed using a Bioanalyzer system (Agilent Technologies, USA). Sequence accuracy was confirmed by reverse transcription followed by Sanger sequencing.

Preparation and characterization of mRNA-LNPs

mRNA was encapsulated into lipid nanoparticles (LNPs) using GenVoy-ILM reagents and a NanoAssemblr Ignite microfluidic mixing platform (Precision Nanosystems, Canada). Lipids were diluted in ethanol, and mRNA was diluted in acetate buffer (pH 4.0) to achieve a lipid-to-mRNA molar ratio of 4:1 (ionizable lipid: nucleotide). The solutions were combined at a 3:1 (v/v) ratio with a total flow rate of 12 mL/min. Formulated mRNA-LNPs were diluted in phosphate-buffered saline (PBS, pH 7.4) and concentrated using centrifugal filters. Empty LNPs, used as a formulation

control for immunofluorescence analysis and animal immunization, were prepared using the same microfluidic mixing procedure as the mRNA-LNPs, but without mRNA.

Particle size, polydispersity index (PDI), and zeta potential of the mRNA-LNP formulation were measured by dynamic light scattering using a Zetasizer instrument (Malvern Panalytical, UK). Samples were diluted in PBS (200-fold for particle size and PDI, 30-fold for zeta potential) and measured at 25 °C. Encapsulation efficiency (%EE) was determined using a RiboGreen assay in the presence or absence of 1% Triton X-100 to distinguish total and unencapsulated mRNA. Fluorescence was measured, and %EE was calculated as:

$$\%EE = [(\text{total mRNA-free mRNA})/\text{total mRNA}]\times 100$$

All characterization measurements of the mRNA-LNP formulation were performed in triplicate.

In vitro expression analysis

FreeStyle 293F cells were cultured in suspension and used for *in vitro* expression analysis. Expression of EV71 VP1 from the IVT mRNA was examined by western blot following transfection reagent-mediated delivery and by immunofluorescence analysis following mRNA-LNP treatment.

For western blot analysis, cells were seeded at 2×10^6 cells per well and incubated with 2.0 µg of IVT mRNA using an RNA transfection reagent (Takara Bio, Japan) for 72 hrs. Culture supernatants were collected, concentrated, and analyzed by western blot using a mouse monoclonal anti-EV71 VP1 antibody (clone 1A2-C1, GeneTex, USA), followed by an HRP-conjugated secondary antibody (DakoCytomation, Denmark). Immunoreactive bands were detected using a chemiluminescent substrate (Thermo Fisher Scientific, USA). Untreated cells served as negative controls.

For immunofluorescence analysis, cells were seeded at 1×10^6 cells per well and assigned to four groups: untreated negative control, empty LNPs, unformulated EV71 VP1 mRNA, and EV71 VP1 mRNA-LNPs. The EV71 VP1 mRNA-LNPs group was treated with 8.0 µg of mRNA in LNP formulation (72.4 µL), the empty LNPs group was treated with an equivalent volume of empty LNPs (72.4 µL), and the unformulated EV71 VP1 mRNA group was treated with 8.0 µg of mRNA without LNP formulation. After incubation for 72 hrs, cells were fixed with 40% methanol in acetone at -20 °C overnight and stained with the same primary antibody, followed by staining with a FITC-conjugated secondary antibody (DakoCytomation, Denmark). Nuclei were counterstained with DAPI, and corresponding bright-field and fluorescence images were acquired using a fluorescence microscope (Nikon Instruments, USA).

Animal immunization and sample collection

Female BALB/c mice (6-8 weeks old) were obtained from Nomura Siam International Co., Ltd. and housed under specific-pathogen-free conditions. All animal procedures were conducted in accordance with institutional guidelines for animal care and use and were approved by the Thammasat University Animal Care and Use Committee (approval No. 017/2025).

Mice were randomly assigned to three groups (N=6 per group): EV71 VP1 mRNA-LNPs group, empty LNPs control group, and PBS control group. Animals were immunized intramuscularly with 5 µg mRNA in a total volume of 50 µL on day 0 (prime) and day 28 (boost). Control groups received either empty LNPs as a formulation control or PBS. Blood samples were collected prior to immunization, at 4 weeks post-prime via submandibular vein puncture, and at 4 weeks post-boost via cardiac puncture under thiopental anesthesia. Mice were euthanized following terminal blood collection. Serum samples were separated and stored at -80 °C until analysis.

Measurement of EV71 VP1-specific IgG

VP1-specific IgG responses were measured by ELISA. Plates were coated with recombinant EV71 VP1 protein (0.1 µg/well; Antibody System, France) and blocked with 3% bovine serum albumin in PBS. Serum samples were diluted 1:1000 and incubated on coated plates. Bound antibodies were detected using an HRP-conjugated rabbit anti-mouse secondary antibody (1:2000), followed by color development with TMB substrate (SeraCare Life Sciences, USA). Absorbance was measured at 450 nm with a reference wavelength of 620 nm.

Endpoint IgG titers were determined for sera from mice immunized with EV71 VP1 mRNA-LNPs using the same ELISA protocol with twofold serial dilutions (1:250-1:8,192,000). Titers were defined as the reciprocal of the highest dilution exceeding the cutoff value, calculated from pre-immunization sera using a prediction-limit method at 99% confidence level.[12]

Statistical analysis

Physicochemical properties of the mRNA-LNP formulation, including particle size, PDI, zeta potential, and encapsulation efficiency, are presented as mean±standard deviation (SD). EV71 VP1-specific IgG levels (OD450) are presented as mean±SD, whereas endpoint antibody titers are presented as geometric mean titers (GMT) with 95% confidence intervals. Statistical analyses were performed using GraphPad Prism version 8 (GraphPad Software, USA). Differences among groups were analyzed using one-way analysis of variance (ANOVA) with Tukey's multiple comparisons test. Changes over time were analyzed using repeated-measures ANOVA with Tukey's multiple comparisons test. A $p<0.05$ was considered statistically significant.

Results

Design and preparation of EV71 VP1 mRNA construct

The EV71 VP1 mRNA construct was designed

based on a Thai subgenotype B5 isolate (SiICRC04/TH/2012), as schematically illustrated in Figure 1. Codon optimization increased the codon adaptation index (CAI) to >0.9. Secondary structure prediction using RNAfold yielded a minimum free energy of -421.13 kcal/mol, suggesting stable folding of the optimized mRNA transcript.

The optimized VP1 coding sequence was successfully cloned into the Linearized Template Vector (BspQ I), and correct insertion was confirmed by colony PCR and Sanger sequencing. The resulting plasmid was efficiently linearized, yielding a DNA template suitable for *in vitro* transcription.

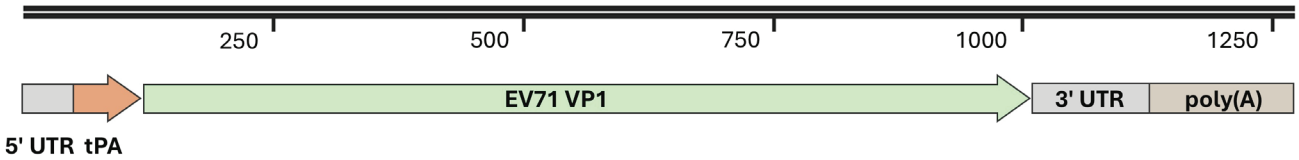


Figure 1. *Schematic representation of the EV71 VP1 mRNA construct, including the 5' untranslated region (UTR) with a Kozak sequence, the tissue plasminogen activator (tPA) signal peptide, the VP1 coding sequence, the 3' UTR, and the poly(A) tail.*

Characterization of in vitro-transcribed EV71 VP1 mRNA

Capped, N1-methylpseudouridine-modified ($m^1\Psi$) EV71 VP1 mRNA was successfully synthesized by IVT. The reaction yielded 862 µg of mRNA from 140 µL, corresponding to a concentration of 6.1 mg/mL.

PCR analysis of DNase-treated samples showed no detectable amplification (Figure 2A), indicating removal of residual plasmid DNA. RNA integrity analysis using the Agilent TapeStation system showed a single band at approximately 1,192 nt, consistent with the expected transcript length including the poly(A) tail (Figure 2B). Sequence verification by reverse transcription followed by Sanger sequencing confirmed concordance with the designed VP1 coding sequence.

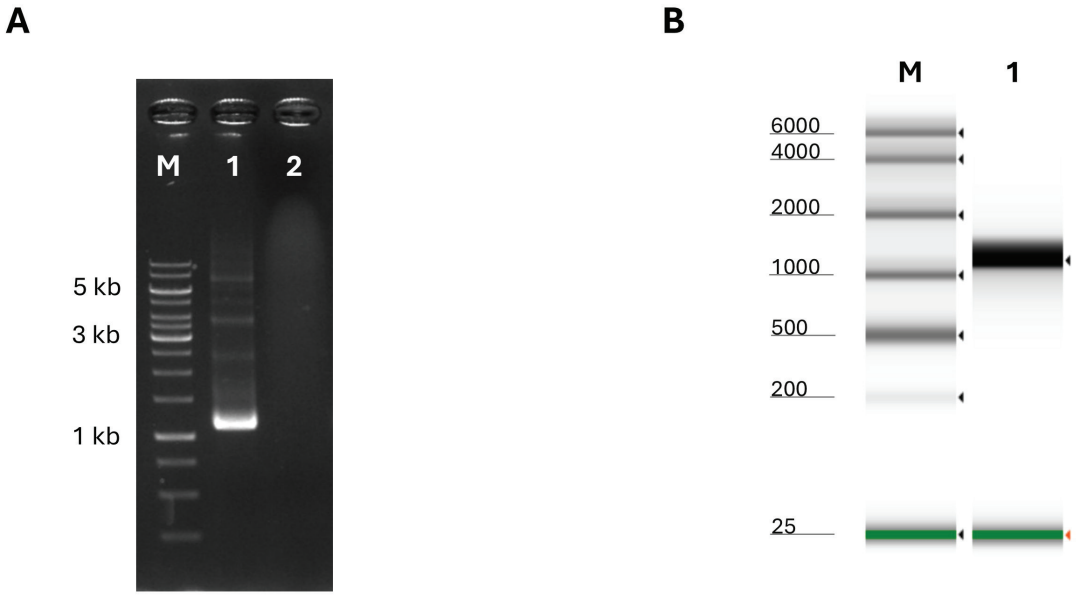


Figure 2. *Quality assessment of in vitro-transcribed EV71 VP1 mRNA. A: PCR analysis of DNase-treated mRNA samples confirming removal of residual plasmid DNA; Lane M: DNA ladder (1 kb), Lane 1: PCR from linearized pDNA (positive control), Lane 2: PCR from DNase-treated mRNA preparation; B: RNA integrity analysis using Agilent TapeStation, showing a single band corresponding to the expected transcript size, Lane M: DNA ladder (1 kb), Lane 1: EV71 VP1 mRNA (~1,192 nt).*

Preparation and characterization of EV71 VP1 mRNA-LNPs

The *in vitro*-transcribed EV71 VP1 mRNA was encapsulated in lipid nanoparticles using a microfluidic mixing approach, yielding a clear, homogeneous formulation.

The mRNA-LNPs exhibited a mean hydrodynamic diameter of 99.24±0.46 nm, with a low polydispersity index (0.041±0.004) and a slightly negative zeta potential (-20.54±0.98 mV). The encapsulation efficiency was 92.56±1.53% (Table 1).

Table 1. Physicochemical characteristics of EV71 VP1 mRNA-LNP formulation.

Formulation	**Size** (nm)	**PDI**	**Zeta potential** (mV)	**%EE**
EV71 VP1 mRNA-LNPs	99.24±0.46	0.041±0.004	-20.54±0.98	92.56±1.53

In vitro expression of EV71 VP1 following mRNA transfection and mRNA-LNP treatment

VP1 expression was first examined by Western blot after transfection reagent-mediated delivery of EV71 VP1 mRNA into FreeStyle 293F cells. A VP1-specific immunoreactive band was detected in the transfected sample, but not in the non-transfected control, indicating that the designed mRNA could direct VP1 protein expression. The band migrated at an apparent molecular weight of approximately 45-55 kDa (Figure 3).

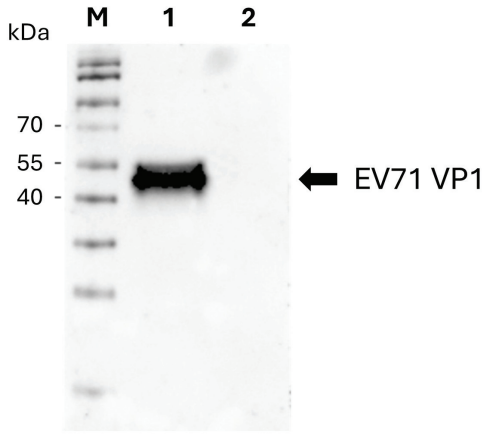


Figure 3. *Western blot analysis of EV71 VP1 expression in FreeStyle 293F cells. VP1 expression was examined in culture supernatants after transfection reagent-mediated delivery of EV71 VP1 mRNA. Lane M: protein ladder, Lane 1: transfected sample, Lane 2: non-transfected control. The arrow indicates the VP1-specific immunoreactive band at approximately 45-55 kDa.*

Immunofluorescence analysis was then performed to assess VP1 expression following mRNA-LNP treatment. VP1-specific FITC fluorescence was observed in EV71 VP1 mRNA-LNP-treated cells, whereas no specific fluorescence was detected in untreated cells, empty LNP-treated cells, or cells treated with unformulated EV71 VP1 mRNA. DAPI and bright-field images confirmed the presence of cells in the examined fields (Figure 4).

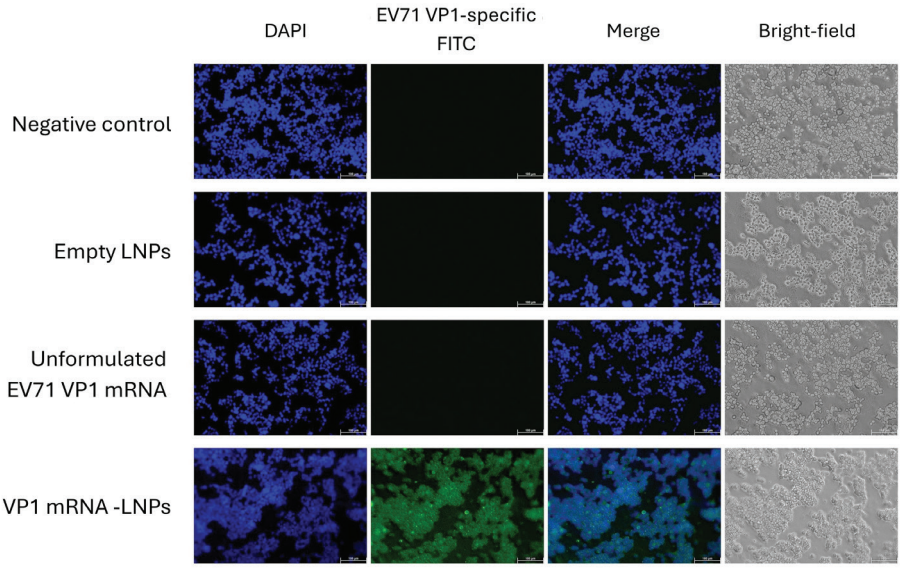


Figure 4. *Immunofluorescence analysis of EV71 VP1 expression in FreeStyle 293F cells following mRNA-LNP treatment. Four groups were examined: untreated negative control, empty LNPs, unformulated EV71 VP1 mRNA, and EV71 VP1 mRNA-LNPs. FITC signals indicate EV71 VP1-specific staining, DAPI indicates nuclei, and merged FITC/DAPI and bright-field images are shown. VP1-specific FITC fluorescence was observed only in the EV71 VP1 mRNA-LNPs group.*

EV71 VP1-specific humoral immune responses in mice

EV71 VP1-specific IgG responses were measured by ELISA using sera collected prior to immunization, at 4 weeks post-prime, and at 4 weeks post-boost. Antibody responses in the EV71 VP1 mRNA-LNPs group were compared with those in the empty LNP formulation control and PBS control groups. One mouse in the PBS group died during the study and was excluded from analysis.

At the pre-immune time point, absorbance values were low and comparable among groups (EV71 VP1 mRNA-LNPs: 0.020±0.008, empty LNPs 0.015±0.002, PBS: 0.015±0.001), indicating the absence of pre-existing VP1-specific antibodies. Following the prime dose, mice immunized with EV71 VP1 mRNA-LNPs developed detectable antibody responses (1.254±0.775), which were significantly higher than those in both control groups (empty LNPs: 0.015±0.002, PBS: 0.017±0.002, $p<0.01$). Antibody levels further increased after booster immunization (4.225±0.322) and remained significantly higher than controls ($p < 0.001$). In contrast, mice receiving empty LNPs showed background-level signals comparable to PBS throughout the study (Figure 5A).

Endpoint titer analysis showed a geometric mean titer (GMT) of 2,000 (95% CI: 901.50-4,437.00) after the prime dose. Following booster immunization, GMT increased to 256,000, with all samples showing titers at this level (Figure 5B). Endpoint titers at both the post-prime and post-boost time points were significantly higher than those at the pre-immune time point ($p<0.05$ and $p<0.001$, respectively).

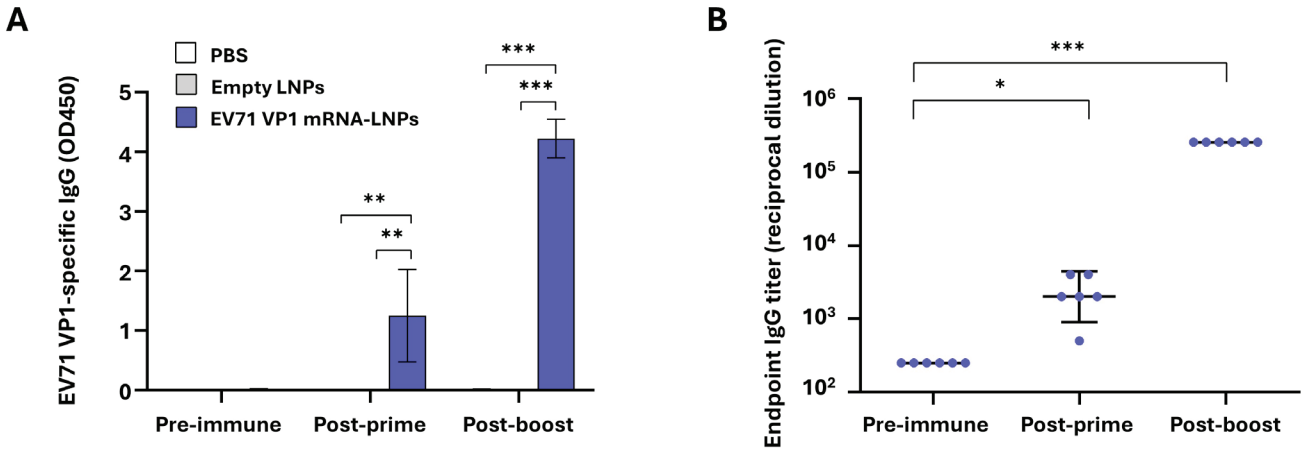


Figure 5. *Humoral immune responses induced by EV71 VP1 mRNA-LNP immunization in mice. A: EV71 VP1-specific IgG responses measured by ELISA at pre-immune, 4 weeks post-prime, and 4 weeks post-boost. Data are presented as mean ± SD, B: Endpoint IgG titers against EV71 VP1. Data are expressed as geometric mean titers (GMT) with 95% confidence intervals. *p<0.05, **p<0.01; ***p<0.001.*

Discussion

This study demonstrates the development of an EV71 VP1-encoding mRNA-LNP construct and its ability to induce antigen-specific humoral immune responses in a mouse model. The integrated workflow, from mRNA design to *in vitro* expression and *in vivo* humoral immune responses, supports its use for EV71 antigen delivery. While existing vaccines and experimental approaches rely on multi-step manufacturing and protein-based antigen production, the mRNA-LNP system demonstrated here enables streamlined design and antigen expression, which may facilitate adaptation to EV71 subgenotypes with genetic diversity.

The selection of a VP1 sequence derived from a circulating Thai subgenotype B5 strain supports the epidemiological relevance of this construct. Subgenotype B5 has been associated with multiple HFMD outbreaks in Thailand and remains in circulation.[3,4,13] The selected sequence retains the SP70 region, a conserved neutralizing epitope within the surface-exposed GH loop, a major antigenic site.[11,14] The preservation of this region supports its role in antibody recognition and suggests its potential relevance across EV71 subgenotypes. In addition, the presence of the SP2 region associated with T-cell responses suggests broader immunological relevance.[15]

The mRNA construct incorporated key structural elements and modified nucleosides consistent with current mRNA vaccine design strategies to enhance translational efficiency while reducing innate immune sensing at the RNA level.[9,16] The predicted secondary structure and minimum free energy suggest a thermodynamically stable transcript, and RNA structure has been reported to influence mRNA stability and translational efficiency.[17] Such structural optimization may support sustained antigen expression, thereby influencing downstream immune responses. The observed yield, absence of detectable residual DNA, and presence of a single dominant transcript band support the quality of the IVT mRNA preparation used for downstream LNP formulation. Given that IVT reactions can generate byproducts, including dsRNA, abortive RNAs, and RNA:DNA hybrids, stringent quality control is important before formulation.[18]

The formulated mRNA-LNPs exhibited physicochemical properties, including particle size, polydispersity index, zeta potential, and encapsulation efficiency, within ranges associated with stable and effective delivery systems. Particle sizes around 100 nm are associated with lymphatic transport and uptake by antigen-presenting cells, which may facilitate antigen delivery and presentation. The slightly negative surface charge may also influence cellular uptake, although its *in vivo* impact remains context dependent.[10,19,20]

In vitro expression analyses confirmed that the designed EV71 VP1 mRNA directed VP1 expression in FreeStyle 293F cells. The apparent molecular weight shift observed by western blot may reflect altered electrophoretic mobility or post-translational processing during ER/Golgi trafficking following secretory pathway targeting by the incorporated tPA signal sequence.[21,22] The exact cause of this shift warrants further investigation. Immunofluorescence analysis showed VP1-specific FITC fluorescence only in EV71 VP1 mRNA-LNP-treated cells, but not in untreated cells, empty LNP-treated cells, or cells treated with unformulated EV71 VP1 mRNA. These findings support the functional role of the LNP formulation in facilitating mRNA delivery and enabling detectable antigen expression under the conditions tested.

Immunization with EV71 VP1 mRNA-LNPs induced VP1-specific IgG responses, particularly following a prime-boost regimen. This pattern is consistent with observations from previous EV71 vaccine studies[23,24] and may reflect affinity maturation processes within germinal centers, leading to the generation of higher-affinity antibody responses.[25] Booster immunization has been shown to influence B-cell selection dynamics within germinal centers, contributing to improved antibody magnitude and quality.[26,27]

Empty LNPs were included as a formulation control to account for background signals or nonspecific effects associated with the LNP formulation in the absence of the EV71 VP1 mRNA payload. The background-level VP1-specific IgG responses observed in the empty LNPs group, comparable to PBS, support that the antibody responses in the EV71 VP1 mRNA-LNPs group were associated with VP1 expression from the mRNA payload. More broadly, the immunogenicity of mRNA-LNP vaccines is not limited to antibody responses. Innate immune sensing associated with mRNA vaccination has been implicated in vaccine adjuvanticity and may influence downstream adaptive immune responses.[28] In addition to humoral responses, mRNA vaccines enable antigen presentation via both MHC class I and class II pathways, thereby supporting the activation of $CD8^+$ and $CD4^+$ T cells.[9] In EV71 infection, T-cell-mediated immunity contributes to viral clearance and protective responses.[29] A recent study reported that an mRNA-based EV71 vaccine can induce both neutralizing antibodies and Th1-biased cellular responses, providing protection in preclinical models.[30] Further investigation is required to evaluate these immunological mechanisms in the context of the present construct.

Overall, this study provides a proof-of-concept demonstration of a VP1-based mRNA-LNP construct for EV71 and lays the groundwork for further investigation. From a translational perspective, this approach may support antigen adaptation and scalable production, thereby facilitating iterative updates to match regionally circulating EV71 subgenotypes. However, comprehensive evaluation of neutralizing activity, cross-genotype protection, and cellular immunity will be required to define its protective potential.

Limitations

This study has several limitations. Immunogenicity was evaluated in a small cohort of mice and should be considered preliminary. Only binding antibody responses were assessed; neutralizing activity, cellular immune responses, antibody durability, and cross-genotype neutralization were not evaluated. In addition, the animal study did not include an unformulated EV71 VP1 mRNA group; therefore, direct comparison of *in vivo* immunogenicity between unformulated mRNA and mRNA-LNPs was not performed. Future studies should address these limitations to better define the protective potential of the VP1-encoding mRNA-LNPs evaluated here.

Conclusion

This study demonstrates the development and preliminary evaluation of an EV71 VP1-encoding mRNA-LNP construct. VP1 expression was confirmed in FreeStyle 293F cells, and immunization induced VP1-specific IgG responses in mice, indicating antigen-specific humoral responses. These findings provide proof of concept for an EV71 mRNA-LNP approach and support further studies to evaluate neutralizing activity, cellular immunity, and protective efficacy.

Ethical approval

This study was approved by the Thammasat University Animal Care and Use Committee (ethical approval No. 017/2025).

Funding

This study was supported by the Department of Medical Sciences, Ministry of Public Health, Thailand.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Parnuphan Panyajai: conceptualization, methodology, formal analysis, project administration, writing-review and editing; **Rujiraporn Pitaksalee**: methodology, investigation, writing original draft, visualization; **Apichai Prachasuphap**: methodology, investigation, formal analysis;

Kodcharad Jongpitisub: investigation; **Pantida Treeyoung**: investigation; **Sirawit Wet-osot**: investigation; **Nalinee Saengtong**: investigation; **Nadthanan Pinyosukhee**: methodology, investigation; **Rattanawadee Wichajarn**: investigation; **Kritsamon Sophondilok**: investigation; **Anicha Luengchaichaweng**: investigation; **Panadda Dhepakson**: supervision.

Acknowledgements

The authors gratefully acknowledge the contributions of all colleagues involved in this study.

References

[1] Bifani AM, Tee HK, Antanasijevic A, Clément S, Tapparel C. Enterovirus A71 priorities, challenges, and future opportunities in humoral immunity and vaccine development. npj Vaccines. 2025; 10: 194. doi: 10.1038/s41541-025-01184-z.

[2] Yi EJ, Shin YJ, Kim JH, Kim TG, Chang SY. Enterovirus 71 infection and vaccines. Clin Exp Vaccine Res. 2017; 6(1): 4-14. doi: 10.7774/cevr.2017.6.1.4.

[3] Noisumdaeng P, Sangsiriwut K, Prasertsopon J, Klinmalai C, Payungporn S, Mungaomklang A, et al. Complete genome analysis demonstrates multiple introductions of enterovirus 71 and coxsackievirus A16 recombinant strains into Thailand during the past decade. Emerg Microbes Infect. 2018; 7(1): 1-12. doi: 10.1038/s41426-018-0215-x.

[4] Noisumdaeng P, Puthavathana P. Molecular evolutionary dynamics of enterovirus A71, coxsackievirus A16 and coxsackievirus A6 causing hand, foot and mouth disease in Thailand, 2000-2022. Sci Rep. 2023;13: 17359. doi: 10.1038/s41598-023-44644-z.

[5] Yan X, Liu Y, Chen F, Chang Z, Zhang Z, Pons-Salort M, et al. The efficacy and effectiveness of enterovirus A71 vaccines against hand, foot, and mouth disease: A systematic review and meta-analysis. PLOS One. 2025; 20(5): e0323782. doi: 10.1371/journal.pone.0323782.

[6] Suppasittikulchai A, Takultacha S, Thubthawee P. Hand, foot and mouth disease and enterovirus 71 vaccines. Thai J Pharm Prac. 2024; 16(4): 970-83 (in Thai).

[7] Puenpa J, Chansaenroj J, Auphimai C, Srimuan D, Thatsanathorn T, Poovorawan Y, et al. Neutralizing antibody against enterovirus-A71 in Thai children: A longitudinal study from birth to age 4 years. Vaccine. 2020; 38(48): 7638-44. doi: 10.1016/j.vaccine.2020.10.002.

[8] Leong KY, Tham SK, Poh CL. Revolutionizing immunization: A comprehensive review of mRNA vaccine technology and applications. Virol J. 2025; 22: 71. doi: 10.1186/s12985-025-02645-6.

[9] Pardi N, Hogan MJ, Porter FW, Weissman D. mRNA vaccines - a new era in vaccinology. Nat Rev Drug Discov. 2018; 17(4): 261-79. doi: 10.1038/nrd.2017.243.

[10] John R, Monpara J, Swaminathan S, Kalhapure R. Chemistry and art of developing lipid nanoparticles for biologics delivery: Focus on development and scale-up. Pharmaceutics. 2024; 16(1): 131. doi: 10.3390/pharmaceutics16010131.

[11] Foo DG, Alonso S, Phoon MC, Ramachandran NP, Chow VT, Poh CL. Identification of neutralizing linear epitopes from the VP1 capsid protein of enterovirus 71 using synthetic peptides. Virus Res. 2007; 125(1): 61-8. doi: 10.1016/j.virusres.2006.12.005.

[12] Frey A, Di Canzio J, Zurakowski D. A statistically defined endpoint titer determination method for immunoassays. J Immunol Methods. 1998; 221 (1-2): 35-41. doi: 10.1016/s0022-1759(98)00170-7.

[13] Puenpa J, Auphimai C, Korkong S, Vongpunsawad S, Poovorawan Y. Enterovirus A71 infection, Thailand, 2017. Emerg Infect Dis. 2018; 24(7): 1386-7. doi: 10.3201/eid2407.171923.

[14] Ku Z, Ye X, Shi J, Wang X, Liu Q, Huang Z. Single neutralizing monoclonal antibodies targeting the VP1 GH loop of enterovirus 71 inhibit both virus attachment and internalization during viral entry. J Virol. 2015; 89(23): 12084-95. doi: 10.1128/jvi.02189-15.

[15] Foo DG, Macary PA, Alonso S, Poh CL. Identification of human CD4 T-cell epitopes on the VP1 capsid protein of enterovirus 71. Viral Immunol. 2008; 21(2): 215-24. doi: 10.1089/vim.2007.0089.

[16] Andries O, Mc Cafferty S, De Smedt SC, Weiss R, Sanders NN, Kitada T. N1-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice. J Control Release. 2015; 217: 337-44. doi: 10.1016/j.jconrel.2015.08.051.

[17] Mauger DM, Cabral BJ, Presnyak V, Su SV, Reid DW, Goodman B, et al. mRNA structure regulates protein expression through changes in functional half-life. PNAS. 2019; 116(48): 24075-83. doi: 10.1073/pnas.1908052116.

[18] Lenk R, Kleindienst W, Szabó GT, Baiersdörfer M, Boros G, Keller JM, et al. Understanding the impact of *in vitro* transcription byproducts and contaminants. Front Mol Biosci. 2024; 11: 1426129. doi: 10.3389/fmolb.2024.1426129.

[19] Kong W, Wei Y, Dong Z, Liu W, Zhao J, Huang Y, et al. Role of size, surface charge, and pegylated lipids of lipid nanoparticles (LNPs) on intramuscular delivery of mRNA. J Nanobiotechnology. 2024; 22: 553. doi: 10.1186/s12951-024-02812-x.

[20] Shi R, Liu X, Wang Y, Pan M, Wang S, Shi L, et al. Long-term stability and immunogenicity of lipid nanoparticle COVID-19 mRNA vaccine is affected by particle size. Hum Vaccin Immunother. 2024; 20(1): 2342592. doi: 10.1080/21645515.2024.2342592.

[21] Zhang Y, Zhai S, Huang H, Qin S, Sun M, Chen Y, et al. Efficient signal sequence of mRNA vaccines enhances the antigen expression to expand the immune protection against viral infection. J Nanobiotechnology. 2024; 22: 295. doi: 10.1186/s12951-024-02488-3.

[22] Scheller C, Krebs F, Wiesner R, Wätzig H, Oltmann-Norden I. A comparative study of CE-SDS, SDS-PAGE, and simple western-precision, repeatability, and apparent molecular mass shifts by glycosylation. Electrophoresis. 2021; 42(14-15): 1521-31. doi: 10.1002/elps.202100068.

[23] Li R, Liu L, Mo Z, Wang X, Xia J, Liang Z, et al. An inactivated enterovirus 71 vaccine in healthy children. N Engl J Med. 2014; 370(9): 829-37. doi: 10.1056/nejmoa1303224.

[24] Zhu FC, Meng FY, Li JX, Li XL, Mao QY, Tao H, et al. Efficacy, safety, and immunology of an inactivated alum-adjuvant enterovirus 71 vaccine in children in China: A multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2013; 381(9882): 2024-32. doi: 10.1016/s0140-6736(13)61049-1.

[25] Elsner RA, Shlomchik MJ. Germinal center and extrafollicular B cell responses in vaccination, immunity, and autoimmunity. Immunity. 2020; 53(6): 1136-50. doi: 10.1016/j.immuni.2020.11.006.

[26] Wang X, Cottrell CA, Hu X, Ray R, Bottermann M, Villavicencio PM, et al. mRNA-LNP prime boost evolves precursors toward VRC01-like broadly neutralizing antibodies in preclinical humanized mouse models. Sci Immunol. 2024; 9(95): eadn0622. doi: 10.1126/sciimmunol.adn0622.

[27] Garg AK, Mittal S, Padmanabhan P, Desikan R, Dixit NM. Increased B cell selection stringency in germinal centers can explain improved COVID-19 vaccine efficacies with low dose prime or delayed boost. Front Immunol. 2021; 12: 776933. doi: 10.3389/fimmu.2021.776933.

[28] Kobiyama K, Ishii KJ. Making innate sense of mRNA vaccine adjuvanticity. Nat Immunol. 2022; 23(4): 474-6. doi: 10.1038/s41590-022-01168-4.

[29] Yee PT, Poh CL. T cell immunity to enterovirus 71 infection in humans and implications for vaccine development. Int J Med Sci. 2018; 15(11): 1143-52. doi: 10.7150/ijms.26450.

[30] Chi F, Zhang X, Zhang D, Zhu A, Zhuang Z, Zhang Z, et al. A nucleoside-modified mRNA vaccine prevents enterovirus A71 infection in mouse model. Front Immunol. 2025; 16: 1535758. doi: 10.3389/fimmu.2025.1535758.

Journal of Associated Medical Sciences 2026; 59(3): 205-214

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



Percentile-based infrared thermography for community screening of foot thermal abnormalities

Sirasith Uat-hao[1], Kemika Wongsatan[1], Runchana Kantanat[2], Kanyawee Sathapong[2], Kittichai Wantanajittikul[1], Chatchanok Udomtanakunchai[1]*

[1]Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Nation University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 31 March 2026
Accepted as revised 26 June 2026
Available online 2 July 2026

Keywords:
Infrared thermography, foot skin temperature, ischemia, inflammation, diabetic foot screening

ABSTRACT

Background: Screening support of foot health abnormalities is essential for preventing complications associated with vascular and metabolic disorders. Infrared thermography provides a noninvasive method for assessing spatial temperature variations related to microvascular perfusion and inflammatory processes; however, standardized frameworks for interpretation in community settings remain limited.

Objectives: To develop a percentile-based thermographic framework for evaluating foot thermal patterns and to investigate the relationship between thermographic indices and underlying vascular-ischemic and metabolic-inflammatory processes in a community population.

Materials and methods: Plantar foot temperature images were obtained using infrared thermography. Normative temperature references were established, and thermographic burden indices representing ischemic and inflammatory patterns were derived based on percentile thresholds. Principal component analysis (PCA) and logistic regression were performed to identify thermographic patterns and evaluate predictive performance.

Results: Most participants demonstrated low thermographic burden, while a subset exhibited elevated indices. PCA identified two principal components corresponding to metabolic-inflammatory and vascular-ischemic dimensions. The combined thermographic model demonstrated moderate discriminative performance (AUC=0.714, 95% CI: 0.631-0.790), with a sensitivity of 74.2% and specificity of 65.0% at the optimal cutoff.

Conclusion: A percentile-based thermographic framework enables objective characterization of foot thermal abnormalities and provides insight into vascular-ischemic and metabolic-inflammatory thermal patterns. The integration of thermographic indices with multivariate and predictive analyses supports their potential as noninvasive tools for risk stratification and community-based screening of foot health abnormalities.

Introduction

Peripheral arterial disease and diabetic foot complications remain major contributors to morbidity in aging populations.[1-3] Early identification of microvascular compromise prior to the development of overt ulceration or tissue loss remains a clinical challenge. Conventional vascular assessments, such as ankle-brachial index and Doppler ultrasound, primarily detect

* **Corresponding contributor.**
Author's Address: *Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: chatchanok.u@cmu.ac.th
doi: 10.66285/JAMS.2026.094
E-ISSN: 2539-6056

macrovascular disease and may not fully capture early spatial perfusion instability or metabolic-inflammatory alterations in the foot microenvironment. Infrared thermography provides a non-invasive method for visualizing cutaneous heat distribution, reflecting underlying perfusion dynamics and inflammatory activity.[4-7] However, most thermographic approaches rely on absolute temperature thresholds or focal hot-spot detection, without incorporating spatial gradients or thermal symmetry as physiological reference patterns.[8-10]

Under physiological conditions, the foot exhibits organized thermal gradients when measured in stable ambient environments. Proximal-to-distal temperature decline on the dorsum and central-to-distal variation on the plantar surface reflect intact perfusion regulation.[11] Disruption of these gradients may indicate early ischemic attenuation, whereas localized hyperthermia may indicate inflammatory activation associated with metabolic stress. Previous studies have shown that a temperature difference greater than 2.2 °C between corresponding regions of the left and right feet may serve as an early indicator of foot ulceration risk.[12]

However, most existing approaches rely on fixed thresholds and lack a standardized scoring framework for integrating spatial thermal patterns.

Therefore, this study aimed to develop a percentile-based, zone-oriented thermographic framework to distinguish physiological thermal patterns from vascular-ischemic and metabolic-inflammatory deviations in a community population. A normative reference subgroup was used to establish percentile thresholds for thermographic burden indices, while principal component analysis and predictive modeling were applied to investigate underlying thermographic patterns and their association with clinical variables.

Materials and methods

Study design

This study was designed as a methodological development and field validation study to construct a percentile-based, zone-oriented ischemic-inflammatory scoring system for foot thermography under community-based screening conditions. The framework was developed to reflect real-world deployment scenarios and support applications in remote or resource-limited healthcare settings. Thermographic data were collected during community-based health outreach services. A dult participants (≥18 years) who underwent bilateral foot thermographic assessment were included. Records were excluded if images were incomplete, exhibited motion artifacts, or contained technical acquisition errors. All participants remained in the examination environment for at least 20 minutes before image acquisition to allow thermal acclimatization. Participant recruitment was based on convenience sampling during community-based health outreach activities. No formal sample size calculation was performed because the study's primary objective was the development of a methodological framework and its exploratory validation.

Thermographic acquisition under field conditions

Infrared thermographic imaging was conducted in non-laboratory community environments. Ambient temperature was not held at a fixed value of 25.0 °C; instead, it was recorded at the start of each acquisition session. Participants were positioned at a standardized imaging location, and thermal images of both the dorsal and plantar surfaces of the feet were acquired using a calibrated infrared thermal camera (Testo 890, Testo SE & Co. KGaA, Germany). The camera was positioned at a fixed distance of 1.0 meters from the feet, perpendicular to the imaging plane, to ensure consistent image acquisition across participants.[13] Images were digitally stored for quantitative analysis. This acquisition strategy was intentionally designed to simulate real-world field deployment rather than controlled laboratory conditions. To reduce inter-session environmental variability, all recorded foot temperatures were standardized to a reference temperature of 25.0 °C, prior to analysis. Temperature normalization was performed using a uniform linear correction model adapted from previously reported thermographic standardization approaches.[13,14]

$$T_{corrected} = T_{measured} \pm (25.0 - T_{ambient})$$

This preprocessing step was applied consistently across the dataset before percentile modeling and score computation to maintain methodological consistency while preserving zone-based relative thermal patterns.

Zone-oriented parameter extraction

Each thermographic image was processed to determine the region of interest (ROI) and then divided into predefined anatomical zones. The following parameters are bilaterally isolated as overall mean temperature, midfoot temperature, toe temperature, load-bearing area temperature, proximal-to-distal temperature gradient, and bilateral temperature difference (Δ right–left). All parameters were computed using a structured computational pipeline implemented in an in-house-developed application (ThermoScan_CK v1.2), which was built as a graphical user interface in MATLAB R2025b (The MathWorks, Natick, MA, USA), to ensure reproducibility.

Development of the percentile-based reference model

A normative reference distribution was constructed from the standardized dataset. Percentile thresholds were calculated for each zone-oriented parameter. Lower percentile boundaries were used to define cold deviation (ischemic tendency), whereas upper percentile boundaries defined heat deviation

(inflammatory tendency). This percentile-based approach was selected to identify abnormal thermal patterns, as it is widely used in clinical practice to detect deviations from normative distributions and is robust to non-normal data. Percentile modeling was performed after environmental normalization to ensure comparability across field-acquired sessions.[15]

Thermographic index and burden score computation

Thermographic indices were derived from region-based temperature measurements. The 10th (P10) and 90th (P90) percentiles were selected to identify relatively extreme thermal deviations while maintaining adequate sample representation and robustness against outliers. For each anatomical region, values below the P10 of the normative reference were classified as the ischemic index, and values above the P90 as the inflammatory index.[16-18]

Thermal asymmetry was defined as the absolute temperature difference between corresponding regions of the left and right feet, with values exceeding the P90 considered abnormal.

Ischemic index, inflammatory index, and thermal asymmetry were calculated separately by aggregating abnormal regions across all anatomical locations. Each index was independently expressed on a scale of 0-100, representing the overall burden of ischemic, inflammatory, and asymmetry-related thermal deviation, respectively.

Severity tier classification

Burden percentages were categorized into four risk levels based on predefined thresholds: low (0-<20), mild (20-<40), moderate (40-<60), and high (≥60).

Statistical analysis

Descriptive statistics were used to summarize participant characteristics and thermographic parameters. Distributional properties were assessed prior to percentile computation. Associations between ischemic and inflammatory burden scores were examined using correlation analysis. All statistical analyses were performed using MATLAB R2025b (MathWorks, Natick, MA, USA), and statistical significance was set at $p<0.05$.

Principal component analysis (PCA) was performed using data from all participants to explore multivariate relationships among demographic, metabolic, and thermographic variables. Variables included age, body mass index (BMI), fasting blood glucose (FBS), ischemic burden, inflammatory burden, and thermal asymmetry. PCA was used as an exploratory technique to identify underlying thermographic patterns and visualize variable clustering.

Logistic regression models were constructed to evaluate the discriminative performance of thermographic indices. Model performance was assessed using receiver operating characteristic (ROC) analysis, and the area under the ROC curve (AUC) was calculated. The optimal cutoff value was determined using the Youden index.

For logistic regression and ROC analyses, participants were classified into two groups according to predefined normative criteria. Individuals in the normative reference group were defined as the negative class, whereas those who did not meet the normative criteria were classified as the positive class (non-normative participants). Thermographic burden indices were evaluated for their ability to discriminate between normative and non-normative participants.

Results

The study was conducted under ambient temperatures ranging from 24.0 to 30.0 °C. A total of 163 participants were included with a mean age of 63.20±8.17 years, comprising 38 males and 125 females. The mean body mass index (BMI) was 23.95±4.29 kg/m^2, and the mean fasting blood sugar (FBS) level was 101.77±23.25 mg/dL. (Table 1).

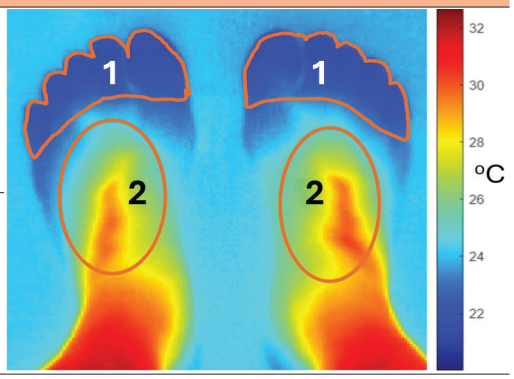
Table 1. Participant characteristics of all participants and the normative reference group.

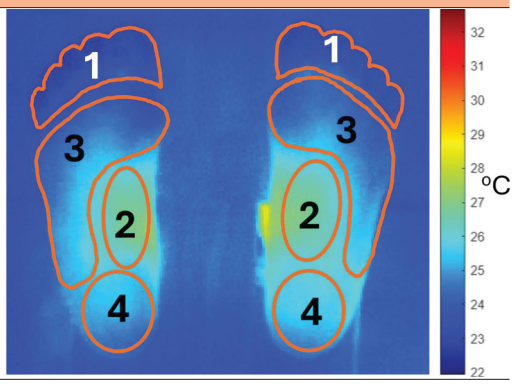
Variable	Mean±SD	
	All participants (N=163)	Normative group (N=66)
Age (years)	63.20±8.17	63.79±6.82
BMI (kg/m^2)	23.95±4.29	22.50±3.34
Fasting blood glucose (mg/dL)	101.77±23.25	92.15±4.74
Total cholesterol (mg/dL)	210.59±49.04	202.98±45.69

Foot temperature was analyzed from both dorsal and plantar thermal images. For the dorsal view, regions of interest (ROIs) included the toe (ROI1), midfoot (ROI2), overall dorsal temperature, the midfoot-to-toe temperature gradient, and the overall dorsal temperature difference (ΔT, right-left).

For the plantar view, ROIs included the toe (ROI1), midfoot (ROI2), forefoot (ROI3), heel (ROI4), overall plantar temperature, the forefoot-to-heel gradient, the load-bearing-to-toe gradient, and the overall plantar temperature difference (ΔT, right-left). Mean temperatures for the right foot, left foot, and bilateral averages are presented in Table 2.

Table 2. Summarizes the regional foot temperature measurements obtained from dorsal and plantar thermographic images. Mean temperatures and temperature gradients were calculated for both feet, including bilateral averages and right-left differences.

Foot temperature (°C) of all participants				
Dorsal region	**Mean±SD**			**Dorsum**
	Both	Right	Left	
Overall dorsal	25.6±1.6	26.8±1.5	24.4±2.3	
Toe (ROI1)	24.3±1.9	26.0±1.7	22.6±3.0	
Midfoot (ROI2)	28.9±1.2	27.0±1.6	28.9±1.3	
ΔT Midfoot-Toe	4.5±1.5	2.8±1.0	6.2±2.8	
ΔT Overall (Right-Left)		2.6±1.9		

Plantar Region	**Mean±SD**			**Plantar**
	Both	Right	Left	
Overall plantar	26.6±1.6	26.7±1.6	26.6±1.6	
Toe (ROI1)	25.4±1.7	25.4±1.7	25.4±1.7	
Midfoot (ROI2)	28.1±1.7	28.1±1.7	28.0±1.8	
Forefoot (ROI3)	26.7±1.6	26.7±1.6	26.7±1.7	
Heel (ROI4)	26.2±1.6	26.2±1.6	26.1±1.7	
ΔT Forefoot-heel	0.7±0.5	0.7±0.5	0.7±0.5	
ΔT Load-bearing-toe	1.5±0.7	1.2±0.6	1.2±0.7	
ΔT Overall (Right-Left)		0.5±0.4		

Normative percentile thresholds

A normative reference group of 66 individuals was selected based on the following criteria: BMI 18.5-25.0 kg/m^2, no history of diabetes, FBS ≤100 mg/dL, and absence of foot ulcers or visible abnormalities (Table 1). Thermal data from the dorsal and plantar regions were analyzed to calculate the 5th, 10th, 90th, and 95th percentiles. Percentiles were computed for the right foot, left foot, and bilateral mean temperatures; however, only the bilateral mean values are presented in Table 3 for clarity.

These percentile values were used to establish temperature-based reference thresholds for identifying potential ischemic and inflammatory patterns. The 10th and 90th percentiles were selected as primary screening thresholds because they represent physiologically meaningful deviations while remaining robust to extreme outliers.

Table 3. Presents normative percentile reference values for dorsal and plantar foot temperatures.

Foot temperature (°C) of the normative reference group							
Dorsal region	**Mean±SD**			**Percentile (both)**			
	Both	Right	Left	P5	P10	P90	P95
Overall dorsal	25.5±1.4	26.6±1.3	24.3±2.1	23.30	23.79	27.22	28.05
Toe (ROI1)	28.7±1.2	26.7±1.4	28.8±1.3	21.76	22.14	26.00	26.91
Midfoot (ROI2)	24.2±1.7	25.8±1.4	22.5±2.8	27.01	27.26	30.34	30.69
ΔT Midfoot-toe	4.6±1.5	2.9±0.9	6.2±2.8	2.47	2.56	6.56	7.04
ΔT Overall (RT-LT)	2.6±1.8			0.15	0.28	4.69	4.93

Table 3. Presents normative percentile reference values for dorsal and plantar foot temperatures (continued).

Plantar region	Mean±SD			Percentile (both)			
	Both	Right	Left	P5	P10	P90	P95
Overall plantar	26.5±1.4	26.5±1.5	26.4±1.5	24.39	24.59	28.45	29.02
Toe (ROI1)	25.2±1.4	25.1±1.5	25.2±1.5	23.30	23.92	26.86	27.36
Midfoot (ROI2)	27.9±1.6	28±1.6	27.8±1.8	25.60	25.78	30.61	30.86
Forefoot (ROI3)	26.5±1.5	26.5±1.5	26.5±1.6	24.44	24.67	28.79	29.30
Heel (ROI4)	26.1±1.4	26.2±1.4	26.1±1.5	23.99	24.28	27.93	28.47
ΔT forefoot-heel	0.6±0.5	0.6±0.4	0.6±0.6	0.04	0.06	1.30	1.50
ΔT Load-bearing-toe	1.6±0.6	1.3±0.5	1.2±0.6	0.53	0.69	2.39	2.81
ΔT Overall (RT-LT)	0.6±0.5			0.06	0.11	1.34	1.41

Note: *P10 and P90 thresholds were used to define hypothermic (ischemic) and hyperthermic (inflammatory) patterns.*

Group comparison and correlation analysis

Comparisons between the normative and non-normative groups were performed using either the independent t-test or the Mann-Whitney U test, depending on data distribution. No significant difference in age was observed between the groups (p=0.447). In contrast, BMI (p=0.0001), fasting blood sugar (p<0.001), and inflammatory thermal burden (p=0.010) were significantly higher in the non-normative group. Ischemic burden (p=0.140) and thermal asymmetry (p=0.164) did not differ significantly between groups.

Correlation analysis showed no significant association between thermal asymmetry and ischemic burden (Spearman r=0.010, p=0.902). A weak negative correlation was observed between thermal asymmetry and inflammatory burden (Spearman r=-0.135, p=0.086), although this did not reach statistical significance.

The distribution of thermographic burden indices was highly skewed, with a substantial proportion of participants exhibiting zero values. In the overall cohort, zero ischemic burden was observed in 43% of participants, while inflammatory burden and thermal asymmetry were absent in 49% and 47% of cases, respectively. Median values were low across all indices (ischemic: 2.86%, inflammatory: 1.43%, asymmetry: 4.54% on a 0-100 scale), indicating that most individuals had minimal thermal deviation.

However, the upper tails of the distribution were markedly elevated, with the 90th percentiles reaching 22.86% for ischemic burden, 28.57% for inflammatory burden, and 40.91% for thermal asymmetry. A similar pattern was observed in the normative subgroup, where median values were lower, and the proportion of zero values was higher, while upper-percentile values remained elevated.

These findings indicate that thermographic burden indices follow a zero-inflated and right-skewed distribution, supporting the use of percentile-based thresholds for classification rather than mean-based criteria. (Table 4 and Figure 1).

Table 4. *Distribution of thermographic burden indices, median (%) and P90 (%).*

Variable	**All participants** (N=163)			**Normative group** (N=66)		
	Median	Zero values N (%)	P90	Median	Zero values N (%)	P90
Ischemic burden	2.86	70 (43%)	22.86	1.43	32 (48%)	17.85
Inflammatory burden	1.43	79 (49%)	28.57	0	41 (62%)	21.43
Thermal asymmetry	4.54	77 (47%)	40.91	9.09	25 (38%)	40.91

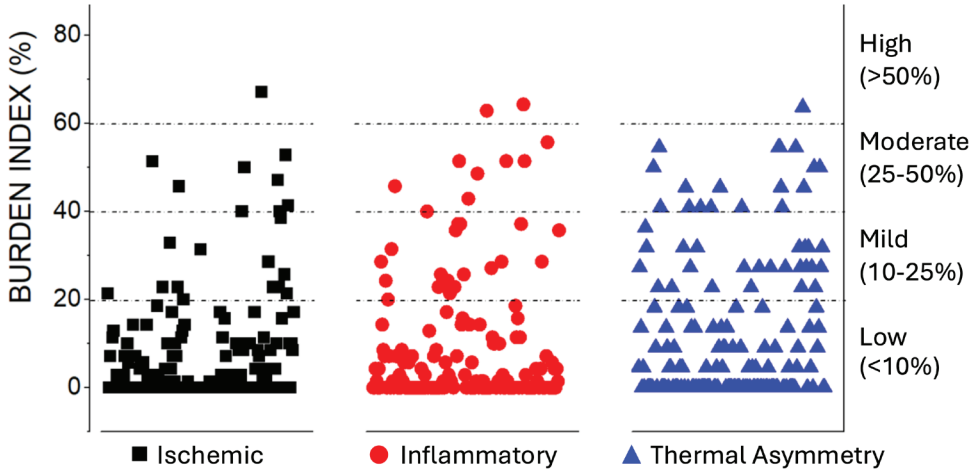


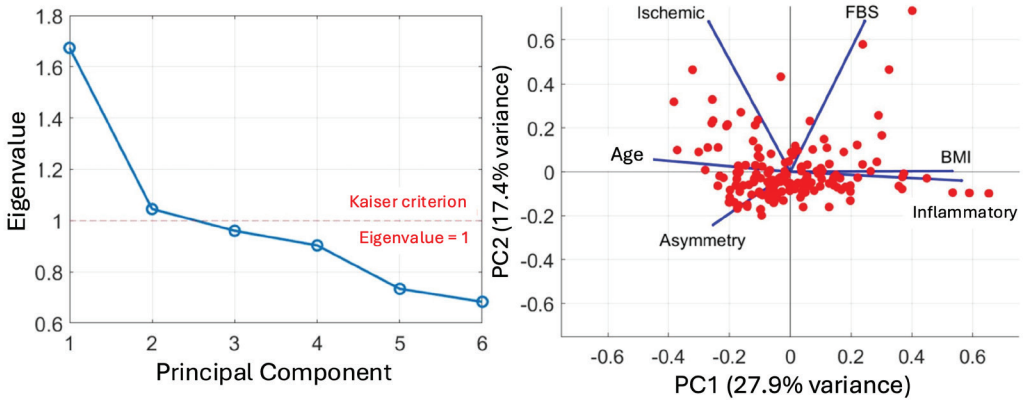
Figure 1. *Distribution of thermographic burden indices among all participants, categorized into low, mild, moderate, and high levels.*

Principal component analysis

Principal component analysis (PCA) with varimax rotation identified two principal components with eigenvalues greater than 1, together explaining 45.32% of the total variance. The first component (PC1) accounted for 27.90% of the variance and was characterized by strong positive loadings for body mass index (BMI) (0.53) and inflammatory thermal burden (0.56), along with a moderate negative loading for age (-0.45). These variables were aligned along a similar axis in the biplot, indicating a shared contribution to this component. PC1 was therefore interpreted as a metabolic-inflammatory dimension.

The second component (PC2) explained 17.42% of the total variance and was primarily defined by strong positive loadings of ischemic burden (0.683) and fasting blood glucose (0.685), with a smaller contribution from thermal asymmetry. In the biplot representation, these variables clustered along a distinct axis, suggesting a dimension related to vascular and perfusion-related variation with metabolic involvement. PC2 was interpreted as a vascular-ischemic dimension with metabolic contribution.

Following varimax rotation, two components were retained. The first component was associated with BMI and inflammatory thermal burden, representing a metabolic-inflammatory pattern, whereas the second component was characterized by ischemic burden and fasting blood glucose, reflecting a vascular-ischemic pattern with metabolic influence. Thermal asymmetry showed a weaker contribution to both components (Figure 2).



PC	Eigenvalues	Variance (%)	Variables	PC1	PC2
PC1	1.67	27.90	Age	-0.452	0.056
PC2	1.05	17.42	Thermal asymmetry	-0.254	-0.242
PC3	0.96	16.01	Inflammatory burden	0.563	-0.041
PC4	0.90	15.05	Ischemic burden	-0.268	0.683
PC5	0.73	12.23	BMI	0.532	0.003
PC6	0.68	11.39	FBS	0.245	0.685

Figure 2. *Principal component analysis (PCA) of thermographic and clinical variables. Two principal components were identified. PC1: metabolic-inflammatory pattern, PC2: vascular-ischemic pattern.*

Logistic regression and combined thermographic index

Four logistic regression models were evaluated to assess the discriminative performance of thermographic parameters. Model 1 included thermal asymmetry as a single predictor. Model 2 incorporated asymmetry and ischemic burden. Model 3 further included inflammatory burden, while Model 4 included all thermographic parameters along with body mass index.

Model performance was assessed using receiver operating characteristic (ROC) analysis to evaluate the ability of thermographic burden indices to discriminate between normative and non-normative participants. Performance improved progressively with the inclusion of additional variables. The model incorporating thermal asymmetry alone showed limited discriminative ability (AUC=0.581, 95% CI: 0.497-0.664). The addition of ischemic and inflammatory burden resulted in modest improvements (AUC=0.602, 95% CI: 0.511-0.686; AUC=0.614, 95% CI: 0.522-0.701, respectively). The full model, including thermal asymmetry, ischemic burden, inflammatory burden, and body mass index, achieved the highest discrimination performance (AUC=0.714, 95% CI: 0.631-0.790). The optimal cutoff value was 0.408, yielding a sensitivity of 74.2% and a specificity of 65.0% (Figure 3).

The optimal cutoff value for the combined thermographic index (TMI) was determined using the Youden index. An optimal cutoff of 0.408 yielded a sensitivity of 74.24% and a specificity of 64.95%. Based on this threshold, participants were primarily classified into low- and moderate-risk groups, with no cases observed in the high-risk category.

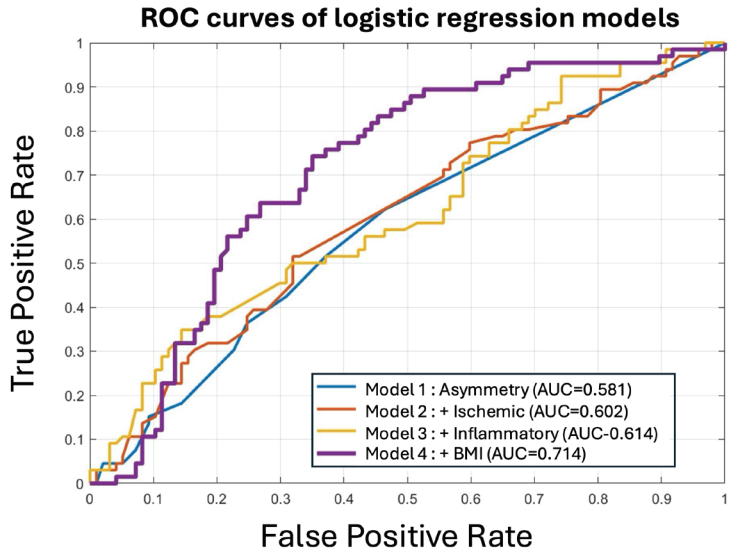


Figure 3. *Receiver operating characteristic (ROC) curves of logistic regression models. The full model (Model 4) showed the highest discriminative performance (AUC=0.714, 95% CI: 0.631-0.790).*

Discussion

Previous investigations have shown that localized increases in plantar temperature may reflect inflammatory activity and tissue stress, particularly in individuals at risk of diabetic foot complications.[2] Likewise, temperature asymmetry between feet has been proposed as an indicator of impaired perfusion or early ischemic changes.[10] Although infrared thermography has demonstrated potential in the prevention and management of diabetic foot disease, its clinical adoption remains limited.[19] However, most previous studies focused on temperature asymmetry or localized hot-spot detection, whereas standardized frameworks for interpreting complex thermographic patterns in community populations remain limited.

The present study proposes a percentile-based thermographic framework for assessing foot thermal patterns in a community population.[15] By establishing normative temperature references and deriving thermographic burden indices, the findings demonstrate that infrared thermography can quantify spatial thermal variations potentially related to vascular-ischemic and metabolic-inflammatory processes.[14,20] The integration of multivariate analysis and predictive modeling further supports its potential utility for risk stratification and community screening.

Principal component analysis revealed a multi-dimensional structure underlying thermographic and clinical variables. Two principal dimensions were identified, corresponding to metabolic-inflammatory and vascular-ischemic patterns. The metabolic-inflammatory component was primarily associated with body mass index and inflammatory thermal burden, suggesting that systemic metabolic factors may influence peripheral thermal distribution, consistent with previous observations that metabolic and inflammatory processes can alter cutaneous temperature patterns.[5,20] In contrast, the vascular-ischemic component was characterized by ischemic burden and fasting blood glucose levels, indicating that perfusion-related thermal variation may be linked to metabolic dysregulation and vascular impairment,

which may affect microcirculatory blood flow and regional heat distribution.[1,3,20] Thermal asymmetry demonstrated a weaker contribution, suggesting a more limited role in explaining overall variance.

The distribution of thermographic burden indices was highly skewed, with a large proportion of individuals exhibiting minimal or zero values and a smaller subset demonstrating markedly elevated indices. This zero-inflated, right-skewed pattern supports using percentile-based thresholds for classification rather than relying on mean-based interpretation, particularly in population-based screening contexts.[15]

From a physiological perspective, thermographic patterns may reflect interactions among metabolic stress, inflammatory activity, and microvascular perfusion. The alignment of ischemic burden and fasting blood glucose within the same component suggests a potential interaction between vascular compromise and metabolic dysregulation. Although causal relationships cannot be established due to the cross-sectional design, the observed clustering supports the notion that thermography captures physiologically meaningful patterns of early tissue stress.

The combined thermographic index demonstrated improved discriminative performance compared with individual parameters. The full model achieved an AUC of 0.714 (95% CI: 0.631-0.790), indicating moderate discrimination between normative and non-normative participants. At the optimal cutoff value of 0.408, the model yielded a sensitivity of 74.2% and a specificity of 65.0%. The relatively higher sensitivity suggests that the framework may be useful for identifying individuals with abnormal thermographic patterns in community settings, whereas the moderate specificity indicates that additional clinical assessment may be required to reduce false-positive findings. Therefore, the present framework should be regarded as a risk stratification and screening-support tool rather than a diagnostic method.

Although the first two principal components explained 45.32% of the total variance, PCA was applied primarily as an exploratory technique for visualization and pattern identification rather than for predictive modeling. The first two components revealed clinically interpretable metabolic-inflammatory and vascular-ischemic patterns and were therefore retained for graphical representation and physiological interpretation. The remaining variance likely reflects the multifactorial nature of foot thermoregulation and biological heterogeneity within the study population.

The absence of high-risk cases in this cohort suggests that the model primarily captures early or subclinical changes rather than advanced pathology. This finding is consistent with the role of thermography as a noninvasive screening modality. Compared with conventional vascular assessments, thermography offers a rapid, contact-free, and radiation-free approach that may be suitable for large-scale or community-based screening programs. The integration of thermographic indices with metabolic variables may further enhance risk stratification and support preventive care strategies.

This study has several strengths, including the establishment of normative thermographic reference values, the development of quantitative burden indices, and the application of multivariate analysis to provide physiological insight. However, several limitations should be considered. First, the cross-sectional design precludes causal inference and does not permit evaluation of longitudinal clinical outcomes. Second, established clinical reference standards, including the ankle-brachial index (ABI), Doppler ultrasound, and neuropathy assessments, were unavailable for validation. Third, data collection was conducted in community-based field settings rather than in laboratory-controlled environments. Although temperature normalization and acclimatization procedures were applied, residual environmental influences may have affected thermographic measurements. Finally, the normative reference database remains relatively small and may require expansion to improve generalizability across broader populations.

Future studies with larger cohorts and longitudinal follow-up are warranted to validate the predictive value of thermographic indices. The integration of thermography with advanced analytical approaches, such as machine learning and mobile health platforms, may further enhance its applicability in automated risk assessment and remote screening.

The conceptual framework presented in Figure 4 integrates the relationships between vascular-ischemic and metabolic-inflammatory processes, thermographic assessment, and community-based screening applications. It illustrates how these underlying mechanisms may give rise to distinct thermographic patterns detectable through infrared imaging. By combining percentile-based references, thermographic burden indices, and multivariate analysis, this framework provides a structured approach for translating complex thermal patterns into clinically interpretable information.

Overall, this integrative approach highlights the potential of thermography not only to characterize underlying physiological variation but also to support risk stratification in community settings. The identification of subtle thermal alterations may facilitate earlier recognition of individuals at risk for vascular or metabolic stress and support preventive foot care strategies. However, further validation against established clinical assessments, including the ankle-brachial index (ABI), Doppler ultrasound, neuropathy testing, and longitudinal clinical outcomes, is required before clinical implementation.

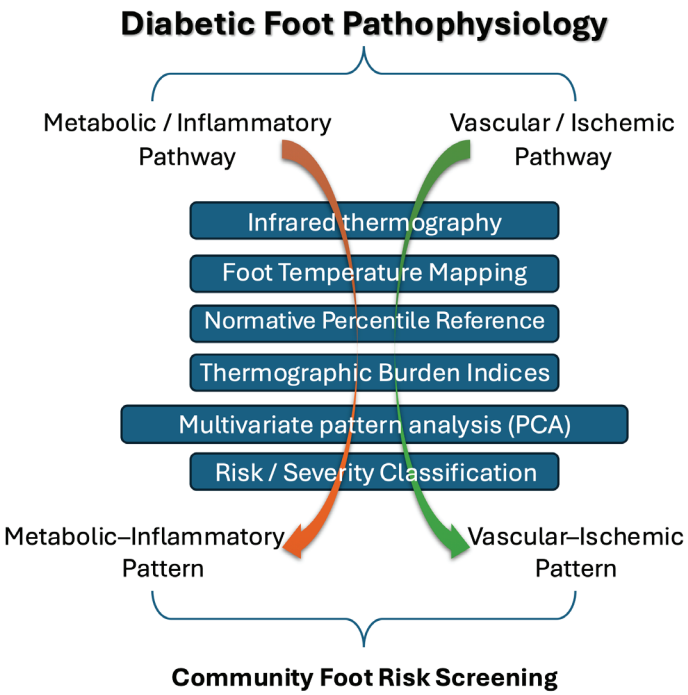


Figure 4. *Conceptual framework linking vascular-ischemic and metabolic-inflammatory processes, thermographic patterns, and community-based risk screening.*

Limitations

The study has several limitations, including the lack of validation against established clinical reference standards (e.g., ABI, Doppler ultrasound, and neuropathy assessments), potential environmental influences during community-based data collection, and the relatively limited size of the normative reference database.

Conclusion

In conclusion, a percentile-based thermographic framework was developed to evaluate foot thermal patterns in a community population. The integration of normative references, thermographic burden indices, and multivariate analysis enables objective characterization of thermal abnormalities associated with metabolic-inflammatory and vascular-ischemic processes.

The combined thermographic index demonstrated moderate discriminative performance, supporting its potential utility for identifying individuals with abnormal thermographic patterns in community settings. Overall, this approach highlights the potential of infrared thermography as a practical, noninvasive, and scalable tool for community-based screening and risk stratification of foot health abnormalities.

Ethical approval

This study was approved by the Human Research Ethics Committee, Faculty of Associated Medical Sciences, Chiang Mai University, under ethical approval code AMSEC-67EX-007. Ethical approval was valid from February 2, 2024, to February 1, 2025.

Funding

This research received no external funding.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Chatchanok Udomtanakunchai: conceptualization, methodology, investigation, data curation, formal analysis, visualization, writing: original draft, review and edit; **Sirasith Uat-Hao**: data curation, image analysis, writing: original draft; **Kemika Wongsatan**: data curation, image analysis, writing: original draft; **Kittichai Wantanajittikul**: conceptualization, methodology, software, image analysis, writing: review and edit; **Kanyawee Sathapong**: writing; review and edit; **Runchana Kantanat**: formal analysis.

Acknowledgements

The authors would like to thank the Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand, for kindly providing laboratory data used in this study. The authors acknowledge the Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand, for facilitating access to the infrared thermography equipment used in this research.

References

[1] Soyoye DO, Abiodun OO, Ikem RT, Kolawole BA, Akintomide AO. Diabetes and peripheral artery disease: A review. World J Diabetes. 2021; 12(6): 827-38. doi: 10.4239/wjd.v12.i6.827.

[2] Armstrong DG, Holtz-Neiderer K, Wendel C, Mohler MJ, Kimbriel HR, Lavery LA. Skin temperature monitoring reduces the risk for diabetic foot ulceration in high-risk patients. Am J Med. 2007; 120(12): 1042-6. doi: 10.1016/j.amjmed.2007.06.028.

[3] Ring EF, Ammer K. Infrared thermal imaging in medicine. Physiol Meas. 2012; 33(3): R33-46. doi: 10.1088/0967-3334/33/3/R33.

[4] Brothers NI, Rebman A, Zenilman JM, Aucott J. Use of infrared thermography in visualizing erythema migrans. Cureus. 2025; 17(8): e89242. doi: 10.7759/cureus.89242.

[5] Lahiri BB, Bagavathiappan S, Jayakumar T, Philip J. Medical applications of infrared thermography: A review. Infrared Phys Technol. 2012; 55(4): 221-35. doi: 10.1016/j.infrared.2012.03.007.

[6] Lavery LA, Higgins KR, Lanctot DR, Constantinides GP, Zamorano RG, Armstrong DG, et al. Home monitoring of foot skin temperatures to prevent ulceration. Diabetes Care. 2004; 27(11): 2642-7. doi: 10.2337/diacare.27.11.2642.

[7] Armstrong DG, Lavery LA, Liswood PJ, Todd WF, Tredwell JA. Infrared dermal thermometry for the high-risk diabetic foot. Phys Ther. 1997; 77(2): 169-75; discussion 76-7. doi: 10.1093/ptj/77.2.169.

[8] Faus Camarena M, Izquierdo-Renau M, Julian-Rochina I, Arrebola M, Miralles M. Update on the use of infrared thermography in the early detection of diabetic foot complications: A bibliographic review. Sensors (Basel). 2023; 24(1). doi: 10.3390/s24010252.

[9] Castillo-Morquecho R, Guevara E, Ramirez-GarciaLuna JL, Martinez-Jimenez MA, Medina-Rangel MG, Kolosovas-Machuca ES. Digital infrared thermography and machine learning for diabetic foot assessment: thermal patterns and classification. J Diabetes Metab Disord. 2024; 23(2): 1967-76. doi: 10.1007/s40200-024-01452-0.

[10] Liew H, Tang W, Plassmann P, Machin G, Simpson R, Edmonds ME, et al. Infrared thermography shows that a temperature difference of 2.2 degrees C (4 degrees F) or greater between corresponding sites of neuropathic feet does not always lead to a diabetic foot ulcer. J Diabetes Sci Technol. 2025; 19(6): 1624-34. doi: 10.1177/19322968241249970.

[11] Li PL, Yick KL, Yip J, Ng SP. Influence of upper footwear material properties on foot skin temperature, humidity and perceived comfort of older individuals. Int J Environ Res Public Health. 2022;19(17). doi: 10.3390/ijerph191710861.

[12] Lavery LA, Higgins KR, Lanctot DR, Constantinides GP, Zamorano RG, Athanasiou KA, et al. Preventing diabetic foot ulcer recurrence in high-risk patients: use of temperature monitoring as a self-assessment tool. Diabetes Care. 2007; 30(1): 14-20. doi: 10.2337/dc06-1600.

[13] Politi S, Aloisi A, Jr., Bartoli V, Guglietta A, Magnifica F. Infrared thermography images acquisition for a technical perspective in screening and diagnostic processes: Protocol standardized acquisition. Cureus. 2021; 13(11): e19931. doi: 10.7759/cureus.9931.

[14] Gatt A, Formosa C, Cassar K, Camilleri KP, De Raffaele C, Mizzi A, et al. Thermographic patterns of the upper and lower limbs: baseline data. Int J Vasc Med. 2015; 2015: 831369. doi: 10.1155/2015/831369.

[15] Pepe MS. The statistical evaluation of medical tests for classification and prediction: Oxford University Press; 2003. Available from: https://doi.org/10.1093/oso/9780198509844.001.0001.

[16] Group WHOMGRS. WHO Child Growth Standards based on length/height, weight and age. Acta Paediatr Suppl. 2006; 450: 76-85. doi: 10.1111/j.16512227.2006.tb02378.x.

[17] Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, Flegal KM, Guo SS, Wei R, et al. CDC growth charts: United States. Adv Data. 2000; 314: 1-27.

[18] Altman DG, Bland JM. Statistics notes: the normal distribution. BMJ. 1995; 310(6975): 298. doi: 10.1136/bmj.310.6975.298.

[19] van Netten JJ, van Baal JG, Liu C, van der Heijden F, Bus SA. Infrared thermal imaging for automated detection of diabetic foot complications. J Diabetes Sci Technol. 2013; 7(5): 1122-9. doi: 10.1177/193229681300700504.

[20] Ilo A, Romsi P, Makela J. Infrared thermography and vascular disorders in diabetic feet. J Diabetes Sci Technol. 2020; 14(1): 28-36. doi: 10.1177/1932296819871270.

Journal of Associated Medical Sciences 2026; 59(3): 215-223

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



Exploring the antioxidant and prebiotic potential of fermented Thai jasmine rice beverage: An *in vitro* study

Ran Kitkangplu[1*], Krongpong Monpengpinij[2], Kongkiat Kespechara[3], Woottichai Nachaiwieng[1,4], Niwed Kullawong[1,4]

[1]*School of Health Science, Mae Fah Luang University, Muang, Chiang Rai Province, Thailand.*
[2]*Siriraj Hospital, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.*
[3]*Sooksatharana (Social Enterprise) Co., Ltd., Muang, Phuket Province, Thailand.*
[4]*Biomedical Technology Research Group for Vulnerable Populations, Mae Fah Luang University, Muang, Chiang Rai Province, Thailand.*

ARTICLE INFO

Article history:
Received 23 March 2026
Accepted as revised 10 June 2026
Available online 2 July 2026

Keywords:
Koji amazake, antioxidant, prebiotic, jasmine rice, Japanese rice, isomaltooligosaccharide

ABSTRACT

Background: Amazake is a traditional Japanese non-alcoholic rice beverage produced through koji fermentation and recognized as a functional food. Thai jasmine rice (*Oryza sativa* L. KDML 105) is rich in health-promoting nutrients and bioactive compounds, yet no studies have explored the health benefits of fermented Thai jasmine rice beverages.

Objectives: To investigate the health-promoting properties of fermented Jasmine rice amazake (JAS-AMA), focusing on phenolic and flavonoid content, antioxidant capacity, prebiotic effects, and antipathogenic activity against bacteria.

Materials and methods: An *in vitro* study compared JAS-AMA with non-fermented Thai jasmine rice (JAS-SYR) and traditional Japanese amazake (JP-AMA). Total phenolic content, total flavonoid content, DPPH and ABTS antioxidant assays, and prebiotic activity against *Lactobacillus plantarum (L. plantarum)* and *Escherichia coli (E. coli)* were evaluated using isomaltooligosaccharide (IMO) as a prebiotic reference.

Results: The total phenolic content of JAS-AMA was significantly higher than that of JAS-SYR (28±0.44 compared with $3.0\pm0.24\times10^{-3}$ mg GAE/µL extract, $p<0.05$), but comparable to that of JP-AMA ($34\pm1.3\times10^{-3}$ mg GAE/µL extract). However, the total flavonoid content of both JAS-AMA and JAS-SYR was significantly lower than JP-AMA (63±6.80 and 12±1.7, respectively, compared with $128\pm22\times10^{-7}$ mg CE/µL extract, $p<0.05$). DPPH assay showed comparable antioxidant activity between JAS-AMA and JP-AMA (70.34% and 65.71% inhibition, respectively), both significantly higher than JAS-SYR (20.14% inhibition, $p<0.05$), whereas ABTS assay revealed significantly lower antioxidant activity in both JAS-AMA and JAS-SYR, lower than JP-AMA (53.57% and 66.05% compared with 95.42% inhibition, $p<0.05$). The maximum OD600 of *L. plantarum* grown with JAS-AMA (2.0025±0.10) was comparable to that of glucose (1.8050±0.1882) and significantly higher than that of IMO alone (1.4213±0.0925, $p<0.05$), suggesting a notable prebiotic effect. No significant differences in *E. coli* growth were observed across any substrate condition ($p>0.05$).

Conclusion: Jasmine rice amazake demonstrates comparable phenolic content and partial antioxidant activity to traditional Japanese amazake, supporting its potential as a functional beverage and encouraging further development using other rice varieties.

* **Corresponding contributor.**
Author's Address: *School of Health Science, Mae Fah Luang University, Muang, Chiang Rai Province, Thailand.*
E-mail address: 6451811006@lamduan.mfu.ac.th
doi: 10.66285/JAMS.2026.095
E-ISSN: 2539-6056

Introduction

Traditional Japanese diets are widely associated with longevity and good health in the Japanese population. A key element of Washoku is rice as the primary carbohydrate source, commonly consumed alongside a variety of fermented foods produced with *Aspergillus oryzae (A. oryzae)* (Koji).[1,2] One representative fermented product is amazake, a sweet, non-alcoholic rice beverage made by fermenting cooked rice with koji. Traditionally, Japanese white rice (*Oryza sativa japonica*) is used for amazake production, during which koji enzymes promote starch saccharification, resulting in the formation of over 300 compounds, including metabolites derived from Koji.[3,4] Due to its rich content of fermentation-derived carbohydrates and bioactive compounds, amazake has gained attention as a potential functional food, particularly amid the rapidly expanding global prebiotic market.[5] Previous studies have identified prebiotic compounds in koji amazake, such as isomaltooligosaccharides (IMO), isomaltotriose, and monohexosyl-ceramides, which have been reported to beneficially modulate intestinal microbiota and improve gastrointestinal health. In addition, amazake contains essential vitamins that contribute to overall human health.[2,4,6] Despite these recognized benefits, current research has primarily focused on amazake produced from Japanese white rice, leaving a critical knowledge gap regarding how different rice varieties influence fermentation outcomes, prebiotic composition, and functional properties. Therefore, further investigation into amazake is necessary to deepen scientific understanding and support the development of diverse, evidence-based prebiotic products.

Thai jasmine rice (*Oryza sativa* L. KDML 105), along with Japanese white rice (*Oryza sativa* japonica), is one of the two main subspecies of Asian cultivated rice, both of which are staple foods for half of the world's population. These subspecies exhibit distinct structural, biochemical, and physiological characteristics, as evidenced by differences in starch microstructure and protein composition, which influence starch digestibility and metabolic behavior.[7,8] Such intrinsic compositional differences are likely to affect enzymatic starch breakdown and saccharification efficiency during fermentation with *A. oryzae*. Consequently, variations in rice subspecies may influence the fermentation dynamics and the profile of bioactive compounds produced in amazake. Despite its widespread consumption and economic importance across Asia, Thai jasmine rice remains underexplored as a substrate for fermented products such as amazake.

The objective of this study is to explore the potential health-promoting properties of fermented Thai jasmine rice beverage or Thai jasmine rice amazake. These properties include antioxidant potential, assessed by measuring phenolic content and antioxidant capacity as previously studied in traditional amazake,[6] and prebiotic potential, evaluated *in vitro* using selected probiotic and pathogenic bacterial strains to assess the selective effects of isomaltooligosaccharides (IMO), serving as a prebiotic reference for comparison with the amazake candidate. *L. plantarum* was included as a well-established probiotic with documented IMO-mediated growth stimulation, serving as a reference for prebiotic activity.[4,9] This *in vitro* study has examined the growth of pathogenic bacteria, specifically *E.coli*, which are expected to proliferate poorly on oligosaccharide carbohydrate sources and jasmine rice-based amazake, to assess the antipathogenic potential.[10]

Materials and methods

Chemicals

Gallic acid (Hi-media, India, Hi-AR grade), trolox (Sigma-Aldrich, US, AR grade purity≥98%), folin-ciocalteu reagent (Sigma-Aldrich, US, AR grade purity ≥98%), 1,1-diphenyl-2-picrylhydrazyl (DPPH) (Sigma-Aldrich, US, AR grade purity ≥98%), diammonium salt of 2,2-azino-bis-(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) (Hi-media, India, Hi-AR grade), sodium nitrite ($NaNO_2$) (KEMAUS, Australia, AR grade purity ≥98%), potassium persulphate ($K_2S_2O_8$)(Hi-media, India, Hi-AR grade), aluminium chloride ($AlCl_3$) (KEMAUS, Australia, AR grade purity ≥98%), sodium hydroxide (NaOH)(RCI Labscan, Thailand, AR grade), sodium carbonate (Na_2CO_3)(Hi-media, India, Hi-AR grade), isopropanol (Sigma-Aldrich, US, AR grade purity ≥98%), Isomaltoligosaccharide (IMO) (Wako, Japan, Food grade).

Sample preparation

The amazake used in this study was kindly provided by YoRice Café (Chiang Mai, Thailand), a producer and distributor of YoRice amazake beverages made from Thai jasmine rice (*Oryza sativa* L., KDML 105). The production process involved preparing the koji starter by inoculating steamed Japanese rice (*Oryza sativa* japonica) with *A. oryzae* and incubating it at 40 °C for 3 days. The resulting starter was then mixed with steamed jasmine rice and clean water, followed by 24 hrs of fermentation at 50 °C with continuous stirring. This production process was developed through iterative trial-and-error optimization of the product. The fermented mixture was subsequently pasteurized and distributed.

The modified amazake samples used in this study consisted of three groups: 1) amazake made from 0.5 kg of cooked Thai jasmine rice fermented with 0.5 kg of koji (1:1 ratio) homogenized with 1 liter of drinking water and pasteurized (JAS-AMA), 2) cooked jasmine rice 0.5 kilogram extract mixed with the sucrose syrup low calories 200 mL/kg of jasmine rice homogenized with 1 liter drinking water to replace the sugar that would be created after fermentation, but not subjected to fermentation or pasteurization (JAS-SYR), and 3) a commercial Japanese amazake product prepared from Japanese rice (JP-AMA), with the trade name

withheld to avoid potential impact on product image. All samples were transported at 7 °C and stored at -20 °C until extraction and subsequent analysis. The sample extraction procedure was based on Oguro *et al.'s* study.[11] Briefly, 10.0 mL of the sample was centrifuged in a Universal 320 centrifuge (Hettich) at 2,300 × g for 5 min to separate into layers. The solution supernatant was transferred into a 250-mL evaporating flask and dried using a Rotary Evaporator Hel-Vop (Heidolph, Germany) until the sample was completely dry. The dried sample was then dissolved in 1 mL of 50% isopropanol. The sample was then diluted fivefold with distilled deionized water (1:4 sample-to-water ratio). The sample was then filtered with a 0.22-micron Nylon syringe filter, obtaining approximately 4.0-4.5 mL of filtrate. The filtered sample was aliquoted and stored in a clean 1 mL glass vial at -20 °C before analysis

Determination of total flavonoid content (TFC)

The sample extract was determined for total flavonoid content (TFC) using the aluminum chloride method, following Mehmood *et al.'s* modifications.[12] Briefly, 20 µL of the amazake extract was added to a 96-well plate, followed by 80 µL of distilled water and 6 µL of a 5% (w/v) $NaNO_2$ solution. The mixture was incubated in the dark at room temperature for 6 min. Subsequently, 6 µL of a 10% (w/v) $AlCl_3$ solution was added, and the mixture was incubated for an additional 6 min under the same conditions. Finally, 40 µL of 1M NaOH was added, and the absorbance was measured at 510 nm. The TFC was expressed as milligrams of catechin equivalents per gram of extract (mg CAE/gm extract).

Determination of total phenolic content (TPC)

The sample extract was determined for total phenolic content (TPC) using the Folin-Ciocalteu method based on Mehmood *et al.* modifications.[12] Briefly, 15 µL of the extract amazake sample and 125 µL of 10% Folin-Ciocalteu reagent were added to a 96-well plate and incubated in the dark at room temperature for 5 min, followed by the addition of 125 µL of 1.89 M Na_2CO_3 solution (w/v). The mixtures were thoroughly mixed and incubated in the dark at room temperature for 90 min. Absorbance was then measured at 725 nm. The TPC was expressed as milligrams of gallic acid equivalents per gram of extract (mg GAE/gm extract).

Determination of DPPH free radical scavenging assay

The DPPH (2,2-diphenyl-1-picrylhydrazyl) radical-scavenging activity was determined using a colorimetric method with trolox as the standard, following the modifications of Zengin *et al.*[13] A 20 µL volume of the amazake extract was added to a 96-well plate, followed by 180 µL of an 80 µM DPPH methanol solution. The mixtures were thoroughly mixed and incubated in the dark at room temperature for 5 min. Following incubation, the absorbance was measured at 490 nm. The DPPH free radical scavenging activity was calculated using the equation:

DPPH %inhibition = [(OD of sample - OD of blank)/ OD of control] ×100

Determination of ABTS free radical scavenging assay

The ABTS (2,2'-azino-bis-3-ethylbenzothiazoline-6-sulfonic acid) radical-scavenging activity was determined using a colorimetric method with trolox as the standard, based on Re *et al.* modifications.[14] The ABTS solution was prepared by mixing 2 mM ABTS with 2.45 mM K2S2O8 in a 1:1 ratio, then kept in the dark at room temperature for 16 hrs. Then the ABTS solution was diluted with 50 mM phosphate buffer (pH 7.4) to achieve an absorbance of 0.70±0.02 at 734 nm. A 20 µL volume of the amazake extract was added to a 96-well plate, followed by 200 µL of ABTS solution. The mixtures were thoroughly mixed and incubated in the dark at room temperature for 6 min. Following incubation, the absorbance was measured at 734 nm. The ABTS free radical scavenging activity was calculated using the equation:

ABTS %inhibition = [(OD of sample - OD of blank)/ OD of control] ×100

Probiotic growth stimulation and pathogenic inhibition stimulation

Probiotic growth stimulation and pathogenic inhibition were assessed following the method of Rurangwa *et al.*, specifically targeting the probiotic bacterium *L. plantarum* (TISTR-2074) and the pathogenic bacterium *E. coli* (TISTR-527).[15-17] They were grown in modified MRS (wMRS) broth supplemented with mixture of 20 gm/L beef extract and proteose peptone, 5 gm/L yeast malt broth, 1 gm/L tween 20, 1 gm/L citric acid monohydrate, 1 gm/L ammonium acetate, 5 gm/L sodium acetate, 2 gm/L di-sodium hydrogen phosphate, 0.2 gm/L magnesium sulfate anhydrous, and 0.05 gm/L manganese (II) sulfate monohydrate, dissolved in distilled water with content 20 gm/L different substrates (in ratio 2:3 of content wMRS) including glucose (GLU), isomaltooligo-saccharide (IMO), and Thai jasmine rice amazake (The solution was approximated using the phenol-sulfuric acid method compare with glucose 20 gm/L). After semi-anaerobic incubation (sealed with a cotton plug) at 37 °C for 48 hrs, an approximate cell concentration of 3×10^8 cells/ mL was obtained in different MRS broths supplemented with various substrates (inoculum: 1% of the broth). For detecting the microbe in broth, we referenced the light wavelengths that should be used from the Montesinos-Cruz and Somerville study.[18] Samples were collected to determine growth or inhibition by monitoring the optical density at 600 nm in the GENESYS™ 180 UV-Vis spectrophotometer (Thermo Scientific, United States). at 600 nm.

Statistical analysis

One-way analysis of variance (ANOVA) followed

by Tukey's test was performed using the Statistical Package for the Social Sciences (SPSS; IBM, United States). Differences with a $p<0.05$ were considered significant. All data are reported as the means of triplicate measurements ± standard deviation (SD).

Results

Total phenolic content, total flavonoid content, and antioxidant activity

Standard curves for all four assays demonstrated excellent linearity across the tested concentration ranges (Figure 1). For the TFC assay, catechin was used as the reference standard (0.0004-0.0020 mg/mL), yielding a linear equation of y=73.217x+0.0007 with an R^2 of 0.9967. For the TPC assay, gallic acid served as the standard (0.0098-1.2500 mg/mL), with a regression equation of y=2.1717x+0.0466 and an R^2 of 0.9981. For the DPPH radical-scavenging assay, Trolox was used as the reference standard (0.0020-0.0310 mg/mL), yielding a linear equation of y=1842.8x+18.804 with an R^2 of 0.9979. Similarly, the ABTS assay standard curve using Trolox (0.0020-0.5000 mg/mL) yielded a regression equation of y=128.92x+5.3052 with an R^2 of 0.9998. All R^2 values exceeded 0.99, confirming high linearity and reliability for each calibration curve across the measured concentration ranges.

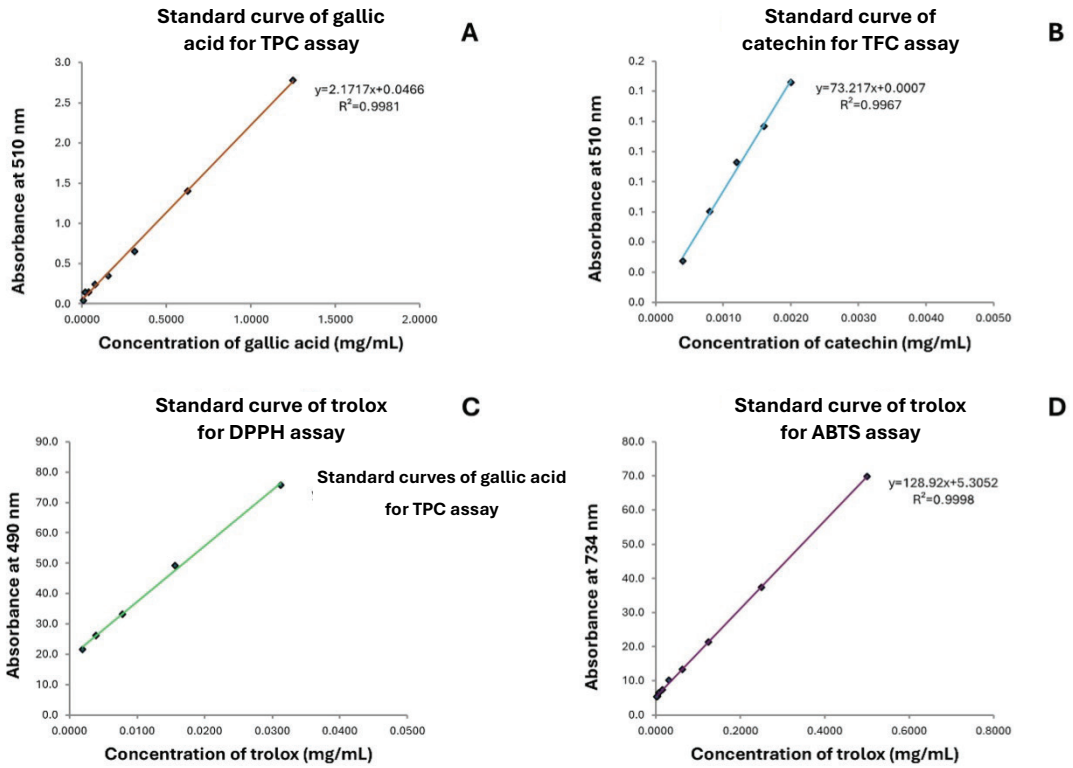


Figure 1. *Standard curves for catechin in the TPC assay (A), gallic acid in the TFC assay (B), trolox in the DPPH (C) and ABTS (D) for radical-scavenging assay. Each data point represents the average of triplicate measurements. Linear regression equations and coefficients of determination ($R^2>0.99$) are shown for each curve.*

TPC and TFC ranked in descending order as JP-AMA > JAS-AMA > JAS-SYR (Figure 2). Koji fermentation significantly enhanced both TPC (28.33±0.44×10^{-3} mg GAE/µL) and TFC (63.51±6.80×10^{-7} mg CE/µL) in JAS-AMA compared to JAS-SYR (3.09±0.24×10^{-3} mg GAE/µL and 12.75±1.7×10^{-7} mg CE/µL, respectively, p<0.05), though both remained significantly lower than JP-AMA (34.6±1.3×10^{-3} mg GAE/µL and 128.84±22×10^{-7} mg CE/µL, respectively, p<0.05).

For antioxidant activity (Figure 3), DPPH scavenging by JAS-AMA (70.34±0.42%) was significantly higher than that of JP-AMA (65.71±1.16%), with both markedly exceeding JAS-SYR (20.13±2.22%, p<0.05). Conversely, ABTS scavenging followed the order JP-AMA (95.42±0.28%) > JAS-SYR (66.05±0.62%) > JAS-AMA (53.57±0.89%), with all groups significantly different from each other (p<0.05).

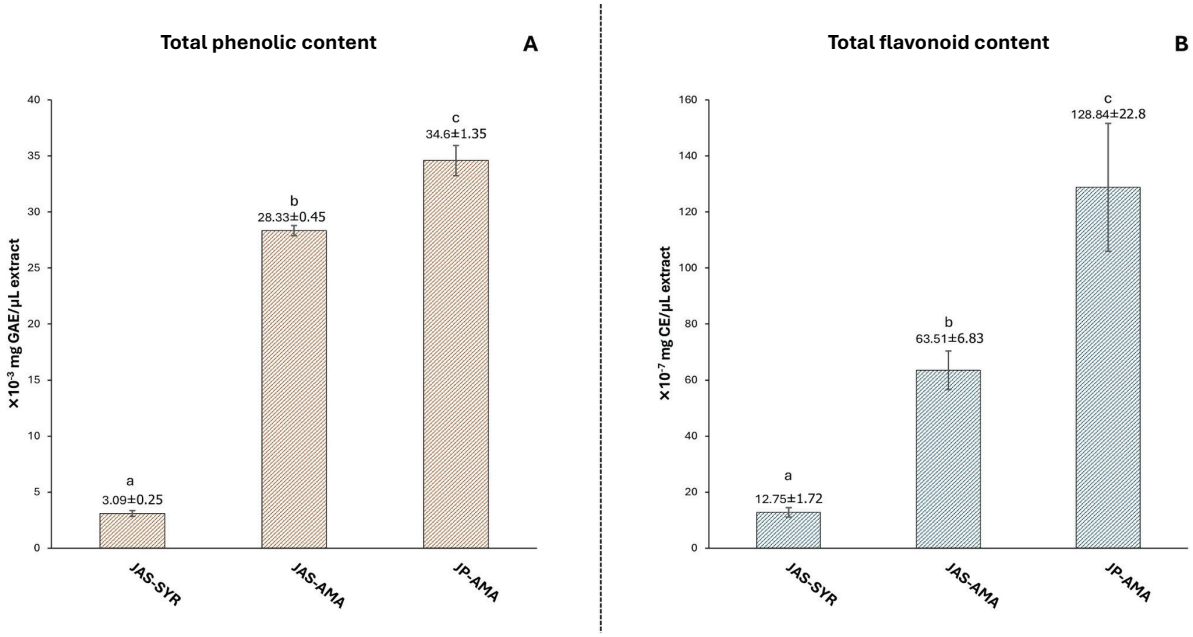


Figure 2. *Total phenolic content (A) and total flavonoid content (B) of non-fermented Thai jasmine rice (JAS-SYR), Thai jasmine rice amazake (JAS-AMA), and traditional Japanese amazake (JP-AMA). Values are expressed as mean±standard deviation (N=3). TPC is expressed as mg gallic acid equivalents (GAE)/µL, and TFC as mg catechin equivalents (CE)/µL. Values followed by different lowercase letters are significantly different (p<0.05, Tukey's test).*

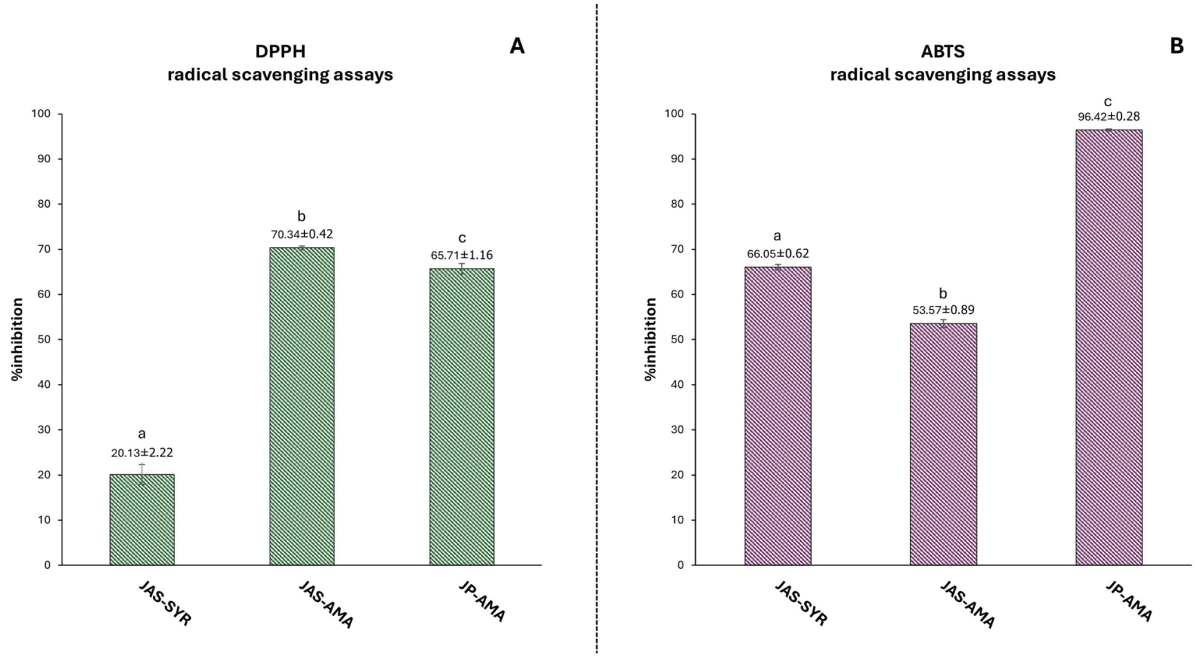


Figure 3. *Antioxidant activity of non-fermented Thai jasmine rice (JAS-SYR), Thai jasmine rice amazake (JAS-AMA), and traditional Japanese amazake (JP-AMA) as measured by DPPH (A) and ABTS (B) radical scavenging assays. Values are expressed as mean±SD (N=3). Antioxidant activity is reported as a percentage inhibition. Values followed by different lowercase letters are significantly different (p<0.05, Tukey's test).*

Probiotic growth stimulation and pathogenic growth inhibition

To evaluate the prebiotic potential of JAS-AMA, the growth of *L. plantarum* and *E. coli* was monitored over 48 hrs by measuring optical density at 600 nm (OD600), with the 36-hour time point used for endpoint comparisons (Figure 4). Four substrate conditions were tested for each bacterium: a carbohydrate-free modified MRS broth as a negative control (LP_wMRS / EC_wMRS), glucose-supplemented MRS broth as a positive control (LP_GLU / EC_GLU), isomaltooligosaccharide-supplemented MRS broth as a prebiotic reference (LP_IMO / EC_IMO), and Thai jasmine rice amazake extract-supplemented MRS broth as the test substrate (LP_JAS_AMA / EC_JAS_AMA).

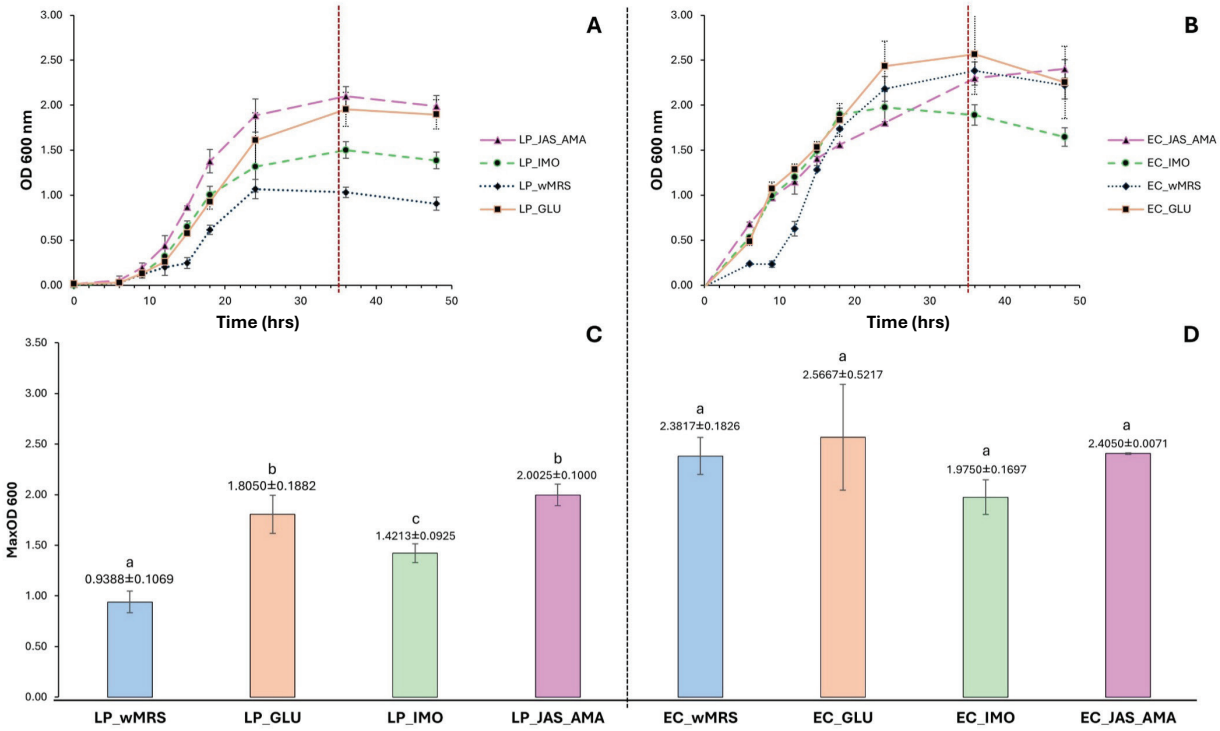


Figure 4. *Growth of L. plantarum (LP) (A, C) and E. coli (EC) (B, D) in the presence of different amazake substrates. (A, B) Optical density (OD_{600}) measurements over time in media supplemented with wMRS (control), GLU (glucose), IMO (isomaltooligosaccharide), and JAS-AMA (Thai jasmine rice amazake). The dashed vertical line at 36 hrs indicates the time point used for endpoint comparisons with mean OD_{600}±SD at 36 hrs. (N=3) below figure (C, D). Values followed by different lowercase letters are significantly different ($p<0.05$, Tukey's test).*

For *L. plantarum*, maximum OD600 differed significantly among substrate groups ($p<0.05$). LP_JAS_AMA (2.0025±0.10) and LP_GLU (1.8050±0.1882) achieved comparable maximum growth and were not significantly different from each other (both group b), while LP_IMO (1.4213±0.0925, group c) and LP_wMRS (0.9388±0.1069, group a) showed progressively lower growth. The finding that LP_JAS_AMA supported *L. plantarum* growth at least as effectively as glucose, and markedly better than IMO alone, suggests a notable prebiotic effect of JAS-AMA on this probiotic strain.

In contrast, no significant differences in maximum OD600 were observed among *E. coli* groups ($p>0.05$), with all groups sharing the same statistical label (group A) and values ranging from EC_IMO (1.9750±0.1697) to EC_GLU (2.5667±0.5217). The inability of JAS-AMA to selectively suppress *E. coli* growth under the tested conditions suggests that the antipathogenic potential of JAS-AMA may not act through direct carbohydrate-mediated growth inhibition; further investigation into other inhibitory mechanisms, such as antimicrobial metabolites produced during koji fermentation, is warranted.

Discussion

Fermentation by *A. oryzae* releases carbohydrate-degrading enzymes, including α-amylase, xylanase, and glucosidase, that break the cross-links of bound phenolic compounds. Rather than fermentation temperature, which aligns with the optimal activity range (~50 °C) of key *Aspergillus* enzymes,[19,20] the composition of the rice cell wall likely governs the concentration and types of phenolics released, explaining the observed varietal differences in TPC and TFC between JAS-AMA and JP-AMA.[21]

Antioxidant capacity generally showed a positive association with phenolic content, consistent with previous reports.[22,23] However, JAS-AMA exhibited

higher DPPH scavenging activity than JP-AMA despite lower TPC, suggesting additional contributions from fermentation-derived compounds such as gamma-oryzanols, tocopherols, and tocotrienols.[23,24] The DPPH and ABTS assays differ in their sensitivity to antioxidants. The DPPH primarily reflects hydrophobic activity, whereas ABTS captures both hydrophilic and lipophilic antioxidants.[25] This distinction is relevant to interpreting the unexpectedly elevated ABTS activity in JAS-SYR, which likely reflects uncharacterized antioxidant constituents in the added syrup rather than phenolic or flavonoid content, representing a key limitation of this study.

Beyond antioxidant activity, phenolic compounds and fermentation-derived oligosaccharides, including isomaltooligosaccharides (IMO), cellobiose, and mannobiose, have been proposed as prebiotic substrates that selectively stimulate beneficial gut bacteria, such as *Lactobacillus* spp., thereby conferring health benefits on the host.[26-28] In the present study, JAS-AMA supported *L. plantarum* growth to levels comparable to those with glucose and significantly greater than those with IMO alone, suggesting a synergistic prebiotic effect from the combined presence of fermentation-derived oligosaccharides and phenolic compounds. Mechanistically, prebiotic fermentation promotes the production of short-chain fatty acids (SCFAs), including butyrate and propionate, which lower colonic pH and exert anti-inflammatory effects.[29]

No significant inhibition of *E. coli* growth was observed under any substrate condition ($p>0.05$), suggesting that the antipathogenic potential of JAS-AMA does not operate through carbohydrate-mediated growth restriction alone. Antimicrobial effects are more likely attributable to secondary metabolites produced during koji fermentation, as metabolic profiling of *A. oryzae* has identified coumarins, fatty acids, and quinone compounds with documented activity against *E. coli*.[30,31] Whether these compounds reach effective concentrations in JAS-AMA warrants further investigation.

Limitations

Several limitations should be acknowledged in this study. Firstly, the chemical composition of the sucrose syrup used in JAS-SYR preparation was not characterized, which prevents a full interpretation of the unexpectedly elevated ABTS radical-scavenging activity observed in that sample, as unquantified antioxidant constituents in the syrup may have independently contributed to this response. Secondly, the study encompassed only three sample types and did not involve in-depth characterization, limiting the generalizability of the findings across a wider range of rice varieties and fermentation conditions. Thirdly, the DPPH assay, being conducted primarily in organic solvents, may have underrepresented the contribution of hydrophilic antioxidants, and the specific enzymatic mechanisms and fermentation parameters responsible for the observed differences between JP-AMA and JAS-AMA were not systematically investigated. Lastly, the prebiotic assessment was conducted exclusively *in vitro*, and the observed growth stimulation of *L. plantarum* and the lack of significant inhibition of *E. coli* under these conditions may not fully reflect the complex ecological interactions that occur within the human gastrointestinal tract, where additional host-derived and microbial factors may modulate the antipathogenic response. These limitations should be considered when interpreting the findings of the present study and highlighting important directions for future research.

Conclusion

In conclusion, this study provides new evidence supporting Thai jasmine rice-based amazake as a novel functional beverage with combined antioxidant and prebiotic properties. To our knowledge, this work demonstrates that koji fermentation of Thai jasmine rice significantly increases phenolic and flavonoid contents, thereby enhancing antioxidant capacity. The observed promotion of *L. plantarum* growth, comparable to that of glucose and superior to IMO alone, highlights the prebiotic potential of jasmine rice amazake in selectively stimulating beneficial gut bacteria. Although significant inhibition of *Escherichia coli* was not demonstrated under the *in vitro* conditions tested, the antipathogenic potential of JAS-AMA warrants further investigation, particularly regarding fermentation-derived antimicrobial metabolites such as coumarins, fatty acids, and quinones.

These findings extend current knowledge beyond traditionally studied Japanese rice-based amazake and establish Thai jasmine rice as a promising alternative substrate for prebiotic beverage development.

Furthermore, the results suggest that varietal diversity within *A. oryza*, including white and pigmented jasmine rice varieties, may differentially influence fermentation outcomes and bioactive compound profiles. This opens new avenues for targeted fermentation strategies to maximize health benefits. Future clinical validation is warranted to confirm the impacts on gut health, immune modulation, and overall well-being, thereby supporting the translation of jasmine rice amazake into evidence-based functional food applications.

Availability of data and materials

The data are available from the corresponding author on reasonable request.

Funding

This research was supported by the 2021 Fundamental Fund of the Program Management Unit for Competitiveness (PMUC), the 2023 master's degree Research Project Fund, and laboratory, equipment, and instrument support from Mae Fah Luang University.

The YoRice amazake samples were kindly provided by YoRice Café.

CRediT authorship contribution statement

Ran Kitkangplu: conceptualization, methodology, resources, validation, laboratory operator, formal analysis, visualization, writing: original draft and edit; **Krongpong Monpengpinij**: conceptualization, validation, formal analysis, project administration, writing: original draft; **Kongkiat Kespechara:** conceptualization, resources, validation, formal analysis, project administration, writing: original draft; **Niwed Kullawong** and **Woottichai Nachaiwieng:** conceptualization, methodology, resources, validation, formal analysis, project administration.

Conflict of interest

The authors declare no competing interests.

Declaration of AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT to improve language and readability. After using this tool/service, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

Acknowledgements

The authors wish to thank Dr. Niwed Kullawong and Dr. Woottichai Nachaiwieng for academic guidance and supervision throughout this study. The authors would like to thank YoRice Café (Chiang Mai, Thailand) for kindly providing the Thai jasmine rice amazake samples used in this study. The authors also gratefully acknowledge Mae Fah Luang University for providing laboratory facilities, equipment, and instrumental support essential to conducting this research.

References

[1] Gabriel AS, Ninomiya K, Uneyama H. The role of the Japanese traditional diet in healthy and sustainable dietary patterns around the world. Nutrients. 2018; 10(2): 173. doi: 10.3390/nu10020173.

[2] Hamajima H, Matsunaga H, Fujikawa A, et al. Japanese traditional dietary fungus koji Aspergillus oryzae functions as a prebiotic for Blautia coccoides through glycosylceramide: Japanese dietary fungus koji is a new prebiotic. SpringerPlus. 2016; 5(1): 1321. doi: 10.1186/s40064-016-2950-6.

[3] Hamajima H, Fujikawa A, Yamashiro M, et al. Chemical analysis of the sugar moiety of monohexosylceramide contained in koji, Japanese traditional rice fermented with Aspergillus. Fermentation. 2016; 2(1): 2. doi: 10.3390/fermentation2010002.

[4] Kurahashi A. Ingredients, functionality, and safety of the Japanese traditional sweet drink amazake. J Fungi (Basel). 2021; 7(6): 469. doi: 10.3390/jof7060469.

[5] hoareaup. The prebiotics market forecast to 2023 and beyond [Internet]. Gnosis by Lesaffre; 2022 Mar 9 [cited 2025 Dec 27]. Available from: https://gnosisbylesaffre.com/blog/the-prebiotics-market-forecast-to-2023-and-beyond/.

[6] Oguro Y, Nishiwaki T, Shinada R, Kobayashi K, Kurahashi A. Metabolite profile of koji amazake and its lactic acid fermentation product by Lactobacillus sakei UONUMA. J Biosci Bioeng. 2017; 124(2): 178-83. doi: 10.1016/j.jbiosc.2017.03.011.

[7] Schwanz Goebel JT, Kaur L, Colussi R, Elias MC, Singh J. Microstructure of indica and japonica rice influences their starch digestibility: a study using a human digestion simulator. Food Hydrocolloids. 2019; 94: 191-8. doi: 10.1016/j.foodhyd.2019.02.038.

[8] Yang Y, Zhu K, Xia H, Chen L, Chen K. Comparative proteomic analysis of indica and japonica rice varieties. Genet Mol Biol. 2014; 37(4): 652-61. doi: 10.1590/S1415-47572014005000015.

[9] Soto C. Effect of isomaltooligosaccharide and gentiooligosaccharide on the growth and fatty acid profile of Lactobacillus plantarum. Electron J Biotechnol [Internet]. 2013 [cited 2025 Sep 10];16(4). Available from: https://www.ejbiotechnology.info/index.php/ejbiotechnology/article/view/v16n4-9.

[10] Chantorn S, Piyapittayanun C, Dangpram P. Bioconversion of agricultural wastes to mannooligosaccharides and their prebiotic potential. Chiang Mai J Sci. 2018; 45(1): 60-7.

[11] Oguro Y, Nishiwaki T, Shinada R, Kobayashi K, Kurahashi A. Metabolite profile of koji amazake and its lactic acid fermentation product by Lactobacillus sakei UONUMA. J Biosci Bioeng. 2017; 124(2): 178-83. doi: 10.1016/j.jbiosc.2017.03.011.

[12] Mehmood A, Javid S, Khan MF, Ahmad KS, Mustafa A. In vitro total phenolics, total flavonoids, antioxidant and antibacterial activities of selected medicinal plants using different solvent systems. BMC Chem. 2022; 16(1): 64. doi: 10.1186/s13065-022-00858-2.

[13] Zengin G, Sarikurkcu C, Uyar P, et al. Crepis foetida L. subsp. rhoeadifolia (Bieb.) Celak. as a source of multifunctional agents: cytotoxic and phytochemical evaluation. J Funct Foods. 2015; 17: 698-708. doi: 10.1016/j.jff.2015.06.041.

[14] Re R, Pellegrini N, Proteggente A, Pannala A, Yang M, Rice-Evans C. Antioxidant activity applying an improved ABTS radical cation decolorization assay. Free Radic Biol Med. 1999; 26(9-10): 1231-7. doi: 10.1016/S0891-5849(98)00315-3.

[15] Rurangwa E, Laranja JL, Van Houdt R, et al. Selected nondigestible carbohydrates and prebiotics support the growth of probiotic fish bacteria mono-cultures in vitro. J Appl Microbiol.

2009; 106(3): 932-40. doi: 10.1111/j.1365-2672.2008.04034.x.

[16] Kerdsup P, Hattayapichat P, Silva JL, Tantratian S. Survival of potential probiotic isolated from fermented tea leaf and encapsulated in multilayer beads stored in makiang (Cleistocalyx nervosum var. paniala) juice. Food Biosci. 2022; 50: 102015. doi: 10.1016/j.fbio.2022.102015.

[17] Saelim K. Transfer of Escherichia coli TISTR527 from surface to fresh-cut cantaloupe [Internet]. Published online 2020. doi: 10.14458/RSU.RES.2020.223.

[18] Montesinos-Cruz V, Somerville GA. Shining a light on spectrophotometry in bacteriology. Antibiotics. 2024; 13(12): 1164. doi: 10.3390/antibiotics13121164.

[19] Yan H. Biocatalytic aldol addition reactions by lipase from Aspergillus oryzae. Chiang Mai J Sci. 2025; 52(2). doi: 10.12982/CMJS.2025.019.

[20] Kumar S, Tissopi T, Mutturi S. The successful synthesis of industrial isomaltooligosaccharides lies in the use of transglycosylating -glucosidases: a review. Carbohydr Polym Technol Appl. 2023; 5: 100325. doi: 10.1016/j.carpta.2023.100325.

[21] Saharan P, Sadh PK, Singh Duhan J. Comparative assessment of effect of fermentation on phenolics, flavanoids and free radical scavenging activity of commonly used cereals. Biocatal Agric Biotechnol. 2017; 12: 236-40. doi: 10.1016/j.bcab.2017.10.013.

[22] Cai Y, Luo Q, Sun M, Corke H. Antioxidant activity and phenolic compounds of 112 traditional Chinese medicinal plants associated with anticancer. Life Sci. 2004; 74(17): 2157-84. doi: 10.1016/j.lfs.2003.09.047.

[23] Choi Y, Jeong HS, Lee J. Antioxidant activity of methanolic extracts from some grains consumed in Korea. Food Chem. 2007; 103(1): 130-8. doi: 10.1016/j.foodchem.2006.08.004.

[24] Xu Z, Hua N, Godber JS. Antioxidant activity of tocopherols, tocotrienols, and -oryzanol components from rice bran against cholesterol oxidation accelerated by 2,2'-azobis(2-methyl propionamidine) dihydrochloride. J Agric Food Chem. 2001; 49(4): 2077-81. doi: 10.1021/jf0012852.

[25] Rumpf J, Burger R, Schulze M. Statistical evaluation of DPPH, ABTS, FRAP, and Folin-Ciocalteu assays to assess the antioxidant capacity of lignins. Int J Biol Macromol. 2023; 233: 123470. doi: 10.1016/j.ijbiomac.2023.123470.

[26] Plamada D, Vodnar DC. Polyphenols—gut microbiota interrelationship: a transition to a new generation of prebiotics. Nutrients. 2021; 14(1): 137. doi: 10.3390/nu14010137.

[27] Sawangwan T, Saman P. Prebiotic synthesis from rice using Aspergillus oryzae with solid state fermentation. Agric Nat Resour. 2016; 50(4): 227-31. doi: 10.1016/j.anres.2016.02.004.

[28] Fuloria S, Mehta J, Talukdar MP, et al. Synbiotic effects of fermented rice on human health and wellness: a natural beverage that boosts immunity. Front Microbiol. 2022; 13: 950913. doi: 10.3389/fmicb.2022.950913.

[29] Davani-Davari D, Negahdaripour M, Karimzadeh I, et al. Prebiotics: definition, types, sources, mechanisms, and clinical applications. Foods. 2019; 8(3): 92. doi: 10.3390/foods8030092.

[30] Son SY, Lee S, Singh D, Lee NR, Lee DY, Lee CH. Comprehensive secondary metabolite profiling toward delineating the solid and submerged-state fermentation of Aspergillus oryzae KCCM 12698. Front Microbiol. 2018; 9: 1076. doi: 10.3389/fmicb.2018.01076.

[31] Tu Z, Choi D, Chen Y, Yu JH, Huynh TN. The food fermentation fungus Aspergillus oryzae is a source of natural antimicrobials against Listeria monocytogenes. J Dairy Sci. 2025; 108(4): 3444-54. doi: 10.3168/jds.2024-25719.

Journal of Associated Medical Sciences 2026; 59(3): 224-237

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



Deep learning approaches for CT image reconstruction: Advancements and clinical implications

Rita Yadav, Rajesh Kumar*, Kavita Singh

Department of Electronics and Communication, University of Allahabad, India.

ARTICLE INFO

Article history:
Received 24 March 2026
Accepted as revised 24 June 2026
Available online 3 July 2026

Keywords:
CT image reconstruction, deep learning, neural networks, medical imaging, filtered back projection, iterative reconstruction

ABSTRACT

Background: CT is a trusted imaging modality for diagnosing disease. The quality of the CT image plays the most important role in disease prediction. Images will be degraded by noise, blur, and other artifacts. Therefore, the enhancement and reconstruction of CT images are required before processing for prediction. Enhancements are only cosmetic changes in images, but reconstruction provides well-defined models for reconstruction. In this paper, we have compared the performance of these four CT brain reconstruction and enhancement techniques: Filtered back projection (FBP), Iterative reconstruction, Generative Adversarial Network (GAN), and CNN autoencoder, by adding Speckle and Poisson noise.

Objectives: The possible objective statements of this study will cover a broad spectrum of developments in medical imaging technology. These are aimed at enhancing diagnostic test quality, patient safety, and overall care outcomes through innovative imaging methods and data analysis. The suggested objectives involve integrating an artificial intelligence algorithm to enhance them, thereby improving patient care and diagnostic accuracy across numerous medical fields. In this experiment, the CT brain stroke dataset was obtained from the Kaggle database and evaluated using performance metrics (PSNR, SSIM, and MSE) across the normal, ischemia, bleeding, and external test folders, with Poisson and Speckle noise added using the GAN, CNN autoencoder, iterative, and FBP techniques.

Materials and methods: The CT brain stroke dataset contains CT images organized in four folders: normal, ischemia, bleeding, and external test. These folders can be used to evaluate performance metrics. The methods applied include: Convolutional Neural Network (CNN) autoencoder, Generative Adversarial Network (GAN), iterative reconstruction, and FBP, each implemented with hyper-tuned parameters. Experimental results calculate PSNR, SSIM, and MSE after adding Poisson and Speckle noise to test the above deep learning algorithms.

Results: The lowest image quality was achieved for FBP for all datasets and both noise types. When the noise was Poisson, the performance of iterative reconstruction was the best, with PSNR exceeding 37 dB, SSIM exceeding 0.972, and MSE as low as 0.0002. On adding Poisson noise, the GAN achieved higher PSNR values and SSIM values up to 32.54 dB and 0.938, respectively, and with low errors in reconstruction, except for the external test folder. The CNN autoencoder (CNN-AE) also yielded significant improvements over FBP, achieving competitive PSNR and SSIM scores across all datasets.

Conclusion: The results indicate that the quality of CT images can be significantly improved using advanced deep learning and Iterative reconstruction methods compared with conventional FBP. It is observed that iterative reconstruction yields the best results for images corrupted by Poisson noise, whereas GANs perform best for images corrupted by Speckle noise. Therefore, reconstruction performance depends on the nature of the noise present in the CT images.

***Corresponding contributor.**
Author's Address: *Department of Electronics and Communication, University of Allahabad, India.*
E-mail address: rajeshkumar@allduniv.ac.in
doi: 10.66285/JAMS.2026.096
E-ISSN: 2539-6056

Introduction

The use of computed tomography (CT) imaging has transformed medical diagnostics by providing detailed 3-D images of organs. It is a complex imaging process that begins with X-ray emission, tissue attenuation, and data acquisition. The final and most significant process for converting raw data into meaningful information is image reconstruction. Visual information of a clinician-interpretable kind. Traditional restorative methods use mathematical algorithms to facilitate this transformation, generating invaluable data for healthcare professionals to support proper diagnosis and treatment planning.[1]

The development of CT imaging technology has greatly improved image quality, acquisition speed, and radiation dose. Modern CT scanners use multidetector arrays and sophisticated software algorithms to generate high-resolution images with fine detail. Such developments have expanded the uses of CT imaging in certain medical subspecialties, including oncology, cardiology, and neurology.[2] Also, the combination of artificial intelligence and machine learning in image reconstruction has opened new possibilities for enhancing diagnostic accuracy and efficiency. With further development of CT technology, it will become an even more important part of personalized medicine, enabling more specific, targeted interventions using detailed anatomical and functional data about patients. Deep learning has more than a basic image-improvement effect on CT image reconstruction.

More complex neural network models, such as Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), have played a key role in expanding the possibilities in medical imaging.[3] Such advanced algorithms can identify complex patterns and features in large datasets of CT images, enabling them to produce high-quality images from incomplete or noisy data.[3] This capacity has led to significant improvements in sparse-view and limited-angle CT reconstruction, where conventional algorithms often fail to produce diagnostic images.

With the implementation of deep learning in CT imaging, new opportunities to make imaging protocols more adaptive and personalized have emerged.[4] These algorithms can optimize reconstruction methods in real-time by performing patient-specific analyses and adjusting imaging parameters to tailor the imaging process to each individual patient. This will not only improve image quality but also reduce radiation exposure and enhance patient safety. In addition, deep learning-based CT reconstruction has demonstrated potential in multimodal imaging, synthesizing data from multiple modalities to provide a more detailed perspective on anatomy and physiology.

As research in this field advances, the potential for deep learning to transform the medical imaging profession and improve patient treatment becomes increasingly evident. One of the most significant changes in medical imaging technology is the integration of deep learning models and the classical reconstruction in CT imaging. This synergy is based on the merits of both approaches, combining deep learning with the robustness of the tested reconstruction algorithms and the flexibility of deep learning. This implies that the quality of the images obtained will improve, radiation exposure will be reduced, and more accurate diagnoses are likely to be achieved by the researcher and clinician. The outcome of such improvements can be more sensitive imaging protocols that detect subtle abnormalities not easily detected with traditional imaging.[5] Moreover, the current development of this combined strategy is set to transform personalized medicine. Using patient-specific data and leveraging the potential of artificial intelligence, the CT imaging tools may be able to provide customized diagnostic and treatment prescriptions. This individualized treatment could lead to more accurate interventions, the maximization of a treatment plan, and, consequently, improved patient outcomes. Also, the more efficient and precise nature of these new imaging methods may simplify work processes in healthcare organizations, potentially reducing costs and providing more people with access to high-quality medical imaging than previously studied.[6]

Filtered Back Projection (FBP) and iterative reconstruction methods. There is a significant difference between the principles and performance properties of these two methods for reconstructing the CT Image.[7] The complexity and speed of FBP arise from the straightforward nature of the mathematical model, which consists of applying a filter to the projection data and then back-projecting it to create the image. This approach is a set of conditions under which ideal imaging is considered, making its real-world application problematic due to imperfect geometry and the lack of noiseless data.[1] The performance of FBP can also be damaged in the low-dose CT scans that have a more dominant noise, leading to streaks and poor image quality. Despite these disadvantages, FBP remains highly applicable in clinical environments due to its computational power and ability to provide immediate results, particularly in emergencies and high-throughput settings.[8]

Iterative reconstruction methods, on the other hand, provide a more advanced method of image reconstruction. These procedures are achieved through successive improvements of a pre-estimation of the original image, comparison with actual data, and comparison of results to implement corrections. In this manner, complex physical models can be incorporated to account for various aspects of the imaging system, such as detector response, beam hardening, and scattering. Also, statistical noise models can be incorporated into iterative methods, allowing noise to be more effectively suppressed and potentially enabling dose reduction without degrading image quality. The adaptability of iterative methods also allows the incorporation of prior knowledge about the object being

imaged, which further improves image quality. Better image quality, however, is obtained at the expense of greater computational complexity and leads to longer think time. This trade-off between image quality and reconstruction speed is a major consideration in the current development and optimization of iterative reconstruction algorithms for clinical use.[9]

The two approaches to CT image reconstruction, FBP and iterative techniques, each present considerable challenges that affect image quality and diagnostic accuracy. Increased sensitivity to noise might cause artifacts and poor image quality, whereas a possible loss of spatial resolution might obscure fine-detail tips, which are important for making an accurate diagnosis. The reconstruction process is also complicated by incomplete or inconsistent data, which is frequently encountered in clinical settings. These constraints have motivated scientists to investigate hybrid schemes and develop superior algorithms that methodically average both FBP and iterative schemes while addressing their drawbacks.[10]

The dynamic nature of the field and the significance of CT image reconstruction to medical imaging are evident in current developments in this area. Scientists and practitioners are working around the clock to achieve the best image quality while addressing key issues, including reducing radiation doses and improving reconstruction efficiency. Those attempts have resulted in new reconstruction algorithms that use machine learning and artificial intelligence methods, and adaptive ones that modify reconstruction parameters according to a particular imaging situation. These new techniques are set to increase diagnostic accuracy, improve the quality of patient care, and even expand the use of CT imaging in other medical areas as they continue to improve. Further advances in the field underscore the importance of image reconstruction in maximizing the benefits of CT technology and its impact on healthcare outcomes in deep medical image processing.[11]

In medical imaging, neural network architectures have been significantly improved, and each type offers distinct advantages for clinical use. Due to their ability to automatically extract hierarchical features from raw pixel data, CNNs have become a staple of image analysis. This feature enables CNNs to be especially effective in lesion detection, organ segmentation, and disease classification across different imaging modalities, including X-rays, CT scans, and MRIs. They are very useful in the analysis of 2D and 3D medical images, as their convolutional layers and pooling operations enable them to capture both local and global features, which is characteristic of CNNs.[12]

Recurrent Neural Networks (RNNs) and Generative Adversarial Networks (GANs) have expanded the potential of medical imaging beyond single-image analysis. RNNs, due to their ability to process sequential data, have been used to analyze dynamic medical images (e.g., cardiac MRI sequences) and longitudinal patient data. This dynamic nature enables one to understand the disease course and treatment response over time. On the other hand, GANs have succumbed to one of the most critical issues in medical imaging, namely, the lack of labeled data. This is because GANs can produce synthetic but realistic medical images, which, in turn, allow data augmentation strategies to enhance the resilience and generalization of other neural network models. Also, GANs have demonstrated utility in image enhancement and noise removal and can make future disease predictions using current imaging data.[13]

Medical imaging with deep learning has transformed the field, delivering levels of accuracy and efficiency never achieved before. These superior algorithms have demonstrated extraordinary capabilities in radiography, CT, MRI, ultrasound, and pathology, enabling more informed decisions and better patient outcomes. Lesion detection, organ segmentation, disease classification, and image reconstruction are other areas where deep learning has been particularly effective. As an illustration, deep learning systems have demonstrated quite impressive performance in identifying diabetic retinopathy in retinal fundus images and malignant lung nodules in chest computed tomography, and, in many cases, outperform human experts.[14]

Although the implementation of deep learning in medical image analysis has made these improvements, there are major challenges associated with the process. The most important issue is patient data privacy, since medical images are sensitive data that ought to be secured. Also, creating powerful deep learning models requires large, annotated datasets, which can be time-consuming and costly. The problem of model interpretability remains a critical concern because healthcare professionals need to understand the rationale behind AI-generated decisions to trust them and comply with regulations. Moreover, it is necessary to ensure that deep learning models can generalize across patient groups and imaging modalities to be used widely. The success of introducing deep learning technologies into clinical care largely depends on overcoming these obstacles, which can ultimately lead to higher diagnostic accuracy, more efficient workflows, and better patient care.[15]

CT reconstruction based on deep learning has transformed the field of medical imaging by enabling vast improvements in imaging and the potential to reduce radiation dose. They operate on complex neural network designs used to derive relationships between low- and high-quality CT images. CNNs have been especially successful in noise and artifact removal, leveraging their hierarchical feature extraction capabilities to exploit the data they process in images. Skip connections and encoder- decoder architecture U-Net variants are highly competitive in fine anatomy

and structural information, which is why they are highly applicable to tasks that need fine segmentation and localization.[16]

The use of GANs has become one of the most promising tools for CT reconstruction, demonstrating strong results in producing high-quality images from low-dose acquisitions. Using a generator-discriminator architecture, GANs can generate realistic, high-fidelity CT images that closely resemble standard-dose scans. Such a method not only enhances image quality but can also significantly minimize radiation exposure in patients. Recently, there have also been studies examining hybrid architectures that integrate the benefits of different types of networks, e.g., CNN-GAN or U-Net-GAN models, to further improve reconstruction quality and address specific issues in CT imaging.[17]

CT reconstruction quality is assessed using a holistic method that is both objective-based and subjective. Quantitative measures give numerical data that can be used to make standard comparisons between the various reconstruction methods. The PSNR compares the maximum possible signal power to the noise power, providing information about the overall quality of the image. The SSIM measures the structural similarity between the reconstructed and reference images, reflecting perceptual differences. The MSE measures the mean squared reconstruction error. These measures are important but not necessarily directly correlated with clinical utility, which is why a qualitative assessment must be added.[18]

Expert radiologists with expertise in qualitative assessment are important in the clinical assessment of CT reconstruction quality.[5,19] These tests are aimed at factors that directly influence diagnostic accuracy, including the clarity of anatomical structures, the appearance and severity of artifacts, and overall image quality with respect to a given diagnostic task. Task-based measurements represent the transition between quantitative and qualitative measurements, in which the performance of the reconstructed image is assessed in a clinical scenario. As an example, lesion detectability assesses whether an abnormality and its characterization can be detected, and segmentation accuracy measures the accuracy of delineating anatomical or pathological regions.[20] These decisions on which evaluation metrics to use are dependent on the reconstruction algorithm and the planned clinical use, and the unique issues relating to that anatomical region or pathology.

This research aims to assess and compare the performance of conventional and deep learning-based CT image reconstruction and enhancement techniques for improving the quality of brain-degraded CT images with Poisson and Speckle noise. The study focuses on the FBP, IR, GANs, and CNN-AE on the public dataset of CT brain stroke images.

The objective of the research is to compare quantitative image quality metrics across reconstruction methods under various noise settings and to identify the best method to improve diagnostic accuracy and support reliable clinical decision-making in medical imaging.

Materials and methods

Dataset used

Brain-stroke-CT dataset downloaded from Kaggle. It contains four folders named Normal (4428 PNG images), Ischemia (1131 PNG images), Bleeding (1094 PNG images), and external test (200 PNG images). We execute such folders using Python code with Google Colab in online mode.[21]

Methods

Firstly, there is dataset preparation for processing. The dataset contains 6,853 clean CT brain images in the folders normal, ischemia, bleeding, and external test, sourced from the Kaggle database. The images were resized and normalized to achieve uniform resolution, which is crucial for input. Then, to obtain a noisy image that simulates realistic low-quality CT images, we inject synthesis noise: Poisson and Speckle (Variance=0.04). These noisy images act as the inputs to reconstruction models (Figure 1). To train deep learning models, GAN and CNN-AE, the dataset was split into training, validation, and testing sets in a 70-15-15 ratio for each dataset folder. Then we applied image reconstruction techniques: GAN, CNN-AE, iterative, and FBP successively at this point. The fifteen captured images were used as the output, and the performance evaluation metrics were PSNR, SSIM, and MSE, calculated to assess the image quality of these models. The performance evaluation metrics of these fifteen images were averaged to achieve the average performance evaluation metrics.

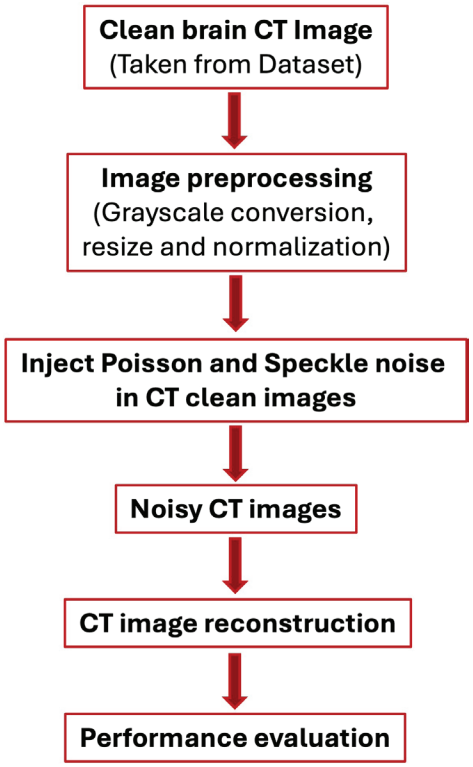


***Figure 1**. Workflow of the model.*

Proposed deep learning methodology

Convolutional Neural Network Autoencoder (CNN autoencoder)

An unsupervised machine learning method is an autoencoder, which learns an approximation of the identity function, with the output being like the input data. A CAE uses a simple autoencoder architecture and replaces the hidden layers with convolutional layers, combining the features of an autoencoder and a CNN. The objective of end-to-end training is to teach CNNs to learn feature representations for classification, whereas CAEs are trained only to learn feature representations of the data that can be used to reconstruct the input. A CNN-AE is a deep learning model that learns compact feature representations and reconstructs images by minimizing reconstruction error. In this work, we use a CNN-AE to remove noise and reconstruct high-quality brain CT images from noisy input images. The CNN-AE has two components.[22]

1. Encoder

The encoder converts the input image into a highly compressed representation using convolutional and max-pooling layers. Feature maps of the convolutional layers in the encoder block were expressed as

$$F^{(i,j)}_{encoder} = \sigma(W^{(j)}_{encoder} F^{(i-1,\ \forall j)}_{encoder} + b^{(j)}_{encoder}) \quad (1)$$

Where ; $F^{(i,j)}_{encoder}$ output feature map, σ represents rectified linear unit (ReLU); $b^{(j)}_{encoder}$ the bias for the jth feature map of the ith layer, and $W^{(j)}_{encoder}$ the jth filter or kernel of the ith layer.

2. Decoder

The decoder is used to recreate the original image using deconvolution and up-sampling of the outputs in low-dimensional feature maps of decoder block. Feature maps of the convolutional layer in the decoder block were expressed as

$$F^{(i,j)}_{decoder} = \sigma(W^{(j)}_{decoder} F^{(i-1,\ \forall j)}_{decoder} + b^{(j)}_{decoder}) \quad (2)$$

Where, $b^{(j)}_{decoder}$ represents bias for the jth feature map of the ith layer; and $W^{(j)}_{decoder}$ is the jth filter or kernel of the ith layer.

Generative Adversarial Network (GAN)

GANs produce fake data that, in theory, cannot be distinguished from real data. They use two parallel networks: the generator (G) and the discriminator (D), pitted against each other in a minimax game. G produces new data, which dread reads, and dread must decide whether the data it has received is synthetic (generated by G) or real. GAN discriminator (D) is trained by using a set of real images, whereas generator(G), which did not consider the training dataset, is forced to create images that resemble real ones mostly, based on the feedback that it gets from discriminator (D), which knows what real images look like. Simultaneously, as G produces images, D would be continually trained to distinguish between real and generated (fake) images. This is done repeatedly; thus, the two networks keep improving their performance, and the aim is to produce images that are not recognizable as the originals. By doing so, the generator will be trained to generate

increasingly realistic images. GANs are widely used in medical imaging through either their generative or discriminative roles. Figure 2 shows the functioning of a GANs.[23]

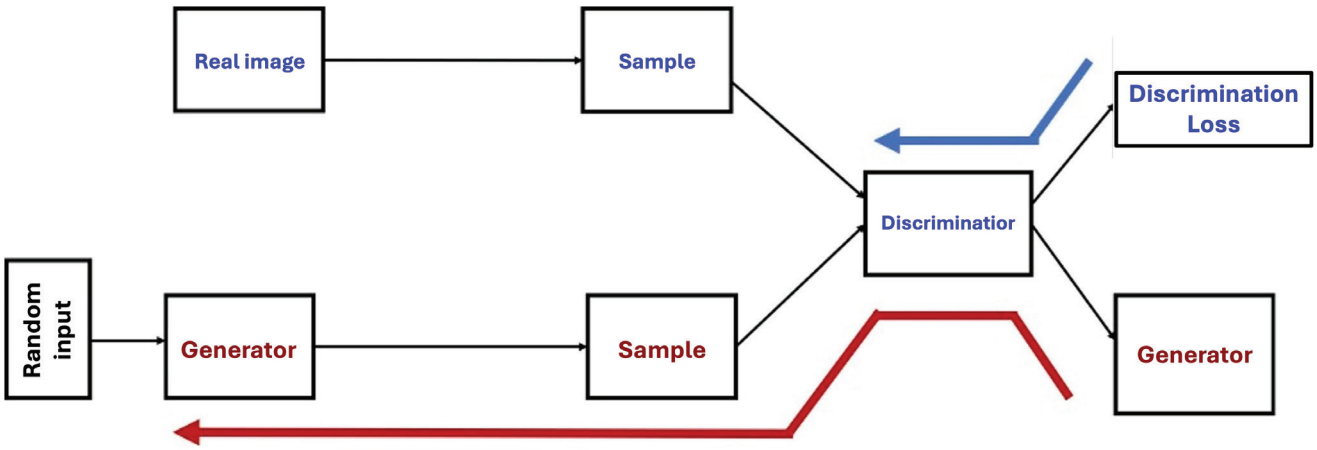


Figure 2. *Functioning of GANs. The blue arrow shows the discriminators' loss backpropagation, and the red arrow shows the generator's loss backpropagation. The generator creates synthetic images from random noise, while the discriminator must distinguish between real and synthetic images.*

Training process

One of the most important parts of training a machine learning model is choosing the loss function and optimization algorithm. The loss function measures the gap between the actual and predicted values and guides the model's learning. Regular loss functions are mean squared error for regression and cross-entropy for classification. A loss-minimizing algorithm, based on an optimization method (e.g., stochastic gradient descent [SGD] or Adam), determines how to update the model parameters. The choice of an appropriate combination of loss function and optimizer is also crucial for efficient, effective model training. Hyperparameter tuning and model selection involve selecting the most suitable combination of the model architecture and training parameters.

Methods that commonly involve this process and operate on different combinations of hyperparameters, such as learning rate, batch size, and network depth, include grid search, random search, and Bayesian optimization. In model selection, various architectures or algorithms are compared to identify the best-performing model for a given task. The training plan includes methods such as transfer learning, in which pre-trained models are fine-tuned to perform new tasks, and fine-tuning, in which a pre-trained model is further trained on task-related data. The strategies may greatly enhance model performance, especially when using small datasets or complicated tasks. The hyperparameters are a learning rate of 0.001 for the CNN autoencoder and 0.0001 for the GAN model. The batch size is 16 (Figure 1).

Evaluation metrics

Quantitative measures of reconstruction quality are objective indicators of the accuracy and fidelity of the reconstructed data or images. Common metrics include the mean squared error (MSE), peak signal -to-noise ratio (PSNR), and structural similarity index (SSIM). These measures compare the reconstructed output with the input, measure disparities in pixel values, provide a general comparison, and assess the conservation of structural information. Besides that, task-related metrics, e.g., segmentation error or object detection rate, may be used to estimate the usefulness of the reconstructed information in practice.

Qualitative assessment procedures involve subjective evaluation by human observers. Such a methodology considers aspects that are not easily measured numerically, including visual appeal, perceptual quality, and the availability of artifacts in the picture. The qualitative evaluations are usually based on ratings, preference studies, or expert assessments to determine the overall quality and usability of the reconstructed data. Comparison with baseline methods is important for contextualizing the performance of new reconstruction techniques. This involves applying and testing the pre-existing algorithms and the suggested method on the same data, using the same assessment criteria. These comparisons help demonstrate improvements in reconstruction quality, computational efficiency, and the new method's applicability to specific domains, providing a comprehensive view of its advantages and constraints.

Results

The experimental results show that a CNN with an autoencoder provides better reconstruction capability. Statistical analysis is conducted to interpret and assess the significance of the data. This was done by using statistical tests to establish the reliability and validity of the results. The methods researchers can use to evaluate the statistical significance of

their findings include hypothesis testing, confidence intervals, and analysis of variance (ANOVA). Besides, they can investigate relationships among variables, identify trends, and assess measurement uncertainty. It is important to compare state-of-the-art methods to situate research within the current body of knowledge. This is done by comparing the proposed approach with known methods and highlighting advancements or new contributions. Comparative analysis is presented with performance indicators side by side, underscoring where the approach is superior or provides special benefits, as shown in Figure 3 and Figure 4.

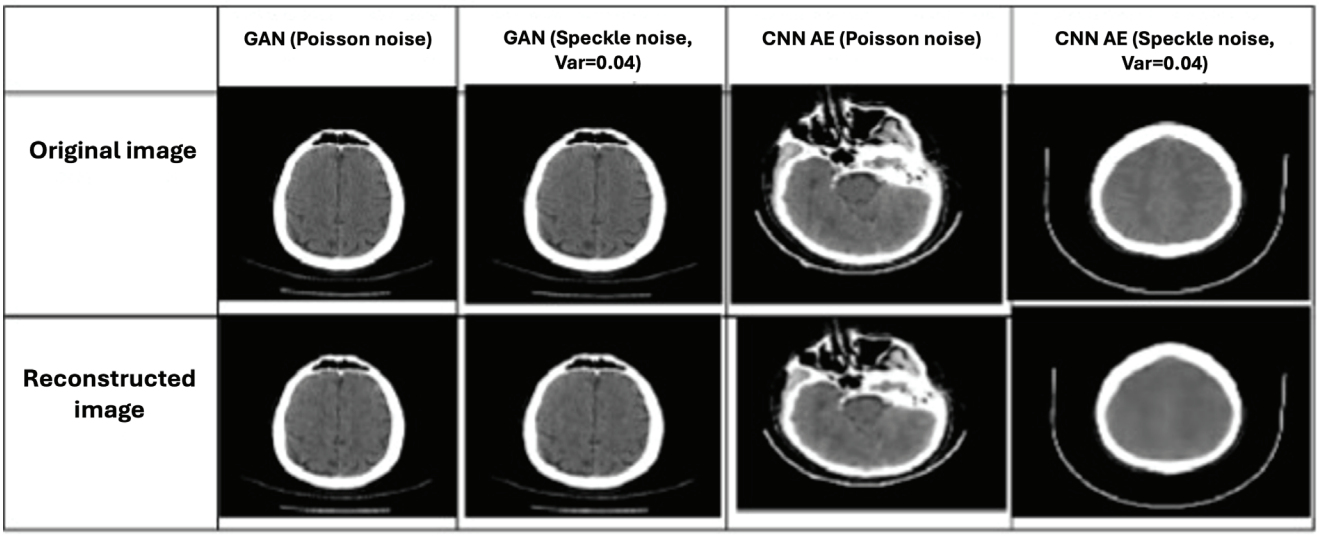


Figure 3. *Comparative figure of GAN and CNN autoencoder method, original image vs reconstructed image.*

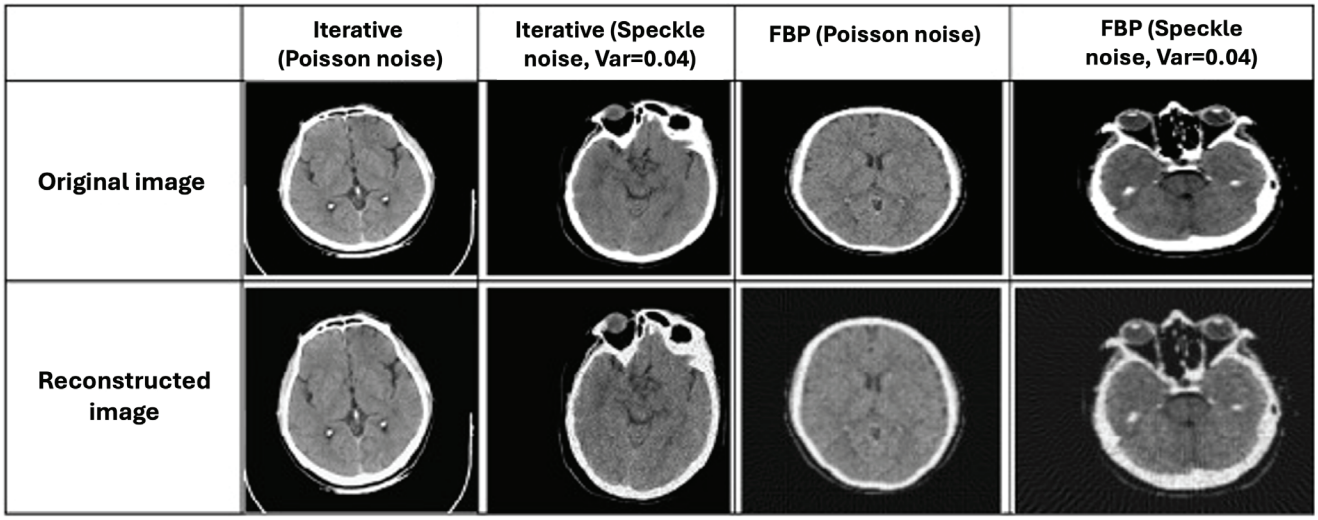


Figure 4. *Comparative figure of FBP and the iterative method, original image vs reconstructed image.*

The parametric evaluation of the results is presented in the presence of GAN and CNN-AE; FBP and iterative methods with Poisson and Speckle noise are reported in Table 1 and Table 2.

Figure 5-Figure10 present the values of MSE PSNR of different methods in presence of Poisson noise and Speckle noise.

Table 1. Average PSNR, SSIM, and MSE value of 15 CT images using GAN and CNN autoencoder, FBP, and iterative methods with Poisson and Speckle noise.

Noise type	Dataset folder	Method	PSNR (dB)↑	SSIM↑	MSE↓
Speckle (Var=0.04)	Normal	FBP	23.25	0.634	0.0049
		Iterative	27.77	0.890	0.0017
		GAN	**32.54**	**0.938**	**0.0006**
		CNN-AE	29.02	0.926	0.0016
	Ischemia	FBP	23.71	0.651	0.0045
		Iterative	28.88	0.917	0.0013
		GAN	**32.17**	**0.937**	**0.0006**
		CNN-AE	26.46	0.897	0.0028
	Bleeding	FBP	23.19	0.644	0.0050
		Iterative	27.45	0.899	0.0019
		GAN	**31.25**	**0.938**	**0.0008**
		CNN-AE	26.87	0.913	0.0024
	External test	FBP	22.66	0.625	0.0057
		Iterative	**28.11**	**0.897**	**0.0016**
		GAN	27.66	0.870	0.0018
		CNN-AE	24.54	0.864	0.0042
Poisson	Normal	FBP	25.10	0.722	0.0033
		Iterative	**37.00**	**0.967**	**0.0002**
		GAN	32.88	0.937	0.0006
		CNN-AE	30.84	0.947	0.0011
	Ischemia	FBP	25.54	0.730	0.0031
		Iterative	**37.34**	**0.972**	**0.0002**
		GAN	31.43	0.929	0.0008
		CNN-AE	30.06	0.941	0.0012
	Bleeding	FBP	24.90	0.723	0.0035
		Iterative	**36.12**	**0.970**	**0.0003**
		GAN	31.76	0.937	0.0007
		CNN-AE	29.71	0.952	0.0012
	External test	FBP	24.29	0.701	0.0041
		Iterative	**37.10**	**0.970**	**0.0002**
		GAN	28.16	0.843	0.0016
		CNN-AE	30.37	0.947	0.0011

Table 2. Comparison of the proposed model with existing models.

Author	Method	PSNR (dB)	SSIM	MSE
Runkel C, *et al.*[25]	FBP	16.81	0.0602	-
Xiaobao Liu [9]	GAN	28.34	0.945	-
Qing Li, *et al.* [24]	CNN-AE	29.91	0.909	-
Proposed	FBP	25.10	0.722	0.0033
	Iterative	37.34	0.972	0.0002
	CNN-AE	30.37	0.947	0.0011
	GAN	32.54	0.938	0.0006

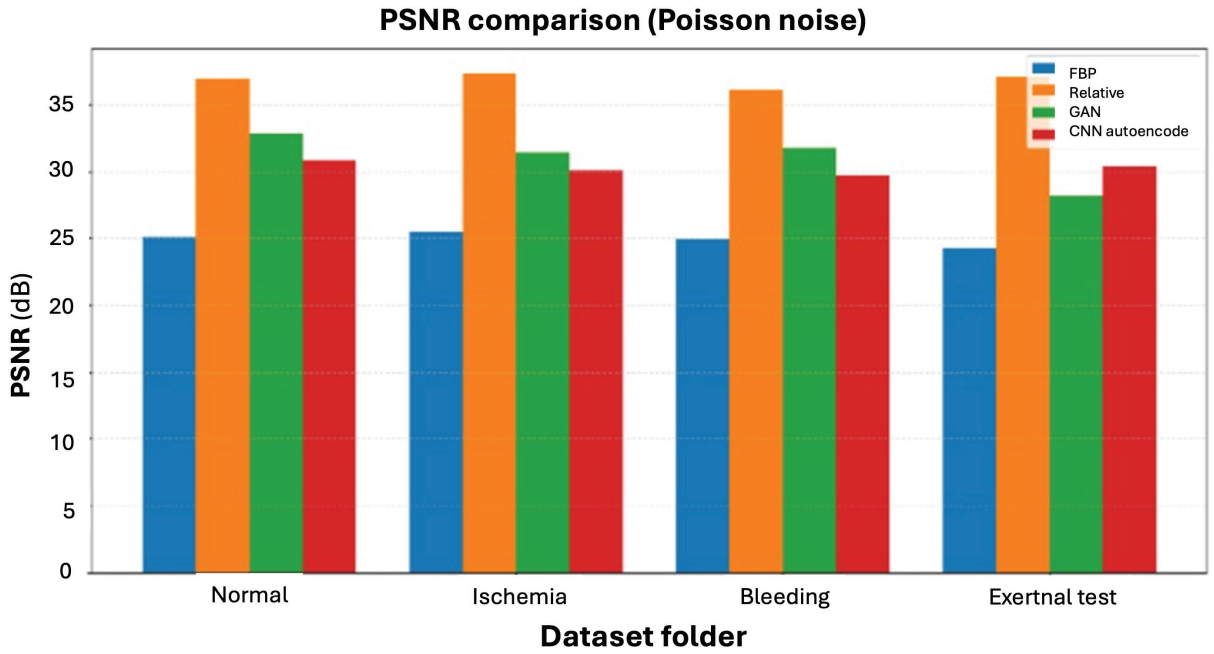


Figure 5. *Comparison of PSNR of GAN and CNN-AE (with Poisson noise).*

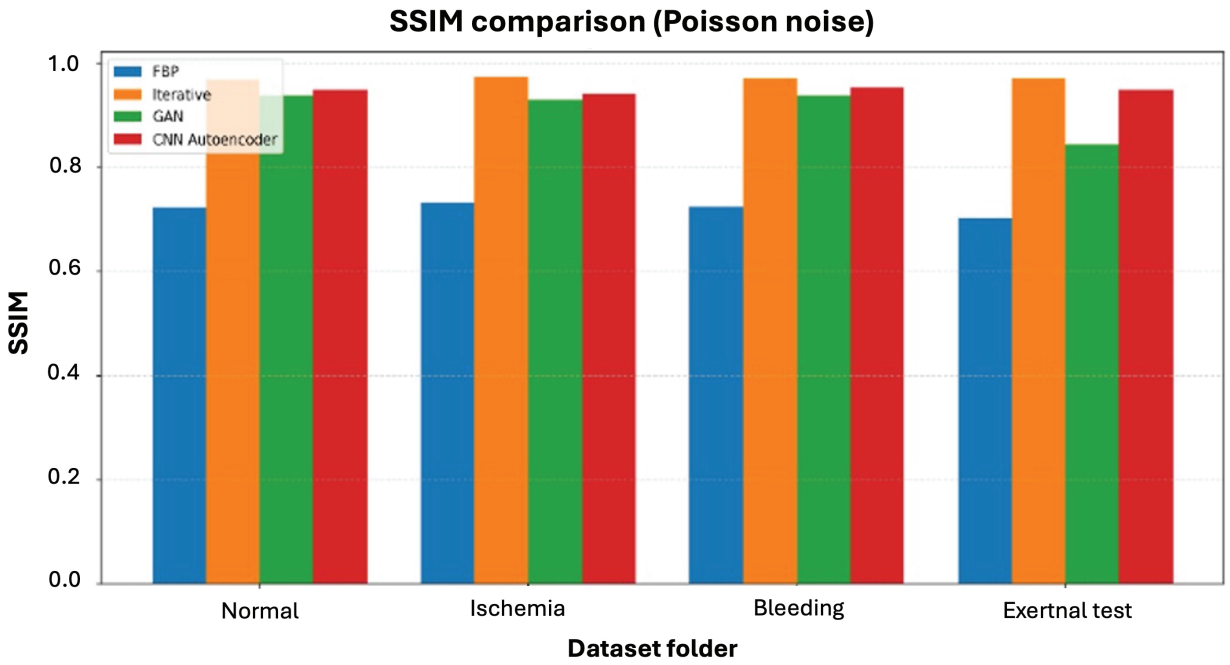


Figure 6. *Comparison of SSIM of FBP, iterative, GAN, and CNN-AE (with Poisson noise).*

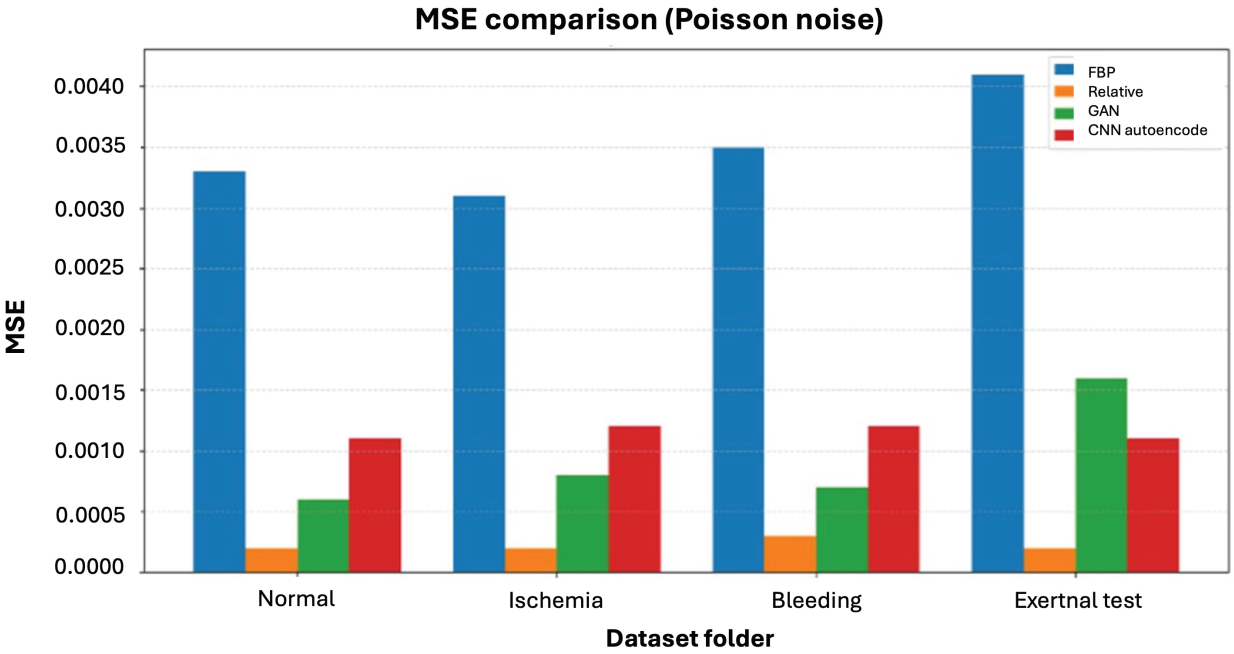


Figure 7. *Comparison of MSE of FBP, Iterative, GAN, and CNN-AE (with Poisson noise).*

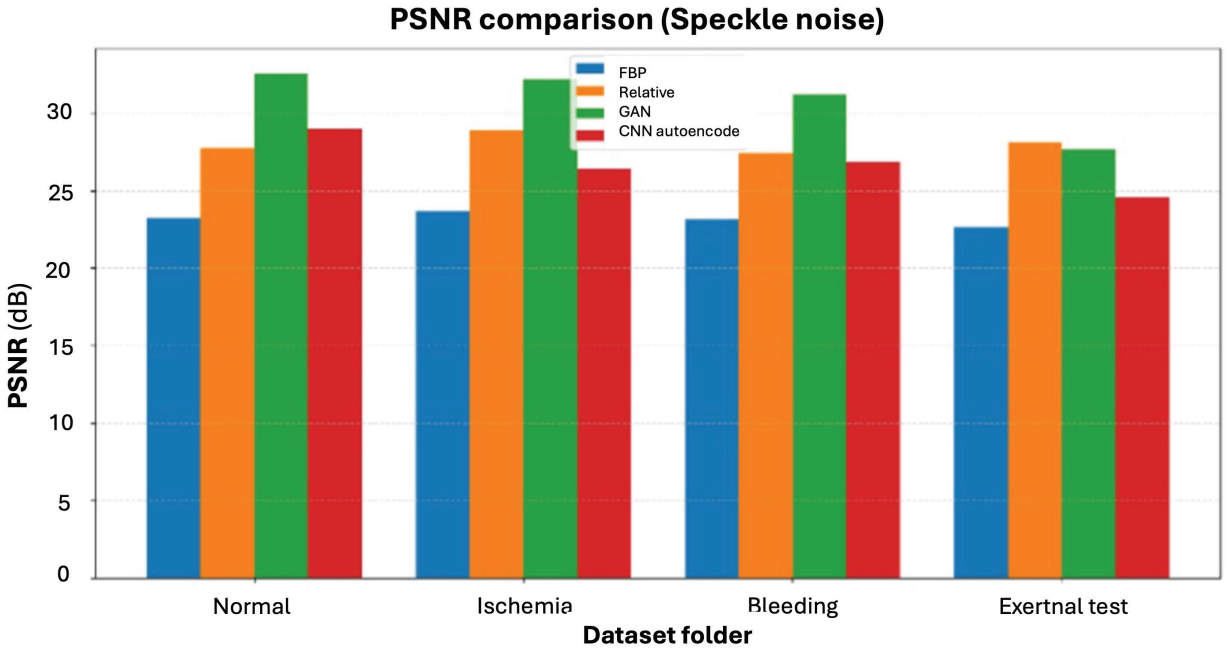


Figure 8. *Comparison of PSNR of FBP, Iterative, GAN, and CNN-AE (with Speckle noise).*

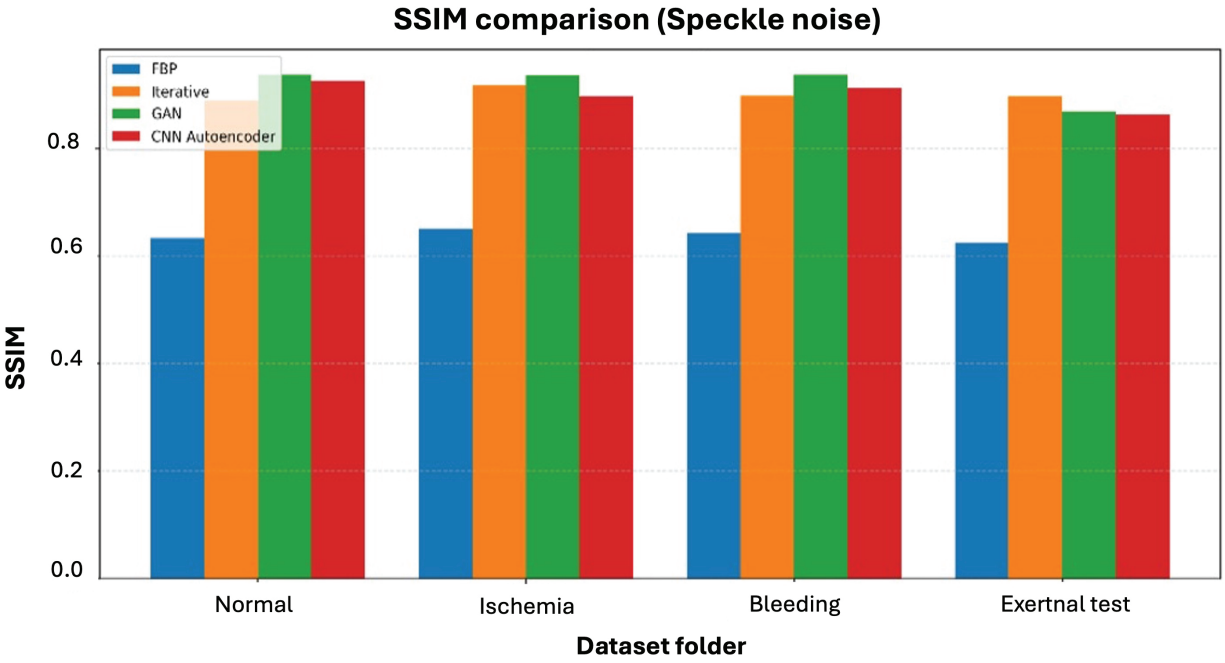


Figure 9. *Comparison of SSIM of FBP, Iterative, GAN, and CNN-AE (with Speckle noise).*

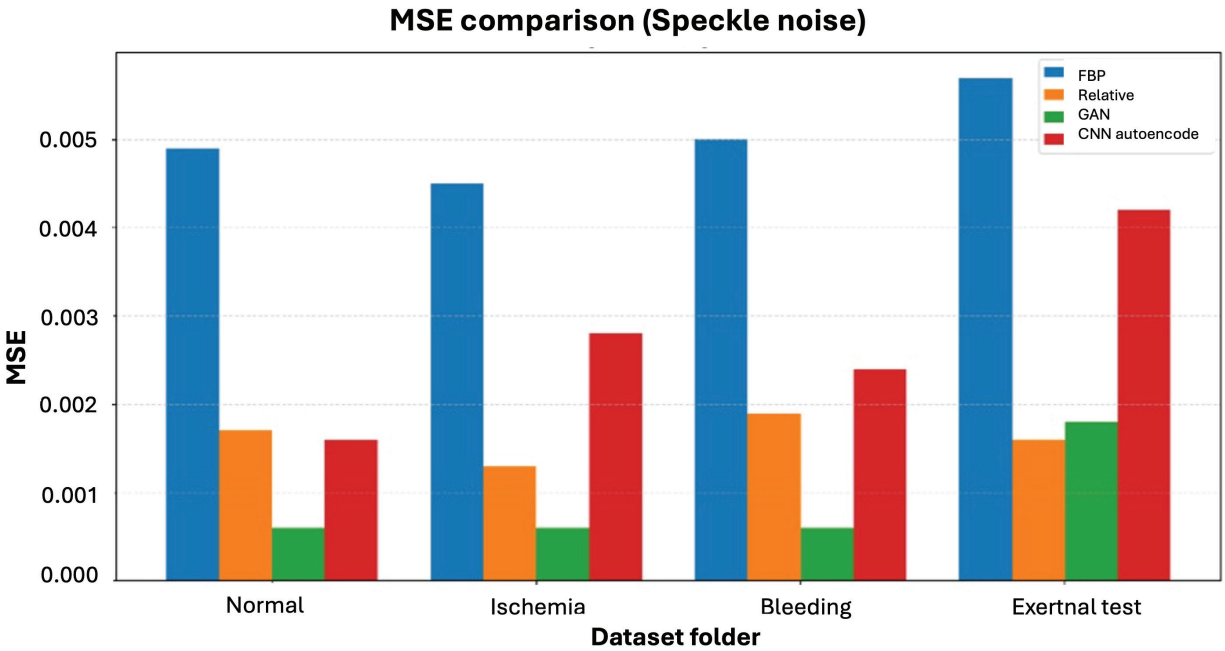


Figure 10. *Comparison of MSE of FBP, Iterative, GAN, and CNN-AE (with Speckle noise).*

The experimental evaluation is indicated in Table 1. We have compared the results of FBP, iterative, GAN, and CNN-AE-based reconstruction on two noise models, i.e., Poisson and Speckle. The assessment metrics used are peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), and mean squared error (MSE) across four dataset folders, i.e., normal, ischemia, external-test, and bleeding.

Analysis of the results

Table 1 shows the results of four CT image reconstruction methods: FBP, IR, GAN, and CNN-AE, with the addition of two types of noise, Speckle and Poisson, in four dataset folders named normal, ischemia, bleeding, and external test folder. The performance parameters evaluated are PSNR, SSIM, and MSE.

1. Performance on Speckle noise

When applying Speckle noise to the four CT image datasets, the GAN approach yields the highest-quality images across all datasets. In the case of the Normal dataset folder, the GAN achieves the highest PSNR of 32.54 dB and SSIM of 0.938, while also having the lowest MSE of 0.0006, indicating that its reconstruction is most like other techniques. However, FBP, a conventional reconstruction approach, performed the worst, achieving a PSNR of 23.71 dB and an SSIM of 0.650. This is also true in the Ischemia dataset folder, where the GAN method achieved a PSNR of 32.17 dB, an SSIM of 0.937, and an MSE of 0.0006, which are better than those of the Iterative and CNN Autoencoder methods. The CNN autoencoder is more successful than the iterative method yet still lower than the GAN approach.

In the bleeding dataset folder, the method achieves the best reconstruction with a PSNR of 31.25 dB and an SSIM of 0.938, whereas the FBP method yields the worst PSNR and the highest MSE. For the external test dataset folder, GAN performance was reduced, yielding PSNR 27.66 dB and SSIM 0.870, indicating reduced generalization to unseen data.

2. Performance on Poisson noise

When Poisson noise is applied to the dataset folder, the iterative method yields the best PSNR of 37.00 dB and SSIM of 0.967.

The CNN-AE and GAN algorithms yielded higher image quality than FBP, but the iterative method performed best in the presence of Poisson noise. CNN-AE produced higher SSIM values compared to GAN on the external dataset, for CNN-AE (0.947) and GAN (0.843), indicating better preservation of structural information on unseen data. FBP again yields the poor performance, with PSNR values 25.54.

Based on the above results, we suggest that reconstruction performance depends on the type of noise.

Discussion

There are differences in the number of CT images across the folders (normal, ischemia, bleeding, and external test). With 4,428 images, more than the ischemia folder (1,131 images), the bleeding folder (1,094 images), and the external test folder (200 CT images). This imbalance may bias model training, which could affect the generalization capacity of reconstruction models. To overcome this, training, validation, and testing were conducted independently for each data folder. Image quality was then assessed using the performance metrics of PSNR, SSIM, and MSE.

Limitation

The brain-stroke-CT dataset was taken from Kaggle. It's not real clinical data, and it does not have the same number of images. Therefore, the experimental results obtained using the above methods may differ. There was synthetic noise (Speckle and Poisson noises), but real-world CT images can have complex noise and artifacts. There was no emphasis on clinical diagnosis or radiologist assessment; only image quality metrics were assessed. These limitations should be addressed in future research.

Conclusion

This paper shows that GAN, iterative, and CNN-AE approaches substantially outperform the conventional approach, FBP, in all datasets and noise addition. When adding Speckle noise, the GAN yields the highest performance matrices (PSNR=32.54, SSIM=0.938, MSE=0.0006). For the external dataset with Speckle noise, IR gives the best performance, but slightly higher than GAN. Under Poisson noise, the IR technique yields the best performance on all dataset folders. GAN and CNN-AE yield improved image quality compared with the traditional method, FBP, but their performance is below that of the Iterative method, even with the addition of Poisson noise.

Finally, conclude that IR is most effective for Poisson-noise-corrupted CT images. Whereas the GAN reconstruction method provides the best performance for Speckle noise addition. Thus, it demonstrates the potential of deep learning and Iterative methods for enhancing image quality and removing noise while preserving anatomical structures.

Ethical approval

This research did not involve any human participants, human data, human tissue, or animal experiments. Therefore, ethical approval was not required.

Funding

The authors declare that no financial support or funding was received from any funding agency,

commercial organization, or not-for-profit organization for the conduct of this research.

Conflict of interest

The authors declare no conflicts of interest regarding this manuscript. No commercial or financial relationships were present in the study that could be considered a possible conflict of interest. The authors also state that there is no conflict of interest in this work. There are no personal, professional, or financial associations that affected the behavior or the results of this study.

CRediT authorship contribution statement

Rita Yadav: conceptualization, investigation, data curation, software, formal analysis, validation, visualization, writing: original draft; **Rajesh Kumar**: research idea, formulated the problem, conceptualization, methodology, supervision, project administration, writing: critically review and edit; **Kavita Singh**: visualization, data curation, writing: manuscript draft, review and edit.

Acknowledgements

The authors sincerely thank Prof. NK Shukla, Head of the Department, Dr. Sudhanshu Kumar Jha, and Mr. Shyam Babu for their valuable support and assistance, directly and indirectly, during the preparation of this manuscript.

References

[1] Zhang M, Shi Y, & Gu S. The use of deep learning methods in low-dose computed tomography image reconstruction: a systematic review. Complex Intell Syst. 2022; 8(6): 5545-61. doi: 10.1007/s40747-022-00724-7.

[2] Fortuni F, Franchin L, Gabrielli D, Vitale E, Conte E, Spinelli A, et al. Advancements and applications of artificial intelligence in cardiovascular imaging: a comprehensive review. Eur Heart J Imaging Methods Pract. 2024; 2(4). doi: 10.1093/ehjimp/qyae136.

[3] Li Z, Tang C, Yan B, Cai A, Zhang W, Liang N, et al. Promising generative adversarial network based sinogram inpainting method for ultra-limited-angle computed tomography imaging. Sensors. 2019; 19(18): 3941. doi: 10.3390/s19183941.

[4] Immonen E, Wong J, Nieminen M, Kekkonen L, Roine S, Törnroos S, et al. The use of deep learning towards dose optimization in low-dose computed tomography: a scoping review. Radiography (Lond). 2021; 28(1): 208-14. doi: 10.1016/j.radi.2021.07.010.

[5] Mese I, Demirsoy U, AltintasMese C, & Anik Y. Innovative advances in pediatric radiology: computed tomography reconstruction techniques, photon-counting detector computed tomography, and beyond. Pediatr Radiol. 2023; 54(1): 1-11. doi: 10.1007/s00247-023-05823-2.

[6] Mitra U, & Rehman SU. Leveraging AI and machine learning for next-generation clinical decision support systems (CDSS; pp 83-112). IGI Global. doi: 10.4018/979-8-3693-7277-7.ch003.

[7] Southard RN, Martinot CA, Temkit MH, Augustyn RA, Bardo DME, & Thorkelson MA. Comparison of iterative model reconstruction versus filtered back-projection in pediatric emergency head CT: dose, image quality, and image-reconstruction times. AJNR Am J Neuroradiol. 2019; 40(5): 866-71. doi: 10.3174/ajnr.a6034.

[8] Pontana F, Faivre JB, Remy-Jardin M, Hachulla AL, Duhamel A, Pagniez J, et al. Chest computed tomography using iterative reconstruction vs filtered back projection (Part 2): image quality of low-dose CT examinations in 80 patients. Eur Radiol. 2010; 21(3): 636-43. doi: 10.1007/s00330-010-1991-4.

[9] Shi L, Karellas A, Vedantham S, & Zhu L. Library based x-ray scatter correction for dedicated cone beam breast CT. Med Phys. 2016; 43(8): 4529-44. doi: 10.1118/1.4955121.

[10] Terzis R, Reimer RP, Nelles C, Celik E, Caldeira L, Heidenreich A, et al. Deep-learning-based image denoising in imaging of urolithiasis: assessment of image quality and comparison to state-of-the-art iterative reconstructions. Diagnostics. 2023; 13(17): 2821. doi: 10.3390/diagnostics13172821.

[11] Sadia RT, Zhang J, & Chen J. CT image denoising methods for image quality improvement and radiation dose reduction. J Appl Clin Med Phys. 2024; 25(2). doi: 10.1002/acm2.14270.

[12] Radiuk P. Applying 3D U-Net architecture to the task of multi-organ segmentation in computed tomography. Appl Comput Syst. 2020; 25(1): 43-50. doi: 10.2478/acss-2020-0005.

[13] Han C, Furukawa Y, Nakayama H, Hayashi H, Araki R, Rundo L, et al. Infinite brain MR images: PGGAN-based data augmentation for tumor detection (Vol. 151; pp 291-303). Springer Singapore. doi: 10.1007/978-981-13-8950-4_27.

[14] Lavudiya NS, & Prasad CVPR. Enhancing ophthalmological diagnoses: an adaptive ensemble learning approach using fundus and OCT imaging. Int J Comput Exp Sci Eng. 2024; 10(4). doi: 10.22399/ijcesen.678.

[15] Hasan AM, Kahtan H, Al-Ahmad AS, Jalab HA, & Meziane F. Combining deep and handcrafted image features for MRI brain scan classification. IEEE Access. 2019; 7: 79959-67. doi: 10.1109/access.2019.292269.

[16] Rahman H, Khan AR, Sadiq T, Farooqi AH, Khan IU, & Lim WH. A systematic literature review of 3D deep learning techniques in computed tomography reconstruction. Tomography. 2023; 9(6): 2158-89. doi: 10.3390/tomography9060169.

[17] Liu X, Gu W, Yao T, Su S, Mo Y, & Shen J. Super-resolution reconstruction of CT images based on multi-scale information fused generative

adversarial networks. Ann Biomed Eng. 2023; 52(1): 57-70. doi: 10.1007/s10439-023-03412-w.

[18] Ohashi K, Nagatani Y, Yoshigoe M, Iwai K, Tsuchiya K, Hino A, et al. Applicability evaluation of full-reference image quality assessment methods for computed tomography images. J Digit Imaging. 2023; 36(6): 2623-34. doi: 10.1007/s10278-023-00875-0.

[19] Bornet PA, Blum A, Germain E, Gondim Teixeira PA, Lombard C, Villani N, et al. Clinical acceptance of deep learning reconstruction for abdominal CT imaging: objective and subjective image quality and low-contrast detectability assessment. Eur Radiol. 2022; 32(5): 3161-72. doi: 10.1007/s00330-021-08410-x.

[20] Parakh A, Cao J, Kambadakone AR, Blake MA, Pierce TT, & Savage CA. Sinogram-based deep learning image reconstruction technique in abdominal CT: image quality considerations. Eur Radiol. 2021; 31(11): 8342-53. doi: 10.1007/s00330-021-07952-4.

[21] Koç U, Akçapınar Sezer E, Alper Özkaya Y, et al. Artificial intelligence in healthcare competition (TEKNOFEST-2021): stroke data set. Eurasian J Med. 2022; 54(3): 248-58.

[22] Nageswara Rao B, Acharya UR, Tan RS, Dash P, Mohapatra M, Sabut S. Convolutional autoencoder-based deep learning for intracerebral hemorrhage classification using brain CT images. Cogn Neurodyn. 2025 Dec; 19(1): 77.

[23] Laino ME, Cancian P, Politi LS, Della Porta MG, Saba L, Savevski V. Generative adversarial networks in brain imaging: a narrative review. J Imaging. 2022 Mar 23; 8(4): 83.

[24] Li Q, Li S, Li R, Wu W, Dong Y, Zhao J, et al. Low-dose computed tomography image reconstruction via a multistage convolutional neural network with autoencoder perceptual loss network. Quant Imaging Med Surg. 2022 Mar; 12(3): 1929.

[25] Runkel C, Biguri A, Schönlieb CB. Continuous learned primal dual. In: 2024 IEEE 34th International Workshop on Machine Learning for Signal Processing (MLSP). 2024 Sep 22; pp 1-6. IEEE.

Journal of Associated Medical Sciences 2026; 59(3): 238-246

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



Proof-of-concept development of a cost-effective PACE-based assay for *CYP2C19* genotyping

Chonticha Sirikul, Auttachai Saejia, Nampeung Anukul*

Division of Transfusion Science, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 18 February 2026
Accepted as revised 4 July 2026
Available online 13 July 2026

Keywords:
Health, pharmacogenomics, drug metabolism, PACE-PCR, CYP2C19

ABSTRACT

Background: Pharmacogenomic testing of the *CYP2C19* gene is a fundamental component of precision medicine, enabling personalized dose adjustments for critical medications, including antiplatelet agents and antidepressants. However, the high cost of currently available commercial genotyping kits creates a financial burden that limits patient accessibility to these essential tests.

Objectives: This study aimed to preliminarily develop and analytically evaluate a novel PCR Allelic Competitive Extension (PACE) assay for genotyping *CYP2C19*2*, **3*, and **17*.

Materials and methods: Thirty DNA samples extracted from buffy coat were analyzed using specifically designed allele-specific primers for *CYP2C19*2*, **3*, and **17*. The genotyping results were reported as zygosity status and evaluated for concordance with the TaqMan™ Drug Metabolism Genotyping Assay.

Results: The analysis identified *1/*1 (53.3%), *1/*2 (30.0%), *1/*3 (10.0%), and *2/*2 (6.7%) genotypes, while the allele frequencies were 73.3% for *1, 21.7% for *2, 5.0% for *3, and 0% for *17. The results from the PACE technique demonstrated 100% concordance with those obtained from the commercial TaqMan™ assay and produced clear, well-separated data points in the allelic discrimination plot analysis.

Conclusion: The novel PACE technique provides high accuracy, equivalent to commercially available test kits, while offering a more cost-effective approach, with a reagent cost of only USD 4.34 (135 Thai baht) per sample for *CYP2C19*2*, **3,* and **17* genotyping. It represents an alternative to CYP2C19 genotyping, supporting the advancement of personalized medicine and reducing the long-term financial burden on healthcare.

Introduction

Pharmacogenomic testing has emerged as a fundamental component of precision medicine, enabling the integration of individual genetic variability into predictions of drug response, therapeutic efficacy, and the risk of adverse drug reactions. Genetic polymorphisms, particularly in the cytochrome P450 (CYP450) gene family, which is involved in drug metabolism, transport, and pharmacological targets, substantially affect drug pharmacokinetics and significantly influence plasma concentrations and clinical outcomes. The implementation of

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: 10.66285/JAMS.2026.097
E-ISSN: 2539-6056

pharmacogenomics into clinical practice supports the selection of appropriate medications and individualized dose adjustments tailored to each patient's genetic profile. This personalized therapy enhances treatment effectiveness, reduces the incidence of adverse drug reactions and drug-related toxicity. Moreover, integrating pharmacogenomics-guided strategies into healthcare systems has the potential to reduce long-term healthcare costs by reducing inappropriate medication use and preventing avoidable drug-related complications. Therefore, pharmacogenomics testing is considered a critical tool for supporting clinical decision-making, enhancing medication safety, and advancing the delivery of personalized patient care.

Cytochrome P450 (CYP450) enzymes are encoded by *CYP* genes and are responsible for drug metabolism. These enzymes play a critical role in catalyzing chemical modifications of pharmaceutical compounds, thereby facilitating their elimination from the body. Genetic variability within *CYP* genes significantly influences enzyme activity and contributes to individual differences in drug response.[1] Based on variations in the *CYP* gene, patients can be categorized into four major phenotypes. There are poor metabolizers (PM), intermediate metabolizers (IM), extensive metabolizers (EM), and ultra-rapid metabolizers (UM), each associated with distinct drug-response risks. PM usually contains homozygous null-functional or deletion of *CYP* alleles, exhibiting elevated plasma drug concentrations, leading to an increased risk of drug toxicity and adverse drug reactions. In contrast, UM with two or more functional *CYP* alleles may experience subtherapeutic drug levels, resulting in diminished treatment efficacy and potential therapeutic failure.[2]

CYP2C19 is a major enzyme in the cytochrome P450 family encoded by the *CYP2C19* gene and is essential for the metabolic activation of prodrug antiplatelet agents, such as clopidogrel, antidepressant drugs, and proton pump inhibitors.[3,4] The distribution of *CYP2C19* polymorphisms varies among ancestral and geographic regions, resulting in inter-ethnic differences in drug metabolism phenotypes and pharmacological response. Loss-of-function alleles, especially *CYP2C19*2* and **3*, are more frequent in Asian populations than in European and African populations, leading to higher frequencies of intermediate- and poor-metabolizer phenotypes.[5,6] In contrast, the gain-of-function *CYP2C19*17* allele is common among European and Middle Eastern populations and associated with rapid and ultrarapid metabolizer phenotypes.[5,6] These genetic differences specific to ethnicity have important implications in pharmacogenomics-guided therapy, especially for drugs metabolized through the CYP2C19 pathway.

Among the Thai population, the *CYP2C19* allele frequencies were 25.60%, 2.50%, and 1.80% for the *CYP2C19*2*, *CYP2C19*3*, and *CYP2C19*17* alleles, respectively.[4,7] The importance of *CYP2C19* genotyping lies in its ability to predict varied drug responses and optimize therapeutic outcomes. *The CYP2C19*1/*1* genotype is the most prevalent, corresponding to normal metabolizers (NM) that show a normal drug response. Two *CYP2C19*2* or *CYP2C19*3* alleles result in the PM phenotype, which significantly reduces or abolishes CYP2C19 activity, diminishing drug response and possibly leading to drug toxicity. While the *CYP2C19*17* allele is associated with ultra-rapid metabolizer (UM) status due to increased enzyme activity, the drug dose may need to be increased to achieve adequate therapeutic efficacy.[8,9] Therefore, genotyping of clinically relevant *CYP2C19* alleles is essential for tailoring personalized dosing, which prevents drug toxicity in poor metabolizers and therapeutic failure in ultra-rapid metabolizers. PCR-based techniques, such as real-time PCR assays, have been widely used to provide rapid and accurate detection of star alleles (*2, *3, *17) for clinical pharmacogenomic applications. Validation studies have demonstrated concordance with DNA sequencing, the analytical gold standard.[10,11] However, currently available commercial *CYP2C19* genotyping kits remain relatively expensive, limiting access for patients who benefit from pharmacogenomic testing due to the associated financial burden. To address this limitation, this study develops an alternative genotyping approach that would reduce testing costs and improve patient access to personalized treatment. This was achieved through the design and implementation of a *CYP2C19* genotyping assay using the PCR Allelic Competitive Extension (PACE) technique.

PACE is an advanced genotyping technology designed for the highly accurate detection of DNA sequence variations. This method utilizes unlabeled allele-specific primers in combination with a universal fluorescent reporting cassette, and the result indicates zygosity status; this technique enables cost-effective, high-throughput genotyping. The PACE platform has been widely used in genetic research, molecular breeding, and pathogen detection due to its high sensitivity, specificity, scalability, and operational efficiency.[12-14] Although this technique has not been reported for use in human gene variation analysis, particularly for single-nucleotide polymorphism (SNP) detection, previous studies have demonstrated successful development of SNP assay designs for plant disease detection, with consistent experimental performance.

Materials and methods

Study design and sample recruitment

This study aims to develop a *CYP2C19* genotyping assay based on the PCR Allelic Competitive Extension (PACE) technique and to compare the results with the commercial *CYP2C19* genotyping kits (TaqMan Drug Metabolism Genotyping Assays, Applied Biosystems™, USA). Thirty DNA samples were obtained from residual clinical buffy coat specimens from patients

archived in the Molecular Biology Laboratory, Division of Transfusion Science, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University. These residual specimens were originally collected for pharmacogenomics routine testing in this laboratory and were subsequently used for research under institutional ethical approval.

DNA extraction and preparation

DNA was extracted from the buffy coat using a QIAamp® Blood Mini Kit (QIAGEN, Germany). Purity and dsDNA concentration were measured by a Qubit fluorometer (Invitrogen, USA). Then, DNA concentration was adjusted to 4 ng/µL and stored at -20 °C until further analysis.

Primer and control template design for PACE and modification

The primer sequences were designed as described in Table 1. Two allele-specific forward primers, each terminating with a nucleotide specific to a given allele, and one common reverse primer were designed based on the rs identifiers rs4244285, rs4986893, and rs12248560, corresponding to the *CYP2C19*2, CYP2C19**3, and *CYP2C19*17* alleles, respectively. In addition, approximately 500 bp DNA templates representing wild-type and mutant alleles were synthesized based on the gene *CYP2C19* sequence (Homo sapiens, chromosome 10, GRCh38.p14, NCBI Reference Sequence: NC_000010.11). The genomic regions spanning positions 94,781,609-94,782,109 for *CYP2C19*2, 94,780,403-94,780,903 for CYP2C19*3*, and 94,761,650-94,762,150 for *CYP2C19*17* were selected. These synthetic templates were used as reference controls for zygosity determination.

Analysis of CYP2C19 alleles using TaqMan™ Drug Metabolism Genotyping Assays

PCR reaction consisted of 5 µL of TaqPath™ ProAmp™ Master Mix, 0.5 µL of each TaqMan™ Drug Metabolism Genotyping Assay (rs4244285, rs4986893, and rs12248560), and 5 µL of genomic DNA (4 ng/µL). Nuclease-free water was added to adjust the total reaction volume to 10 µL.

PCR amplification was performed using a real-time PCR system (QuantStudio™ 5, Applied Biosystems) with the following thermal cycling conditions: a pre-read step at 60 °C for 30 seconds, followed by enzyme activation at 95 °C for 5 minutes. This was followed by 40 amplification cycles consisting of DNA denaturation at 95 °C for 15 seconds and extension at 60 °C for 90 seconds. A final post-read step was performed at 60 °C for 30 seconds. Fluorescence signals were collected during the pre-read, extension, and post-read steps.

Table 1. Allele-specific primers designed for PACE-PCR genotyping of *CYP2C19*2, CYP2C19**3, and *CYP2C19*17*.

ID		**Sequence** (5'→3')
rs12248560	Allele-specific primer 1	GAAGGTGACCAAGTTCATGCTCAAATTTGTGTCTTCTGTTCTCAAAGC
rs12248560	Allele-specific primer 2	GAAGGTCGGAGTCAACGGATTTCAAATTTGTGTCTTCTGTTCTCAAAGT
rs12248560	Common, Reverse primer	CCATCGTGGCGCATTATCTCTTACAT
rs4986893	Allele-specific primer 1	GAAGGTGACCAAGTTCATGCTCAAAAAACTTGGCCTTACCTGGATC
rs4986893	Allele-specific primer 2	GAAGGTCGGAGTCAACGGATTGCAAAAAACTTGGCCTTACCTGGATT
rs4986893	Common, Reverse primer	AACATCAGGATTGTAAGCACCCCCT
rs4244285	Allele-specific primer 1	GAAGGTGACCAAGTTCATGCTTTCCCACTATCATTGATTATTTCCCG
rs4244285	Allele-specific primer 2	GAAGGTCGGAGTCAACGGATTAATTTTCCCACTATCATTGATTATTTCCCA
rs4244285	Common, Reverse primer	GCAAGGTTTTTAAGTAATTTGTTATGGGTT

Analysis of CYP2C19 alleles using PCR Allelic Competitive Extension assay

For *CYP2C19* allele detection using PACE, the assay was prepared as follows. First, a PACE assay primer mix was prepared, consisting of two allele-specific forward primers, each at a final concentration of 12 µM, and one common reverse primer at a final concentration of 30 µM. Next, the PACE genotyping reaction master mix was prepared by combining 5 µL of 2×PACE Genotyping Master Mix (GMM) with 0.138 µL of the PACE assay primer mix. Finally, 5 µL of genomic DNA (4 ng/µL) was added to each reaction, resulting in a total reaction volume of 10 µL.

PCR amplification was conducted using a real-time PCR system (CFX Opus 96, Bio-Rad) with the following cycling parameters: DNA denaturation at 94 °C for 15 minutes, followed by 9 cycles of 94 °C for 20 seconds and 65 °C for 1 minute. followed by 30 cycles of 94 °C for 20 seconds and 57 °C for 1 minute, and a final hold at 37 °C for 1 minute. The fluorescent signal was measured at the endpoint.

Statistical analysis

Genotyping results obtained from the PACE assay were compared with those generated by the commercial *CYP2C19* genotyping kit. The overall

concordance rate between the two methods was calculated as the proportion of identical genotype calls across all samples. The 95% confidence interval (CI) for concordance was estimated using the Wilson score method.[15] Agreement between the two genotyping approaches was evaluated using Cohen's kappa coefficient (κ). The strength of agreement was interpreted as follows: values <0.20 indicate poor agreement; 0.20-0.40, fair agreement; 0.41-0.60, moderate agreement; 0.61-0.80, substantial agreement; and >0.80, almost perfect agreement.[16] Allele and genotype frequencies of *CYP2C19* variants were calculated by direct counting from the genotyping results of the 30 DNA samples and expressed as percentages.

Results

CYP2C19 genotyping using the PACE technique was performed, and allelic discrimination plots for *CYP2C19*2*, *CYP2C19*3,* and *CYP2C19*17* alleles are shown in Figures 1-Figure 3, respectively. The genotyping results from all three alleles were subsequently analyzed together. A total of 30 DNA samples were examined, of which 16 samples were identified as *CYP2C19*1/*1*, 9 samples as *CYP2C19*1/*2*, 3 samples as *CYP2C19*1/*3*, and 2 samples as *CYP2C19*2/*2* (Table 2). The genotyping results obtained using the PACE method demonstrated complete agreement with those generated by the TaqMan™ Drug Metabolism Genotyping Assay, yielding an overall concordance rate of 100% (95%) (CI: 88.7-100.0%) using the Wilson score method. Agreement between the two genotyping approaches showed an almost perfect level of agreement, with a Cohen's kappa coefficient (K) of 1.00. Among the samples analyzed, *CYP2C19*1/*1* was the most prevalent genotype (53.3%), followed by *CYP2C19*1/*2* (30.0%), *CYP2C19*1/*3* (10.0%), and *CYP2C19*2/*2* (6.7%). Notably, no *CYP2C19*17* variant was detected in this study population. Based on direct counting, the allele frequencies were 73.3% for *CYP2C19*1*, 21.7% for *CYP2C19*2*, and 5.0% for *CYP2C19*3*, while *CYP2C19*17* was not observed. No *CYP2C19*17*-positive clinical samples were identified in this cohort. Therefore, evaluation of *CYP2C19*17* detection was limited to synthetic homozygous mutant, heterozygous mutant, and wild-type control templates.

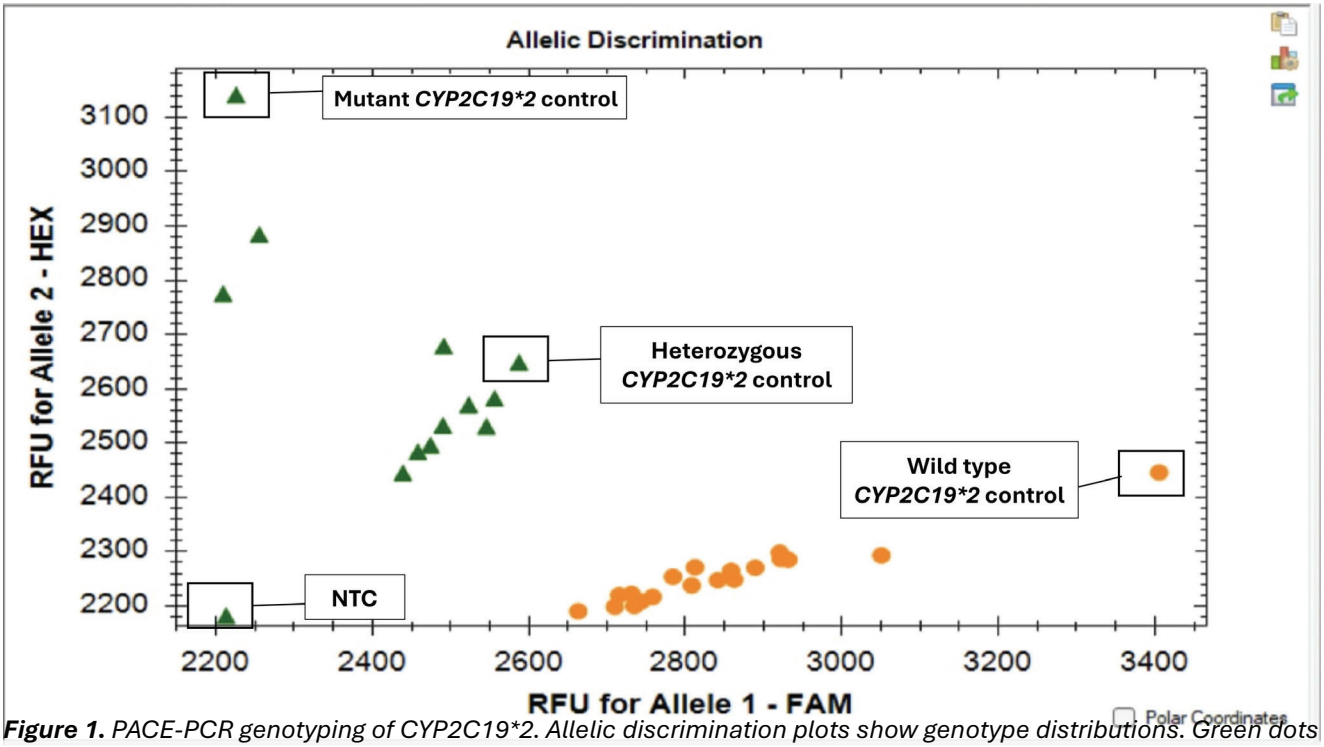


Figure 1. *PACE-PCR genotyping of CYP2C19*2. Allelic discrimination plots show genotype distributions. Green dots represent samples carrying mutant alleles, including homozygous and heterozygous genotypes, while orange dots represent wild-type samples.*

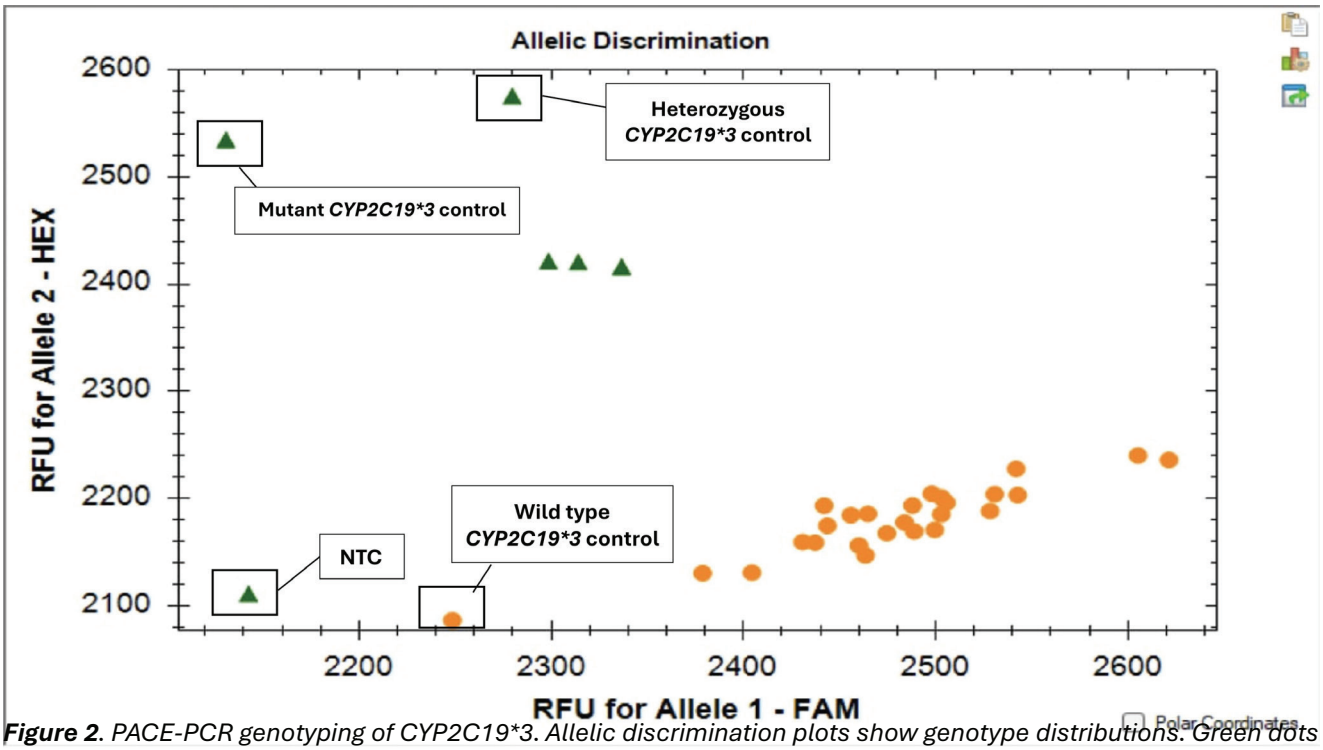


Figure 2. *PACE-PCR genotyping of CYP2C19*3. Allelic discrimination plots show genotype distributions. Green dots represent samples carrying mutant alleles, including homozygous and heterozygous genotypes, while orange dots represent wild-type samples.*

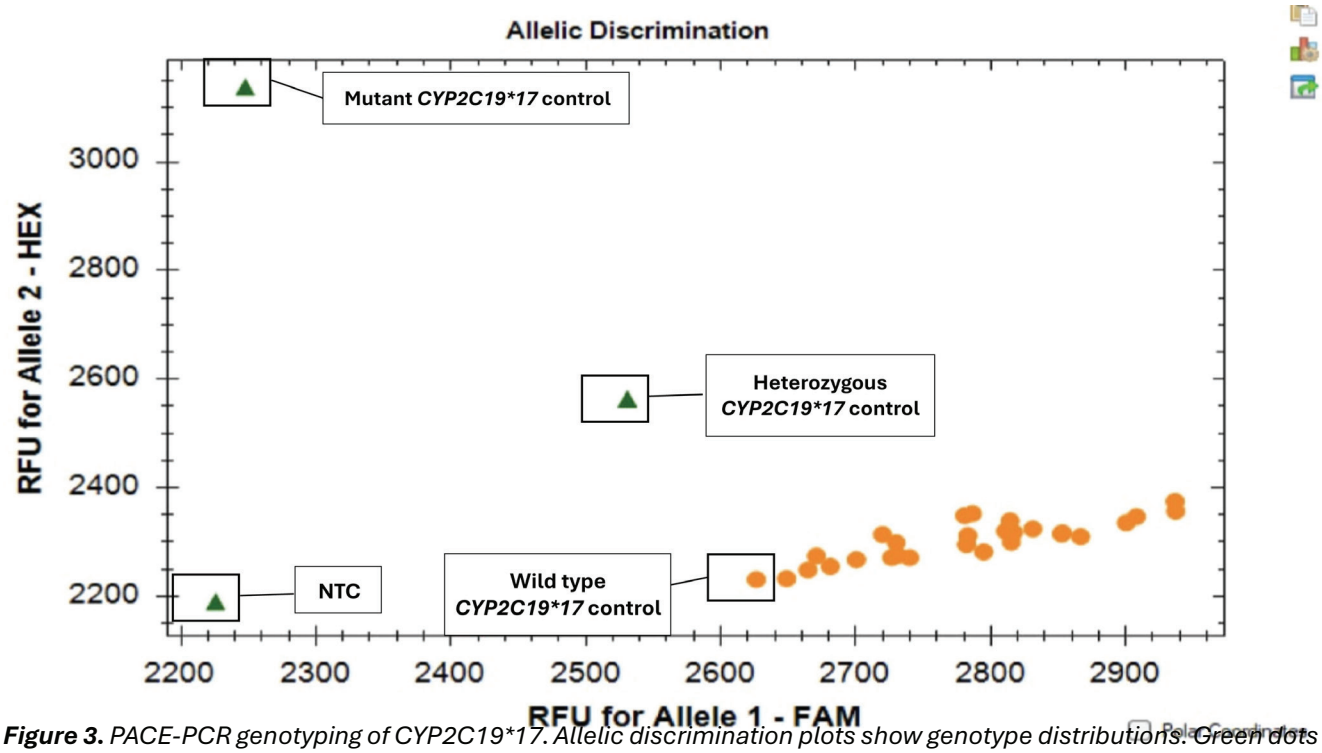


Figure 3. *PACE-PCR genotyping of CYP2C19*17. Allelic discrimination plots show genotype distributions. Green dots represent samples carrying mutant alleles, including homozygous and heterozygous genotypes, while orange dots represent wild-type samples.*

Table 2. Summary of *CYP2C19* genotyping results from *CYP2C19*2, CYP2C19*3* and *CYP2C19*17* allelic discrimination plots.

Sample ID	*CYP2C19* Diplotype	*CYP2C19*2* (G>A)	*CYP2C19*3* (G>A)	*CYP2C19*17* (C>T)
S001	*1/*2	A/G	G/G	C/C
S002	*2/*2	A/A	G/G	C/C
S003	*1/*2	A/G	G/G	C/C
S004	*1/*1	G/G	G/G	C/C
S65056	*1/*3	G/G	G/A	C/C
S65073	*1/*1	G/G	G/G	C/C
S65145	*1/*1	G/G	G/G	C/C
S65157	*1/*1	G/G	G/G	C/C
S65160	*1/*1	G/G	G/G	C/C
S66027	*1/*2	A/G	G/G	C/C
S66087	*1/*2	A/G	G/G	C/C
S66101	*1/*2	A/G	G/G	C/C
S66116	*1/*1	G/G	G/G	C/C
S66117	*1/*1	G/G	G/G	C/C
S66118	*1/*1	G/G	G/G	C/C
S66124	*1/*1	G/G	G/G	C/C
S66138	*1/*2	A/G	G/G	C/C
S66155	*2/*2	A/A	G/G	C/C
S66166	*1/*3	G/G	G/A	C/C
S67018	*1/*1	G/G	G/G	C/C
S67046	*1/*2	A/G	G/G	C/C
S67066	*1/*1	G/G	G/G	C/C
S67109	*1/*1	G/G	G/G	C/C
S67178	*1/*1	G/G	G/G	C/C
S67179	*1/*3	G/G	G/A	C/C
S67196	*1/*3	G/G	G/A	C/C
S68011	*1/*1	G/G	G/G	C/C
S68083	*1/*2	A/G	G/G	C/C
S68087	*1/*2	A/G	G/G	C/C
S68094	*1/*2	A/G	G/G	C/C

Discussion

This study successfully developed a *CYP2C19* genotyping assay based on the PACE technique and compared it against a commercially available TaqMan™ Drug Metabolism Genotyping Assay. The PACE assay presented perfect concordance (100%) with the commercial method, with a Cohen's kappa coefficient of 1.00. These findings confirm that the PACE assay provides highly accurate, reliable genotype calls for the *CYP2C19* variants and clear allelic discrimination. However, the current study presents only an initial proof-of-concept evaluation. Comprehensive analytical characteristics, including limit of detection, specificity, intra- and inter-assay precision, reproducibility across operators and instruments, and robustness under varying experimental conditions, were not fully investigated and should be addressed in future studies prior to routine clinical implementation. Together, without confirmation by sequencing-based methods, the current study should be considered as a comparative analytical evaluation rather than a complete method validation.

Compare PACE to other commercial real-time methods in terms of time-consuming and cost per assay sample. PCR hands-on time is the same, but reagent costs are USD 4.34 (135 Thai baht) per sample for PACE, which may be lower than those of other commercial real-time genotyping assays. This estimated assay cost represents only direct reagent consumable expenses. The calculation did not include labor costs, equipment acquisition or depreciation, maintenance, infrastructure, or other operational expenses. Therefore, the present analysis should not be interpreted as a formal cost-effectiveness evaluation. Nevertheless, the lower reagent cost observed for the PACE assay suggests potential economic feasibility for pharmacogenomics screening applications, particularly in resource-limited settings.

Using this new approach, the genotype and allele frequency distributions observed in this study are consistent with those reported in larger Thai populations. Previous studies have consistently shown that the wild-type allele *CYP2C19*1* has the highest frequency, typically ranging from approximately 70-73%, while *CYP2C19*2* is the most prevalent loss-of-function variant, with approximately 21-25% of the population carrying it, representing the major genetic determinant of the poor metabolizer phenotype.[17,18] The predominance of the *CYP2C19*1/*1* genotype observed in this study (53.3%) is highly comparable to that reported by Nakhonsri *et al*. (48.82%), suggesting that approximately half of the Thai population exhibits extensive metabolizer (EM) phenotype.[18] For intermediate metabolizers (IM) phenotype, *CYP2C19*1/*2* and *CYP2C19*1/*3* genotypes have been previously reported as 36.02% and 6.63%, respectively.[18] Similarly, those were observed in our study as 30.0% and 10.0%, respectively. For the poor metabolizer (PM) phenotype, although the allele frequency of *CYP2C19*3* in this study (5%) was slightly higher than that reported by Areesinpitak *et al*. (3%),[17] it was comparable to the findings of Nakhonsri *et al*. (4.5%).[18] Regional variation among study populations, such as differences between northeastern and northern Thai cohorts, may partly explain the observed variation in *CYP2C19*3* allele frequency across studies. However, we observed only *CYP2C19*2/*2* genotype (6.7%), while Nakhonsri *et al*. showed both *CYP2C19*2/*2* (6.16%) and *CYP2C19*2/*3* (2.37%).[17,18] Since *CYP2C19*17*, the gain-of-function variant, is known to occur at a lower frequency in Asian populations, approximately 1.8-2% of Thai individuals, the absence of this allele in this study is unexpected. This discrepancy is likely attributable to the relatively small sample size (N=30), which reduces the probability of detecting this allele in smaller cohorts.[17,18] Although complete concordance (100%) and κ=1.00 were observed between the PACE assay and the commercial TaqMan™ assay, these findings should be interpreted with caution due to the relatively small sample size and limited genotype diversity.

Therefore, this study should be considered as a preliminary analytical validation rather than a complete assessment of assay performance.

The observed genotype distribution in this study also highlights the potential clinical relevance of *CYP2C19* pharmacogenomics in the Thai population. The clinical relevance in translating *CYP2C19* genotype into predicted metabolizer phenotype for pharmacogenomics-guided therapy was shown in Table 3. Loss-of-function alleles (*2 and *3) of *CYP2C19* are associated with reduced enzymatic activity and impaired conversion of clopidogrel to its active metabolite, which may increase the risk of adverse cardiovascular events during antiplatelet therapy.[20] Additionally, these variants could affect the metabolism of proton pump inhibitors and some antidepressants.[21] In contrast, the *CYP2C19*17* gain-of-function allele has been associated with increased enzyme activity, which may contribute to altered therapeutic response and reduced plasma drug concentrations. Thus, accurate *CYP2C19* genotyping may enable the implementation of personalized pharmacotherapy and precision medicine.

Future studies should include a larger sample size and a more genetically diverse population to further validate the analytical performance of this newly developed PACE assay, along with confirmation by a sequencing-based method. Expanding the panel to include additional clinically relevant or rare *CYP2C19* variants and evaluating reproducibility and inter-laboratory robustness are recommended to support the clinical implementation of this assay.

Table 3. *CYP2C19* genotype–phenotype associations and potential pharmacogenomic implications.[19]

CYP2C19 genotype	Predicted phenotype	Expected enzyme activity	Potential clinical implication
*1/*1	Normal metabolizer (NM)	Normal	Expected standard drug response
*1/*2, *1/*3	Intermediate metabolizer (IM)	Reduced	Reduced clopidogrel activation; altered response to PPIs/ antidepressants
*2/*2, *2/*3, *3/*3	Poor metabolizer (PM)	Markedly reduced or absent	Increased risk of clopidogrel therapeutic failure and drug toxicity
*1/*17	Rapid metabolizer (RM)	Increased	Potential reduced drug exposure
*17/*17	Ultrarapid metabolizer (UM)	Significantly increased	Increased metabolism and potential therapeutic failure for some drugs

Note: *Source: PharmGKB (https://www.clinpgx.org/page/cyp2c19RefMaterials (accessed on 29 May 2026))*

Limitations

Although this study showed 100% concordance between the newly developed PACE and the commercial TaqMan™ genotyping assay, the PACE still has some limitations. Firstly, the sample size was small (N=30), which could limit the statistical robustness of concordance analysis and allele frequency estimation. Consequently, the observed 100% concordance and κ=1.00 may overestimate assay performance due to limited genetic variability within the study cohort. Secondly, although complete concordance between the PACE assay and the commercial TaqMan™ genotyping assay was observed, the comparator method was limited to real-time PCR genotyping without confirmation by a gold-standard sequencing-based approach, such as Sanger sequencing. Therefore, potential shared systematic errors between the two PCR-based methods cannot be completely excluded. Thirdly, the study evaluated only three *CYP2C19* variants (*2, *3, and *17), and other clinically relevant alleles reported in the PharmVar database (URL https://www.pharmvar.org/) were not assessed. Finally, no *CYP2C19*17*-positive clinical samples were included in this sample set, preventing direct evaluation of assay sensitivity for this allele in patient specimens.

Therefore, the current study demonstrated the technical feasibility of detecting *CYP2C19*17* using synthetic control templates and is considered a preliminary study.

Conclusion

As an affordable alternative to commercial genotyping platforms, this study successfully developed a *CYP2C19* genotyping assay based on the PACE technique. Complete concordance with the TaqMan™ Drug Metabolism Genotyping Assay was observed with nearly perfect agreement (κ=1.00). The newly developed PACE assay showed promising preliminary analytical performance and identified the *CYP2C19*2*, *3, and *17* variants. This assay demonstrated lower direct reagent costs and may represent a potentially affordable alternative for *CYP2C19* genotyping. However, further comprehensive analytical and clinical validation studies, together with a comprehensive economic evaluation, are required before conclusions regarding cost-effectiveness can be established and implemented in routine pharmacogenomic testing.

Ethical approval

This research was granted exemption from ethical review by the Research Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University (Research Project Code AMSEC-67EM-014; date of approval: 19 March 2024).

Funding

This study was funded by the Faculty of Associated Medical Sciences, Chiang Mai University, Thailand (R67IN00058).

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Chonticha Sirikul: conceptualization, methodology, data curation, formal analysis, writing: original draft, funding acquisition; **Auttachai Saejia**: methodology, data analysis guidance, sample preparation; **Nampeung Anukul**: conceptualization, supervision, validation, writing: original draft, review and editing.

References

[1] Abdelmonem HB, Abdelaal NM, Anwer EKE, Rashwan AA, Hussein MA, Ahmed YF, et al. Decoding the role of CYP450 enzymes in metabolism and disease: A comprehensive review. Biomedicines. 2024; 12(7): 1467. doi: 10.3390/biomedicines12071467.

[2] Zhao M, Ma J, Li M, Zhang Y, Jiang B, Zhao X, et al. Cytochrome P450 enzymes and drug metabolism in humans. Int J Mol Sci. 2021; 22(23): 12808. doi: 10.3390/ijms222312808.

[3] Samer CF, Lorenzini KI, Rollason V, Daali Y, Desmeules JA. Applications of CYP450 testing in

the clinical setting. Mol Diagn Ther. 2013; 17(3): 165-84. doi: 10.1007/s40291-013-0028-5.

[4] Kittipanprayoon S, Vorasayan P, Chutinet A, Chariyavilaskul P, Suwanwela NC. The prevalence of CYP2C19 polymorphism in patients with symptomatic intracranial atherosclerosis. Cerebrovasc Dis Extra. 2025; 15(1): 68-72. doi: 10.1159/000543331.

[5] Zhou Y, Lauschke VM. Population pharmacogenomics: An update on ethnogeographic differences and opportunities for precision public health. Hum Genet. 2022; 141(6): 1113-36. doi: 10.1007/s00439-021-02385-x.

[6] Nieh HV, Roman YM. Major allele frequencies in CYP2C9 and CYP2C19 in Asian and European populations: A case study to disaggregate data among large racial categories. J Pers Med. 2025; 15(7): 274. doi: 10.3390/jpm15070274.

[7] Sukprasong R, Chuwongwattana S, Koomdee N, Jantararoungtong T, Prommas S, Jinda P, et al. Allele frequencies of single nucleotide polymorphisms of clinically important drug-metabolizing enzymes CYP2C9, CYP2C19, and CYP3A4 in a Thai population. Sci Rep. 2021; 11(1): 12343. doi: 10.1038/s41598-021-90969-y.

[8] Rollinson V, Turner R, Pirmohamed M. Pharmacogenomics for primary care: An overview. Genes (Basel). 2020; 11(11): 1337. doi: 10.3390/genes11111337.

[9] Shubbar Q, Alchakee A, Issa KW, Adi AJ, Shorbagi AI, Saber-Ayad M. From genes to drugs: CYP2C19 and pharmacogenetics in clinical practice. Front Pharmacol. 2024; 15: 1326776. doi: 10.3389/fphar.2024.1326776.

[10] Larsen JB, Rasmussen JB. Pharmacogenetic testing revisited: 5' nuclease real-time polymerase chain reaction test panels for genotyping CYP2D6 and CYP2C19. Pharmgenomics Pers Med. 2017; 10: 115-28. doi: 10.2147/PGPM.S131580.

[11] Cervinski MA, Schwab MC, Lefferts JA, Lewis LD, Lebel KA, Tyropolis AM, et al. Establishment of a CYP2C19 genotyping assay for clinical use. Am J Clin Pathol. 2013; 139(2): 202-7. doi: 10.1309/AJCP9K2KDOCPCBSV.

[12] Lopez-Ortiz C, Reddy UK, Zhang C, Natarajan P, Nimmakayala P, Benedito VA, et al. QTL and PACE analyses identify candidate genes for anthracnose resistance in tomato. Front Plant Sci. 2023; 14: 1200999. doi: 10.3389/fpls.2023.1200999.

[13] Ali GS, Eltaher S, Li J, Freeman B, Singh S. GWAS identifies a polyembryony locus in mango: Development of KASP and PACE markers for marker-assisted breeding. Front Plant Sci. 2025; 16: 1508027. doi: 10.3389/fpls.2025.1508027.

[14] Rudell EC, Cutti L, Turra GM, Angonese PS, Tasca VF, Dos Santos OD, et al. Variability and spatial distribution of ALS-inhibitor resistance mechanisms in Brazilian Echinochloa crus-galli. Plant Physiol Biochem. 2025; 228: 110237. doi: 10.1016/j.plaphy.2025.110237.

[15] Newcombe RG. Two-sided confidence intervals for the single proportion: Comparison of seven methods. Stat Med. 1998; 17(8): 857-72. doi:10.1002/(sici)10970258(19980430)17:8<857::aid-sim777>3.0.co;2-e.

[16] Wang J, Yang Y, Xia B. A simplified Cohen's Kappa for use in binary classification data annotation tasks. IEEE Access. 2019; 7: 164386-97. doi: 10.1109/ACCESS.2019.2953104.

[17] Areesinpitak T, Kanjanawart S, Nakkam N, Tassaneeyakul W, Vannaphasaht S. Prevalence of CYP2C19, CYP3A4 and FMO3 genetic polymorphisms in healthy northeastern Thai volunteers. ScienceAsia. 2020; 46(4): 397-402. doi: 10.2306/scienceasia1513-1874.2020.050.

[18] Nakhonsri V, John S, Panumasmontol H, Jantorn M, Chanthot P, Hanpramukkun N, et al. The diversity of CYP2C19 polymorphisms in the Thai population: Implications for precision medicine. Ther Clin Risk Manag. 2024; 20: 95-105. doi: 10.2147/TCRM.S443815.

[19] Caudle KE, Whirl-Carrillo M, Relling MV, Hoffman JM, Donnelly RS, Haidar CE, et al. Advancing clinical pharmacogenomics worldwide through the Clinical Pharmacetics Implementation Consortium (CPIC). Clin Pharmacol Ther. 2025; 118(6): 1512-22. doi: 10.1002/cpt.3685.

[20] Lee CR, Luzum JA, Sangkuhl K, Gammal RS, Sabatine MS, Stein CM, et al. Clinical Pharmacogenetics Implementation Consortium Guideline for CYP2C19 genotype and clopidogrel therapy: 2022 update. Clin Pharmacol Ther. 2022; 112(5): 959-67. doi: 10.1002/cpt.2526.

[21] Lima JJ, Thomas CD, Barbarino J, Desta Z, Van Driest SL, El Rouby N, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2C19 and proton pump inhibitor dosing. Clin Pharmacol Ther. 2021; 109(6): 1417-23. doi: 10.1002/cpt.2015.

Journal of Associated Medical Sciences 2026; 59(3): 247-255

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



Diagnostic potential of circulating Meteorin-like protein in Graves' disease with subclinical hyperthyroidism subgroup analysis

Sajjad Kadhim Ashour[1*], Mohammed Imran Hamzah[1], Mahmood Shakir Khudhair[2]

[1]Department of Chemistry and Biochemistry, College of Medicine, Al-Nahrain University, Baghdad, Iraq.
[2]Department of Internal Medicine, College of Medicine, Al-Nahrain University, Baghdad, Iraq.

ARTICLE INFO

Article history:
Received 22 February 2026
Accepted as revised 10 July 2026
Available online 14 July 2026

Keywords:
Metrnl, Graves' disease, subclinical hyperthyroidism, endocrine disorders.

ABSTRACT

Background: Meteorin-like protein (Metrnl) is an adipokine mostly produced and secreted by adipose tissue. Metrnl is an adipokine that enhances insulin sensitivity, promotes the browning of adipose tissue, and increases energy expenditure.

Objectives: This study aimed to investigate alterations in serum Metrnl levels in patients with Graves' disease (GD) and subclinical hyperthyroidism (SCH), with the objective of assessing its potential as a diagnostic biomarker.

Materials and methods: This case-control study was matched for age, sex, and BMI and included 40 patients with GD, 40 with SCH, and 80 controls. The average ages of the participants ranged from 22 to 59 years. Samples were obtained from July 2025 to January 2026. Metrnl levels were measured using a sandwich enzyme-linked immunosorbent assay. T4 and TSH levels were evaluated using the VIDAS system (bioMérieux), whereas thyroid-stimulating hormone receptor antibodies (TRAb) were detected using the Roche Cobas e411 Analyzer. The Roche Cobas c111 Analyzer was employed to evaluate lipid profiles.

Results: The mean serum Metrnl levels in the control group were 11.7±1.9 ng/mL. In contrast, patients with GD had a significantly elevated mean Metrnl level of 22.1±2.3 ng/mL (p<0.001). The SCH group had a markedly elevated mean concentration of 22.4±2.8 ng/mL (p<0.001) compared to the control group. Serum concentrations of TSH, total cholesterol, triglycerides, LDL, and VLDL were markedly reduced in both GD and SCH relative to controls (p<0.001). T4 and TRAb levels were significantly elevated in GD and SCH compared with controls (p<0.001).

Conclusion: Elevated levels of Metrnl in Graves' disease suggest its potential as a new biomarker for diagnosis and evaluation of metabolic risks, alongside conventional thyroid function tests.

Introduction

Thyroid hormones are essential regulators of metabolism, growth, and development, and thyroid dysfunction can disrupt multiple physiological systems. [1] Graves' disease (GD) is the most common cause of hyperthyroidism and is an autoimmune disorder characterized by stimulation of the thyroid gland and excessive thyroid hormone production, with potential involvement of extra-thyroidal tissues such as the eyes and skin.[2] It is more common in women aged 30-60 years and results from genetic and environmental

* **Corresponding contributor.**
Author's Address: *Department of Chemistry and Biochemistry, College of Medicine, Al-Nahrain University, Baghdad, Iraq.*
E-mail address: sajjad.kadhim@ced.nahrainuniv.edu.iq
doi: 10.66285/JAMS.2026.098
E-ISSN: 2539-6056

factors, with immune-related (T-cell) genes playing a key role.[3] Subclinical hyperthyroidism (SCH) is defined by suppressed TSH levels with normal free T4 and free T3 concentrations, often with minimal or absent clinical symptoms. Both conditions represent important stages of thyroid dysfunction with clinical and metabolic significance.[4]

Meteorin-like protein (Metrnl) is an adipokine mostly synthesized and released by adipose tissue. Metrnl was first recognized in an early study as an adipokine, demonstrating pleiotropic effects on glucose homeostasis. These results included enhanced insulin sensitivity, increased browning of adipose tissue, and elevated energy expenditure.[5] Metrnl, which is encoded by 936 base pairs and contains 311 amino acids, is located in regions of human chromosome 17q25.3. The small (~27 kDa) secreted protein is encoded by a gene called Meteorin-like (*METRNL*).[6] Thyroid hormone is a vital regulator of energy balance, plays a key role in regulating basal metabolic rate, enables adaptive thermogenesis, and manages body weight through modulating food intake and energy expenditure.[7] Excess thyroid hormone induces a hypermetabolic state marked by increased lipid mobilization, diminished cholesterol levels, weight loss, heat intolerance, and heightened energy expenditure. Thyroid hormones predominantly regulate metabolism by binding to thyroid hormone receptors α or β, thereby affecting white and brown adipose tissue. Moreover, research involving both humans and animals has demonstrated that elevated levels of thyroid hormone increase the expression of the uncoupling protein-1 (UCP-1) gene, thereby augmenting brown adipose tissue activity and facilitating the browning of adipose tissue.[8]

Recent studies have indicated that there is a complex link between adipokines, such as the newer adipokine Metrnl, and thyroid hormones in regulating metabolism. Thyroid hormones are known to control thermogenesis and basal metabolic rate, while adipokines are endocrine mediators that connect the liver, skeletal muscle, and adipose tissue to energy homeostasis.[9] Accumulating evidence suggests that thyroid dysfunction may influence circulating Metrnl concentrations, potentially contributing to the metabolic alterations associated with hyperthyroidism. Metrnl is predominantly expressed in skeletal muscle and adipose tissue, and its secretion is regulated through several transcriptional pathways, particularly peroxisome proliferator-activated receptor gamma (PPARγ) and PPARγ coactivator-1 alpha (PGC-1α). In hyperthyroid states, elevated thyroid hormone activity may stimulate PGC-1α and PPARγ signaling pathways, thereby promoting a hypermetabolic and catabolic state characterized by increased energy expenditure, enhanced lipid mobilization, and altered metabolic homeostasis. These metabolic adaptations may contribute to increased Metrnl expression.[10] At the tissue level, thyroid hormone excess induces thermogenic programming via upregulation of uncoupling protein-1 (UCP-1) and type 2 iodothyronine deiodinase (DIO2), leading to browning of white adipose tissue. This process increases mitochondrial density and enhances fatty acid oxidation, thereby accelerating whole-body energy turnover. Although these adaptations are generally associated with improved glucose utilization, persistent metabolic stress and altered adipokine signaling, including changes in Metrnl activity, may contribute to variability in insulin sensitivity.[11] Metrnl exerts several metabolic effects that may contribute to adipose tissue remodeling and energy homeostasis. Experimental evidence suggests that Metrnl enhances glucose uptake, promotes fatty acid oxidation, and improves insulin sensitivity in skeletal muscle through activation of peroxisome proliferator-activated receptor delta (PPARδ)-dependent pathways and upregulation of AMP-activated protein kinase (AMPK), a key regulator of cellular energy metabolism. These metabolic actions share similarities with the physiological effects of thyroid hormones, which are known to increase thermogenic activity, stimulate oxidative phosphorylation, and enhance mitochondrial biogenesis.[12]

The relationship between thyroid hormone and Metrnl functions indicates a potential feedback or synergistic interaction, wherein changes in Metrnl expression and activity may partially regulate or enhance the thermogenic processes linked to hyperthyroidism. The weight loss commonly seen in hyperthyroid patients is due to increased energy expenditure and the associated stimulation of adipose tissue browning.[13] However, persistent activation of these metabolically demanding pathways may aggravate protein-energy disequilibrium, thereby sustaining the hypermetabolic state characteristic of hyperthyroidism and promoting progressive catabolism of both skeletal muscle and adipose tissue. Within this metabolic context, the elevated circulating Metrnl concentrations observed in hyperthyroid conditions may reflect a compensatory adaptive response intended to mitigate metabolic stress and preserve insulin sensitivity. This interpretation is supported by experimental evidence demonstrating that Metrnl enhances insulin signaling, improves glucose homeostasis, and attenuates insulin resistance, as evidenced by favorable modulation of surrogate indices, such as the homeostatic model assessment of insulin resistance (HOMA-IR), in metabolic disease models.

Furthermore, Metrnl has been implicated in the regulation of lipid metabolism through stimulation of fatty acid oxidation and enhancement of lipid utilization pathways, suggesting a mechanistic association between thyroid hormone excess, altered lipid handling, and insulin resistance-related metabolic disturbances. Nevertheless, the precise molecular mechanisms underlying thyroid hormone-mediated regulation of Metrnl expression and

secretion remain insufficiently elucidated and require further mechanistic investigation.[5,15] Thyroid hormone receptor-mediated transcriptional regulation and the indirect modulation of upstream regulators, such as PGC-1α and inflammatory cytokines, are two potential mechanisms through which thyroid affects tissues that produce Metrnl. The elevation in metabolism may also arise from circulating catecholamines and heightened sympathetic nervous system activity associated with hyperthyroidism, both of which are recognized to enhance thermogenic gene expression.[16]

Serum levels of TSH, thyroid hormones, and TRAb serve as the mainstays for diagnosing GD and SCH; however, these biomarkers primarily reflect thyroid functional status instead of immune-metabolic alterations during disease development. In addition, measuring thyroid hormones is susceptible to physiological variation, hypothalamic-pituitary-thyroid feedback regulation, and analytical interference such as from high-dose biotin supplementation. Consequently, there is an increasing interest in identifying complementary circulating biomarkers that more accurately reflect disease-specific alterations in metabolism and inflammation. Metrnl, an immunometabolic adipokine may well be a potential candidate to supplement standard thyroid function tests and further elucidate the growing field of autoimmune thyroid disease.

Therefore, this study aims to comprehensively evaluate serum Metrnl levels in patients with Graves' disease and subclinical hyperthyroidism compared with healthy controls, to determine its association with thyroid function parameters, and to assess its diagnostic performance, as a supportive diagnostic indicator, in differentiating Graves' disease from healthy individuals using regression-based analysis. Additionally, the study explores its potential involvement in the early metabolic alterations observed in subclinical hyperthyroidism.

Materials and methods

Study design and participants

This case-control study was conducted following approval from the Institutional Review Board (IRB) of a medical college (Date: 10 March 2025; Ref. No. 217/3/2). All procedures were performed in accordance with institutional ethical standards and the principles of the Declaration of Helsinki. Verbal informed consent was obtained from all participants prior to enrollment. Sample collection was carried out at a tertiary care medical center, and laboratory analyses were performed in an academic biochemistry laboratory between July 2025 and January 2026. Participants were aged 22-59 years. The study included 40 patients diagnosed with Graves' disease and 40 patients with subclinical hyperthyroidism, recruited by an endocrinologist. Additionally, 80 age- and gender-matched individuals were enrolled as a control group.

Inclusion criteria

Participants were eligible for inclusion if they matched the following criteria:

1. GD was diagnosed based on suppressed TSH levels, elevated total thyroxine (T4), and positive TRAb. Only patients presenting with overt hyperthyroidism at the time of evaluation were included. All cases were newly diagnosed and had not received any prior treatment.
2. SCH is characterized by suppressed TSH levels with total T4 within the laboratory reference range and clinical signs and symptoms. In the present study, SCH was classified based on TSH and total T4 due to resource limitations and the unavailability of free thyroid hormone measurements. All included patients were newly diagnosed and had not received any prior treatment at the time of enrollment.
3. The control group exhibited normal thyroid function testing, with TSH and T4 levels within reference limits, and no known thyroid disease.

Exclusion criteria

Participants were excluded if they had diabetes mellitus, obesity, or other endocrine disorders; autoimmune diseases excluding thyroid disorders; chronic inflammatory, pulmonary (e.g., COPD), cardiac, renal, or hepatic diseases; malignancy or a history of thyroid cancer; pregnancy; current use of antithyroid, immunosuppressive, or anti-inflammatory medications, or mineral supplements; or an inability to provide informed consent.

Sample size calculation

The sample size was calculated a priori using G*Power software (version 3.1.9.4; Heinrich Heine University Düsseldorf, Germany).[17] A one-way ANOVA (fixed effects, omnibus) was conducted, considering a modest effect size (f=0.30) in accordance with Cohen's guidelines.[18] The required minimum sample size, with a significance level of 0.05 and 90% statistical power, was 144 people. To account for any dropouts and missing data, an extra 10% was included, giving a final target sample of 160 participants.

Calculation of body mass index

Body mass index (BMI) is a widely used measure of weight-for-height, calculated as weight (kg)/height (m^2) and classified according to WHO criteria.[19]

Sample collection and biochemical analysis

Morning fasting venous blood samples (5 mL) were obtained from each participant and then placed into gel tubes. Samples were centrifuged at 3,000 rpm for 10 minutes to isolate serum. One mL of serum was used for lipid profile analysis, while 1.5 mL was allocated to thyroid function testing, including TSH, T4, and TRAb.

The residual serum was divided into microcentrifuge tubes and stored at -20 °C until analysis. TSH and total T4 were selected because they represent the routine thyroid function tests available in our institution during the study period.

Serum Metrnl concentrations were measured using a commercially available sandwich enzyme-linked immunosorbent assay (ELISA) kit (MyBioSource, USA) according to the manufacturer's instructions. The standard calibration curve range was 0.05-15 ng/mL, with a minimum detectable concentration (analytical sensitivity) of 0.023 ng/mL. To ensure measurement accuracy and assay linear range, serum samples exceeding the upper limit of the standard curve were appropriately diluted and re-measured. Final concentrations were calculated by applying the corresponding dilution factors. All samples were analyzed in duplicate, and internal quality control procedures recommended by the manufacturer were followed to ensure assay precision and reproducibility.

TSH and T4 were assessed using the VIDAS automated immunoassay system (bioMérieux, France), based on a solid-phase enzyme-linked fluorescent immunoassay (ELFA) with fluorescence signal detection following antigen-antibody binding. TRAb levels were measured using the Roche Cobas e411 Analyzer (Roche Diagnostics, Germany), which employs a chemiluminescent immunoassay (CLIA) with signal generation through chemiluminescent reaction.

Lipid profile parameters, including total cholesterol (TC), triglycerides (TG), HDL-C, and LDL-C, were analyzed using enzymatic colorimetric methods on the Roche Cobas c111 Analyzer (Roche Diagnostics, Germany). VLDL-C was estimated using the Friedewald-derived formula (VLDL-C=TG/5, mg/dL).

Statistical analysis

Statistical analysis was conducted using the IBM SPSS statistical software program (version 25.0; IBM Corp., Armonk, NY, USA). The Shapiro-Wilk test was applied to evaluate the normality of data distribution. Descriptive statistics were employed to summarize the data. Comparisons between the three study groups (GD, SCH, and control) were performed using one-way analysis of variance (ANOVA) for normally distributed variables, followed by Tukey's post hoc test for multiple comparisons. Spearman's correlation coefficient (r) was determined to assess the relationships between variables. Statistical significance was established at $p<0.05$.

Results

No significant differences were seen in age, sex distribution, or BMI among the GD, SCH, and control groups (all $p>0.05$), providing suitable baseline comparability between the groups. Serum T4 and TRAb levels were significantly higher in both GD and SCH patients than in healthy controls (both $p<0.001$), whereas TSH levels were significantly decreased in the patient groups ($p<0.001$). Regarding lipid metabolism, total cholesterol (TC), triglycerides (TG), HDL-C, LDL-C, and VLDL-C levels were significantly lower in patients with GD and SCH than in controls (all $p<0.001$). Metrnl levels were significantly elevated in both the GD and SCH groups compared to controls ($p<0.001$). No significant difference was seen between GD and SCH patients ($p=0.812$) (Figure and Table 1).

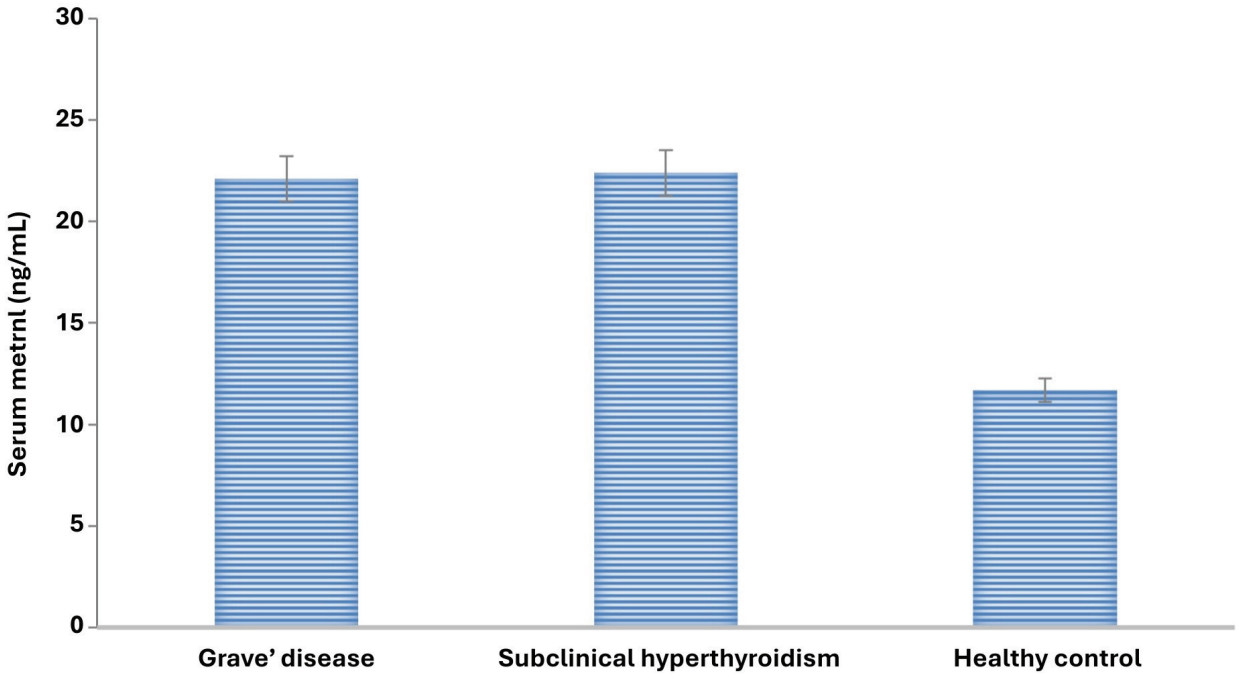


Figure 1. *Serum Metrnl concentrations in GD, SCH, and control groups.*

Table 1. Demographic, biochemical parameters and serum Metrnl levels in GD, SCH, and control groups.

Parameter	GD (N=40)	SCH (N=40)	Control (N=80)	*p* value
Sex (M/F)	19/21	18/22	34/46	0.980
Age (years)	39.20±8.10*	41.40±8.30*	39.80±10.80**	0.975
BMI (kg/m^2)	23.90±2.80*	24.20±3.10**	23.70±2.90*	0.680
T4 (nmol/L)	204.00±17.20*	147.50±10.20**	81.50±9.10***	<0.001
TSH (mIU/L)	0.02±0.01*	0.15±0.04**	2.28±0.73***	<0.001
TRAb (IU/L)	11.78±5.76*	2.33±0.52**	0.54±0.33***	<0.001
TC (mg/dL)	160.60±8.90*	166.80±6.70**	194.00±17.40***	<0.001
TG (mg/dL)	92.60±11.30*	98.40±6.10**	114.70±16.70***	<0.001
HDL-C (mg/dL)	38.70±4.00*	41.50±3.10**	51.80±5.80***	<0.001
LDL-C (mg/dL)	105.30±12.20*	108.10±7.70*	124.20±8.30**	<0.001
VLDL-C (mg/dL)	13.80±1.00*	19.60±1.30**	23.00±3.40***	<0.001
Metrnl (ng/mL)	22.10±2.30*	22.40±2.80*	11.70±1.90**	<0.001

Note: *Data are presented as mean±SD or number (N). Group comparisons were performed using one-way analysis of variance (ANOVA) followed by Tukey's post hoc test for continuous variables and the chi-square test for categorical variables. *,**,***statistically significant differences between groups (p<0.05). A p<0.05 was considered statistically significant. BMI: body mass index, T4: thyroxine, TSH: thyroid-stimulating hormone, TRAb: thyroid-stimulating hormone receptor antibodies, TC: total cholesterol, TG: triglycerides, HDL-C: high-density lipoprotein cholesterol, LDL-C: low-density lipoprotein cholesterol, VLDL-C: very-low-density lipoprotein cholesterol, Metrnl: meteorin-like protein.*

Spearman correlation analysis showed no statistically significant associations between serum Metrnl and thyroid function tests (all $p>0.05$), indicating that Metrnl reflects the hyperthyroid state independently of hormonal levels.

Serum Metrnl levels showed no statistically significant association with SCH (OR=1.13, 95% CI: 0.99-1.31, $p=0.081$). However, Metrnl showed a significant positive association with GD, with an Odds Ratio of 1.706 (95% CI: 1.35-2.15, $p<0.001$), as shown in Table 4.

Table 4. Logistic regression analysis of serum Metrnl and thyroid function markers as predictors of subclinical hyperthyroidism and Graves' disease.

Variable	SCH		GD	
	OR (95% CI)	*p* value	OR (95% CI)	*p* value
Metrnl (ng/mL)	1.13 (0.99-1.31)	0.081	1.706 (1.35-2.15)	<0.001
T4 (nmol/L)	1.10 (1.06-1.14)	<0.001	1.745 (1.35-2.25)	<0.001
TSH (mIU/L)	0.09 (0.03-0.32)	<0.001	2.40 (1.62-3.55)	<0.001
TRAb (IU/L)	4.77 (2.05-11.08)	<0.001	1.186 (1.10-1.28)	<0.001

ROC curve analysis showed that serum Metrnl differentiated Graves' disease patients from healthy controls more clearly than SCH patients (Figures 2 and Figure 3). In Graves' disease, the AUC was 0.989 (95% CI: 0.950-0.999, $p<0.0001$), with a sensitivity of 100% and specificity of 97.5% at a cutoff value of >14.82 ng/mL. In SCH patients, the AUC was 0.913 (95% CI: 0.847-0.957, $p<0.0001$), with a sensitivity of 92.5% and specificity of 85.0% at a cutoff value of >14.50 ng/mL.

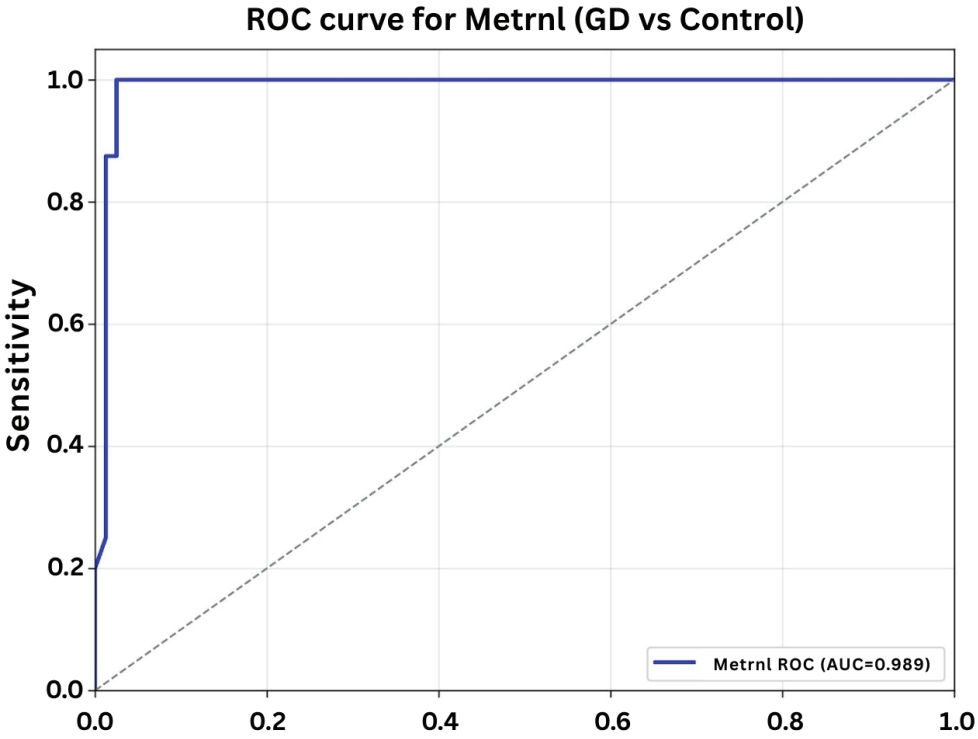


Figure 2. *ROC curve of Metrnl for discrimination between GD and control groups.*

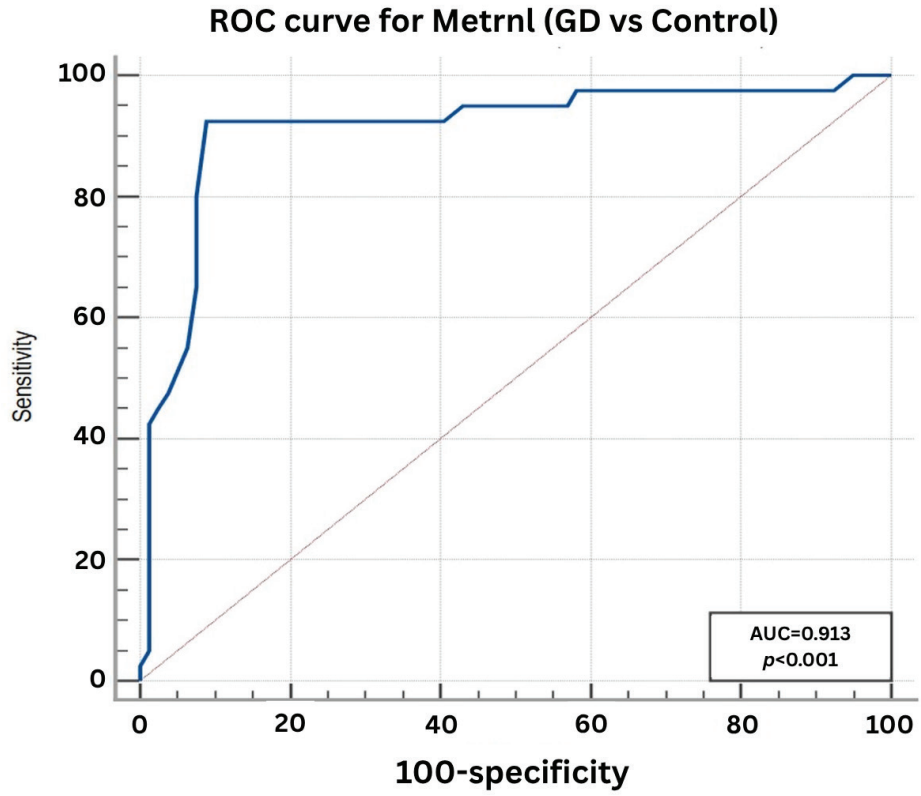


Figure 3. *ROC curve of Metrnl for discrimination between SCH and control groups.*

Discussion

The present study shows that circulating Metrnl levels are significantly elevated in patients with Graves' disease and subclinical hyperthyroidism compared to healthy controls. Significantly, no significant difference was detected between GD and SCH, indicating that Metrnl increase can appear early in thyroid overactivity and isn't just affected by the autoimmune aspect of Graves' disease.

These findings largely agree with those of Wen *et al*., who observed elevated circulating Metrnl levels in newly diagnosed, untreated hyperthyroid patients, showing strong correlations with FT3 and FT4 and an inverse relationship with TSH.[13] In contrast to their findings, the present study did not identify significant relationships between Metrnl and thyroid function markers (T4, TSH, or TRAb). This gap might result from methodological variations, specifically the inclusion of total T4 instead of free thyroid hormone fractions, which more closely represent biologically active hormone levels. Differences in study demographics, disease severity, sample size, and uncontrolled contributing factors could clarify the varying findings.

The absence of significant correlations in the current study indicates that Metrnl does not directly represent circulating thyroid hormone levels. It could indicate metabolic or inflammatory changes associated with hyperthyroidism, especially in Graves' disease. Additional research is required to explain the conditions under which Metrnl and thyroid function parameters may be mechanistically associated.

The lack of direct associations between Metrnl and circulating thyroid hormones in the present study suggests that Metrnl may not directly reflect hormonal excess, but rather a secondary metabolic and inflammatory response associated with hyperthyroid states. The elevated circulating Metrnl levels observed in GD and SCH may therefore be attributed to metabolic and inflammatory alterations induced by thyroid hormone excess. Thyroid hormones are key regulators of thermogenesis, mitochondrial function, and lipid metabolism, and may enhance the activation of PGC-1α and PPAR-related signaling pathways that regulate Metrnl expression in skeletal muscle and adipose tissue. Moreover, thyroid hormone excess promotes adipose tissue browning and adaptive thermogenesis through increased mitochondrial biogenesis and upregulation of thermogenic mediators, processes in which Metrnl has also been implicated. In addition, the chronic low-grade inflammatory milieu characteristic of GD may further contribute to altered Metrnl secretion through immune-metabolic interactions.

The observed elevation of Metrnl in SCH further suggests that these metabolic alterations may occur during the early stages of thyroid overactivity, even prior to the development of overt clinical manifestations. Nevertheless, the precise molecular mechanisms underlying thyroid hormone-mediated regulation of Metrnl expression and secretion remain incompletely elucidated and warrant further investigation.

This interpretation is supported by the comprehensive conceptual framework described by Dong *et al*., which focuses on the relationship among Metrnl signaling, thyroid hormone pathways, and the hypothalamic-pituitary-thyroid axis in the regulation of thermogenesis and systemic energy expenditure.[20]

Gong *et al*. demonstrated significantly reduced Metrnl levels in patients with Graves' disease compared to healthy controls, making it the only study to report an inverse trend.[21] Variations in immunological activity, treatment status, demographic variables, or test variability could be responsible for this variance. Considering the dual metabolic and immunomodulatory characteristics of Metrnl, it is likely that its circulating levels are affected by the fluctuating equilibrium between metabolic activity and autoimmune-mediated inflammation in thyroid disease.

Collectively, the present findings suggest that circulating Metrnl levels are altered in hyperthyroid states and may reflect broader metabolic and inflammatory disturbances rather than thyroid hormone excess alone. The elevation of serum Metrnl observed in both Graves' disease and subclinical hyperthyroidism may indicate that alterations in Metrnl secretion occur during the early phases of thyroid overactivity. Given the recognized role of Metrnl in energy homeostasis and lipid metabolism, the observed changes in circulating Metrnl may represent an adaptive metabolic response to the hypermetabolic and inflammatory milieu associated with thyroid overactivity.

Therefore, serum Metrnl may serve as a complementary biomarker for distinguishing patients with Graves' disease from healthy individuals, while also providing additional insight into the metabolic and inflammatory alterations associated with subclinical hyperthyroidism.

Limitations

Several limitations must be considered. The limited sample size and cross-sectional design prevent causal inferences between circulating Metrnl levels and hyperthyroid conditions, such as Graves' disease. The assessment of longitudinal changes in Metrnl post-treatment was not conducted. Another limitation of the present study is that free thyroid hormones (fT4 and fT3) were not measured due to laboratory and resource constraints, and only total T4 was available for analysis. This may have reduced the precision of thyroid functional assessment, particularly in identifying isolated T3-toxicosis or mild thyrotoxic states. Nevertheless, the diagnosis of Graves' disease and subclinical hyperthyroidism was established based on clinical evaluation in combination with TRAb measurements and suppressed TSH levels, which strengthened the diagnostic classification of the study groups. Body composition measures, such as body fat percent-

age, were not assessed, despite Metrnl acting as an adipomyokine and potentially being affected by metabolic tissue distribution. The single-center design may restrict the generalizability of these results. Consequently, extensive multicenter longitudinal investigations with thorough metabolic and hormonal assessments are necessary to confirm these findings.

Conclusion

The results of this study indicate that circulating Metrnl may serve as a potential biomarker for differentiating patients with Graves' disease and subclinical hyperthyroidism from healthy individuals. The observed variations in circulating Metrnl levels in these conditions indicate a possible association between thyroid dysfunction and Metrnl regulation. Due to its established involvement in energy metabolism, adipose tissue function, and inflammatory signaling, Metrnl may indicate underlying metabolic and immunological changes associated with thyroid hormone dysregulation. However, larger prospective and mechanistic studies are required to further validate its diagnostic utility and to clarify the biological pathways linking Metrnl to thyroid dysfunction.

Ethical approval

The study was performed according to the ethical standards of the Declaration of Helsinki and received approval from the Research Ethics Committee/ Institutional Review Board (IRB) of a medical college (Date: 10 March 2025; Ref. No. 217/3/2). All participants provided verbal informed consent prior to their involvement in the study.

Funding

This study received no funding.

Data availability

Data can be provided upon request from the authors.

Declaration of generative AI and AI-assisted technologies

In the writing of this paper, the authors utilized Gemini (Google LLC, Mountain View, CA, USA) to enhance language and readability. The tool was used exclusively in the introduction and discussion phases. After utilizing this tool, the authors checked and updated the content as required and accepted full responsibility for the finished article.

Conflict of interest

The authors declare that they personally have no competing interests, whether financial or otherwise.

CRediT authorship contribution statement

Sajjad Kadhim Ashour: conceptualization, methodology, formal analysis, investigation, data curation, visualization, writing: original draft, review and edit; **Mohammed Imran Hamzah:** supervision, conceptualization, methodology, validation, writing: review and editing; **Mahmood Shakir Khudhair**: clinical supervision, patient diagnosis, investigation, resources, and writing: review and edit.

Acknowledgements

The authors express appreciation to Al-Imamain Al-Kadhimain Medical City for permitting the collection of clinical samples and access to patient data for this study. The authors acknowledge the Department of Chemistry and Biochemistry at Al-Nahrain University for facilitating toaccess of its research laboratories for biochemical testing.

References

[1] Salman AG, Mahdi IA-J, Mukhlef AK, abd alsattar Mohammad R, Zaghir MSH, muatez Wadaa'a N. Physiological aspects of thyroid disorders: Anatomy, hormones, diagnosis and management. Current Clinical and Medical Education. 2024; 2(5): 17-32. doi: 10.5455/ccme.2024.04.001.

[2] Phagoora J, Bakilwal S, Hamzehpour A, et al. Graves disease-A comprehensive review. Physician's Journal of Medicine. 2024; 3(1). doi: 10.59585/pjm.v3i1.341.

[3] Antonelli A, Ferrari SM, Ragusa F, et al. Graves' disease: Epidemiology, genetic and environmental risk factors and viruses. Best Pract Res Clin Endocrinol Metab. 2020; 34(1): 101387. doi: 10.1016/j.beem.2020.101387.

[4] Tsai K, Leung AM. Subclinical hyperthyroidism: A review of the clinical literature. Endocr Pract. 2021; 27(3): 254-6. doi: 10.1016/j.eprac.2021.02.007.

[5] Li Z, Gao Z, Sun T, et al. Meteorin-like/Metrnl, a novel secreted protein implicated in inflammation, immunology, and metabolism: A comprehensive review of preclinical and clinical studies. Front Immunol. 2023; 14: 1098570. doi: 10.3389/fimmu.2023.1098570.

[6] Garcia-Beltran C, Navarro-Gascon A, Lopez-Bermejo A, et al. Meteorin-like levels are associated with active brown adipose tissue in early infancy. Front Endocrinol (Lausanne). 2023; 14: 1136245. doi: 10.3389/fendo.2023.1136245.

[7] Walczak K, Sieminska L. Obesity and thyroid axis. Int J Environ Res Public Health. 2021; 18(18): 9434. doi: 10.3390/ijerph18189434.

[8] Seifert J, Chen Y, Schöning W, et al. Hepatic energy metabolism under the local control of the thyroid hormone system. Int J Mol Sci. 2023; 24(5): 4861. doi: 10.3390/ijms24054861.

[9] Machado SA, Pasquarelli-do-Nascimento G, Da Silva DS, et al. Browning of the white adipose tissue regulation: New insights into nutritional and metabolic relevance in health and diseases. Nutr Metab (Lond). 2022; 19(1): 61. doi: 10.1186/s12986-022-00694-x.

[10] Alizadeh H. Myokine-mediated exercise effects: The role of myokine meteorin-like hormone (Metrnl). Growth Factors. 2021; 39(1-6): 71-8. doi: 10.1080/08977194.2021.1989417.

[11] Satheesan A, Kumar J, Leela KV, et al. The multifaceted

regulation of white adipose tissue browning and their therapeutic potential. J Physiol Biochem. 2025: 1-23. doi: 10.1007/s13105-025-01050-1.
[12] Tayebi SM, Golmohammadi M, Eslami R, Shakiba N, Costa PB. The effects of eight weeks of circuit resistance training on serum METRNL levels and insulin resistance in individuals with type 2 diabetes. J Diabetes Metab Disord. 2023; 22(2): 1151-8. doi: 10.1007/s40200-023-01239-0.
[13] Wen X, Ding X, Chang X, et al. Elevated serum Meteorin-like levels in patients with hyperthyroidism. BMC Endocr Disord. 2022; 22(1): 304. doi: 10.1186/s12902-022-01231-6.
[14] Ibrahim AY, Neamah NF. Serum Meteorin-Like (Metrnl) as a potential biomarker of obesity-related dyslipidemia and insulin resistance. Biomed Pharmacol J. 2025; 18(3): 2407-18. doi: 10.13005/bpj/3104.
[15] Alizadeh H. Meteorin-like protein (Metrnl): A metabolic syndrome biomarker and an exercise mediator. Cytokine. 2022; 157: 155952. doi: 10.1016/j.cyto.2022.155952.
[16] Townsend LK, Steinberg GR. AMPK and the endocrine control of metabolism. Endocr Rev. 2023; 44(5): 910-33. doi: 10.1210/endrev/bnad012.
[17] Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behav Res Methods. 2009; 41(4): 1149-60. doi: 10.3758/BRM.41.4.1149.
[18] Cohen J. Statistical power analysis for the behavioral sciences. Lawrence Erlbaum Associates; 1988. doi: 10.4324/9780203771587.
[19] World Health Organization. Body mass index (BMI) classification. Web report / Guideline. WHO; 2022.
[20] Dong W-s, Hu C, Hu M, et al. Metrnl: A promising biomarker and therapeutic target for cardiovascular and metabolic diseases. Cell Commun Signal. 2024; 22(1): 389. doi: 10.1186/s12964-024-01768-4.
[21] Gong L, Huang G, Weng L, et al. Decreased serum interleukin-41/Metrnl levels in patients with Graves' disease. J Clin Lab Anal. 2022; 36(10): e24676. doi: 10.1002/jcla.24676.

Journal of Associated Medical Sciences 2026; 59(3): 256-267

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



Method comparison and clinical classification agreement between Freelite and N-Latex serum-free light chain assays for clinical laboratory implementation

Dararat Dankai[1], Phichayut Phinyo[2], Wiyada Dankai[3*]

[1]Immunology Laboratory, Diagnostic Laboratory, Maharaj Nakorn Chiang Mai Hospital, Faculty of Medicine, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Department of Biomedical Information and Clinical Epidemiology, Faculty of Medicine, Chiang Mai University, Chiang Mai Province, Thailand.
[3]Department of Pathology, Faculty of Medicine, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 6 March 2026
Accepted as revised 10 July 2026
Available online 14 July 2026

Keywords:
Serum free light chains, analytical validation, method comparison, κ/λ ratio

ABSTRACT

Background: Measurement of serum free light chains (SFLC) is an essential component in the diagnosis and monitoring of plasma cell dyscrasias. Two analytical platforms are widely used in clinical laboratories: the Freelite assay, which employs polyclonal antibodies with turbidimetric detection, and the N-Latex assay, which uses monoclonal antibodies with nephelometric measurement. Because these assays differ in antibody design, calibration, and analytical detection principles, numerical concentrations and κ/λ ratios may not be directly interchangeable across platforms. This study aimed to evaluate the analytical agreement and clinical classification concordance between the Freelite and N-Latex assays to assess the feasibility of implementing the N-Latex method in routine laboratory practice.

Materials and methods: A cross-sectional method-comparison study was performed using 33 paired archived residual serum samples obtained from routine clinical testing. Free κ and λ concentrations were measured using both Freelite and N-Latex assays on the same specimens. Analytical agreement was assessed using the Wilcoxon signed-rank test, Spearman correlation, Passing–Bablok regression, and Bland-Altman analysis. κ/λ ratios were interpreted according to assay-specific reference intervals, and categorical agreement was evaluated using Cohen's kappa coefficient.

Results: Freelite produced significantly higher κ concentrations than N-Latex (median 43.0 vs 38.9 mg/L, p=0.0002), whereas λ concentrations did not differ significantly. Strong correlations were observed for both κ (ρ=0.897) and λ (ρ=0.951); however, correlation alone was not sufficient to demonstrate quantitative interchangeability. Passing-Bablok regression demonstrated proportional bias for κ measurements, with Freelite yielding higher values at elevated concentrations and the 95% confidence interval of the slope not including 1.0. Bland-Altman analysis revealed wide limits of agreement for κ values at extreme concentrations, indicating clinically relevant numerical differences between methods. When κ/λ ratios were interpreted categorically, using assay-specific reference intervals, concordant classification was observed in 30 of 33 samples (91%), corresponding to substantial agreement (κ=0.79).

Conclusions: The Freelite and N-Latex assays demonstrate strong correlation but exhibit systematic quantitative differences, particularly for κ measurements at high concentrations. Although κ/λ ratio interpretation is largely concordant, the two assays should not be considered directly interchangeable for absolute numerical interpretation or longitudinal monitoring. Implementation of the

* **Corresponding contributor.**
Author's Address: *Department of Pathology, Faculty of Medicine, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: wiyada.dankai@cmu.ac.th
doi: 10.66285/JAMS.2026.099
E-ISSN: 2539-6056

N-Latex assay is feasible in routine clinical laboratories provided that assay-specific reference intervals, appropriate dilution protocols, and antigen-excess precautions are applied, and clinicians are informed about expected differences when transitioning between analytical platforms.

Introduction

Quantification of serum-free light chains (SFLC), comprising free kappa (κ) and lambda (λ) immunoglobulin light chains, plays an important role in the screening, diagnosis, and monitoring of plasma cell dyscrasias, including multiple myeloma (MM), monoclonal gammopathy of undetermined significance (MGUS), and AL amyloidosis.[1,2] These disorders are characterized by clonal proliferation of plasma cells that produce excessive amounts of either κ or λ light chains, resulting in an imbalance of the κ/λ ratio. The International Myeloma Working Group (IMWG) recommends SFLC measurement as part of the standard diagnostic algorithm, particularly in non-secretory or light-chain-only myeloma, where conventional electrophoretic techniques may fail to detect monoclonal protein production.[3,4]

The κ/λ ratio is especially useful for distinguishing monoclonal from polyclonal immunoglobulin production and therefore serves as an important indicator of clonality. An elevated ratio suggests κ-restricted monoclonality, whereas a decreased ratio indicates λ-restricted monoclonality. In routine clinical practice, SFLC results are interpreted alongside serum protein electrophoresis, immunofixation electrophoresis, and bone marrow examination to establish a diagnosis or to monitor disease progression and treatment response.[5] Furthermore, SFLC measurements may support disease monitoring and treatment-response assessment in selected plasma cell disorders, particularly oligosecretory multiple myeloma, nonsecretory multiple myeloma, Bence Jones multiple myeloma, and AL amyloidosis.[6]

Among the commercially available assays for SFLC measurement, the Freelite assay and the N-Latex assay are the two platforms most widely used in clinical laboratories. The Freelite assay employs polyclonal antibodies with turbidimetric detection, whereas the N-Latex assay uses monoclonal antibodies measured by nephelometry.[7,8] Because these assays differ in antibody composition, epitope recognition, and analytical detection methods, numerical concentrations of κ and λ may vary between platforms, which can subsequently influence the calculated κ/λ ratio.[9] Polyclonal antibodies used in Freelite may bind multiple epitopes on free light chains, whereas monoclonal antibodies used in N-Latex provide greater analytical specificity but may fail to recognize certain conformational variants.

Several comparative studies have reported limited numerical agreement between the two assays, although clinical interpretation based on κ/λ ratios is often broadly concordant.[10-12] The lack of assay standardization remains a major challenge in SFLC testing. At present, no internationally accepted reference standard for SFLC quantification exists, which prevents direct comparability or interchangeability of results across analytical platforms.[13,14] Consequently, the Clinical and Laboratory Standards Institute (CLSI) recommends that laboratories perform local method comparison and validation studies before implementing a new measurement procedure, as outlined in guideline EP09-A3.[15]

The clinical implications of assay variability should not be underestimated. Differences in SFLC values between analytical platforms may lead to discrepancies in disease interpretation, particularly when results fall close to diagnostic or therapeutic thresholds. These challenges may be further amplified in patients with renal impairment, in whom altered clearance of free light chains may affect assay performance and the interpretation of the κ/λ ratio.[16,17] Therefore, careful interpretation of SFLC results is required, especially when laboratories transition from one assay platform to another.

At our institution, SFLC testing has historically been performed using the Freelite™ assay through a reagent-rental arrangement that included analyzer support. However, recent procurement policy changes have made continuation of this model increasingly difficult, prompting evaluation of alternative assays that could provide comparable analytical performance while remaining compatible with the laboratory's existing analytical infrastructure.

The N-Latex FLC assay has emerged as a potential alternative platform. Its monoclonal antibody-based design may offer advantages in analytical specificity and reproducibility.[18] Nevertheless, before introducing a new assay into routine clinical practice, it is essential to determine whether the candidate method produces results that are analytically comparable with those generated by the existing platform and whether clinical interpretation remains consistent.

Importantly, this challenge is not unique to a single laboratory. Clinical laboratories worldwide may need to replace or transition between SFLC assay platforms due to procurement policies, instrument availability, or infrastructure changes. A recent systematic review by Ho *et al*. highlighted substantial variability in SFLC assay performance across studies and emphasized the importance of local validation before introducing a new platform into routine clinical use.[19] However, data evaluating assay comparability during real-world

laboratory transitions remain limited.

Given the central role of SFLC testing in the diagnosis and monitoring of plasma cell disorders, and the potential clinical impact of assay-related discrepancies, we conducted this study to evaluate the analytical agreement and clinical classification concordance between the established Freelite assay and the N-Latex assay. The aim of this study was to determine whether the N-Latex platform could be implemented as an alternative method in routine clinical laboratory practice while maintaining reliable interpretation of κ/λ ratios.

Materials and methods

Study design

This study was designed as a cross-sectional method-comparison study using archived residual serum specimens to evaluate the analytical agreement and clinical classification concordance between two serum free light chain (SFLC) assay systems.

Sample selection and handling

Residual serum specimens obtained from routine diagnostic testing at Maharaj Nakorn Chiang Mai Hospital, Faculty of Medicine, Chiang Mai University, were used for the analysis. Specimens were originally collected during standard phlebotomy procedures, and the remaining aliquots were stored at -20 °C after completion of the requested clinical testing, in accordance with routine laboratory storage protocols. Samples were thawed once for analysis and were not subjected to repeated freeze-thaw cycles. For the purposes of this validation study, all specimens were anonymized and de-identified before analysis. No additional blood was drawn from patients, and no direct patient contact occurred.

Samples were eligible for inclusion if sufficient serum volume was available to perform paired κ and λ SFLC measurements using both assays on the same specimen. Hemolyzed samples and specimens with insufficient residual volume were excluded. Based on Freelite κ/λ ratio interpretation, which represented the routine method used at our institution, 25 of 33 samples were classified as normal and 8 as abnormal. Thus, most included specimens represented routine clinical samples with normal κ/λ ratio classification, while a smaller subset represented abnormal or borderline classifications.

Sample size justification

A total of 33 paired serum samples were available during the validation period and were therefore included in the study. The Clinical and Laboratory Standards Institute (CLSI) guideline EP09-A3 recommends using 40-100 patient samples for method-comparison studies when feasible. Therefore, the present study included fewer samples than recommended for a comprehensive method-comparison validation. The limited sample number reflected the specialized nature and relatively low routine testing volume of SFLC assays, as well as the requirement for sufficient archived residual serum volume to perform paired measurements on two analytical platforms. Although the number of specimens was below the recommended CLSI range, the included samples covered a wide analytical concentration range, including physiological and markedly elevated free light chain values. Accordingly, this study should be interpreted as a preliminary real-world local implementation validation rather than definitive evidence of full analytical equivalence between the two assays.

Assay methods and instruments

Paired measurements of serum free κ and λ light chains were performed on the same serum specimens using two automated immunoassay platforms according to the respective manufacturers' instructions.

The existing method was the Freelite assay (The Binding Site Ltd., Birmingham, UK), which employs polyclonal antibodies in a turbidimetric format and was performed on the Optilite analyzer. Calibration and internal quality control procedures were conducted according to routine laboratory operating protocols. Samples exceeding the analytical measurement range were automatically diluted by the analyzer and reanalyzed to obtain reportable concentrations.

The candidate method was the N Latex FLC assay (Siemens Healthineers, Germany), which utilizes monoclonal antibodies with nephelometric detection and was performed on the BN II Nephelometer. Manufacturer-provided reference intervals were applied (κ 6.7-22.4 mg/L, λ 8.3-27.0 mg/L, and κ/λ ratio 0.31-1.56). Specimens exceeding the reportable range were manually diluted using the manufacturer-supplied diluent and reanalyzed to ensure that measured values fell within the analytical measurement range.

For results below the analytical measurement range, values were not numerically extrapolated beyond the validated lower limit. Such results were reported according to the laboratory reporting convention as below the lower reportable limit, with interpretation based on the κ/λ ratio and the paired light-chain result. For results above the analytical measurement range, dilution and repeat measurement were performed before final reporting. Results were released only when the final measured value after dilution was within the validated reportable range and dilution recovery was considered acceptable according to routine quality procedures.

To minimize pre-analytical variability, both assays were performed on the same set of paired serum specimens. Internal quality control was assessed using commercial free light chain control materials at both normal and pathological concentration levels on each platform. During the validation period, the observed

coefficients of variation for internal quality control materials were <5% for both assays.

Detection and management of antigen excess or Hook effect

Because antigen excess can produce falsely low immunoassay results in specimens with very high analyte concentrations, laboratory procedures included specific steps for identifying and managing suspected hook-effect specimens. Antigen excess was considered when the measured SFLC concentration was unexpectedly low or inconsistent with the κ/λ ratio, serum protein electrophoresis, immunofixation findings, total immunoglobulin results, renal status, or clinical suspicion of plasma cell dyscrasia.

When antigen excess was suspected, the specimen was repeated after appropriate dilution using manufacturer-recommended diluent. For N-Latex testing on the BN II platform, built-in antigen-excess protection was used where available, and manual or serial dilution was performed for specimens with unexpectedly low results, markedly abnormal κ/λ ratios, or values near or above the measuring interval. The diluted result was multiplied by the dilution factor and accepted only if the measured concentration fell within the validated analytical range. If dilution nonlinearity or clinical inconsistency was observed, the result was flagged for cautious interpretation, and confirmatory testing using an alternative or orthogonal method was considered.

Statistical analysis

Statistical analyses were performed using MedCalc (MedCalc Software Ltd., Ostend, Belgium) and STATA version 16.0 (StataCorp, College Station, TX, USA). The distribution of κ and λ free light chain concentrations was evaluated using the Shapiro–Wilk test. Because the data were not normally distributed, non-parametric statistical methods were applied. Continuous variables were summarized using medians and interquartile ranges (IQR). Paired κ and λ concentrations obtained from the two assay methods were compared using the Wilcoxon signed-rank test. The association between measurements obtained from the two assays was evaluated using Spearman's rank correlation coefficient (ρ). Analytical agreement between the methods was further assessed using Passing-Bablok regression and Bland-Altman analysis.

Passing-Bablok regression, a non-parametric regression method robust to outliers, was used to evaluate potential systematic bias between the two measurement procedures. Regression equations and 95% confidence intervals (CI) for slope and intercept were calculated. For interpretation, a slope confidence interval including 1.0 and an intercept confidence interval including 0.0 were considered supportive of absence of proportional and constant bias, respectively. Because SFLC assays provide quantitative results, correlation alone was not considered sufficient evidence of agreement or interchangeability.

Bland-Altman analysis was performed by plotting the difference between methods (Freelite minus N-Latex) against the mean of the two measurements. Mean bias and 95% limits of agreement (mean±1.96 standard deviations) were used to assess the magnitude of systematic and random variation between methods. Because an extremely high-concentration specimen affected the visual scaling of scatter and Bland-Altman plots, close-up plots excluding the highest-concentration specimen were prepared to better visualize the lower and mid-concentration ranges while retaining the full-range analysis for overall interpretation.

For clinical interpretation, each sample's κ/λ ratio was classified as normal or abnormal according to the method-specific reference intervals (0.26-1.65 for Freelite and 0.31-1.56 for N-Latex). Concordance in classification between the two assays was evaluated using Cohen's kappa coefficient (κ), which measures agreement beyond chance. Kappa values were interpreted according to conventional thresholds: values greater than 0.80 indicate near-perfect agreement, 0.61-0.80 substantial agreement, 0.41-0.60 moderate agreement, 0.21-0.40 fair agreement, and ≤0.20 slight agreement.

McNemar's test was additionally used to evaluate whether systematic directional disagreement in classification was present. All statistical tests were two-tailed, and a $p<0.05$ was considered statistically significant.

Data protection and ethical considerations

This study involved the secondary use of anonymized leftover clinical serum specimens and therefore posed no more than minimal risk to participants. No direct patient contact occurred, and no identifiable personal or clinical information was accessed. All laboratory data were coded and handled in accordance with institutional data protection procedures. Based on these conditions, the study qualified for a waiver of informed consent under ethical guidelines governing the secondary use of human biological materials.

This study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Research Ethics Committee of the Faculty of Medicine, Chiang Mai University (Study Code: PAT-2568-0329).

Results

A total of 33 paired serum samples (N=33) were analyzed using both the Freelite and N-Latex assays. The measured concentrations of free κ and λ light chains covered a wide analytical range, reflecting the inclusion of specimens with both physiological and markedly elevated values. Freelite κ concentrations ranged from 15.8 to 7,130 mg/L, whereas N-Latex κ concentrations

ranged from 12.6 to 2,690 mg/L. For λ measurements, Freelite values ranged from 7.3 to 1,170 mg/L, while N-Latex values ranged from 8.3 to 607 mg/L.

The distribution of free light chain concentrations measured by the two assays is summarized in Table 1. The median κ concentration measured by Freelite was 43.0 mg/L (IQR 34.1-92.0), which was significantly higher than the median κ concentration measured by N-Latex (38.9 mg/L; IQR 16.7–68.7) (Wilcoxon signed-rank test, p=0.0002). In contrast, the median λ concentration measured by Freelite (50.3 mg/L; IQR 25.0-89.1) was slightly lower than that obtained using N-Latex (55.5 mg/L; IQR 29.2-111.5), although the difference was not statistically significant (p=0.117).

Table 1. Distribution of serum-free light chain concentrations measured by Freelite and N-Latex assays (N=33).

Analyte	**Median (IQR)** mg/L		***p* value***
	Freelite	N-Latex	
Free κ FLC	43.0 (34.1-92.0)	38.9 (16.7-68.7)	0.0002
Free λ FLC	50.3 (25.0-89.1)	55.5 (29.2-111.5)	0.117

Note: **values compared using the Wilcoxon signed-rank test.*

Despite the differences in absolute concentrations, strong correlations were observed between the two assays. The Spearman rank correlation coefficient was ρ=0.897 for κ and ρ=0.951 for λ (both p<0.001), indicating strong monotonic relationships between the measurements obtained from the two analytical platforms. However, the correlation coefficients did not exceed 0.975, and correlation alone was not considered sufficient to establish quantitative interchangeability between assays.

Passing-Bablok regression analysis was performed to further evaluate the relationship between methods. For κ measurements, the regression equation was N-Latex=-13.7+0.76×Freelite (slope 95% confidence interval, CI: 0.56-0.90). Because the 95% CI for the κ slope did not include 1.0, proportional bias was observed, with N-Latex yielding lower κ values than Freelite, particularly at higher concentrations. For λ measurements, the regression equation was N-Latex=3.9+0.96×Freelite (slope 95% CI:0.84-1.09). The 95% CI of the λ slope included 1.0, suggesting better quantitative alignment for λ than for κ measurements. Scatter plots illustrating these relationships are shown in Figure 1.

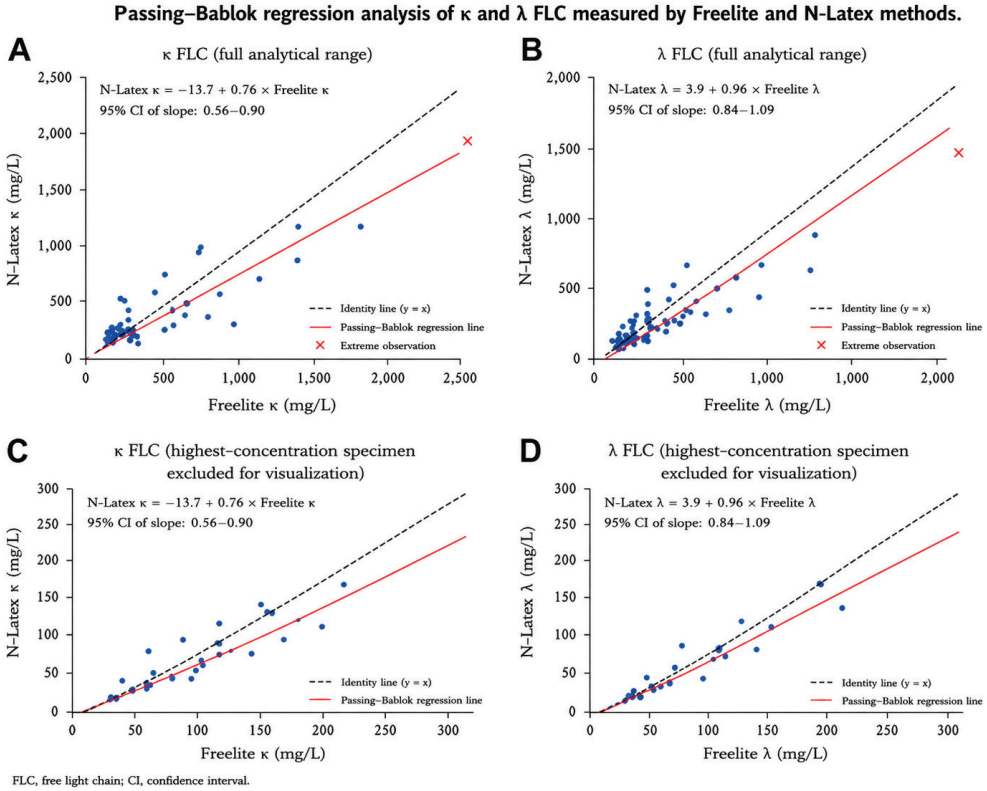


Figure 1. *Passing-Bablok regression plots comparing Freelite and N-Latex serum free light chain measurements. A and B; κ and λ measurements across the full analytical range, C and D: excluding the highest-concentration specimen for visualization of the lower and mid-concentration ranges. The identity line represents y=x, and the regression line represents the Passing-Bablok fit.*

Agreement between the two measurement methods was further evaluated using Bland-Altman analysis. For κ measurements, the mean difference (Freelite minus N-Latex) was +236 mg/L, with 95% limits of agreement ranging from -1,703 to +2,176 mg/L. For λ measurements, the mean bias was +20.7 mg/L, with 95% limits of agreement ranging from -233 to +274 mg/L. The wide limits of agreement, particularly for κ, indicate that individual numerical results may differ substantially between assays and should not be used interchangeably for patient monitoring. Bland-Altman plots illustrating these results are presented in Figure 2.

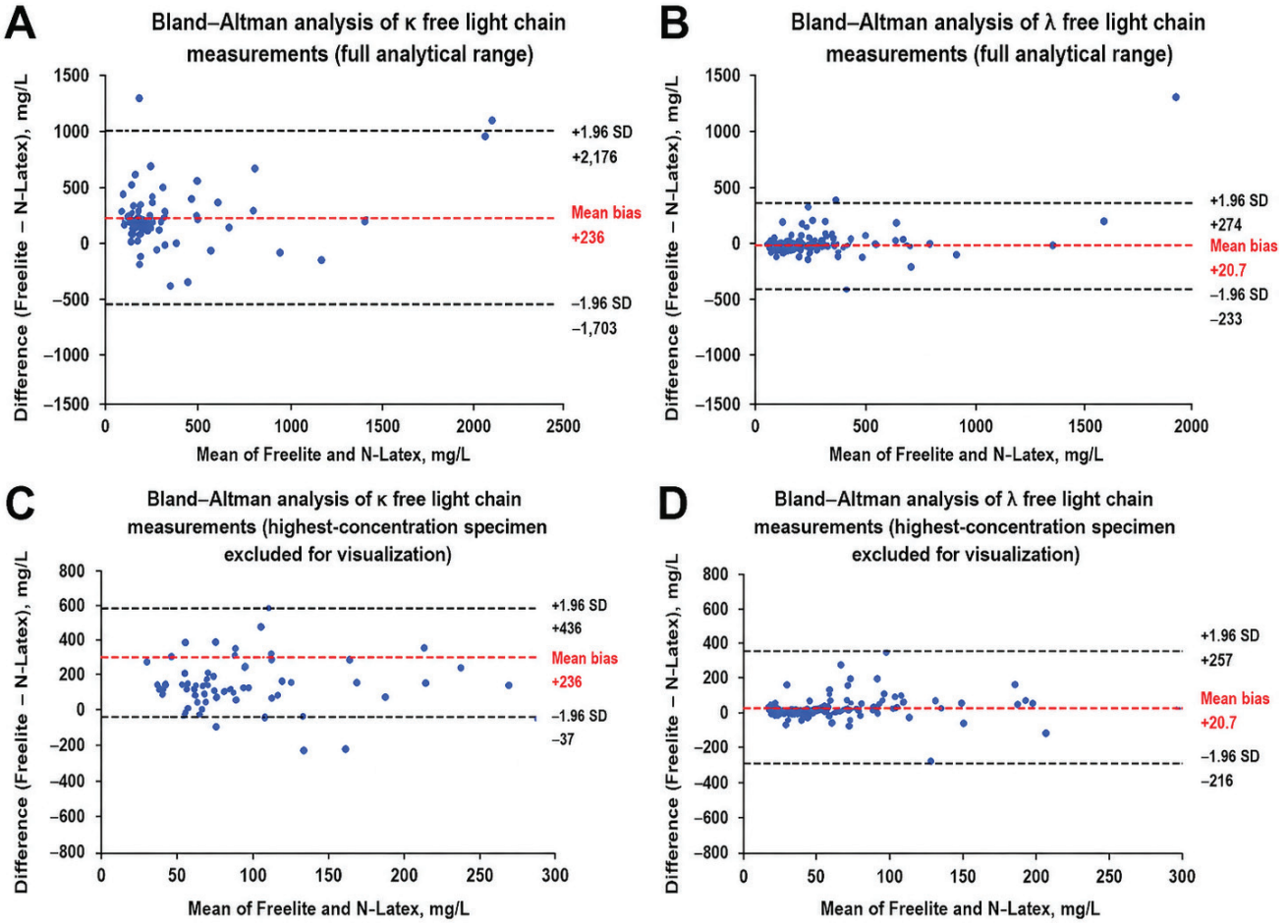


Figure 2. *Bland-Altman plots comparing Freelite and N-Latex serum free light chain measurements. A: κ free light chain measurements across the full analytical range, B: λ free light chain measurements across the full analytical range, C: κ free light chain measurements after exclusion of the highest-concentration specimen for visualization purposes only, D: λ free light chain measurements after exclusion of the highest-concentration specimen for visualization purposes only. Red dashed lines indicate the mean bias, and black dashed lines indicate the 95% limits of agreement.*

For clinical interpretation, κ/λ ratios from both assays were categorized as normal or abnormal according to assay-specific reference intervals. Using Freelite reference limits, 25 of 33 samples (75.8%) were classified as normal and 8 (24.2%) as abnormal. Using N-Latex reference limits, 26 samples (78.8%) were classified as normal and 7 (21.2%) as abnormal.

Overall concordance between the two assays was observed in 30 of 33 samples (91%). The Cohen's kappa coefficient was 0.79, indicating substantial agreement in categorical interpretation. McNemar's test showed no statistically significant difference in classification between the two methods (p>0.99). The contingency table for classification concordance is shown in Table 2.

Table 2. Concordance of κ/λ ratio classification between Freelite and N-Latex assays (N=33).

	N-Latex normal	N-Latex abnormal	Total
Freelite normal	24	1	25
Freelite abnormal	2	6	8
Total	26	7	33

Three samples showed discordant κ/λ ratio classifications between the two assays. Because both assays were performed on the same serum specimen, these discordances are unlikely to represent true biological variation and most likely reflect assay-related analytical or classification differences. In all cases, discrepancies occurred in samples with κ/λ ratios close to the respective reference interval thresholds or in specimens with elevated free light chain concentrations.

Details of the discordant cases are summarized in Table 3. In *Case A*, the Freelite κ/λ ratio was 1.669 (classified as abnormal), whereas the N-Latex ratio was 1.126 (within the reference range). In *Case B*, Freelite reported a κ/λ ratio of 1.71 (abnormal) while N-Latex reported 0.82 (normal). In *Case C*, Freelite produced a ratio of 0.375 (normal), whereas N-Latex produced a ratio of 0.248 (below the reference limit). No cases were observed in which N-Latex indicated a markedly elevated κ/λ ratio while Freelite classified the ratio as normal. These findings suggest that discordant classifications occurred mainly near assay-specific decision limits rather than as major categorical disagreement.

In all three discordant cases, paired measurements were performed on the same serum specimens; therefore, the discordant classifications were unlikely to represent true biological variation. Case A and Case B were classified as abnormal by Freelite but normal by N-Latex because the N-Latex κ/λ ratios shifted into the reference interval. In Case B, the discrepancy was mainly driven by substantially lower κ concentration measured by N-Latex. In Case C, the N-Latex ratio shifted below the lower reference limit, mainly due to lower κ and higher λ values compared with Freelite. All discordant-case values were within the reportable analytical range.

Table 3. Discordant κ/λ ratio classifications between Freelite and N-Latex assays.

Case	Freelite κ (mg/L)	Freelite λ (mg/L)	Freelite κ/λ ratio	Freelite classification	N-Latex κ (mg/L)	N-Latex λ (mg/L)	N-Latex κ/λ ratio	N-Latex classification	Main driver of discordance	Measurement range status
A	43.4	26.0	1.669	**Abnormal**	33.9	30.1	1.126	Normal	Lower κ and higher λ by N-Latex shifted the ratio into the normal range	Within reportable range
B	117.0	68.4	1.711	**Abnormal**	42.0	51.3	0.819	Normal	Markedly lower κ by N-Latex was the main driver of normalization	Within reportable range
C	33.3	88.8	0.375	Normal	24.5	98.6	0.248	**Abnormal**	Lower κ and higher λ by N-Latex shifted the ratio below the lower reference limit	Within reportable range

Note: *Freelite κ/λ ratio reference interval: 0.26-1.65. N-Latex κ/λ ratio reference interval: 0.31-1.56. Measurement range status indicates whether final reported values were within the reportable analytical range.*

Discussion

In this method-comparison validation study, we evaluated the analytical and clinical agreement between the Freelite polyclonal serum free light chain (SFLC) assay and the N-Latex monoclonal SFLC assay. The results demonstrate strong overall correlation between the two methods but also reveal systematic quantitative differences, particularly for κ measurements at higher concentrations. In addition, a small proportion of samples showed discordant κ/λ ratio classifications. These findings are relevant for laboratories evaluating the interchangeability of SFLC assays during platform transitions.

From an analytical perspective, the Freelite assay tended to yield higher κ concentrations than the N-Latex assay, particularly at elevated levels. This pattern was reflected by the positive bias observed in the Bland-Altman analysis and by the regression slope below unity in the Passing-Bablok model. In contrast, λ measurements showed smaller differences between methods, with Freelite values tending to be slightly lower than those obtained with the N-Latex assay. Similar analytical patterns have been reported in previous comparative studies. For example, Daves *et al*. reported higher κ concentrations with Freelite and slightly higher λ concentrations with N-Latex, suggesting that differences in antibody composition and assay calibration may contribute to systematic variation between platforms.[20] The additional contribution of the present study is its focus on real-world local implementation during transition from a Freelite/Optilite workflow to an N-Latex/BN II workflow, including practical considerations for discordant classification, dilution procedures, antigen-excess management, and clinician communication.

The clinical relevance of the observed bias depends on the intended use of SFLC testing. For longitudinal monitoring, the mean κ difference of +236 mg/L and the wide limits of agreement are clinically important because an apparent change after switching assay platforms may reflect analytical bias rather than true biological change. The smaller λ bias of +20.7 mg/L is less likely to affect interpretation in many routine samples, but it may still be relevant when results are close to clinical decision thresholds or when small serial changes are being assessed.

The observed analytical differences likely reflect the distinct antibody designs used in the two assays. The Freelite assay uses polyclonal antibodies that recognize multiple epitopes on free light chains. This broader epitope recognition may allow detection of heterogeneous or polymerized light chain forms and may contribute to higher measured κ concentrations. In contrast, the N-Latex assay employs monoclonal antibodies directed against specific epitopes, which may provide greater analytical specificity but may detect certain light chain conformations differently. Differences in calibrator assignment and assay architecture may also contribute to variation in measured concentrations.

The difference in detection principle may also influence assay performance. Turbidimetry measures reduction in transmitted light caused by immune-complex formation, whereas nephelometry measures light scattered by immune complexes. Nephelometry is generally considered more analytically sensitive for small particles and low concentrations because scattered light can be detected against a low background signal. This potential sensitivity advantage may support reproducibility in lower concentration ranges. However, it does not establish direct interchangeability, because SFLC assay performance is also influenced by antibody specificity, calibration, dilution linearity, and antigen-excess protection.

Another analytical consideration relates to the measurement of very high free light chain concentrations. Antigen excess (Hook effect) can lead to falsely low results when samples exceed the analytical range and are not appropriately diluted. On the Optilite platform used for the Freelite assay, samples exceeding the measurement range are automatically diluted and reanalyzed. In contrast, nephelometric platforms such as the BN II analyzer may require manual dilution procedures to detect antigen excess. In the present dataset, the largest discrepancy in κ concentrations occurred in a sample with extremely elevated values, which may reflect such analytical limitations. These findings highlight the importance of establishing appropriate dilution protocols and verification procedures when implementing new SFLC assays. In practice, antigen excess should be suspected when SFLC results are unexpectedly low or inconsistent with the κ/λ ratio, serum protein electrophoresis, immunofixation findings, total immunoglobulin results, renal status, or clinical suspicion. In such cases, repeat testing with appropriate manual or serial dilution should be performed, and results should be accepted only when the measured concentration falls within the validated analytical range after dilution correction.

Despite the observed differences in absolute quantitation, the correlations between the two assays were strong. For most samples within typical clinical ranges, the two methods demonstrated consistent monotonic relationships, indicating that higher values measured by one assay generally corresponded to higher values measured by the other. Nevertheless, the proportional differences observed indicate that the two assays should not be considered directly interchangeable on an absolute numerical scale. Consistent with current laboratory practice guidelines, results obtained using different SFLC assays should not be used interchangeably for longitudinal monitoring of individual patients. Instead, laboratories should apply assay-specific reference intervals and maintain methodological consistency whenever possible during patient follow-up.

From a clinical perspective, agreement in κ/λ ratio classification was observed in 91% of samples, corresponding to substantial concordance (κ=0.79). The discordant cases observed in this study occurred in samples with κ/λ ratios close to the respective reference interval thresholds. In such borderline situations, relatively small analytical differences in κ or λ concentrations may shift the calculated ratio across the diagnostic cutoff. Similar findings have been reported in previous studies, where discrepancies between Freelite and N-Latex assays were observed in samples with ratios near decision limits.

In the present study, two discordant cases involved elevated κ/λ ratios detected by the Freelite assay but not by N-Latex, whereas one case involved a reduced ratio detected only by N-Latex. These observations are consistent with previous reports indicating that neither assay consistently identifies all abnormalities detected by orthogonal diagnostic methods such as immunofixation electrophoresis. Without independent confirmation using a reference method, it is not possible to determine which assay more accurately reflects the underlying disease status in these cases. Because paired measurements were performed on the same serum specimens, these discordant classifications are unlikely to represent true biological variation. They are more plausibly explained by assay-related analytical or classification differences near reference interval thresholds.

Importantly, none of the discordant cases in our cohort involved markedly abnormal free light chain concentrations that would meet established myeloma-defining thresholds. Instead, discrepancies occurred in samples with κ/λ ratios close to diagnostic cutoffs. Laboratories transitioning between assay platforms should therefore anticipate that a small proportion of results may shift between normal and abnormal classification, particularly among samples with borderline values. The expanded review of discordant cases showed that discrepancies occurred near assay-specific decision limits rather than as major categorical disagreements.

The presence of an extremely high-concentration specimen also influenced visualization of the scatter and Bland-Altman plots. Full-range plots were retained because high SFLC concentrations are clinically relevant in plasma cell disorders, while close-up plots excluding the highest-concentration specimen were added to improve visualization of the lower and mid-concentration ranges. This approach allows both complete analytical-range assessment and clearer visualization of the distribution of most specimens.

From an implementation perspective, our findings suggest that the N-Latex assay represents a feasible alternative to the Freelite assay within our laboratory environment. The assay can be performed on the BN II nephelometric platform already available in our laboratory, allowing integration into existing workflows. However, clinicians should be informed that quantitative differences may occur when transitioning between assays. When a new assay is introduced, previously established patient baselines may shift due to analytical differences rather than biological change. Laboratories may therefore consider establishing new baseline values for patients undergoing longitudinal monitoring after the transition.

Limitations

This study has several limitations. Firstly, the sample size was relatively small (N=33), which was constrained by the availability of paired residual samples during the validation period. This number was below the 40-100 patient samples recommended by CLSI EP09-A3 for comprehensive method-comparison studies when feasible. Although the included specimens covered a wide analytical concentration range, the limited sample size reduces the ability to detect subtle analytical differences and may increase the influence of individual outliers. In addition, most samples had normal κ/λ ratio classifications, and abnormal specimens were underrepresented. Secondly, the study used leftover diagnostic samples; therefore, the sample population may not fully represent all clinical scenarios encountered in routine practice. Thirdly, discrepant results were not independently verified using reference diagnostic methods such as immunofixation electrophoresis or urine Bence-Jones protein analysis. Consequently, the study evaluates analytical agreement between assays rather than diagnostic accuracy relative to a clinical reference standard. Finally, the study focused primarily on comparative accuracy, and other analytical performance characteristics, including long-term precision, analytical sensitivity, lot-to-lot variation, dilution linearity, antigen-excess performance, and cost-effectiveness, were not evaluated in detail.

Conclusion

This method-comparison study demonstrates that the Freelite and N-Latex serum free light chain assays show strong overall correlation but exhibit systematic quantitative differences, particularly for κ measurements. These differences may influence κ/λ ratio interpretation when values are close to diagnostic thresholds. Nevertheless, categorical agreement between assays was substantial, with most samples yielding concordant clinical classifications. Our findings support the feasibility of implementing the N-Latex assay for routine SFLC testing, provided that assay-specific reference intervals, appropriate dilution protocols, and antigen-excess precautions are applied, and clinicians are informed about potential quantitative differences during the transition between analytical platforms. However, the two assays should not be considered directly interchangeable for absolute numerical interpretation or longitudinal monitoring,

and new patient baselines may be needed after platform transition.

Ethical approval

This study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Research Ethics Committee of the Faculty of Medicine, Chiang Mai University (Study Code: PAT-2568-0329).

Funding

The study received no external funding

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Dararat Dankai: conceptualization, investigation, methodology, validation, data curation, writing: original draft; **Phichayut Phinyo**: methodology, formal analysis, software, visualization, writing: review and edit; **Wiyada Dankai:** conceptualization, methodology, supervision, project administration, resources, validation, writing: review and edit. All authors approved the final version of the manuscript.

Acknowledgements

The authors gratefully acknowledge the support provided by the Faculty of Medicine, Chiang Mai University, in the completion of this study.

References

[1] Dispenzieri A, Kyle RA, Katzmann JA, Therneau TM, Larson DR, Benson JT, et al. Immunoglobulin free light chain ratio is an independent risk factor for progression of monoclonal gammopathy of undetermined significance. Leukemia. 2009; 23(2): 215-3. doi: 10.1038/leu.2008.311.

[2] Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet Oncol. 2014; 15(12): e538-48. doi: 10.1016/S1470-2045(14)70442-5.

[3] Bradwell AR, Carr-Smith HD, Mead GP, Tang LX, Showell PJ, Drayson MT, et al. Highly sensitive, automated immunoassay for immunoglobulin free light chains in serum and urine. Clin Chem Lab Med. 2003; 41(3): 321-30. doi: 10.1515/CCLM.2003.053.

[4] Kastritis E, Terpos E, Roussou M, Gavriatopoulou M, Migkou M, Pamboucas C, et al. Evaluation of the prognostic value of serum free light chain ratio in patients with symptomatic multiple myeloma. Leukemia. 2015; 29(6): 1337-44. doi: 10.1038/leu.2015.42.

[5] Wu AH. Performance of immunoassays for detection of free light chains in serum and urine. Clin Chim Acta. 2014; 431: 1-7. doi: 10.1016/j.cca.2014.01.025.

[6] Tosi P, Tomassetti S, Merli A, Polli V. Serum free light-chain assay for the detection and monitoring of multiple myeloma and related conditions. Ther Adv Hematol. 2013; 4(1): 37-41. doi: 10.1177/2040620712466863.

[7] Tate JR, Mollee P, Dimeski G, Carter AC, Gill D. Analytical performance of serum free light-chain assay during monitoring of multiple myeloma. Clin Chem. 2012; 58(3): 523-39. doi: 10.1373/clinchem.2011.170662.

[8] Te Velthuis H, Knop I, Stam P, van den Bergh FA, Ooms B, Smit JW, et al. N Latex FLC - new monoclonal high-performance assays for the determination of free light chain kappa and lambda. Clin Chem. 2011; 57(4): 499-506. doi: 10.1373/clinchem.2010.155820.

[9] Palladini G, Merlini G. Serum free light chain assay for monitoring patients with AL amyloidosis. Blood. 2014; 124(4): 535-43. doi: 10.1182/blood-2014-03-535385.

[10] Ladwig PM, Barnidge DR, Snyder MR, Katzmann JA, Murray DL. Quantification of serum kappa free light chains using two different assays: Assessment of precision, limit of quantification, and biological variation. Am J Clin Pathol. 2014; 142(2): 231-9. doi: 10.1309/AJCPOC7U9RCHYVLI.

[11] Zegers I, Van der Heyden S, Sciacovelli L, Secchiero S, Brandslund I, Streichert T, et al. Towards a reference system for serum free light chain measurements. Clin Chem Lab Med. 2019; 57(9): 1301-7. doi: 10.1515/cclm-2019-0145.

[12] Svarstad E, Bostad L, Myklebust R, Harboe T, Hovig T, Langholm R, et al. Light chain nephropathy in patients with multiple myeloma: A clinicopatho logical study of 41 patients. Nephrol Dial Transplant. 2013; 28(5): 1234-41. doi: 10.1093/ndt/gfs470.

[13] Jacobs JF, Hoedemaekers R, Teunesen MA, Willems HL, van den Bergh FA, van der Hoeven JG. Serum free light chain analysis: Reference values in healthy children. Clin Chem. 2011; 57(5): 713-20. doi: 10.1373/clinchem.2010.160135.

[14] Miller WG, Tate JR, Barth JH, Jones GR, Parslow RA, Rifai N, et al. Harmonization of laboratory testing-current achievements and future strategies. Arch Pathol Lab Med. 2014; 138(2): 234-43. doi: 10.5858/arpa.2013-0268-CP.

[15] Clinical and Laboratory Standards Institute (CLSI). Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline - Third Edition (EP09-A3). Wayne, PA: CLSI; 2013.

[16] Hutchison CA, Harding S, Hewins P, Mead GP, Townsend J, Bradwell AR, et al. Quantitative assessment of serum and urinary polyclonal free light chains in patients with chronic kidney disease. Clin J Am Soc Nephrol. 2008; 3(6): 1684-

90. doi: 10.2215/CJN.02290508.
[17] Lippi G, Avanzini P, Pavesi F, Aloe R, Cervellin G. Evaluation of serum free light chains in patients with chronic kidney disease. Clin Chim Acta. 2017; 464: 19-25. doi: 10.1016/j.cca.2016.11.016.
[18] Abraham RS, Clark RJ, Bryant S, Larson T, Kyle RA, Katzmann JA. Correlation of serum immuno globulin free light chain quantification with urinary Bence Jones protein in light chain myeloma. Arch Pathol Lab Med. 2012; 136(5): 563-72. doi: 10.5858/arpa.2011-0317-OA.
[19] Ho C, Pacaud L, Evison F, Pratt G, Harding S, Zaihra T, et al. Multicenter evaluation of the N Latex FLC assay for the detection of monoclonal free light chains and comparison with Freelite and serum protein electrophoresis. Diagnostics (Basel). 2021; 11(8): 1416. doi: 10.3390/diagnostics11081416.
[20] Daves M, Giacomuzzi F, Bergmann M, Cemin R. Analytical comparison of Freelite™ and N Latex FLC assays for quantifying serum free light chains and impact on the diagnosis of monoclonal gammopathies. Clin Chem Lab Med. 2021; 59(1): 117-24. doi: 10.1515/cclm-2020-0585.
[21] Sasson SC, Lawrie C, Rivory LP, Ho PJ. Comparison of two serum free light chain assays for the diagnosis and monitoring of patients with plasma cell dyscrasias. Intern Med J. 2015; 45(1): 60-6. doi: 10.1111/imj.12624.

Journal of Associated Medical Sciences 2026; 59(3): 268-284

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



Avocado juice (*Persea americana*) with an abundance of bioactive substances on coupling with probiotics revealed potential health benefits

Kanchi Ravi Padma[1*], Sankari Mohan[1], Lingareddy Nikitha Reddy[1], Dhanasekar Chitra[2]

[1]Department of Biotechnology, Sri Padmavati Mahila Visvavidyalayam (Women's University), Tirupati, India.
[2]Department of Biotechnology, Sri Sankara Arts and Science College, Enathur, Kanchipuram, India.

ARTICLE INFO

Article history:
Received 23 February 2026
Accepted as revised 5 May 2026
Available online 16 July 2026

Keywords:
P. americana, Lactobacillus, probiotic, antioxidant activity, probiotic juice, antibacterial activity

ABSTRACT

Background: The rising trend in consumption of plant-based functional beverages has spurred research into alternative carriers for probiotics other than dairy products. Avocado (*Persea americana*) is rich in bioactive compounds, antioxidants, vitamins, and healthy lipids, making it suitable for use as a substrate in probiotic fermentation. Adding probiotic *Lactobacillus* strains to avocado juice can make it more nutritious and therapeutic.

Objectives: This research assessed the efficacy of *Persea americana* (avocado) juice as a non-dairy medium for probiotic *Lactobacillus* strains.

Materials and methods: The strains (*L. acidophilus* and *L. plantarum*) were added to pasteurized juice and evaluated for their ability to withstand stress conditions, including varying pH levels, bile salt concentrations (0.1-0.6%), and NaCl concentrations (0.5-3%).

Results: The resulting fermented juice exhibited improved functional characteristics when compared to non-fermented juice. The antioxidant activity in the fermented juice was significantly elevated, with lower IC_{50} values recorded in both ABTS (17.27 µg/mL) and DPPH (24.64 µg/mL) assays, in contrast to pasteurized juice (31.46 µg/mL and 51.95 µg/mL, respectively), and was found to be comparable to gallic acid standards. Biochemical assessments indicated elevated reducing sugar (683.36 mg/mL) and protein levels, suggesting active microbial metabolism during fermentation. Additionally, the fermented juice exhibited moderate antimicrobial activity against selected pathogenic strains, though lower than that of the standard antibiotic control.

Conclusion: These results imply that avocado juice not only supports the viability of probiotics but also enhances bioactive properties through the fermentation process. In conclusion, this study highlights the potential of avocado juice as a plant-based carrier for probiotics, presenting a promising alternative to dairy-based functional beverages and offering additional nutritional and health benefits.

Introduction

At present, individuals regard food as essential not only for their physical well-being but also for their mental health. The primary aim of contemporary meals is to alleviate hunger while also addressing and mitigating chronic diseases, malnutrition, and the deterioration of health that accompanies aging.[1] The rising expenses associated with healthcare, coupled with an increased understanding of the impact of food

* **Corresponding contributor.**
Author's Address: *Department of Biotechnology, Sri Padmavati Mahila Visvavidyalayam (Women's University), Tirupati, India.*
E-mail address: thulasipadi@gmail.com
doi: 10.66285/JAMS.2026.100
E-ISSN: 2539-6056

on long-term health, have led people globally to take an interest in functional foods, which provide health advantages beyond mere nutritional value.

The avocado (*Persea americana* Mill., Lauraceae) is cultivated in tropical and subtropical regions worldwide. Avocados vary by type depending on their growing climate, with the Hass variety being the most prevalent and widely available. During the production of guacamole and essential oils, the avocado processing industry generates a solid waste byproduct consisting of seeds, peels, and residual pulp, which constitutes approximately 21-30% of the fruit. These byproducts are frequently discarded as waste, presenting an environmental challenge; however, they may also contain bioactive compounds. The avocado seed comprises the endocarp and the kernel, and its triterpenoid constituents have been associated with anticancer properties in studies involving oral, breast, and liver cells.[2-4]

Every day, humans are exposed to a significant quantity of pathogens, predominantly comprising bacteria. Probiotic bacteria have been incorporated into various dietary regimens for several decades owing to their advantageous effects on human health.[5] In 2001, the World Health Organization, in conjunction with the United States Department of Agriculture, delineated probiotics as microorganisms that, when administered to a host in adequate concentrations, promote their overall well-being. *Lactobacillus* and *Bifidobacterium are* the two most commonly used genera in probiotics.[6] These bacteria are generally regarded as benign since they can persist within the human body and, in contrast to conventional pathogens, possess the capacity to both ameliorate and avert disease. Furthermore, these microorganisms have historically played a crucial role in the fermentation of dairy products and in the preservation of food.

Innovations in the alimentary sector are currently focused on the novel functions that food assumes, notably in mitigating lifestyle-related health issues, predominantly through the development of dietary supplements that modulate the microbial composition of the gastrointestinal tract. Furthermore, the food and beverage sector has successfully launched and advanced a diverse array of pioneering products recognized as functional foods, a trend driven by consumers' increasing inclination towards products without artificial additives. Foods designated as functional can exert beneficial effects on specific physiological systems in addition to providing adequate nutritional value, thereby contributing to human health in numerous ways.[7] The principal classifications of functional foods encompass prebiotics, probiotics, and, more recently, symbiotics.[8]

Probiotic foods are commonly associated with dairy products, including cheese, sour milk, and yogurt. However, recent developments have introduced various food ingredients as potential probiotic substrates.[9,10] This shift can be attributed to the adverse characteristics of dairy products that influence consumer perceptions, such as lactose intolerance, allergies, and milk's cholesterol content.[11,12] Consequently, alternative food substrates, including cereals, vegetables, and fruits, both fermented and unfermented, have been proposed as optimal carriers for probiotics to mitigate these concerns.[13] Notably, fruit juices are regarded as particularly attractive and suitable substrates for probiotic delivery. They offer several advantages as probiotic "vehicles," including: 1) high nutritional value, 2) beneficial health effects, and 3) broad consumer acceptance and consumption worldwide.[14,15]

Among the diverse range of fruits, *Persea americana,* commonly known as avocado, offers a distinctive potential as a functional food base. Originating in Central America and widely grown in tropical and subtropical regions, the avocado is appreciated for its creamy texture, rich flavor, and remarkable nutrient profile. It is a source of heart-healthy monounsaturated fats, dietary fiber, potassium, and a broad spectrum of vitamins, including E, C, and several B-complex vitamins. Significantly, different parts of the avocado, such as the pulp, peel, and seed, are abundant in polyphenols, flavonoids, and triterpenes, which have shown antioxidant, antimicrobial, and anticancer effects.[2,3]

The global probiotic industry is primarily composed of three main product categories: foods, dietary supplements, and pharmaceutical products. The quality of probiotic products can vary significantly by type and geographical origin.[15] Probiotics predominantly include various species such as *Saccharomyces, Propionibacterium, Bifidobacterium,* and *Lactobacillus*.[16] In recent years, there has been notable progress in the production of milk and other dairy items that contain probiotic bacteria. These products encompass ice cream, frozen dairy desserts, buttermilk, both plain and flavored liquid milk, milk powder, baby food, and a variety of cheeses.[17,18] The primary advantage of probiotic microorganisms is their ability to coexist and help maintain a balance between harmful infections and beneficial bacteria essential for normal bodily functions.[19] Currently, there is a marked increase in the consumption of probiotic microorganisms, particularly through probiotic dairy products. However, for many individuals who are unable or unwilling to incorporate dairy into their diets, fruit-based beverages may serve as an ideal alternative for delivering probiotics.[20]

Numerous studies indicate that consuming fruit juice, which is rich in nutrients and supports the growth of probiotics, is the most effective way to obtain them. Recognizing these advantages, several companies have invested in developing and marketing fruit-flavored probiotic beverages, including pineapple, apple, orange, and grape varieties. The avocado tree, scientifically known as *Persea americana* and part of

the Laurel family, is the focus of this research. Avocados are not only flavorful but also possess a rich, creamy consistency.

However, there is currently a scarcity of data regarding the use of *Lactobacillus* probiotics in conjunction with avocado fruit.[21] By performing this study, we hope to explore if avocado juice can be a new option for growing *Lactobacillus* strains and to assess the beverage's nutritional, antioxidant, and antimicrobial properties for possible use as a probiotic drink.

Materials and methods

Isolation and identification of lactic acid bacteria

Sample collection

The isolation was performed by taking fresh raw milk samples under aseptic conditions and using a local dairy source as the source of lactic acid bacteria. To eliminate contamination, about 50 mL of milk was collected into airtight, sterile containers. The samples were put in an ice box and taken to the laboratory within 2-4 hrs of collection. The samples were kept at 4 °C upon arrival and processed within 24 hrs. Handling procedures were performed using sterile methods to preserve sample integrity and ensure reproducibility.

Isolation of the organism

The sample was inoculated into de Man Rogosa, Sharpe (MRS) medium and cultured for 24 hrs to isolate the microorganisms. A Petri plate containing 100 µL of sample was used. Fifteen mL of MRS agar (HiMedia Laboratories Pvt. Ltd., India) was poured on the petri plate containing the sample and undisturbed for some time to solidify. After solidification, the plates were incubated for 24 hrs. Once bacterial colonies were visible, they were subcultured on MRS media by the streaking method.

The pour plate method under aseptic conditions was used to isolate microorganisms. Sterile physiological saline was used to serially dilute the sample (10^{-1} to 10^{-6}). Aseptically, 1 mL of appropriate dilutions was transferred to sterile Petri plates. Each plate was poured with sterile MRS agar that was allowed to cool at 45-50 °C. The plates were left to solidify at ambient temperature (25-28 °C) for 15-20 minutes, after which they were incubated in an inverted position at 37 °C for 48-24 hrs. All steps were performed in a laminar airflow cabinet with sterile materials. Specific colonies were picked after incubation and subcultured onto new MRS agar plates to obtain pure isolates.

Stock solution preparation

Milk samples were collected, and bacterial isolation was performed on MRS agar at 37 °C for 24 hrs under aseptic conditions. The medium of MRS (gm/L) was as follows: beef extract (10), yeast extract (5), proteose peptone (10), dextrose (20), sodium acetate (5), polysorbate-80 (1), magnesium sulfate (0.1), ammonium citrate (2), dipotassium phosphate (2), manganese sulfate They were sterilized and the pH was adjusted to 6.5±0.2.[22] The isolates were characterized at the species level on the basis of colony morphology on MRS agar, Gram staining, catalase test, and typical physiological and biochemical features, such as growth patterns and fermentation of carbohydrates, which were similar to *Lactobacillus acidophilus* and *Lactobacillus plantarum*. The pure cultures were grown in MRS broth as stock cultures for further use.

Fruit juice preparation

Fresh avocado (*Persea americana*) fruit was purchased from the local market in Bangalore. The fruits were thoroughly washed with tap water and distilled water. Then the fruit was cut into small pieces and made into juice; the juice was transferred to autoclaved conical flasks and covered with a cotton plug to prevent contamination, pasteurized at 75 °C for 30 minute, and allowed to cool to room temperature.

Fermentation of fruit Juice

A volume of 100 mL of pasteurized avocado (*Persea americana*) juice was moved to sterile Erlenmeyer flasks. The juice was inoculated under aseptic conditions with 1% (v/v) actively growing cultures of *Lactobacillus acidophilus and Lactobacillus plantarum* (one culture at a time). The size inoculated was about 10^{7}-10^{8} CFU/mL. The control flasks were kept uninoculated. Incubation at 37 °C was performed at 24-hr intervals, during which 10 mL samples were withdrawn aseptically and analyzed for biochemical parameters, such as pH, titratable acidity, and sugar content. A calibrated pH meter (µ PH systems 361, Systronics, New Delhi, India) was used to measure the pH.

Sensory evaluation

Sensory analysis was conducted to determine the organoleptic characteristics of fermented and non-fermented *Persea americana* juice, including appearance, aroma, taste, texture, and overall acceptability. Semi-trained volunteers (20-35 years old) used to basic sensory evaluation methods were involved in the study (15 in total). Before taking part, all participants signed an informed consent form using a student consent form, and the study procedures were in line with the general ethical standards of sensory evaluation. To avoid bias, freshly made samples (fermented juice (inoculated with Lactobacillus and incubated 48 hrs) and non-fermented pasteurized juice) were served at room temperature (22-25 °C) in coded, odorless (30 mL) plastic cups. The samples were rated on a 5-point hedonic scale (1=dislike extremely, 5=like extremely) by panelists. Palate cleansing between samples was done with water, and assessments were done in an odor-free, neutral environment. One-way ANOVA and post-hoc Tukey test were used to statistically compare the results with $p<0.05$.[23]

Biochemical tests

Biochemical characterization of the isolates was carried out using standard microbiological procedures. All media and reagents were obtained from HiMedia Laboratories Pvt. Ltd. and prepared according to the manufacturer's instructions. All tests were performed under aseptic conditions and incubated at 37 °C for 24-48 hrs.

Indole production test

A loopful of culture was inoculated into 5 mL of sterile tryptone broth and incubated for 24 hrs. After incubation, 0.5 mL of Kovac's reagent was added gently along the sides of the tube. Formation of a red-colored ring at the surface indicated a positive result.

Methyl red (MR) test

The isolate was inoculated into MR-VP broth and incubated for 48 hrs. After incubation, 2-3 drops of methyl red indicator were added. Immediate development of a red indicated a positive result.

Voges-Proskauer (VP) test

After 48 hrs of incubation in MR-VP broth, 0.6 mL of Barritt's reagent A (α-naphthol) and 0.2 mL of reagent B (40% KOH) were added. The tube was shaken and allowed to stand for 15-30 minutes. A pink-to-red color change indicated a positive result.

Citrate utilization test

The isolate was streaked onto Simmons citrate agar slants and incubated. A change from green to blue, along with growth, indicated a positive result.

Hydrogen sulfide (H_2S) test

The culture was inoculated onto TSI agar by stabbing the butt and streaking the slant, then incubated. Black precipitate formation indicated H_2S production.

Nitrate reduction test

The isolate was inoculated into nitrate broth and incubated for 24 hrs. After incubation, 2-3 drops each of sulfanilic acid and α-naphthylamine were added. The development of red indicated nitrate reduction. If no color developed, a small amount of zinc dust was added. the absence of red after zinc addition confirmed a positive result.

Triple sugar iron (TSI) agar test

The culture was stabbed into the butt and streaked over the slant of TSI agar and incubated. Acid production was indicated by yellow coloration, alkaline reaction by red coloration, gas production by cracks or bubbles, and H_2S production by blackening.

Glucose fermentation test

The isolate was inoculated into glucose broth containing a Durham tube and incubated. A change of the medium color indicated acid production, while gas accumulation in the Durham tube indicated gas production.

Casein hydrolysis test

The culture was streaked onto skim milk agar plates and incubated. Formation of a clear zone around the colonies indicated casein hydrolysis.

Cellulose degradation test

The isolate was spot-inoculated onto cellulose agar plates and incubated. After incubation, plates were flooded with a Congo red solution for 15 minutes, then rinsed with 1M NaCl. Clear zones around colonies indicated cellulose degradation.

Urease test

The isolate was streaked onto urea agar slants and incubated. A change from yellow/orange to pink indicated the presence of urease activity.

Catalase test

A small amount of fresh culture was placed on a clean glass slide, and 1-2 drops of 3% hydrogen peroxide were added. Immediate effervescence (bubble formation) indicated a positive result.

Oxidase test

A colony was smeared onto filter paper impregnated with oxidase reagent (tetramethyl-p-phenylenediamine). The development of dark purple within 10-30 sec indicated a positive result.

All observations were recorded and interpreted using standard microbiological identification criteria.

Effect of pH

The effect of pH on the growth of isolates was evaluated by adjusting the pH of pasteurized *Persea americana* juice to 5, 6, 7, 8, 9, and 10 using sterile 1 N NaOH. A volume of 15 mL of juice was dispensed into sterile conical flasks for each pH level. Each flask was inoculated with 2 mL (≈1% v/v) of actively growing cultures of *Lactobacillus acidophilus* (approximately 10^7-10^8 CFU/mL), prepared in MRS broth under aseptic conditions. After inoculation, 200 µL of each sample was aseptically spread onto sterile MRS agar plates (prepared separately and autoclaved). The plates were allowed to solidify, then incubated at 37 °C for 24-48 hrs. Microbial growth at different pH levels was assessed by colony formation and density on agar plates.

Effect of temperature

Fifteen mL of pasteurized *Persea americana* juice and 2 mL of microbial broth were added to each of the 6 conical flasks. After that, each conical flask was exposed to a range of temperatures for roughly 24 hrs, including room temperature, the refrigerator, and 30 °C,

35 °C, 40 °C, and 45 °C. After a 24-hr period, 200 µL of juice was pipetted onto sterile petri plates, followed by the addition of MRS agar, and the plates were solidified. Following that, the plates were incubated at 37 °C for 24-48 hrs.

Effect of bile salt

The tolerance of probiotic isolates to bile salt was evaluated using different concentrations of bile salt (0.1-0.6% w/v). The study was done using bile salt (HiMedia Laboratories Pvt. Ltd., Mumbai, India). Six sterile conical flasks with 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6 (w/v) bile salt solutions were prepared, and one control flask was prepared with no bile salt. Aseptic conditions were maintained by filling each flask with 15 mL of pasteurized *Persea americana* juice. *Lactobacillus acidophilus* with an approximate initial microbial load of 10^7-10^8 CFU/mL, then in 2 mL (about 1% v/v) were inoculated into each flask to achieve homogeneous initial microbial load.[44] The flasks were left at 37 °C for 24 hrs. Each sample was aseptically placed on sterile MRS agar plates (HiMedia Laboratories Pvt. Ltd., India) after incubation. They were left to solidify, after which they were incubated at 37 °C for 24-48 hrs. The growth of microbes was evaluated in terms of colony formation and viability across varying bile salt concentrations.

Effect of sodium chloride

The effect of sodium chloride on the growth of probiotic isolates was evaluated using different NaCl concentrations (0.5-3.0% w/v). Six sterile conical flasks containing NaCl at concentrations of 0.5%, 1.0%, 1.5%, 2.0%, 2.5%, and 3.0% (w/v), along with one control flask without NaCl, were prepared. Each flask was filled with 15 mL of pasteurized *Persea americana* juice under aseptic conditions. The flasks were inoculated with 2 mL (≈1% v/v) of actively growing cultures of *Lactobacillus acidophilus* or *Lactobacillus plantarum* (approximately 10^7-10^8 CFU/mL). The experiments were performed in triplicate to ensure reproducibility and to support statistical analysis. All flasks were incubated at 37 °C for 24 hrs. After incubation, samples were aseptically plated using the pour-plate method onto MRS agar (HiMedia Laboratories Pvt. Ltd.). The plates were allowed to solidify and then incubated at 37 °C for 48-24 hrs. Microbial growth was assessed based on colony formation and viability at different NaCl concentrations. The results obtained from triplicate experiments were used for statistical analysis (ANOVA followed by Tukey's test).[24]

Biochemical assay

Reducing Sugars by the DNS method were estimated using the following methodology. The amount of sugar present was reduced by the use of dinitrosalicylic acid (DNS) method Miller.[29] Preparation of reagents: A glucose stock solution (1 mg/mL) was made by dissolving 10 mg of glucose in 10 mL of distilled water. A standard solution was then made by adding 1 mL of stock solution to 5 mL of distilled water. DNS reagent was prepared by combining 1 gm of 3, 5-dinitrosalicylic acid, 1 gm NaOH, 200 mg phenol, and 50 mg sodium sulfite in 100 mL of distilled water. In another beaker, 40% (w/v) sodium potassium tartrate solution was prepared. To conduct the assay, 1 mL of DNS reagent was added to 1 mL of glucose standard or sample in a test tube that had been appropriately diluted. This mixture was heated in a boiling water bath (90-100 °C) 10-5 minutes to enable color development. Following the heating procedure, 1 mL of 40% sodium potassium tartrate solution was placed to stabilize the color, and the tubes were allowed to cool down to room temperature. The distilled water (8-10 mL) was used to dilute the reaction mixture, and the absorbance was recorded at 540 nm on a spectrophotometer. A standard curve was drawn, and the glucose and reducing sugar content of the samples was calculated.

Estimation of protein

The protein content was assessed through the Lowry method[25]

Preparation of reagents

A stock solution of bovine serum albumin (BSA) at a concentration of 1 mg/mL was created by dissolving 10 mg of BSA in 10 mL of distilled water. A working standard was generated by diluting 1 mL of the stock solution to a final volume of 5 mL with distilled water.

Reagent A was formulated by dissolving 2% sodium carbonate (Na_2CO_3) in 0.1N NaOH. Reagent B was made by dissolving 0.5% copper sulfate in 1% potassium sodium tartrate. Reagent C, known as the alkaline copper reagent, was prepared by combining 50 mL of reagent A with 1 mL of reagent B. The Folin-Ciocalteu reagent, commercially available, was diluted to 1N before use.

Procedure

In the assay, 1 mL of the standard or a suitably diluted sample was placed in a test tube, followed by the addition of 5 mL of reagent C. The mixture was vortexed and incubated at room temperature for 10 minutes. Following the incubation period, 0.5 mL of the diluted Folin-Ciocalteu reagent was added swiftly, and the mixture was thoroughly mixed immediately. The tubes were then incubated in the dark at room temperature for 30 min to allow for color development. The absorbance was recorded at 660 nm using a spectrophotometer. A standard curve was established using BSA, and the protein concentration of the samples was calculated accordingly.

Anti-oxidant assay

ABTS

Gallic acid (10 mg) was dissolved in 10 mL of methanol to produce a stock solution. Methanol (9 mL) was added to 1 mL of stock to generate the working

solution. Three sets of test tubes were used; one set was pipetted with 10-50 µL of standard. In addition, the other two sets of test tubes received 10-50 µL samples of regular juice and of pasteurized probiotic juice. After preparing each test tube solution to a final volume of 100 µL with methanol, 3 mL of the ABTS solution was added. Prior to adding ABTS, a concentrated ABTS solution was added to a diluted ABTS solution to set the absorbance to 1. After 30 minutes of dark incubation, the test tubes were removed, and autozero was used to compute the absorbance at 734 nm. To calculate the percentage (%) of inhibition, a formula was used:

Percentage (%) inhibition = (Abs of control-Abs of Test/ Abs of control)×100

DPPH

To prepare the stock, 10 mg of gallic acid and 10 ml of methanol were combined. Methanol (9 mL) was added to 1 mL of stock to create the working solution. They took three sets of test tubes. An empty tube containing 100 µL of methanol was used as the blank, and one set was pipetted with 10-50 µL of the standard; the remaining two sets contained regular juice and pasteurized probiotic juice. Methanol was added to each standard tube to a capacity of 100 µL. Next, 3 mL of DPPH solution was added to each test tube. Prior to adding DPPH, concentrated DPPH was added to diluted DPPH to set the DPPH solution's absorbance to 1. For 15 minutes, all the tubes were incubated in the dark. At 517 nm, the absorbance was measured following incubation.[26] To calculate the percentage (%) of inhibition, a formula was used:

Percentage (%) inhibition = (Abs of control- Abs of Test/ Abs of control)×100

Anti-microbial activity

Probiotic avocado juice was prepared by blending 100 gm of ripe avocado with 100 mL of distilled water, then pasteurized at 72 °C for 15 sec. A 10% (v/v) probiotic culture was added and incubated at 37 °C for 24 hrs. Luria-Bertani (LB) agar and Sabouraud Dextrose Agar (SDA) were prepared using commercial media (HiMedia Laboratories Pvt. Ltd., Mumbai, India), autoclaved, and poured into sterile Petri plates. After solidification, pathogenic bacteria were swabbed, and wells were filled with either probiotic or regular avocado juice (50-100 µL per well). Each setup was prepared in triplicate. Plates were incubated for 24 hrs at 37 °C and examined for zones of inhibition.

Data analysis

ANOVA with $p<0.05$ and Duncan's Multiple Range Test (DMRT) were used to identify significant differences among mean values. Data are reported as the mean and standard deviation (mean±SD) of triplicate measurements. Version 12 of the SPSS program was used for all statistical analyses. Pearson's correlation coefficient (r) was used to obtain correlations in bivariate linear correlations.

Results

Isolation, screening, and identification of lactic acid bacteria

Lactic acid bacteria (LAB) were isolated using de Man, Rogosa, and Sharpe (MRS) agar medium under anaerobic incubation conditions. The obtained colonies were screened for morphological and biochemical characteristics, including Gram staining, catalase negativity, and acid production. Selected isolates were further evaluated for probiotic properties, including tolerance to acidic pH, bile salts, and NaCl concentrations. The potent isolates showing maximum survivability and functional activity were identified as *Lactobacillus acidophilus and Lactobacillus plantarum*. Identification was confirmed through standard microbiological and biochemical characterization methods.

Isolation, subculturing, and screening of a probiotic strain in avocado juice

This investigation indicates that morphological traits were used to identify the lactic acid bacteria (LAB). In addition, all of the LAB were identified to the generic level by other biochemical tests, which showed negative (data not shown), showing that *Lactobacillus* spp. were whitish, small to large in size, with a circular margin on MRS media (Figure 1), and *Lactococcus* spp. were creamy white to yellowish in color, small to large in size, with a circular margin. To subculture, pure cultures of specific, morphologically distinct bacterial colonies were inoculated onto solidified LB agar plates. To prevent contamination, do this in the laminar air flow chamber. The plates were maintained at 37 °C in an incubator for 24-48 hrs. The juice was screened to ensure sterility. LB agar was prepared and autoclaved before being used for screening. Following autoclaving, 200 µL of pasteurized avocado juice was pipetted onto the Petri plates. Agar was then added to the juice and left to solidify. The plates were incubated at 37 °C for 24 to 48 hrs after solidification.

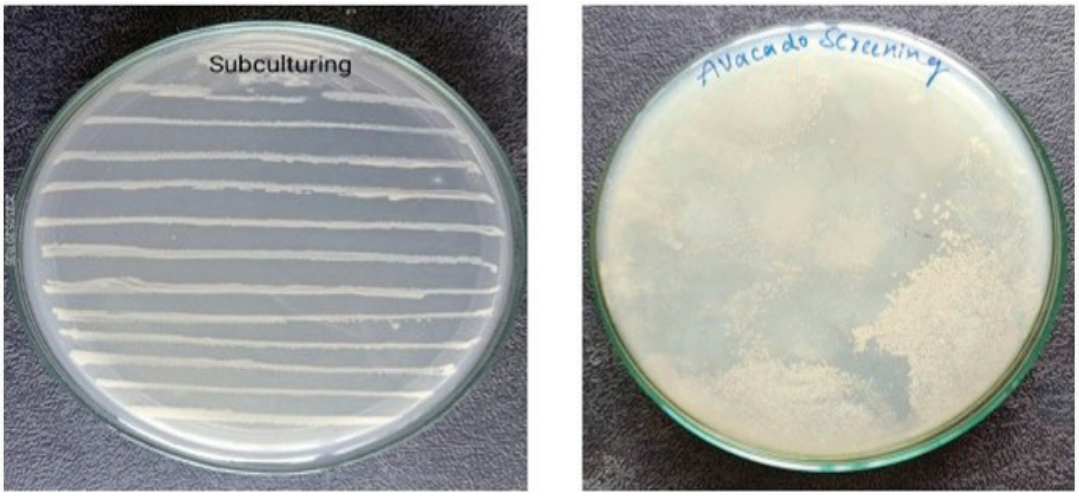


Figure 1. *Isolation, subculturing, and screening of probiotic Lactobacillus strains in avocado juice. A: subculturing of Lactobacillus, B: screening of probiotic strain in avocado juice.*

The organisms were isolated from the milk sample on MRS media by the pour-plate method. The isolated bacterial colonies were identified as Gram-positive rods based on their morphological features. The microscopic view of the Gram-stained organisms is shown in Figure 2.

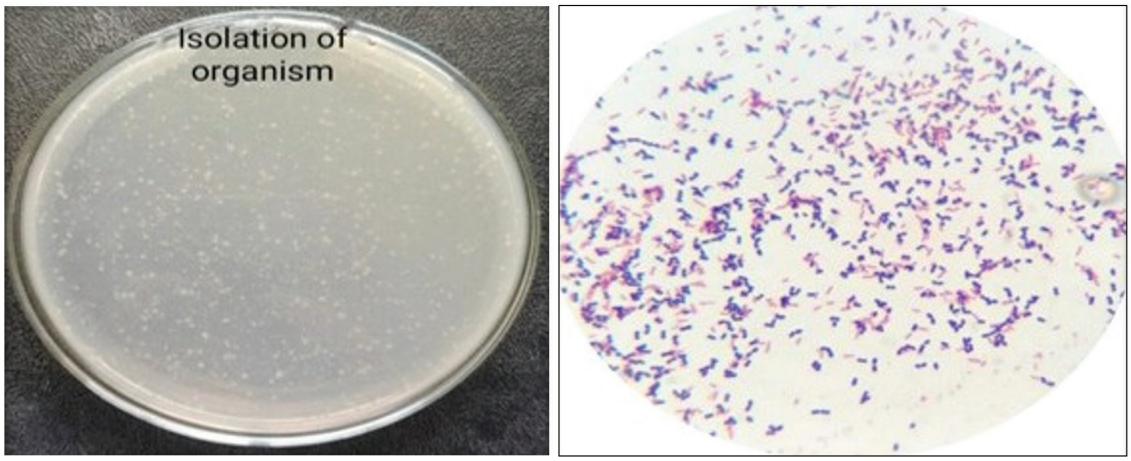


Figure 2. *Isolation of LAB on MRS media and bacterial colonies were identified as Gram-positive rods under a 40X magnification microscope.*

Biochemical characterization of isolates

The biochemical tests were performed to identify and characterize the bacterial isolates, particularly to confirm their affiliation with lactic acid bacteria (LAB) and to support their classification as *Lactobacillus* species. These tests help differentiate bacteria based on metabolic activities, enzyme production, and substrate utilization patterns.

The observed results-specifically, Gram positive (presumed), catalase-negative, oxidase-negative characteristics, along with positive carbohydrate fermentation patterns-are indicative of *Lactobacillus* species. The lack of indole production, citrate utilization, urease activity, and H_2S production further reinforces this classification. Consequently, these biochemical tests were used to verify the identity and functional traits of the isolates, thereby confirming their suitability as probiotic strains for fermentation research.

Probiotic juice using Persea americana (avocado)
Extraction and fermentation of fruit juice

The Juice from fruit was extracted and pasteurized at 72 °C for 30 minutes, and the sterility of the pasteurized juice is depicted in Figure 3. It is claimed that pasteurized juice is free of contamination because no colonies are present on the Petri plate. The juice was given a 7-day fermentation period by adding 20 mL of a broth containing microorganisms to 200 mL of pasteurized juice.

Optimization of parameters

The viability of probiotic *Lactobacillus* at a pH range of 5-10, different temperatures from fridge to room temperature, 0.1-0.5% concentration of bile salt, and 0.5-3% NaCl concentration is shown in Figure 4.

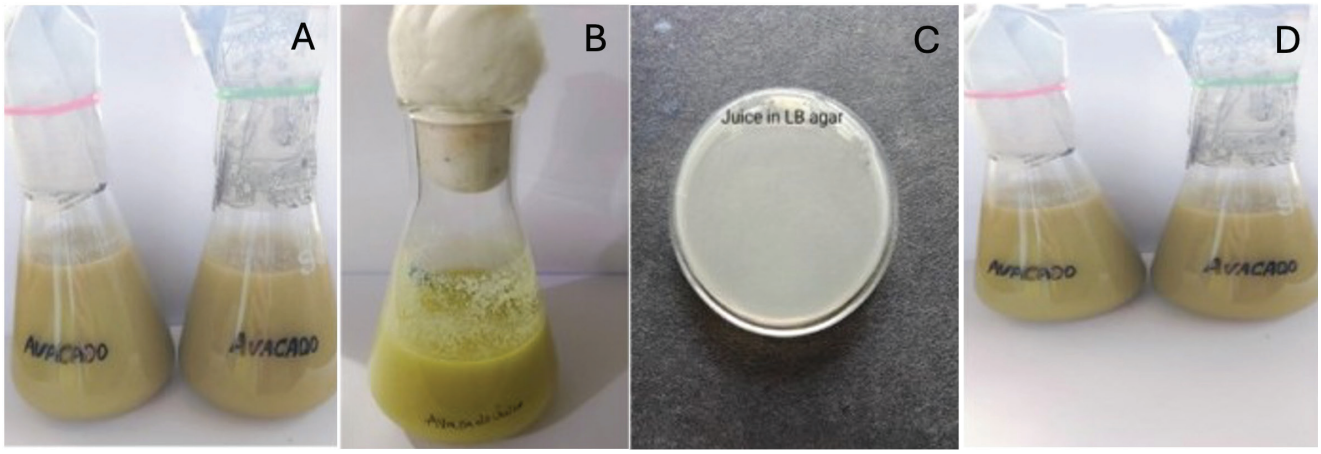


Figure 3. *Preparation and fermentation process of avocado juice.*
A: extraction of fruit juice, B: pasteurization, C: sterility check of the pasteurized milk, D: fermented fruit juice.

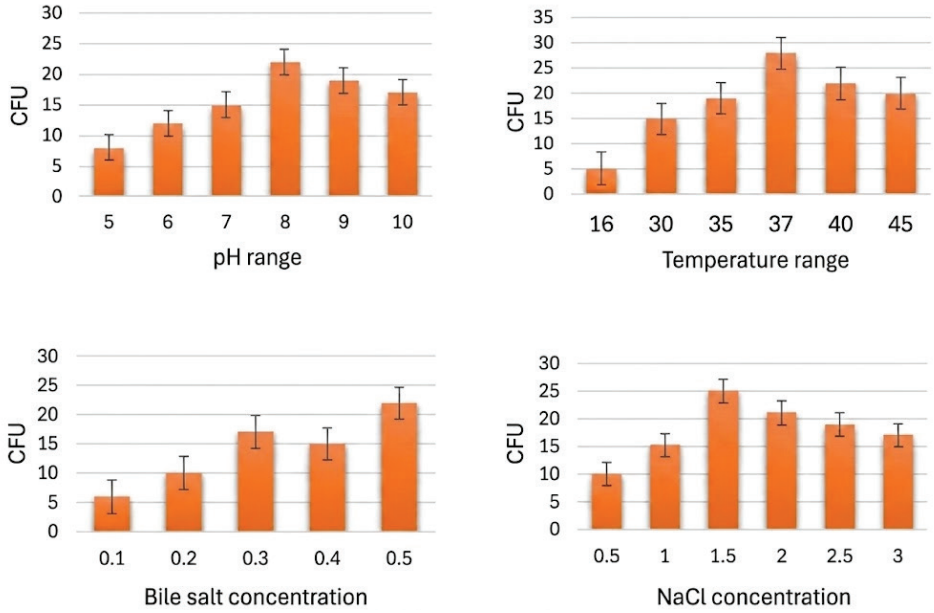


Figure 4. *Enhancement of cultural conditions for the viability of Lactobacillus.*
A: influence of pH (ranging from 5 to 10), B: temperature settings (both refrigerated and ambient temperature), C: concentration of bile salts (0.1 to 0.5% w/v), D: concentration of NaCl (0.5 to 3% w/v) on the survival of probiotic Lactobacillus. Viability is quantified in colony-forming units/mL (CFU/mL). The data presented are mean values derived from triplicate experiments.

Sensory evaluation

The sensory evaluation of both fermented and non-fermented *Persea americana* (avocado) juice was conducted using a 5-point hedonic scale to assess consumer preferences for appearance, aroma, taste, texture, and overall acceptability. The fermented juice, inoculated with *Lactobacillus* and incubated for 48 hrs, consistently received higher scores across all sensory characteristics. It attained mean scores of 4.2±0.5 for taste, 4.1±0.6 for aroma, and 4.3±0.5 for overall acceptability, indicating an improved sensory profile compared with the pasteurized, non-fermented sample, which scored 3.4±0.6. These enhancements may be linked to the breakdown of sugars and the generation of aromatic compounds during the fermentation process, as similarly noted in earlier research involving probiotic fruit beverages.[27,28] Statistical evaluation through one-way ANOVA followed by Tukey's post-hoc test validated that the differences were significant at $p<0.05$. These results highlight the potential of fermented *Persea americana* juice as a more enjoyable and consumer-friendly functional beverage (Figure 5).

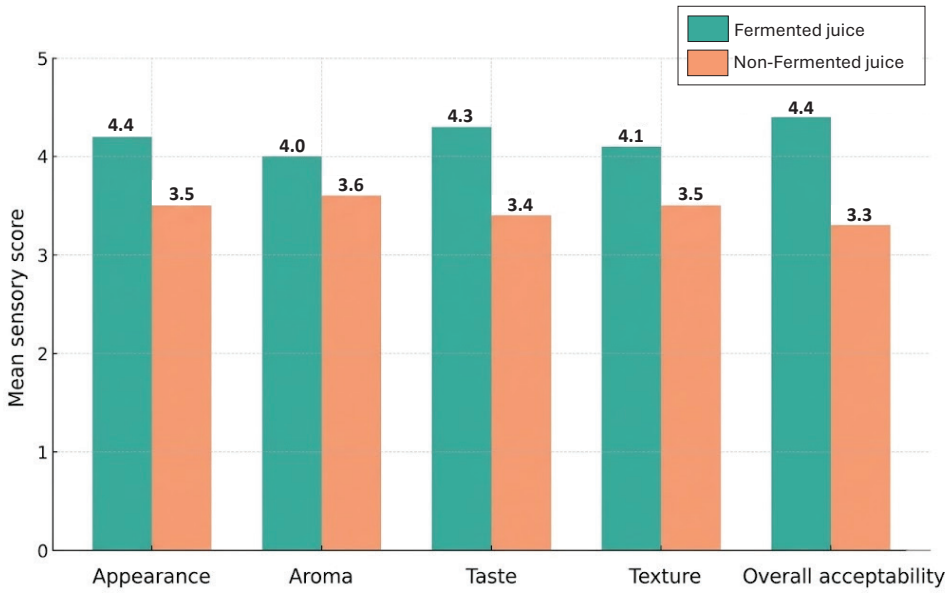


Figure 5. *A histogram was employed to illustrate the sensory evaluation scores of both fermented and non-fermented Persea americana juice. The assessment utilized a 5-point hedonic scale to evaluate appearance, aroma, taste, texture, and overall acceptability. The green bars denote the fermented juice, while the orange bars signify the non-fermented juice. Error bars represent the standard deviations derived from the responses of 15 panelists.*

Biochemical assays

Estimation of reducing sugar by the dinitrosalicylic acid method

The content of reducing sugars (RSC) was determined utilizing the dinitrosalicylic acid (DNS) method as described by Miller and Jain.[29,30] This technique relies on the reduction of 3,5-dinitrosalicylic acid under alkaline conditions by reducing sugars such as glucose and fructose, resulting in the formation of 3-amino-5-nitrosalicylic acid, which exhibits an orange-red color with peak absorbance at 540 nm. The color intensity is directly proportional to the concentration of reducing sugars in the sample. A standard calibration curve was established using D-glucose (Y=0.014x+0.1202, R^2=0.9938), and the reducing sugar content in the samples was calculated and reported in mg/mL. All measurements were conducted in triplicate (N=3) to guarantee accuracy and reproducibility. The results are presented as mean±SD, and the relative standard deviation (RSD) was computed to assess precision. These results are illustrated in Table 1 and Figure 6.

These RSD values indicate good precision, especially since most are under 5%, which is generally acceptable in biochemical assays.

Table 1. Estimation of reducing sugar by dinitrosalicylic acid.

Sample	**O.D.** (540 nm)	**Concentration** (mg/mL)	**SD**	**RSD** (%)
Blank	0	0	0	0
1	0.740	40	0.030	4.05
2	1.345	80	0.040	2.98
3	1.820	120	0.040	2.20
4	2.360	160	0.060	2.54
5	2.860	200	0.060	2.10
Normal juice	0.120	83.50	0.005	4.17
Probiotic juice	0.066	37.36	0.002	3.03

Note: *RSD values indicate good precision, especially since most are under 5%.*

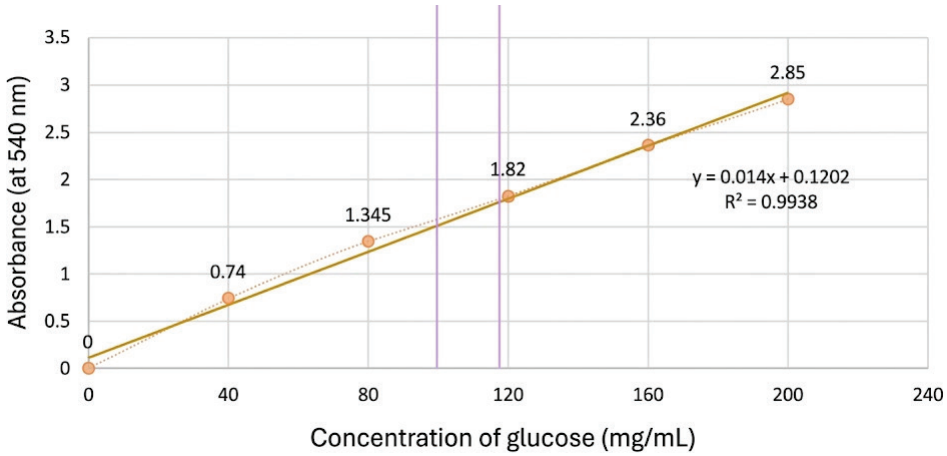


Figure 6. *Estimation of reducing sugar by DNS method.*

Estimation of protein using Lowry's method

The protein content was assessed using the modified Lowry method, which combines the Biuret reaction (formation of a copper–protein complex under alkaline conditions) with the reduction of the Folin-Ciocalteu reagent by aromatic amino acids. This process yields a stable blue-colored complex that can be measured at wavelengths between 650 and 750 nm.[31] A standard calibration curve was established utilizing bovine serum albumin (BSA), and the protein concentration in the samples was calculated accordingly. The findings are illustrated in Table 2 and Figure 7. All measurements were conducted in triplicate (N=3) (technical replicates) to guarantee precision and reproducibility. The results are presented as mean±standard deviation (SD).

Table 2. Estimation of protein by Lowry's method.

Sample	**OD** (540nm)	**Concentration** (mg/mL)	**SD**	**RSD %**
Blank	0	0	0	0
1	0.057	40	0.0025	4.30
2	0.129	80	0.0055	4.25
3	0.182	120	0.0085	4.66
4	0.22	160	0.0102	4.62
5	0.273	200	0.0135	4.94
Normal juice	0.698	493.3571	0.0165	2.37
Probiotic juice	0.964	683.3571	0.0241	2.51

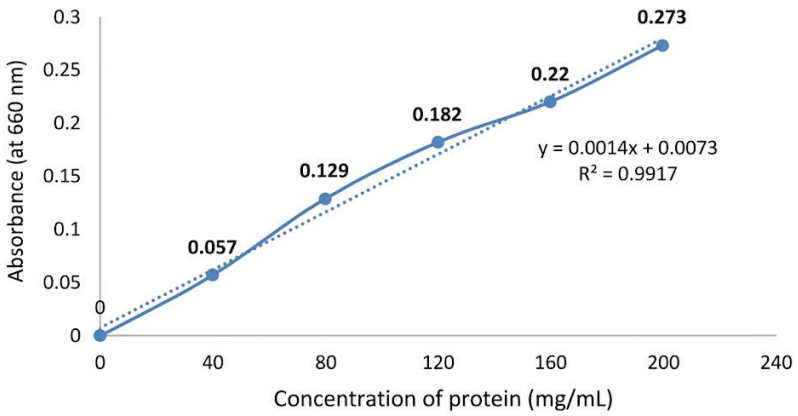


Figure 7. *Standard graph for the estimation of proteins.*

Antioxidant activity

ABTS assay

The antioxidant activity of the samples was assessed using the ABTS [2,2-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)] radical-scavenging assay, measured at 734 nm.[32] This assay relies on the capacity of antioxidants to neutralize the ABTS radical cation, leading to a reduction in absorbance that is proportional to the antioxidant capacity. Various concentrations of the samples (10-50 µg/mL) were evaluated, and the percentage inhibition of ABTS radicals was calculated relative to gallic acid, which served as a standard. The IC_{50} values (the concentration needed to inhibit 50% of radicals) were derived from the inhibition curves. The findings indicated that the fermented juice demonstrated superior antioxidant activity (IC_{50}=17.27 µg/mL) when compared to pasteurized avocado juice (IC_{50}=31.46 µg/mL), and it was similar to the standard gallic acid (IC_{50}=18.73 µg/mL).[33] This increased activity may be due to the presence of bioactive compounds such as phenolics and flavonoids. All experiments were conducted in triplicate (N=3), and the results are presented as mean±SD. The data are presented in Table 3 and Figure 8.

Table 3. ABTS assay of fermented avocado and pasteurized avocado juice samples.

Concentration (µg/mL)	**Gallic acid** (control) Standard	**Fermented avocado juice**	**Pasteurized avocado juice**
10	0.786±0.024	0.597±0.018	0.794±0.024
20	0.610±0.018	0.457±0.014	0.611±0.018
30	0.440±0.013	0.359±0.011	0.493±0.015
40	0.268±0.008	0.263±0.008	0.410±0.012
50	0.129±0.004	0.188±0.006	0.282±0.008
IC_{50} (µg/mL)	**18.73**	**17.27**	**31.46**

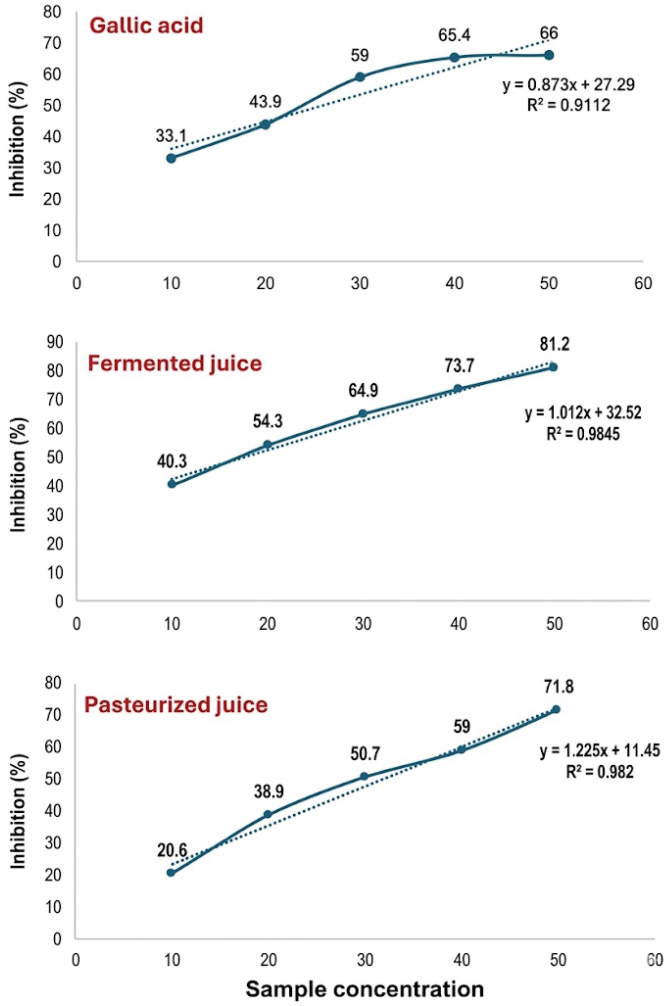


Figure 8. *Antioxidant assay (ABTS method). A: gallic acid (control), B: fermented juice, C: pasteurized juice.*

DPPH radical scavenging assay

The antioxidant properties of the samples were further assessed using the DPPH (2,2-diphenyl-1-picrylhydrazyl) radical-scavenging assay, measured at 517 nm.[33,34] This technique relies on the conversion of the purple-hued DPPH radical into a yellow diphenylpicrylhydrazine in the presence of antioxidants, leading to a reduction in absorbance.

Various concentrations of the samples were evaluated, and the percentage of radical-scavenging activity was calculated relative to gallic acid, which served as the standard. The IC_{50} values (the concentration needed to inhibit 50% of DPPH radicals) were derived from the inhibition curves.

The findings revealed that the fermented juice demonstrated superior antioxidant activity (IC_{50}=24.64 µg/mL) when compared to pasteurized avocado juice (IC_{50}=51.95 µg/mL) and exhibited activity comparable to that of gallic acid (IC_{50}=28.23 µg/mL). The increased antioxidant activity may be due to the presence of bioactive compounds such as phenolics and flavonoids. All experiments were conducted in triplicate (N=3), and the results are presented as the mean±SD. The data are in Table 4 and Figure 9.

Table 4. DPPH aasay of pasteurised juice and fermented *P. americana* juice samples.

Concentration (µg/mL)	**Gallic acid** (control)	**Fermented juice**	**Pasteurized juice**
10	0.728±0.022	0.694±0.021	0.899±0.027
20	0.629±0.019	0.542±0.016	0.772±0.023
30	0.465±0.014	0.421±0.013	0.683±0.020
40	0.332±0.010	0.309±0.009	0.572±0.017
50	0.232±0.007	0.215±0.006	0.553±0.017
IC_{50} (µg/mL)	**28.23**	**24.64**	**51.95**

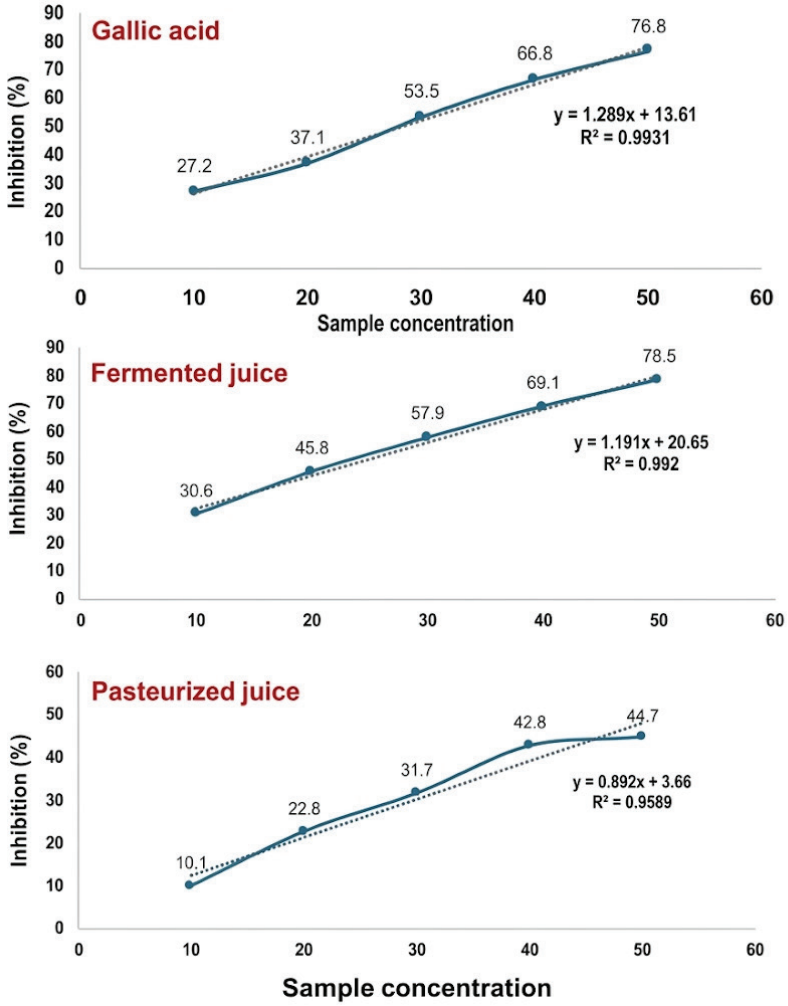


Figure 9. *DPPH assay. A: gallic acid standard, B: fermented juice, C: pasteurized juice.*

Antimicrobial assay

The antibacterial efficacy of fermented and pasteurized *Persea americana* juice was assessed utilizing the agar well diffusion technique.[34] Six pathogenic indicator strains: *E.coli*, *B.cereus*, *S.aureus*, *S.typhimurium*, *S.mutans*, and *P.aeruginosa*, were sourced from the stock cultures of the Dairy Microbiological Laboratory at the National Research Center. Each bacterial strain was activated in tryptone soy broth and incubated at 37 °C for 24 hrs. Approximately 1 mL of the activated culture (≈10^4 CFU/mL) was combined with 20 mL of molten Mueller-Hinton agar (Becton Dickinson) and poured into sterile Petri plates. Following solidification, wells measuring 5 mm in diameter were aseptically created using a sterile cork borer. A volume of 50 µL from each juice sample was introduced into the wells. Penicillin (potassium penicillin, 20%) at concentrations ranging from 50 to 200 mg/mL was used as a positive control, while sterile phosphate buffer served as a negative control. The plates were incubated at 37 °C for 24 hrs. All experiments were conducted in triplicate (N=3) to guarantee reproducibility. The antimicrobial activity was assessed by measuring the zone of inhibition (ZOI, mm) surrounding each well. Results were reported as mean±SD. A larger zone of inhibition signified a more potent antibacterial activity. The data are presented in Table 5 and Figure 10, and statistical analysis was conducted using ANOVA.

Table 5. Antimicrobial activity of pasteurised and fermented *P. americana* juice

Strains	**Control** (mm)	**Pasteurized juice** (mm)	**Fermented juice** (mm)
Salmonella typhi	2.3±0.07	1.2±0.04	1.5±0.05
Pseudomonas aeruginosa	2.8±0.08	1.3±0.04	1.1±0.03
Streptococcus mutans	3.0±0.09	0.0±0.00	0.0±0.00
Staphylococcus aureus	1.8±0.05	2.9±0.09	1.3±0.04
Bacillus cereus	2.5±0.08	2.9±0.09	0.0±0.00
Escherichia coli	4.3±0.13	0.0±0.00	1.3±0.04

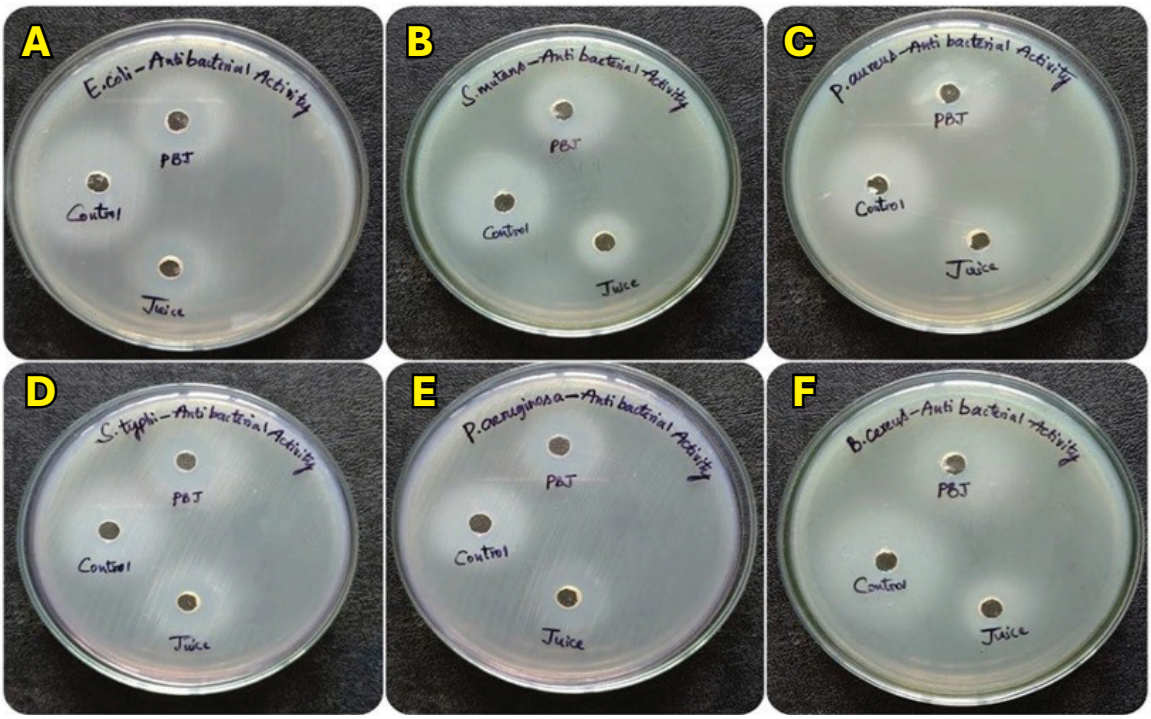


Figure 10. *Antimicrobial property.*
A: E.coli, B: S.mutans, C: S.aureus, D: S.typhi, E: P.aurogenosa, F: B.cereus.

Discussion

This research investigated the feasibility of *Lactobacillus* spp. in pasteurized *Persea americana* (avocado) juice and evaluated its potential as a probiotic medium. The results confirmed that *Lactobacillus* can thrive and proliferate in avocado juice without the need for any specialized growth additives. This aligns with the findings of Anil Panghal *et al.,* who highlighted that fruit juices, rich in dietary fiber, vitamins, minerals, and antioxidants, can serve as effective alternative substrates for the cultivation of probiotics.[35] The successful proliferation of *Lactobacillus* within the juice matrix illustrates the promise of avocado juice as a non-dairy vehicle for probiotics.

Morphological identification via Gram staining validated the existence of Gram-positive rods, which are indicative of *Lactobacillus* species. MRS medium, recognized as a standard growth medium for lactic acid bacteria, facilitated successful cultivation.[36] A variety of environmental factors, including pH, NaCl, and bile salt concentrations, were assessed to identify the conditions most conducive to probiotic survival. The organism exhibited optimal growth at bile salt concentrations of 0.5% and 1.5% NaCl, indicating that the bacteria possess considerable bile and osmotic tolerance, which are crucial for survival within the gastrointestinal tract. Additionally, the analysis of pH ranges showed that pH 8 provided the most favorable conditions for bacterial growth. These findings are consistent with other research indicating that slightly alkaline conditions may be beneficial for certain *Lactobacillus* strains in fruit matrices.

The incubation period significantly influenced bacterial viability. Although 24 hrs of incubation proved inadequate, 48 hrs offered sufficient time for the bacteria to establish and proliferate. The prolonged incubation likely facilitated the accumulation of vital metabolites and adaptation to the juice environment, thereby enhancing colony formation. This observation aligns with findings from other probiotic studies that recommend extended fermentation times to achieve stable growth and functional benefits.

The primary emphasis of the research was on the antioxidant potential of fermented avocado juice. The antioxidant activity was evaluated using DPPH and ABTS spectrophotometric assays, with IC_{50} values indicating that the fermented juice exhibited greater antioxidant capacity than its pasteurized counterpart. This observation supports the findings of Khan *et al.,* who indicated that fermentation frequently enhances the antioxidant potential of foods by generating bioactive compounds such as glutathione, butyrate, and folate.[37] Likewise, Martin *et al*. showed that the fermentation of blueberry juice using a novel bacterium significantly enhanced its antioxidant properties.[38] Additionally, Bujna *et al.* observed increases in antioxidant activity in apricot juice fermented with *Bifidobacterium lactis* and *Lactobacillus acidophilus*.[28] Collectively, these studies underscore the role of probiotic fermentation in enhancing the health-promoting attributes of fruit-based beverages.

Regarding antimicrobial activity, the fermented and pasteurized juice samples did not show any zones of inhibition (ZI) against mutant *Streptococcus* strains in the well diffusion test, whereas the antibiotic control showed significant ZI. This difference is likely due to the greater potency and concentration of synthetic antibiotics compared to the natural metabolites present in fruit juice. Although probiotic bacteria can produce antimicrobial substances such as bacteriocins and organic acids, they may not generate these compounds at sufficient concentrations under the tested conditions to effectively inhibit resistant or mutant bacterial strains.[40] Furthermore, juice matrices may lack the synergistic compounds or cofactors that could enhance antimicrobial efficacy. This finding contrasts with studies where probiotic strains have shown inhibitory effects against pathogens, suggesting that the antimicrobial effect is highly dependent on the specific strain and the conditions under which it is tested.

It is widely recognized that distinct strains of probiotic bacteria possess specific antioxidant characteristics that also contribute to their antibacterial properties. These probiotics are known to generate various metabolites with antioxidant capabilities, including glutathione, butyrate, and folate.[30,39] Martin *et al.* demonstrated that a novel bacterium isolated from the microflora of *Serratia raccinii* fruit significantly enhanced the antioxidant activity of blueberry juice during fermentation.[38] In a similar vein, Bujna *et al*. noted a moderate enhancement in the antioxidant levels of apricot juice fermented with *Bifidobacterium lactis* Bb12 and *Lactobacillus acidophilus* La5.[28] These results emphasize the essential function of *Lactobacillus* and other probiotic strains in the fermentation of fruit juices and highlight their extensive functional applications.[41] Moreover, recent research has indicated that dairy products enriched with probiotics can inhibit the enzymatic functions of amylase and glucosidase, implying a potential hypoglycemic effect in vitro.[42-45]

Another potential explanation for the reduced or nonexistent ZI in the fermented and pasteurized juices may relate to the characteristics of the target organism. The mutant strain of Streptococcus utilized could have an intrinsic resistance to mild antimicrobial agents, which accounts for the larger inhibition zones observed with control antibiotics. Furthermore, as fermentation modifies the juice's pH and metabolite composition, the overall antimicrobial effectiveness may be influenced by fermentation dynamics that were not completely captured within the 48-hr duration of this experiment.

In summary, this research demonstrates that *Lactobacillus spp*. can survive and flourish in pasteurized avocado juice without additional supplementation. The fermentation process significantly enhanced the juice's

antioxidant capacity, although the antimicrobial effects were somewhat limited under the tested conditions. These results indicate that avocado juice is a promising medium for developing probiotic beverages, offering considerable functional advantages. Nonetheless, further refinement of fermentation parameters and selection of strains may improve antimicrobial characteristics and broaden the functional potential of these products. This study adds to the growing body of evidence supporting the use of non-dairy matrices for the delivery of probiotics and the development of functional foods.

Limitations

The study was conducted under laboratory conditions, which may not completely reflect industrial-scale production or real market storage conditions. Only two probiotic strains, *Lactobacillus acidophilus* and *Lactobacillus plantarum*, were evaluated. Therefore, the findings cannot be generalized to all probiotic microorganisms. In sensory evaluation, such as taste, aroma, texture, and consumer acceptability of fermented avocado juice, these aspects were not extensively investigated. The antimicrobial activity observed in the fermented juice was moderate and lower than that of standard antibiotic controls. *In vivo* studies and clinical trials were not conducted; therefore, the direct health benefits of fermented avocado juice in humans require further confirmation.

Conclusion

This research emphasizes the potential of *Persea americana* (avocado) juice as a viable medium for creating a functional probiotic beverage. The effective growth and survival of *Lactobacillus* in pasteurized avocado juice without additional nutrients underscores the juice's suitability as a natural carrier for probiotic cultures. Fermentation significantly improved the juice's antioxidant capacity, as evidenced by DPPH and ABTS assays, and also demonstrated enzyme-inhibitory effects, indicating possible hypoglycemic benefits. Although no zones of inhibition were observed against Streptococcus in antimicrobial evaluations, this may be due to the limited concentrations of bioactive compounds under the tested conditions. The strain exhibited optimal growth at pH 8, 0.5% bile salt, 1.5% NaCl, and after 48 hrs of incubation, confirming its robustness against gastrointestinal-like stress. These results endorse the development of *Lactobacillus*-fermented *P. americana* juice as an innovative, plant-based functional drink with improved nutraceutical properties, presenting promising implications for the food and health sectors.

Ethical approval

Not applicable

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

CRediT authorship contribution statement

Kanchi Ravi Padma: manuscript drafted, supervise; **Lingareddy Nikitha Reddy**: research work; **Sankari Mohan:** supervise.

Acknowledgements

The authors gratefully acknowledge Sri Padmavati Mahila Visvavidyalayam, Tirupati, Andhra Pradesh, India, for constant support and encouragement.

References

[1] Oliveira A, Amaro AL, Pintado M. Impact of food matrix components on nutritional and functional properties of fruit-based products. Curr Opin Food Sci. 2018; 22: 153-9. doi: 10.1016/j.cofs.2018.02.002.

[2] Padilla-Camberos E, Martínez-Velázquez M, Flores-Fernández JM, Villanueva-Rodríguez S. Acute toxicity and genotoxic activity of avocado seed extract (Persea americana Mill., c.v. Hass). Sci World J. 2013; 2013: 245828. doi: 10.1155/2013/245828.

[3] Araújo RG, Rodriguez-Jasso RM, Ruiz HA, Pintado MME, Aguilar CN. Avocado by-products: Nutritional and functional properties. Trends Food Sci Technol. 2018;80:51-60. doi: 10.1016/j.tifs.2018.07.027.

[4] Abubakar ANF, Achmadi SS, Suparto IH. Triterpenoid of avocado (Persea americana) seed and its cytotoxic activity toward breast MCF-7 and liver HepG2 cancer cells. Asian Pac J Trop Biomed. 2017; 7(5): 397-400. doi: 10.1016/j.apjtb.2017.01.012.

[5] Jan T, Negi R, Sharma B, Kour D, Kumar S, Rai AK, et al. Diversity, distribution and role of probiotics for human health: Current research and future challenges. Biocatal Agric Biotechnol. 2023; 52: 102889. doi: 10.1016/j.bcab.2023.102889.

[6] Behrooz AB, Mohammad N. An investigation into the probiotic, anti-bacterial, anti-biofilm, and safety properties of Lactococcus lactis NJ414. LWT. 2024; 201: 116251. doi: 10.1016/j.lwt.2024.116251.

[7] Mantzourani I, Plessas S, Odatzidou M, Alexopoulos A, Galanis A, Bezirtzoglou E, et al.

Effect of a novel Lactobacillus paracasei starter on sourdough bread quality. Food Chem. 2019; 271: 259-65. doi: 10.1016/j.foodchem.2018.07.183.

[8] Wong WY, Chan BD, Leung TW, Chen M, Tai WCS. Beneficial and anti-inflammatory effects of formulated prebiotics, probiotics, and synbiotics in normal and acute colitis mice. J Funct Foods. 2022; 88: 104871. doi: 10.1016/j.jff.2021.104871.

[9] James A, Wang Y. Characterization, health benefits and applications of fruits and vegetable probiotics. CyTA J Food. 2019; 17(1): 770-80. doi: 10.1080/19476337.2019.1652693.

[10] Vitali B, Minervini G, Rizzello CG, Spisni E, Maccaferri S, Brigidi P, et al. Novel probiotic candidates for humans isolated from raw fruits and vegetables. Food Microbiol. 2012; 31(1): 116-25. doi: 10.1016/j.fm.2011.12.027.

[11] Kandylis P, Pissaridi K, Bekatorou A, Kanellaki M, Koutinas AA. Dairy and non-dairy probiotic beverages. Curr Opin Food Sci. 2016; 7: 58-63. doi: 10.1016/j.cofs.2015.11.012.

[12] Ranadheera CS, Vidanarachchi JK, Rocha RS, Cruz AG, Ajlouni S. Probiotic delivery through fermentation: Dairy vs. non-dairy beverages. Fermentation. 2017; 3(4): 67. doi: 10.3390/fermentation3040067.

[13] Ephrem E, Najjar A, Charcosset C, Greige-Gerges H. Encapsulation of natural active compounds, enzymes, and probiotics for fruit juice fortification, preservation, and processing: An overview. J Funct Foods. 2018; 48: 65-84. doi: 10.1016/j.jff.2018.06.010.

[14] Horáčková Š, Rokytová K, Bialasová K, Klojdová I, Sluková M. Fruit juices with probiotics – new type of functional foods. Czech J Food Sci. 2018; 36(4): 284-8. doi: 10.17221/413/2017-CJFS.

[15] Durmaz Y, Kilicli M, Toker OS, Konar N, Palabiyik I, Tamturk F. Using spray-dried microalgae in ice cream formulation as a natural colorant: Effect on physicochemical and functional properties. Algal Res. 2020; 47: 101811. doi: 10.1016/j.algal.2020.101811.

[16] Perricone M, Bevilacqua A, Altieri C, Sinigaglia M, Corbo MR. Challenges for the production of probiotic fruit juices. Beverages. 2015; 1(2): 95-103. doi: 10.3390/beverages1020095.

[17] Mantzourani I, Terpou A, Alexopoulos A, Bezirtzoglou E, Bekatorou A, Plessas S. Production of a potentially synbiotic fermented cornelian cherry (Cornus mas L.) beverage using Lactobacillus paracasei K5 immobilized on wheat bran. Biocatal Agric Biotechnol. 2019; 17: 347-51. doi: 10.1016/j.bcab.2018.12.012.

[18] Anekella K, Orsat V. Optimization of microencapsulation of probiotics in raspberry juice by spray drying. LWT. 2013; 50(1): 17-24. doi: 10.1016/j.lwt.2012.08.003.

[19] Plaza-Diaz J, Ruiz-Ojeda FJ, Gil-Campos M, Gil A. Mechanisms of action of probiotics. Adv Nutr. 2019; 10(Suppl 1): S49-66. doi: 10.1093/advances/nmy063.

[20] Naseem Z, Mir SA, Wani SM, Rouf MA, Bashir I, Zehra A. Probiotic-fortified fruit juices: Health benefits, challenges, and future perspective. Nutrition. 2023; 115: 112154. doi: 10.1016/j.nut.2023.112154.

[21] Alireza V, Behrooz AB, Farideh TY, Seyed AM, Hamid N. Diversity and probiotic potential of lactic acid bacteria isolated from Horreh, a traditional Iranian fermented food. Probiotics Antimicrob Proteins. 2017; 10(2): 258-68. doi: 10.1007/s12602-017-9282-x.

[22] Nagpal R, Kumar A, Kumar M. Fortification and fermentation of fruit juices with probiotic lactobacilli. Ann Microbiol. 2012; 62: 1573-8. doi: 10.1007/s13213-012-0412-5.

[23] Meilgaard M, Civille GV, Carr BT. Sensory evaluation techniques. 5th ed. Boca Raton: CRC Press; 2016. doi: 10.1201/b19493.

[24] Yang E, Fan L, Yan J, Jiang Y, Doucette C, Fillmore S. Antimicrobial activity of bacteriocin-producing Lactobacillus paracasei HD1-7 in Chinese fermented milk. Food Control. 2012; 23(2): 338-44. doi: 10.1016/j.foodcont.2011.07.031.

[25] Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the Folin phenol reagent. J Biol Chem. 1951; 193(1): 265-75. doi: 10.1016/S0021-9258(19)52451-6.

[26] Brand-Williams W, Cuvelier ME, Berset C. Use of a free radical method to evaluate antioxidant activity. LWT. 1995; 28(1): 25-30. doi: 10.1016/S0023-6438(95)80008-5.

[27] Panghal A, Sharma P, Chhikara N, Gat Y. Development of probiotic beverages: A review on the technology, recent trends, and advancements. Nutrire. 2017; 42: 1-12. doi: 10.1186/s41110-017-0021-5.

[28] Bujna E, Nikoletta AF, Anh MT, Mai SD, Quang DN. Lactic acid fermentation of apricot juice by mono- and mixed cultures of probiotic Lactobacillus and Bifidobacterium strains. Food Sci Biotechnol. 2018; 27: 547-55. doi: 10.1007/s10068-017-0291-y.

[29] Miller GL. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem. 1959; 31(3): 426-8. doi: 10.1021/ac60147a030.

[30] Jain A, Jain R, Jain S. Quantitative analysis of reducing sugars by 3, 5-dinitrosalicylic acid (DNSA method). In: Basic Techniques in Biochemistry, Microbiology and Molecular Biology. New York: Humana; 2020: pp 181-3. doi: 10.1007/978-1-4939-9861-6_20.

[31] Wilson K, Walker J. Principles and techniques of practical biochemistry. 7th Ed. Cambridge: Cambridge University Press; 2010. doi: 10.1017/CBO9780511841477.

[32] Castro-López C, Bautista-Hernández I, González-

Hernández MD, Martínez-Ávila GCG, Rojas R, Gutiérrez-Díez A, et al. Polyphenolic profile and antioxidant activity of leaf purified hydroalcoholic extracts from seven Mexican Persea americana cultivars. Molecules. 2019; 24(1): 173. doi: 10.3390/molecules24010173.

[33] Melgar B, Dias MI, Ciric A, Sokovic M, Garcia-Castello EM, Rodriguez-Lopez AD, et al. Bioactive characterization of Persea americana Mill. by-products: A rich source of inherent antioxidants. Ind Crops Prod. 2018; 111: 212-8. doi: 10.1016/j.indcrop.2017.10.027.

[34] Fouad MT, Moustafa A, Hussein L, Romeilah R, Gouda M. In-vitro antioxidant and antimicrobial activities of selected fruit and vegetable juices and fermented dairy products commonly consumed in Egypt. Res J Pharm Biol Chem Sci. 2015; 6(2): 541-50.

[35] Panghal A, Janghu S, Virkar K, Gat Y, Kumar V, Chhikara N. Potential non-dairy probiotic products – A healthy approach. Food Biosci. 2017; 21: doi: 10.1016/j.fbio.2017.12.003.

[36] Lee HM, Lee YA. A differential medium for lactic acid producing bacteria in a mixed culture. Lett Appl Microbiol. 2008; 46(6): 676-81. doi: 10.1111/j.1472-765X.2008.02371.x.

[37] Khan A, Goutam S, Rajat KP. An approach for reduction of false predictions in reverse engineering of gene regulatory networks. J Theor Biol. 2018; 445: 9-30. doi: 10.1016/j.jtbi.2018.02.015.

[38] Martin LJ, Matar C. Increase of antioxidant capacity of the lowbush blueberry (Vaccinium angustifolium) during fermentation by a novel bacterium from the fruit microflora. J Sci Food Agric. 2005; 85(9): 1477-84. doi: 10.1002/jsfa.2142.

[39] Food and Agriculture Organization and World Health Organization. Evaluation of health and nutritional properties of powder milk and live lactic acid bacteria [Internet]. Córdoba: FAO/WHO; 2001 [cited 2005 Sep 8]. Available from: ftp://ftp.fao.org/es/esn/food/probio_report_en.pdf

[40] Mostafa RJ, Behrooz AB, Morteza T, Mohammad Ali H, Omer ST. Evaluation of the probiotic, anti-microbial, anti-biofilm, and safety properties of Levilactobacillus brevis Lb13H. LWT. 2024; 207: 116636. doi: 10.1016/j.lwt.2024.116636.

[41] Shah N, Prajapati JB. Effect of carbon dioxide on sensory attributes, physico-chemical parameters and viability of probiotic L. helveticus MTCC 5463 in fermented milk. J Food Sci Technol. 2013; 51(12): 3886-93. doi: 10.1007/s13197-013-0941-6.

[42] Sivudu SN, Umamahesh K, Reddy OVS. A comparative study on probiotication of mixed watermelon and tomato juice by using probiotic strains of lactobacilli. Int J Curr Microbiol Appl Sci. 2014; 3(11): 977-84.

[43] Anthony F, Edmond R. Perspective: Reductionist nutrition research has meaning only within the framework of holistic and ethical thinking. Adv Nutr. 2018; 9(6): 655-70. doi: 10.1093/advances/nmy044.

[44] Gueimonde M, Margolles A, de los Reyes-Gavilán CG, Salminen S. Competitive exclusion of enteropathogens from human intestinal mucus by Bifidobacterium strains with acquired resistance to bile-a preliminary study. Int J Food Microbiol. 2004; 91(3): 253-60. doi: 10.1016/S0168-1605(03)00371-1.

[45] Shaghayegh E, Mohammad N, Behrooz AB, Salam AI. Assessment of the probiotic, anti-bacterial, and anti-biofilm characteristics of Lacticaseibacillus rhamnosus CWKu-12, along with its potential impact on the expression of virulence genes in Listeria monocytogenes ATCC 19115. LWT. 2024; 210: 116815. doi: 10.1016/j.lwt.2024.116815.

Journal of Associated Medical Sciences 2026; 59(3): 285-293

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



Genotyping-based prediction of Diego blood group phenotypes: Implications for Thai blood donors

Sarisa Chidtrakoon[1], Wiradee Sasikarn[2], Tanaporn Choychimplee[2], Kamphon Intharanut[2], Oytip Nathalang[2*]

[1]Blood Bank, Thammasat University Hospital, Pathum Thani Province, Thailand.
[2]Graduate Program in Biomedical Sciences, Faculty of Allied Health Sciences, Thammasat University, Pathum Thani Province, Thailand.

ARTICLE INFO

Article history:
Received 7 May 2026
Accepted as revised 14 July 2026
Available online 20 July 2026

Keywords:
Diego blood group system, genotyping, predicted phenotypes, Thais

ABSTRACT

Background: The Di^a and Di^b antigens of the Diego (DI) blood group system can cause alloimmunization following transfusion or pregnancy. Although the Di(a-b+) phenotype predominates, the Di(a+) antigen is relatively common in Asian populations. Due to serological limitations, particularly the lack of anti-Di^b, genotyping is increasingly used to predict DI phenotypes and facilitate identifying antigen-negative blood donors for patients with anti-Di^a or anti-Di^b.

Objectives: The study aimed to determine the frequencies of *DI*A* and *D*B* alleles to predict DI phenotypes and establish an inventory of Thai blood donors with the rare Di(a+b-) phenotype.

Materials and methods: A total of 3,500 blood samples from unrelated healthy Thai blood donors were included. All samples were tested with anti-Di^a using the conventional tube technique and genotyped for *DI*A* and *DI*B* alleles by in-house polymerase chain reaction with sequence-specific primer technique. The DI allele frequencies were calculated to estimate the risk of alloimmunization.

Results: The DI genotyping and phenotyping results showed 100% concordance. Among 3,500 Central Thai blood donors, predicted phenotypes were ranked as Di(a-b+), 3,407 (97.34%), followed by Di(a+b+), 92 (2.63%), and Di(a+b-), 1 (0.03%). The *DI*A* and *DI*B* allele frequencies were 0.013 and 0.987. The *DI*A* and *DI*B* frequencies are consistent with those observed in Central and Northern Thai populations and in other Southeast Asian populations, but differ significantly from those observed in Southern Thai, Japanese, Chinese, and Korean populations (p<0.001). Notably, one donor with the rare Di(a+b-) phenotype was observed in this study; this likely reflects the greater likelihood of detecting rare phenotypes when screening a much larger sample size than in related Thai studies.

Conclusion: The implementation of the PCR-SSP technique plays a crucial role in identifying DI blood group antigens, particularly in situations where specific antisera are unavailable. This approach also contributes to establishing an inventory of donors with rare blood phenotypes, thereby enhancing transfusion safety.

* **Corresponding contributor.**
Author's Address: *Graduate Program in Biomedical Sciences, Faculty of Allied Health Sciences, Thammasat University, Pathum Thani Province, Thailand.*
E-mail address: oytip.n@allied.tu.ac.th
doi: 10.66285/JAMS.2026.101
E-ISSN: 2539-6056

Introduction

The Diego (DI) blood group system is clinically important in transfusion practice. The International Society of Blood Transfusion (ISBT) has assigned 23 antigens located in the integral band 3 glycoprotein of the red blood cell (RBC) membrane.[1] Two closely related antigen pairs of the DI system are Di^a/Di^b and Wr^a/Wr^b

Of these, Di^b and Wr^b are high-prevalence antigens, while Di^a and Wr^a are low-prevalence antigens.[2-4] Importantly, the Di(a**+**) antigen is relatively common in Asian populations**.** Its frequency ranges from 2 to 12% in Eastern Asian populations**.**[4] Regarding the data on three DI phenotypes among populations, Di(a–b+) is the most frequent**.** This is followed by Di(a+b+), while Di(a+b+) is a rare phenotype**.** In Thai donor populations, the predicted Di(a+) antigen frequency of only Di(a+b+) phenotype was 3**.**70% among Central and Northern Thais and 0.94% among Southern Thais**.**[5-8] Although anti-Di^a and anti-Di^b are involved in mild to severe cases of hemolytic disease of the fetus and newborn (HDFN) and hemolytic transfusion reactions (HTRs),[2,3] anti-Di^a is more frequently identified among pregnant women and patients with repeated transfusions. This is due to the high prevalence of the Di(a-b+) phenotype across various populations.[2-4,9-11]

In general, to determine Di(a+) and Di(a-) phenotypes, RBCs are tested with anti-Di^a, using the conventional tube technique (CTT) or column agglutination technique (CAT). The phenotypic differentiation between Di(a+b+) and Di(a+b-) is inconclusive with CTT or CAT due to the lack of an anti-Di^b. It is known that the Di^a and Di^b antigens are encoded by the *DI*A* and *DI*B* alleles, respectively, which arise from a single nucleotide variant (SNV), c.2561C>T in exon 19 of the human erythrocyte membrane anion-transporter gene (*SLC4A1*). This polymorphism results in a p.(Pro854Leu) amino acid substitution that distinguishes the Di^a and Di^b antigens. In recent years, molecular blood group typing has become an important tool in transfusion medicine. Related studies from our group in Thailand demonstrated that DI genotyping accurately predicts Di^a and Di^b antigens in the absence of commercially available anti-Di^b reagent, enabling discrimination between the Di(a+b+) and rare Di(a+b-) phenotypes that cannot be distinguished by anti-Di^a serology alone. Hence, different genotyping techniques have been used to determine *DI*A* and *DI*B* alleles to predict accurate DI phenotypes among donors and patients, e.g., PCR with sequence-specific primer (PCR-SSP),[5-8,12] real-time PCR,[13] high-resolution melting curve analysis (HRM),[13,14] microarrays[10,11] and DNA sequencing.[5-8,11-14]

In one patient with Di(a-b+) phenotypes receiving Di(a+) from either Di(a+b+) or Di(a+b-) blood donors, anti-Di^a was produced, making a compatible blood transfusion easy. In contrast, for Di(a+b-) patients who have developed anti-Di^b, finding compatible antigen-negative blood donors can be extremely difficult, potentially leading to significant delays in transfusion support. Hence, correctly detecting DI blood group antigens, particularly among patients receiving chronic transfusion therapy, is essential to prevent further problems of incompatible transfusions and to ensure the appropriate use of blood units with low-prevalence phenotypes; thereby, promoting safe and effective transfusion support.

In Thailand, the availability of blood donors with the rare Di(a+b-) phenotype remains unclear because related studies were limited by the unavailability of anti-Di^b reagent and relatively small sample sizes. Expanding the study population enables a more precise estimation of DI allele and phenotype frequencies and increases the likelihood of identifying individuals with the rare Di(a+b-) phenotype. Establishing an inventory of these rare donors would facilitate timely provision of compatible blood for patients with anti-Di^b; thereby, improving transfusion support for this clinically challenging patient group.[5]

Therefore, this study aimed to determine the frequencies of the *DI*A* and *DI*B* alleles to predict DI phenotypes and establish an inventory of Thai blood donors possessing the rare Di(a+b-) phenotype**.**

Materials and methods

Sample collection

Leftover EDTA-anticoagulated blood samples were collected from 3,500 unrelated, healthy Thai donors at the Blood Bank, Thammasat University Hospital, Pathum Thani, Thailand. Written informed consent was obtained from each subject. The sample size was determined using a single-population proportion formula based on the prevalence of the Di(a+b-) phenotype among Thais (0.6%).[15] Assuming a 95% confidence level and a margin of error of 0.08%, the minimum required sample size was estimated to be 3,500 blood donors. Inclusion criteria for each subject include Thai ethnicity and meeting the donor selection criteria according to the Guidelines of the National Blood Centre, Thai Red Cross Society.[16] Donors with positive infectious marker testing were excluded.

DNA standards

Known DNA samples comprising 1 Di(a+b-), 5 Di(a+b+) and 1 Di(a-b+) phenotypes confirmed by DNA sequencing in an earlier study,[5] were used as controls.

ABO blood group and Di^a antigen testing using serological techniques

All blood samples were tested for ABO cell and serum grouping by CAT using an ORTHO VISION Analyzer (Ortho-Clinical Diagnostics, Raritan, NJ, USA). Di^a antigen detection was performed using the CTT. Briefly, 1 drop of monoclonal human anti-Di^a (National Blood Centre, Thai Red Cross Society, Bangkok, Thailand) was added, followed by 1 drop of a 3-5% red cell suspension in 0.9% normal saline. The test tube was mixed and incubated at room temperature for 10 minutes. After centrifugation, the reactions were examined macroscopically, and agglutination strength was graded as 4+, 3+, 2+, 1+, weakly positive (w+) or negative (0). Negative, Di(a-b+) and positive, Di(a+b+) control cells were also tested in parallel.

DI*A and DI*B allele genotyping

Genomic DNA was extracted from peripheral blood samples using the DNeasy® Blood & Tissue Kit (QIAGEN, Hilden, Germany), according to manufacturer instructions. DNA samples were further stored at -20 °C until use for genotyping. Genotyping of *DI*A* and *DI*B* alleles was performed using an in-house PCR-SSP technique, with modifications.[5] Primers and amplification conditions were identical to a related study.[5] Briefly, 1 µL of genomic DNA (50 ng/µL) was amplified in a total volume of 10 µL using 5 µM of DI-AB-F forward primer 5'-GGTGTGATAGGCACTGACCC-3', 1 µL and 5 µM DI-A-R reverse primer 5'-GGGCCAGGG AGGCCA-3', 1 µL for *DI*A* allele detection and 5 µM of DI-AB-F forward primer, 1 µL for DI-B-R reverse primers 5'-GGGCCAGGGAGGCCG-3', and 1 µL for *DI*B* allele detection. Co-amplification of the *HGH* gene using 6 µM HGH forward primer 5'-TGCCTTCCCAA CCATTCCCTTA-3', 1 µL, and 6 µM HGH reverse primer 5'-CCACTCACGGATTTCTGTTGTGTTTC-3', 1 µL was run as the internal control and 5 µL of 2X PCR reaction mixture (Green Hot Start, Biotechrabbit, Berlin, Germany) in the thermal cycler (Applied Biosystems™ VeritiPro™).

PCR was performed with an initial denaturation at 95 °C for 30 seconds, followed by 10 cycles of 95 °C for 30 seconds and 69 °C for 10 seconds, then 20 cycles of 95 °C for 10 seconds, 62 °C for 50 seconds, and 72 °C for 30 seconds, with a final extension at 72 °C for 5 minutes and a 4 °C hold. PCR products were separated using electrophoresis on a 1.5% agarose gel in 1x TBE buffer at 100 V and visualized under a blue-light transilluminator. The expected amplicon size was 130 bp for both the *DI*A* and *DI*B* alleles, with a 434-bp HGH gene fragment serving as the internal control.

Known DNA samples of Di(a+b-), Di(a+b+) and Di(a-b+) phenotypes were used as controls. A total of 3,500 DNA samples of Thai blood donors were tested for the detection of the *DI*A* and *DI*B* alleles using PCR-SSP.

DNA sequencing

To confirm the PCR-SSP results, DNA samples from 50 genotyped blood donors (1 *DI*A/A*, 4 *DI*A/B,* and 45 *DI*B/B*) were sequenced using U2Bio DNA Sequencing Services (Bangkok, Thailand). A 598-bp fragment encompassing the c.2561C>T SNP was amplified using forward and reverse primers, identical to those used in a related study under the same PCR conditions used for DI genotyping.[5]

Statistical analysis

Gene frequencies were calculated using the gene counting method. The Chi-square (χ^2) test was used to evaluate whether the observed genotype frequencies agreed with the expected ones under the Hardy-Weinberg equilibrium. The χ^2 test of homogeneity was used to determine the difference between Central Thai (this study) and those reported in Thai and other Asian populations,[5-8] including Han Chinese (Shanghai),[17] Chinese (Shen-Zhen),[18] Chinese (Chengdu),[19] Japanese,[10] Korean,[10] Southeast Asian[10] and Filipino,[10] respectively. All statistical analyses were conducted using IBM SPSS Statistics, Version 26.0 (SPSS Inc., Chicago, IL, USA). A p=0.05 or less was considered significant.

In addition, the percentage of predicted Di^a/Di^b incompatibilities was calculated using the formula: [*DI*A/A* (*DI*B/B*+*DI*A/B*) + *DI*B/B* (*DI*A/A*+*DI*A-B*)]; *DI*A/A*, *DI*A/B* and *DI*B/B* representing the genotype frequencies in each population. The estimated risk of Di^a/Di^b alloimmunization was determined by multiplying the probability of being predicted by the Di^a/Di^b negative phenotype frequency by the probability of having a predicted Di^a/Di^b phenotype frequency,[18] i.e., anti-Di^a or anti-Di^b, by considering the Di^a/Di^b pair of antigens under four conditions described below:

1. Di^a/Di^b donor to Di^b/Di^b patient (risk of anti-Di^a production): Di^a/Di^b x Di^b/Di^b
2. Di^a/Di^a donor to Di^b/Di^b patient (risk of anti-Di^a production): Di^a/Di^a x Di^b/Di^b
3. Di^b/Di^b donor to Di^a/Di^a patient (risk of anti-Di^b production): Di^b/Di^b x Di^a/Di^a
4. Di^a/Di^b donor to Di^a/Di^a patient (risk of anti-Di^b production): Di^a/Di^b x Di^a/Di^a

Overall, the global risk of alloimmunization associated with Di^a/Di^b incompatibility is calculated as the sum of the risks across the four conditions.

Results

Participants data

Among 3,500 central Thai blood donors, ages ranged from 18 to 65 years. For the ABO blood group, group O was the most common (1,347, 38.49%) followed by group B (1,118, 31.94%), group A (760, 21.71%), and group AB (275, 7.86%).

Serological testing for Di^a antigen

The distribution Di^a antigen among 3,500 blood donor samples was analyzed. Of these, 93 donors (2.66%) were Di(a+), while 3,407 donors (97.34%) were Di(a-).

Genotyping results of DI*A and DI*B alleles

Genotyping of *DI*A and DI*B* alleles using an in-house PCR-SSP was successfully performed among all participants. Genotypes were determined by the specific product size, as described in the materials and methods section. The genotype results in this study demonstrated that the PCR product sizes corresponded to the three expected DI genotypes *DI*A/A*, *DI*A/B,* and *DI*B/B*, as shown in Figure 1.

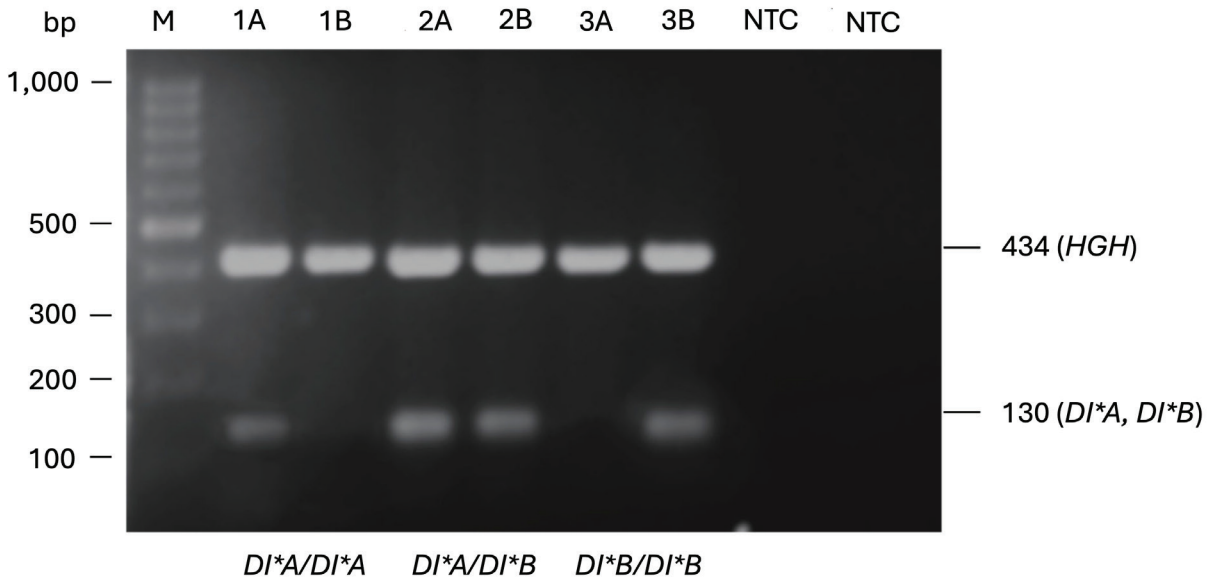


Figure 1. *A representative gel showing Diego blood group genotyping using the PCR-SSP technique. The 434 bp amplification product of the HGH control primer was present in all lanes, confirming successful amplification. Genotypes were determined based on the presence or absence of amplification products specific to alleles of DI*A(A) and DI*B (B). Lanes 1A-1B: DI*A/A, Lanes 2A-2B: DI*A/B, Lanes 3A-3B: DI*B/B, NTC: nontemplate control, Lane M: 100 bp ladder marker (Fermentas, Carlsbad, CA, USA). Arrows indicate the size of DI gene fragments (right): DI*A and DI*B (130 bp), and HGH internal control (434 bp).*

*DI*A* and *DI*B* genotyping results using PCR-SSP among 3,500 donor samples with known Di(a+) and Di(a-) status were used to predict DI phenotypes and agreed with the phenotyping results using serological testing. In addition, the DI genotypes were ranked in order as *DI*B/B* (3,407/3,500, 97.34%), *DI*A/B* (92/3,500, 2.63%), *and DI*A/A* (1/3,500, 0.03%). The determined *DI*A and DI*B* genotypes among 3,500 Thai blood donors were consistent with the Hardy-Weinberg equilibrium (χ^2=0.045, *p*=0.830). Allele frequencies of *DI*A and DI*B* among Central Thai blood donors in this study were 0.013 (94/7,000) and 0.987 (6,906/7,000), respectively (Table 1).

Table 1. *DI*A* and *DI*B* genotype and allele frequencies in Central Thai donor populations (N=3,500).

Allele	**Allele frequency**	**Genotype**	**Observed** (%)	**Expected** (HWE)	χ^2	***p*** (DF=1)
*DI*A*	94 (0.013)	*DI*A/A*	1	1		
*DI*B*	6,906 (0.987)	*DI*A/B*	92	90	0.045	0.830
		*DI*B/B*	3,407	3,409		

Note: DF: Degree of freedom, HWE: Hardy-Weinberg equilibrium.

Moreover, the in-house PCR-SSP assay was validated against serological DI phenotyping, the gold standard for Di[a] antigen determination, and demonstrated complete agreement for all tested samples, yielding 100% sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy. Furthermore, validation of the assay using fifty randomly selected samples representing the *DI*A/A, DI*A/B*, and *DI*B/B* genotypes demonstrated 100% concordance between the PCR-SSP results and DNA sequencing (Figure 2).

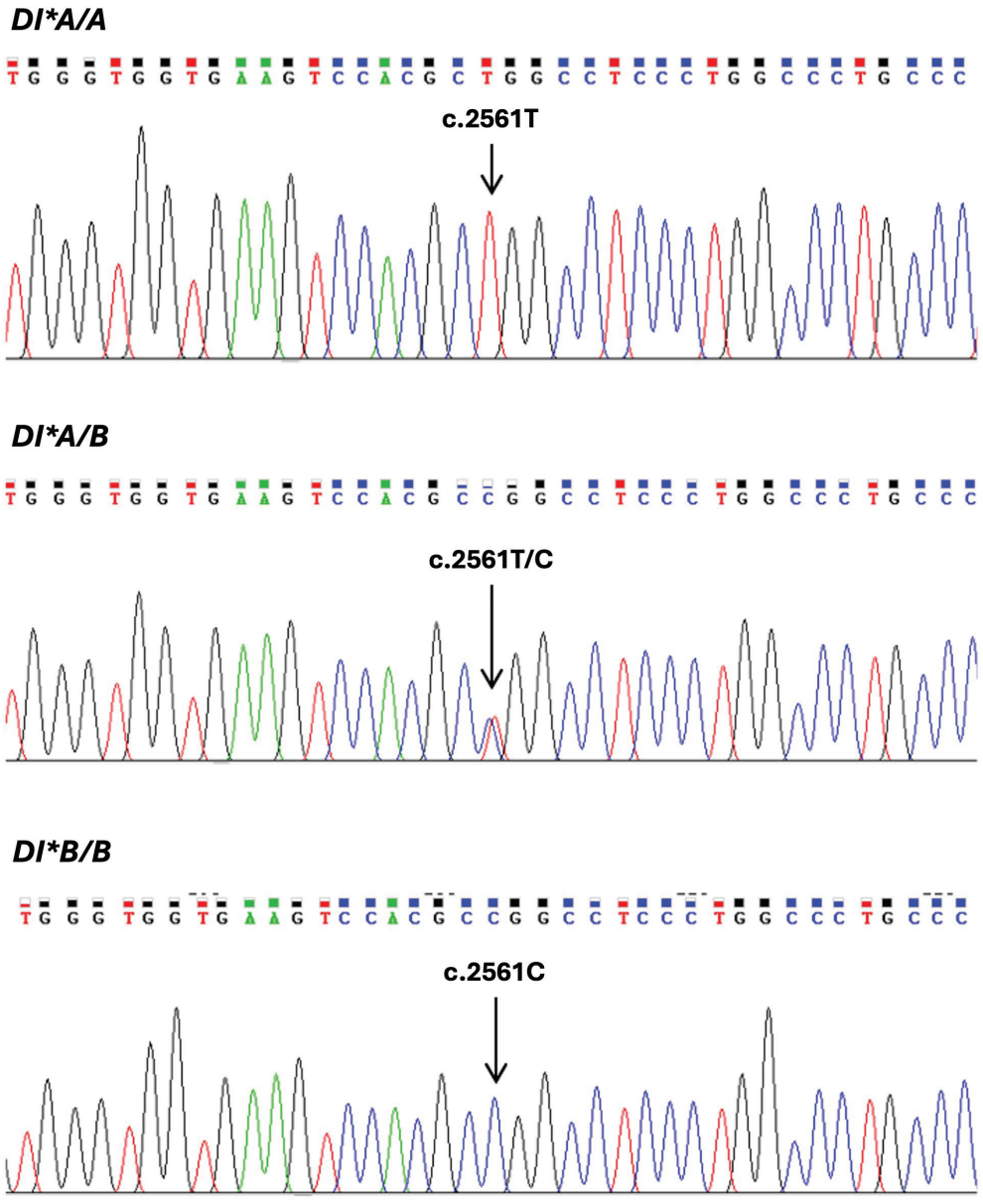


Figure 2. *Electropherograms of the DI gene at the DI*A and DI*B polymorphism (SNP c.2561C>T). DNA sequences shown are 1: DI*A homozygote, 2: DI*A/B heterozygote, 3: DI*B homozygote.*

The *DI*A* and *DI*B* genotype and allele frequencies among 3,500 Central Thai blood donors, as well as other Thai and Asian populations, are presented in Table 2. No significant difference was observed between Central Thai (this study) and the other three Thai populations, including Central Thai,[5] Northern Thai,[6] Northeastern Thai,[7] and Southeast Asian[10] populations (p>0.05). Significant differences in *DI*A* and *DI*B* frequencies were observed when comparing Central Thai blood donors with Southern Thai,[8] Filipino,[10] Han Chinese (Shanghai),[17] Chinese (Shen-Zhen),[18] Chinese (Chengdu),[19] Japanese[10] and Korean[10] populations (p<0.05).

Table 2. *DI*A* and *DI*B* allele frequencies in Asian populations.

Population	Number	Allele frequencies		χ^2	*p* value
		*DI*A*	*DI*B*		
Central Thai*	3,500	94 (0.013)	6,906 (0.987)	-	-
Central Thai[5]	1,011	37 (0.018)	1,985 (0.982)	2.600	0.1068
Northern Thai[6]	300	11 (0.018)	589 (0.982)	0.976	0.3233
Northeastern Thai[7]	345	4 (0.006)	686 (0.994)	2.907	0.0882
Southern Thai[8]	427	4 (0.005)	850 (0.995)	4.724	<0.05
Filipino[10]	1,333	21 (0.008)	2,645 (0.992)	5.062	<0.05
Southeast Asian[10]	942	28 (0.015)	1,856 (0.985)	0.225	0.6351
Han Chinese (Shanghai)[17]	403	36 (0.045)	770 (0.955)	43.063	<0.001
Chinese (Shen-Zhen)[18]	1,766	126 (0.036)	3,406 (0.964)	56.799	<0.001
Chinese (Chengdu)[19]	300	36 (0.060)	564 (0.940)	71.292	<0.001
Japanese[10]	1,022	88 (0.043)	1,956 (0.957)	70.410	<0.001
Korean[10]	1,033	120 (0.058)	1,946 (0.942)	138.014	<0.001

Note: *Central Thai in this study compared to other populations.

The Di^a/Di^b incompatibilities were predicted among Central Thais in this study with related reports in Thai populations (Table 3). The Di^a/Di^b incompatibilities among Central Thais were 2.62%, which were closely related to those of Central Thais[5] (3.53%) and Northern Thais[6] (3.54%). The lowest percentage of predicted Di^a/Di^b incompatibilities was observed in Southern Thais[8] (0.93%), followed by Northeastern Thais (1.15%).[7] In addition, the risk of Di^a and Di^b alloimmunization among different Thai populations was analyzed (Table 3). An important finding of our study was that, according to the DI genotype data, the estimated risk of Di^a alloimmunization in Central and Northern Thai populations was about 2.8-3.8-fold higher than in Southern Thai populations (Table 4). Importantly, the estimated risk of Di^b alloimmunization was very low (0.0003) in the Central Thai population, whereas its risk was 0.00 in other Thai populations.

Table 3. Estimated risks of Di^a and Di^b incompatibility and alloimmunization in Thai populations.

Population	Number	Predicted Di^a Phenotype		Di^a incompatibilities (%)	Alloimmunization risks	
		Negative	Positive		Di^a	Di^b
Central Thai*	3,500	0.973	0.027	2.62	0.026	0.0003
Central Thai[5]	1,011	0.963	0.037	3.53	0.035	0.0000
Northern Thai[6]	399	0.964	0.036	3.54	0.035	0.0000
Northeastern Thai[7]	345	0.988	0.012	1.15	0.012	0.0000
Southern Thai[8]	427	0.991	0.009	0.93	0.009	0.0000

Note: *this study.

Table 4. Predicted DI phenotypes according to ABO blood group among 3,500 blood donors.

DI phenotype	ABO blood group (%)			
	O	B	A	AB
Di(a+b-)	1 (0.07)	0 (0.00)	0 (0.00)	0 (0.00)
Di(a+b+)	32 (2.38)	37 (3.31)	14 (1.84)	9 (3.27)
Di(a-b+)	1,314 (97.55)	1081 (96.69)	746 (98.16)	266 (96.73)
Total	1,347 (100.00)	1,118 (100.00)	760 (100.00)	275 (100.00)

Additionally, the distribution of ABO blood groups and the frequency of Di(a+b+) phenotypes were evaluated among 3,500 blood donors. The highest frequency of Di(a+b+) phenotype was observed in group B (37/1,118, 3.31%), followed by group AB (9/275, 3.27%), group O (33/1,347, 2.45%), and group A (14/760, 1.84%). Although the frequency of the Di(a+b+) phenotype was higher among individuals with blood group B compared with all other blood groups combined (OR=1.41, 95%CI: 0.92-2.16, p=0.11), and lower in those with blood group A (OR=0.64, 95%CI: 0.36-1.13, p=0.13), these differences were not statistically significant.

Discussion

Because the Di(b+) antigen is highly prevalent across populations, anti-Di^a is more frequently detected than anti-Di^b among pregnant women and multi-transfused patients.[3,4,12] In a study conducted at the National Blood Centre, Thai Red Cross Society (NBC-TRC), only anti-Di^a was identified in 1,766 multi-transfused Thai patients (1.20%).[21] A severe case of delayed HTRs and a case series of HDFN caused by anti-Di^b were reported in Japanese populations.[22,23] Recently, a case reported by the NBC-TRC was observed in which a mother with panagglutinating antibodies was ultimately confirmed to have anti-Di^b. She delivered a newborn with a Di(a+b+) phenotype, without evidence of anemia or jaundice (unpublished data). Hence, anti-Di^b among Thai patients may be reported as an unidentified antibody because it agglutinates all reagent RBCs in the antibody identification panel, necessitating the use of additional panel cells with the Di(a+b-) phenotype for accurate identification.

Commercial anti-Di^b reagent is unavailable to distinguish between the Di(a+b+) and Di(a+b-) phenotypes; therefore, genotyping is recommended and has been widely used in population studies.[3,5-8,10-19] Related studies in Thai populations have reported only the Di(a+b+) phenotype among individuals with Di(a+), whereas Di(a+b-) was not observed,[5-8] possibly due to limited sample sizes. Larger studies are required to better assess the risk of alloimmunization and to find matched blood donors for alloimmunized patients. A related study has demonstrated that DI genotyping using the in-house PCR-SSP assay showed complete concordance with serological phenotyping and DNA sequencing results. Furthermore, this method is simple, convenient, and cost-effective compared with commercially available PCR-based assays.[5]

In this study, we performed DI genotyping using in-house PCR-SSP in 3,500 known Di^a antigen phenotype samples; the predicted DI phenotypes revealed 100% concordance with the phenotyping results. The validity of in-house PCR-SSP for DI genotyping was consistent with the determined genotypes by DNA sequencing. Importantly, one homozygous *DI*A/A* group O blood donor was found among 3,500 blood donors. The probability of finding an extended-match donor with the Di(a+b-) phenotype for a patient developing anti-Di^b is 0.03%. Identifying this phenotype should be incorporated in a rare donor registry to enable rapid identification and recall when compatible blood is required. Molecular characterization of rare donors can complement conventional serological screening, particularly where anti-Di^b reagents are unavailable, and facilitate the provision of compatible RBC units for patients with anti-Di^b. Therefore, mass donor screening may help expand the rare donor registry and improve the availability of compatible blood for future transfusion support in the Thai population.

A related study conducted by the NBC-TRC reported the frequencies of Di^a and Di^b antigens among 5,285 Thai blood donors included in the rare blood type inventory, as well as in reagent screening and panel cells.[24] The results showed that 11 donors exhibited the Di(a+b–) phenotype (0.21%). However, the samples were not randomly selected, so the data could not accurately assess the risk of alloimmunization.

No significant differences were found in *DI*A* and *DI*B* frequencies between Central Thai (this study) and the other three Thai populations: Central Thai,[5] Northern Thai[5,] and Northeastern Thai.[7] However, significantly lower frequencies were observed only among Southern Thais,[8] resulting in an increased risk of Di^a incompatibility in Central Thai populations. Similar significant differences in allele frequencies were observed when comparing Central Thai blood donors with Filipino, Chinese, Japanese, and Korean populations (p<0.05).[10,17-19]

Furthermore, the prevalence of the Di(a+b+) phenotype among donors was analyzed by ABO blood group. No statistically significant differences were

observed between groups, suggesting no evidence of an association between the Di(a+) phenotype and ABO blood group in this study.

Limitations

This study encountered several limitations. Firstly, only one donor with the rare Di(a+b-) phenotype was identified, limiting the ability to draw broader conclusions about the prevalence and characteristics. Secondly, this constituted a single-center study including only Central Thai blood donors, which may limit the generalizability of the findings to other regions of the country. Thirdly, DNA sequencing was performed to validate the PCR-SSP results in only a subset of samples rather than the entire study population. Finally, the study did not evaluate clinical outcomes such as the impact of DI genotyping on transfusion management or patient outcomes. Therefore, multicenter studies involving larger and more geographically diverse donor populations, together with comprehensive molecular validation and clinical outcome assessments, are warranted to confirm and extend these findings.

Conclusion

In this study, the PCR-SSP technique was employed, which is particularly valuable to identify Di^b antigen status in the absence of specific antisera; thereby, facilitating accurate identification of rare Di(a+b-) phenotypes in this cohort (approximately 1 in 3,500). This molecular approach may facilitate the identification of compatible donors and could support transfusion safety, although clinical outcomes were not evaluated in the present study.

Ethical approval

This cross-sectional study was approved by the Human Research Ethics Committee of Thammasat University Hospital (HREC-TUH, COE No. 015/2569).

Funding

This research received no external funding.

Conflict of interest

The authors declare they have no conflict of interest.

CRediT authorship contribution statement

Sarisa Chidtrakoo: data and sample collection, method; **Wiradee Sasikarn:** method; **Tanaporn Choychimplee:** method, statistical analysis; **Kamphon Intharanut:** data interpretation, writing and review; **Oytip Nathalang:** conceptualization, method investigation, data interpretation, writing: review and editing.

Acknowledgements

The authors gratefully acknowledge the Blood Bank, Thammasat University Hospital, for providing blood donor samples..

References

[1] Hyland CA, Castilho L, Chen Q, Clausen FB, Denomme GA, Fichou Y, et al. International Society of Blood Transfusion Working Party on Red Cell Immunogenetics and Blood Group Terminology report of Gothenburg, Barcelona and four virtual business meetings: Update on blood group systems. Vox Sang. 2026; 121(2): 202-12. doi: 10.1111/vox.70148.

[2] Figueroa D. The *Diego* blood group system: A review. Immunohematology. 2013; 29(2): 73-81. PMID: 24094240.

[3] Daniels G. Human blood groups. 3rd Ed. Malden, MA: Blackwell Science; 2013.

[4] Reid ME, Lomas-Francis C, Olsson M. The blood group antigen factsbook. 3rd Ed. New York: Academic Press; 2012.

[5] Nathalang O, Panichrum P, Intharanut K, Thattanon P, Nathalang S. Distribution of *DI*A* and *DI*B* allele frequencies and comparisons among central Thai and other populations. PloS One. 2016; 11(10): e0165134. doi: 10.1371/journal.pone.0165134.

[6] Wongkaewpotong C, Pattanapeechai S, Leetrakool N, Nathalang O. Prevalence of *DI*A* and *DI*B* alleles in northern Thai blood donors. J Hematol Transfus Med. 2017; 27(1): 19-26.

[7] Simtong P, Jumnainsong A, Leelayuwat C. Genetic profile of RHCE, Kell, Duffy, Kidd, Diego and MNS hybrid glycophorins blood groups in ethnic northeastern Thais: Alleles, genotypes and risk of alloimmunization. Transfus Med. 2024; 34(4): 287-97. doi: 10.1111/tme.13055.

[8] Chesor M, Mitundee S, Nathalang S, Thattanon P, Intharanut K, Tobunluepop P, et al. *DI*A* and *DI*B* allele frequencies among Southern Thai blood donors. Indian J Hematol Blood Transfus. 2018; 34(3): 506-9. doi: 10.1007/s12288-017-0886-9.

[9] Wei CT, Al-Hassan FM, Naim N, Knight A, Joshi SR. Prevalence of Diego blood group antigen and the antibody in three ethnic population groups in Klang valley of Malaysia. Asian J Transfus Sci. 2013; 7(1): 26-8. doi: 10.4103/0973-6247.106725.

[10] Delaney M, Harris S, Haile A, Johnsen J, Teramura G, Nelson K. Red blood cell antigen genotype analysis for 9087 Asian, Asian American, and Native American blood donors. Transfusion. 2015; 55(10): 2369-75. doi: 10.1111/trf.13163.

[11] Belsito A, Costa D, Fiorito C, De Iorio G, Casamassimi A, Perrotta S, et al. Erythrocyte genotyping for transfusion-dependent patients at the Azienda Universitaria Policlinico of Naples. Transfus Apher Sci. 2015; 52(1): 72-7. doi: 10.1016/j.transci.2014.12.006.

[12] Nathalang O, Khumsuk P, Intharanut K, Choychimplee T, Sasikarn W. Extended blood group genotyping among Thai patients with repeated transfusions. Clin Lab. 2025; 71(8):

1549-55. doi: 10.7754/Clin.Lab.2025.250134.

[13] Novaretti MC, Ruiz AS, Dorlhiac-Llacer PE, Chamone DA. Application of real-time PCR and melting curve analysis in rapid Diego blood group genotyping. Immunohematology. 2010; 26(2): 66-70. PMID: 20932077.

[14] Nathalang O, Intharanut K, Chidtrakoon S. High-resolution melting curve analysis to predict extended blood group phenotypes among Thai donors and patients. Transfus Med Hemother. 2021; 49(3): 163-71. doi: 10.1159/000520713.

[15] Fongsarun J, Nuchprayoon I, Yod-in S, Kupatawintu P, Kidprasirt C. Blood groups in Thai blood donors. [in Thai]. Thai J Hematol Transf Med. 2002; 12(4): 277-86.

[16] Chiewsilp P, Bejrachandra S, Thongthaisin A. Standards for blood banks and transfusion services. [in Thai]. 5th Ed. Bangkok: Udomsuksa; 2026.

[17] Ye L, Zhang J, Wang C, Yang Q, Han S, Zhu Z. Performance of a microarray-based genotyping system for red cell and platelet antigens in China. Blood Transfus. 2015; 13(4): 690-3. doi: 10.2450/2015.0301-14.

[18] Wu GG, Su YQ, Yu Q, Jin SZ, Zhao TM. Development of a DNA-based genotyping method for the Diego blood group system. Transfusion. 2002; 42(12): 1553-6. doi: 10.1046/j.1537-2995.2002. 00248.x.

[19] Gong T, Hong Y, Wang N, Fu X, Zhou C. Validation of a blood group genotyping method based on high-resolution melting curve analysis. Immunohematology. 2014; 30(4): 161-5. PMID: 25831264.

[20] Nathalang O, Ang RM, Kurin B, Limprasert S, Mitundee S, Leetrakool N, et al. Predicted S and s phenotypes from genotyping results among Thai populations to prevent transfusion-induced alloimmunization risks. Transfus Apher Sci. 2018; 57: 582-6. doi: 10.1016/j.transci.2018.07.019.

[21] Kupatawintu P, Emthip M, Sungnoon D, O-vataga P, Manahul V, Limtamaporn S, et al. Unexpected antibodies of patients, blood samples sent for testing at NBC. TRCS. [in Thai]. J Hematol Transfus Med. 2010; 20(4): 255-62.

[22] Hatano Y, Otsuka S, Chousa M, Saito S, Nollet KE, Ohto H, et al. Fatal delayed hemolytic transfusion reaction associated with anti-Di(b) and anti-E. Transfus Apher Sci. 2012; 47(3): 263-8. doi: 10.1016/j.transci.2012.07.020.

[23] Mochizuki K, Ohto H, Hirai S, Ujiie N, Amanuma F, Kikuta A, et al. Hemolytic disease of the newborn due to anti-Di: A case study and review of the literature. Transfusion. 2006; 46(3): 454-60. doi: 10.111/j.1537-2995.2006.00743.x.

[24] Kerdkaewngam K, Tingtong U. A study of the Di^a and Di^b antigen frequencies in blood donors of National Blood Centre, Thai Red Cross Society. [in Thai]. J Hematol Transfus Med. 2021; 31(4): 309-14.

Journal of Associated Medical Sciences 2026; 59(3): 294-304

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



Robusta peaberry coffee attenuates inflammation, regulates apoptosis, and improves testicular cell integrity in diabetic rats

Mochamad Rizal Maulana[1], Teguh Budiharjo[1], Djoko Priyatno[1], Burhannudin Burhannudin[2], Puji Lestari[2], Sofhia Azqilla Nur[1], Yulian Nursaputra Habibina[1], Eko Naning Sofyanita[1*]

[1]Division of Medical Laboratory Technology, Department of Health Analyst, Ministry of Health Polytechnic of Health Semarang. Central Java Province, Indonesia.
[2]Department of Medical Laboratory Technology Ministry of Health Polytechnic of Health Jakarta 3, DKI Jakarta Province, Indonesia.

ARTICLE INFO

Article history:
Received 11 February 2026
Accepted as revised 12 July 2026
Available online 22 July 2026

Keywords:
Diabetes mellitus, robusta peaberry coffee, chlorogenic acid, testicular injury, apoptosis

ABSTRACT

Background: Diabetes mellitus (DM) is a chronic metabolic disorder that can impair male reproductive function through oxidative stress, inflammation, and apoptosis. Robusta peaberry coffee contains chlorogenic acid (CGA) and other polyphenols with antioxidant and anti-inflammatory properties that may protect against diabetes-related testicular injury.

Objectives: This study aimed to evaluate the effects of robusta peaberry coffee on inflammation, apoptosis-related markers, reproductive hormones, and testicular cell integrity in streptozotocin (STZ)-induced diabetic male rats.

Materials and methods: Male Wistar rats (N=28) were randomly assigned to four groups: diabetic control (KN), diabetic rats treated with metformin (KP; 500 mg/kg body weight [BW]/day), diabetic rats treated with robusta peaberry coffee at 100 mg/kg BW/day (P1), and diabetic rats treated with robusta peaberry coffee at 200 mg/kg BW/day (P2). Experimental diabetic rats were induced by a single intraperitoneal injection of STZ (40 mg/kg BW). Rats with fasting blood glucose (FBG) ≥200 mg/dL were included. Treatments were administered orally for 14 consecutive days. At the end of the intervention, serum nitric oxide (NO), tumor necrosis factor-alpha (TNF-α), interleukin-1 beta (IL-1β), caspase-3, testosterone, luteinizing hormone (LH), follicle-stimulating hormone (FSH), and the numbers of Sertoli and Leydig cells were evaluated.

Results: Testosterone, LH, FSH, Sertoli cell count, and Leydig cell count were significantly lower in the KN group than in the KP, P1, and P2 groups (p<0.05). In contrast, NO, TNF-α, IL-1β, and caspase-3 were significantly higher in the KN group than in the treatment groups (p<0.05). The P2 group showed the greatest improvement, with higher reproductive hormone levels and testicular cell counts and lower inflammatory and apoptosis-related markers than P1 (p<0.05).

Conclusion: Robusta peaberry coffee attenuates inflammation, modulates apoptosis-related responses, and improves testicular cell integrity in diabetic rats. These findings suggest that robusta peaberry coffee may serve as a promising complementary natural intervention for mitigating diabetes-induced male reproductive dysfunction.

***Corresponding contributor.**
Author's Address: *Division of Medical Laboratory Technology, Department of Health Analyst, Ministry of Health Polytechnic of Health Semarang. Central Java Province, Indonesia.*
E-mail address: ekonaning@poltekkes-smg.ac.id
doi: 10.66285/JAMS.2026.102
E-ISSN: 2539-6056

Introduction

Diabetes mellitus (DM) is a chronic metabolic disorder that can impair male reproductive function through persistent hyperglycemia, oxidative stress, inflammation, endocrine disruption, and apoptosis.[1,2] In diabetic testes, excessive reactive oxygen species activate inflammatory pathways, including nuclear factor kappa B (NF-κB), thereby increasing nitric oxide (NO), tumor necrosis factor-alpha (TNF-α), and interleukin-1 beta (IL-1β). These mediators disturb the testicular microenvironment, damage Sertoli and Leydig cells, and reduce steroidogenesis and spermatogenesis.[3,4]

Although coffee and coffee-derived polyphenols have been studied for their metabolic and antioxidant effects in diabetes, evidence on robusta peaberry coffee and testicular function remains limited.[5-7] This gap is important because diabetic reproductive injury involves interconnected inflammatory, apoptotic, hormonal, and histological changes.[8,9] Therefore, an intervention should not be evaluated only through a single marker, but through multiple parameters that reflect the inflammatory burden, apoptosis-related response, reproductive hormone status, and testicular cell integrity.

Robusta peaberry coffee was selected because peaberry beans contain a single seed within the coffee cherry and may have a denser and more concentrated bioactive profile than regular beans.[10,11] Robusta coffee is also recognized for its relatively high chlorogenic acid (CGA) content compared with several other coffee types.[12,13] CGA and other phenolic compounds can scavenge free radicals, suppress inflammatory signaling, and reduce cytokine expression. These properties provide a biological rationale for testing robusta peaberry coffee as a potential protective agent against diabetes-induced testicular injury.[6,14,15]

The use of whole coffee extract may also introduce components such as caffeine and other methylxanthines that can cause adverse effects when consumes excessively, including increased sympathetic activity, gastrointestinal discomfort, and sleep disturbance.[16,17] For this reason, the intervention dose should be justified and standardized. In the present study, the 100 and 200 mg/kg BW doses were selected based on previous experimental coffee-extract studies showing antioxidant and metabolic effects in diabetic models, while allowing comparison between a lower and a higher dose.[18,19]

The novelty of this study lies in the simultaneous evaluation of robusta peaberry coffee on inflammatory biomarkers, an apoptosis-related marker, reproductive hormones, and histological indicators of testicular cell integrity in diabetic rats. This integrated approach provides a clearer understanding of whether robusta peaberry coffee can protect testicular function under diabetic conditions.

Materials and methods

Study design

This study was a real experimental laboratory study using a randomized post-test-only control-group design.

Experimental animals, diabetes induction, and materials

Twenty-eight male Wistar rats (Rattus norvegicus), aged 8-10 weeks and weighing approximately 250 gm, were obtained from a certified laboratory animal facility. Animals were acclimatized for one week before experimentation and were housed individually in stainless steel cages under controlled conditions (24±2 °C, 50-60% relative humidity, and a 12-hour light/dark cycle). Standard pellet diet and water were provided ad libitum.

All procedures were approved by the Health Research Ethics Committee of the Polytechnic of Health, Ministry of Health, Semarang (No. 837/EA/F. XXIII.38/2025).

After acclimatization, experimental diabetic rats were induced by a single intraperitoneal injection of streptozotocin (STZ) at 40 mg/kg BW dissolved in 0.1M citrate buffer (pH 4.5). Rats with fasting blood glucose ≥200 mg/dL were considered diabetic and included in the study. Seventy-two hours after STZ administration, FBG was measured from tail-vein blood samples using a glucometer (EasyTouch®, GCUModel: ET-301, Bioptik Technology, Taiwan). The device was checked using manufacturer-recommended control solution before measurement, and test strips were used within the expiration date and stored according to the manufacturer's instructions. Rats with FBG ≥200 mg/dL were considered diabetic and included in the study. Animals that did not reach this threshold were excluded to ensure homogeneity of diabetic status.

Robusta peaberry coffee beans were obtained from a local supplier in Banyuwangi, Indonesia. The beans were ground into fine powder using a dry electric grinder (Yongkang Tianqi Shengshi Industry and Trade Co., Ltd.). STZ, metformin hydrochloride (99.60%; PT Phapros Tbk, Indonesia), ELISA kits for NO, TNF-α, IL-1β, caspase-3, testosterone, LH, and FSH (Wuhan Fine Biotech Co., Ltd. [FineTest], Wuhan, China), and hematoxylin-eosin reagents were used in this study.

Preparation and standardization of robusta peaberry coffee

An aqueous robusta peaberry coffee extract was freshly prepared each day from roasted coffee beans obtained from the same production batch. Coffee powder was mixed with distilled water at a fixed ratio of 1 gm:10 mL (w/v) and extracted by stirring at 90 °C for 15 minutes, followed by filtration through Whatman No. 1 filter paper to remove insoluble residues. The filtrate was diluted to obtain a working concentration

of 25 mg/mL, allowing accurate dose administration according to body weight. Preparation consistency was maintained by using the same coffee batch, extraction temperature, extraction time, powder-to-solvent ratio, filtration procedure, and freshly prepared extract before each administration. Although quantitative phytochemical standardization (e.g., chlorogenic acid or caffeine determination) was not performed for each daily preparation, previous studies have demonstrated that robusta coffee is naturally rich in chlorogenic acid, which is considered one of its principal bioactive constituents.

Experimental groups and treatment

Diabetic rats were randomly allocated into four groups (N=7 per group): 1) diabetic control (KN), 2) diabetic rats treated with metformin at 500 mg/kg BW/day (KP), 3) diabetic rats treated with robusta peaberry coffee at 100 mg/kg BW/day (P1), and 4) diabetic rats treated with robusta peaberry coffee at 200 mg/kg BW/day (P2). Based on an average body weight of 0.25 kg, the daily coffee doses were 25 mg (P1) and 50 mg (P2), corresponding to 1 mL/day and 2 mL/day, respectively, using the 25 mg/mL working solution. All interventions were administered orally by gavage once daily for 14 consecutive days. The KN group received an equivalent volume of distilled water by oral gavage as the vehicle control.

Sample collection

At the end of the treatment period, rats were fasted overnight and anesthetized using ketamine. Blood samples were collected through retro-orbital puncture using a capillary glass tube. Blood was allowed to clot at room temperature and was centrifuged at 3,000 rpm for 15 minutes to obtain serum. Testicular tissues were excised, rinsed with physiological saline, and processed for histopathological analysis.

Biochemical and hormonal analysis

Serum nitric oxide (NO) levels were measured using a commercial ELISA kit (FineTest catalog No. ER0150, Wuhan, China) according to the manufacturer's instructions. Absorbance was recorded at 450 nm, and concentrations were calculated from the standard curve. Serum TNF-α and IL-1β were quantified using sandwich ELISA kits (FineTest catalog No. ER1393 for TNF-α and ER1094 for IL-1β). Absorbance was measured at 450 nm, and cytokine concentrations were interpolated from calibration curves. Serum caspase-3 was measured using a quantitative ELISA kit (FineTest catalog No. ER0143).

In this study, serum caspase-3 was interpreted as an apoptosis-related systemic marker associated with diabetic testicular injury, not as a direct cell-specific measurement of apoptosis within Sertoli or Leydig cells.

Serum testosterone, LH, and FSH were measured with specific ELISA kits (FineTest catalog Nos. EU0400 for testosterone, ER1123 for LH, and ER0960 for FSH) and calculated using standard reference curves generated for each assay.

Histopathological examination

Testicular tissues were fixed in 10% neutral buffered formalin for 24 hours, dehydrated through graded ethanol, cleared in xylene, embedded in paraffin, sectioned at 5 µm, and stained with hematoxylin and eosin. Microscopic examination was performed at 400× magnification. Sertoli and Leydig cells were counted using a systematic random field approach in five non-overlapping fields per section. Cells were identified based on typical morphology and clearly visible nuclei, and the mean value was calculated for each animal. Counting was performed in a blinded manner to reduce observer bias.

Statistical analysis

Data are presented as mean±SD. Statistical comparisons among groups were performed using one-way analysis of variance (ANOVA), followed by Tukey's honestly significant difference (HSD) post-hoc test for multiple comparisons. Statistical analysis was conducted using SPSS version 25 (IBM Corp., USA). A $p<0.05$ was considered statistically significant. Graphs were generated using GraphPad Prism version 9 (GraphPad Software, USA).

Results

Effects of robusta peaberry coffee on serum reproductive hormone levels

Serum testosterone levels were lowest in the KN group and were significantly lower than those in all other groups ($p<0.05$). Serum testosterone levels were significantly higher in the KP, P1, and P2 groups than in the KN group ($p<0.05$). The P2 group demonstrated a significantly higher testosterone level than the P1 group ($p<0.05$) (Figure 1A).

Serum luteinizing hormone (LH) levels were lowest in the diabetic control group (KN). Treatment with metformin (KP) and robusta peaberry coffee (P1 and P2) resulted in significantly higher LH levels than those in the KN group ($p<0.05$). The increase in LH levels was greater with higher coffee doses. The P2 group showed the highest LH level among all experimental groups ($p<0.05$) (Figure 1B).

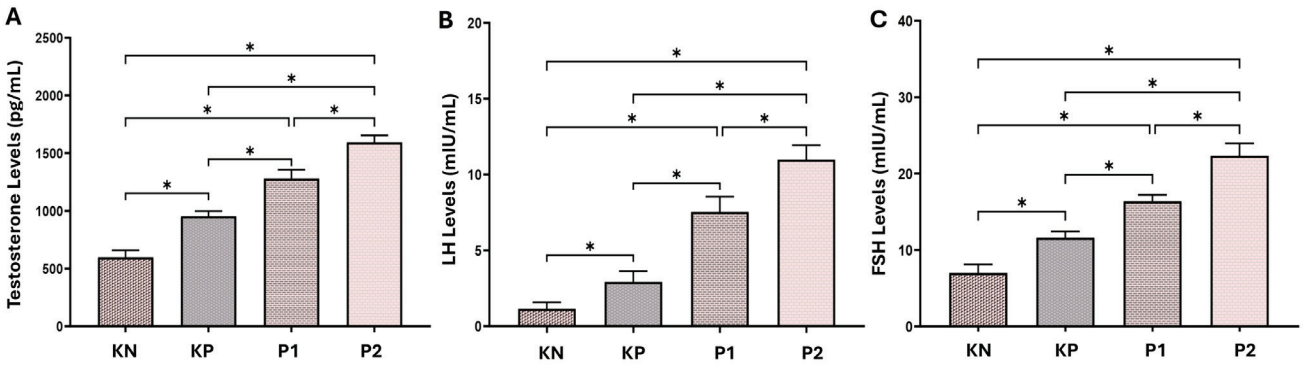


Figure 1. *Reproductive hormone parameters. A: testosterone levels, B: luteinizing hormone (LH) levels, C: follicle-stimulating hormone (FSH) levels, KN: diabetic control, KP: diabetic rats treated with metformin, P1: diabetic rats treated with robusta peaberry coffee at 100 mg/kg BW/day, P2: diabetic rats treated with robusta peaberry coffee at 200 mg/kg BW/day. *p<0.05.*

A similar trend was observed for serum follicle-stimulating hormone (FSH) levels across the experimental groups. The KN group had the lowest FSH level, whereas significantly higher values were found in the KP, P1, and P2 groups (p<0.05). FSH levels increased gradually from KP to P1, peaking in the P2 group. Significant differences were also observed between the coffee-treated groups and the diabetic control group (p<0.05). Compared with the lower coffee dose, administration of 200 mg/kg BW/day produced a greater increase in FSH levels.

Effects of robusta peaberry coffee on Sertoli and Leydig cell integrity

Figure 2A shows the mean Sertoli cell count in the KN, KP, P1, and P2 groups. The KN group had the lowest Sertoli cell count, followed by progressively higher counts in the KP and P1 groups, with the highest count observed in the P2 group (p<0.05). These quantitative findings were supported by the histopathological images in Figure 3, which showed reduced testicular cellular integrity in the KN group and progressive improvement in the KP, P1, and P2 groups, with the most pronounced improvement observed in the P2 group.

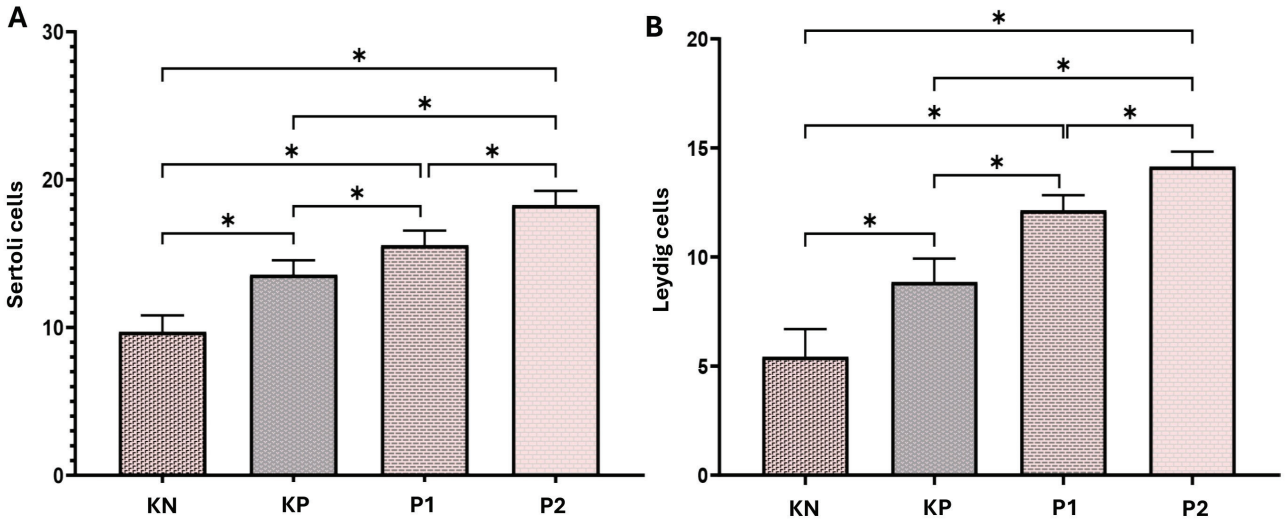


Figure 2. *Testicular cell integrity after treatment. A: Sertoli cell count, B: Leydig cell counts in each group, KN: diabetic control, KP: diabetic rats treated with metformin, P1: diabetic rats treated with robusta peaberry coffee at 100 mg/kg BW/day, P2: diabetic rats treated with robusta peaberry coffee at 200 mg/kg BW/day. *p<0.05.*

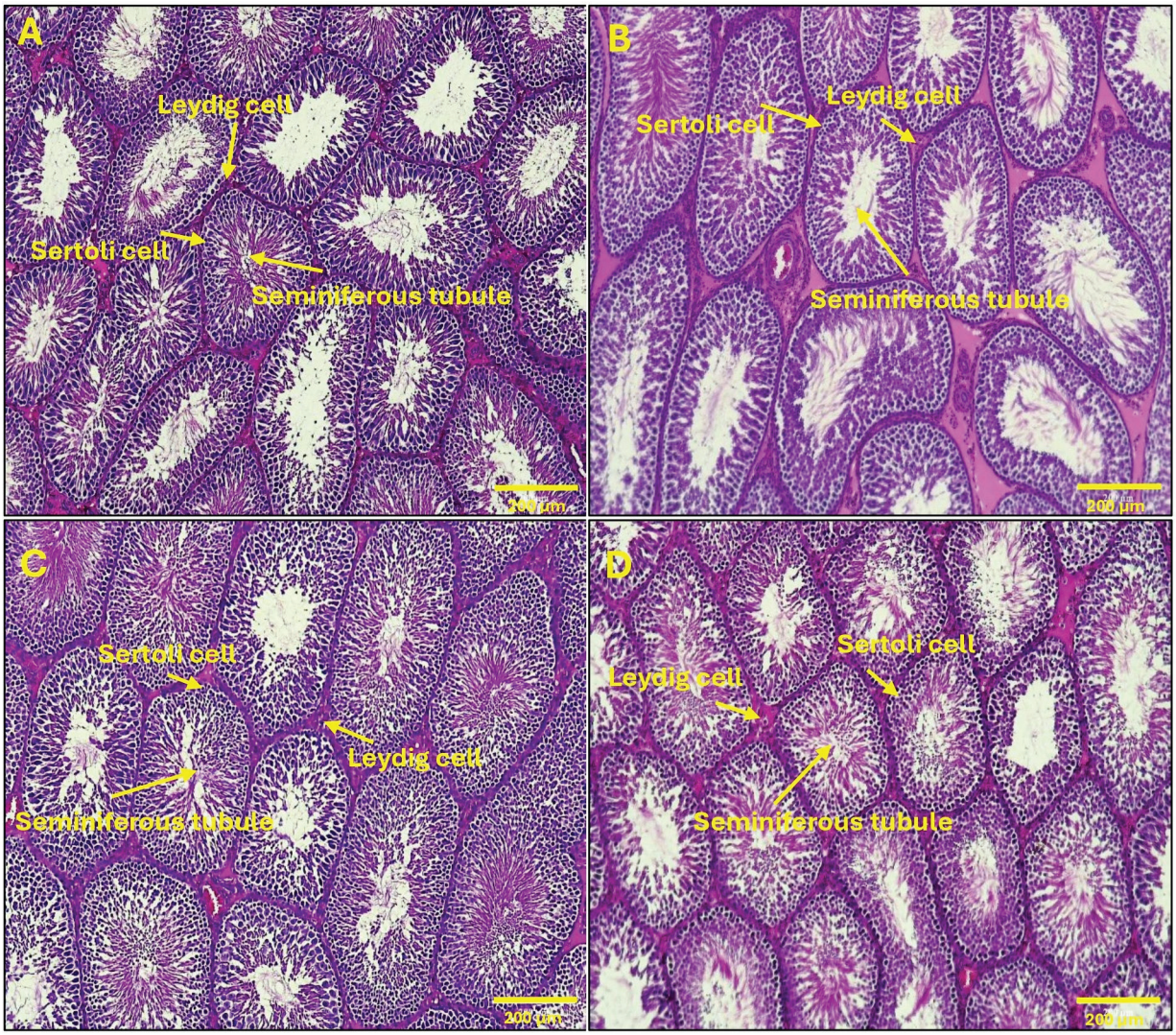


Figure 3. *Histopathological images of testicular tissue from diabetic rats, H&E-stained; scale bar = 200 µm. A: diabetic control group (KN), B: diabetic rats treated with metformin (KP), C: diabetic rats treated with robusta peaberry coffee at 100 mg/kg BW/day (P1), D: diabetic rats treated with robusta peaberry coffee at 200 mg/kg BW/day (P2). The arrows indicate Sertoli cells, Leydig cells, and seminiferous tubules.*

Effects of robusta peaberry coffee on serum NO, TNF-α, IL-1β, and caspase-3

NO levels were highest in the KN group, with a mean of 13.80±0.92 ng/mL, and were significantly higher than those in the KP group, which showed a lower mean of 10.20±0.79 ng/mL (p<0.05). A further decrease was observed in the P1 group, with a mean of 6.04±0.67 ng/mL, which was significantly lower than those in both the KN and KP groups (p<0.05). The P2 group showed the lowest NO level, with a mean of 3.48±0.97 ng/mL, which was significantly lower than those in all other groups (p<0.05). Compared with the KN group, the P2 group showed a reduction of more than 10.32 ng/mL (p<0.05) (Figure 4A).

Figure 4B shows that the TNF-α level was highest in the KN group (79.70±7.58 pg/mL) and was significantly higher than that in the KP group (65.90±7.21 pg/mL) (p<0.05). A greater reduction was observed in the P1 group, with a mean of 53.60±3.01 pg/mL, which was significantly lower than those in the KN and KP groups (p<0.05). The lowest TNF-α level was observed in the P2 group (43.60±2.34 pg/mL), which differed significantly from those of the other groups (p<0.05). Compared with the KN group, the P2 group showed a reduction of more than 36 pg/mL (p<0.05). Greater variability was observed in the KN and KP groups than in the P1 and P2 groups, as indicated by the larger standard deviations.

Figure 4C demonstrates that the IL-1β level was highest in the KN group, with a mean of 93.55±7.69 pg/mL, and was significantly higher than that in the KP group (65.26±3.52 pg/mL) (p<0.05). The P1 group showed a mean IL-1β level of 52.70±3.93 pg/mL, which was significantly lower than those in the KN and KP groups (p<0.05). The lowest IL-1β level was observed in the P2 group, with a mean of 38.02±3.74 pg/mL, which was significantly lower than those in all other groups (p<0.05). Compared with the KN group, the P2 group showed a reduction of more than 55.53 pg/mL (p<0.05).

Caspase-3 activity was highest in the KN group, with a mean of 3.96±0.49 ng/mL, and was significantly higher than that in the KP group (3.14±0.42 ng/mL) (p<0.05). The P1 group showed a mean value of 2.37±0.36 ng/mL, which was significantly lower than those in the KN and KP groups (p<0.05). The lowest caspase-3 activity was observed in the P2 group, with a mean of 1.51±0.28 ng/mL, which was significantly lower than those in all other groups (p<0.05). Compared with the KN group, the P2 group showed a reduction of more than 2.4 ng/mL (p<0.05) (Figure 4D).

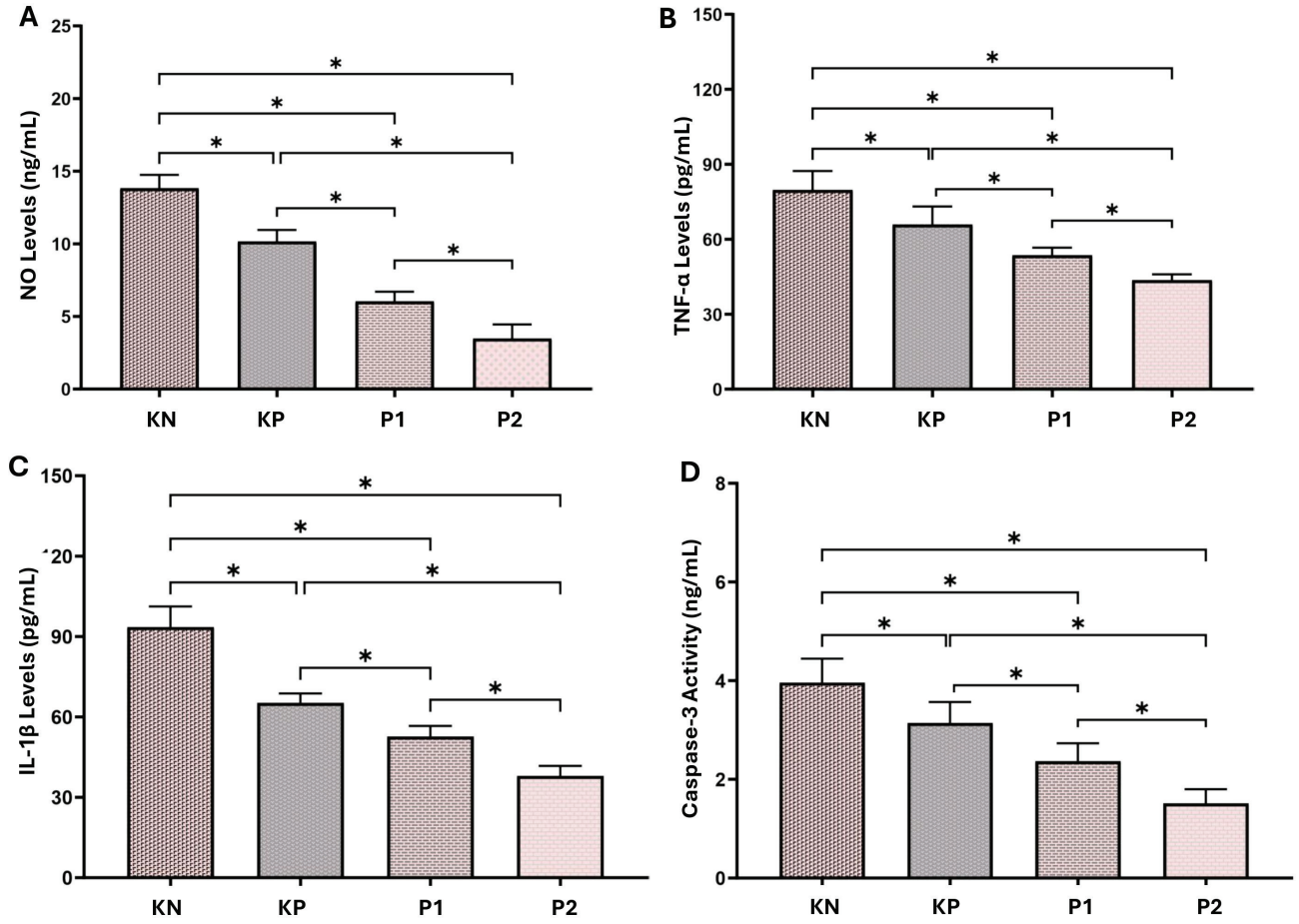


Figure 4. *Effects of robusta peaberry coffee. A: serum nitric oxide (NO), B: tumor necrosis factor-alpha (TNF-α), C: interleukin-1 beta (IL-1β), D: caspase-3, KN: diabetic control, KP: diabetic rats treated with metformin, P1: diabetic rats treated with robusta peaberry coffee at 100 mg/kg BW/day, P2: diabetic rats treated with robusta peaberry coffee at 200 mg/kg BW/day. *p<0.05.*

Correlation between NO, TNF-α, IL-1β, reproductive hormones, caspase-3, Sertoli cells, and Leydig cells

Figure 5 reveals the Pearson correlation heatmap among inflammatory markers, reproductive hormones, the apoptosis marker, and testicular cells. NO and TNF-α showed a strong positive correlation (r=0.92), suggesting that higher NO levels were generally accompanied by higher TNF-α levels in the pooled dataset. However, this relationship should be interpreted as a statistical association rather than evidence of a direct biological mechanism or causal effect. In contrast, NO and TNF-α showed strong negative correlations with testosterone, LH, FSH, Sertoli cells, and Leydig cells, suggesting inverse statistical associations between inflammatory markers and reproductive parameters.

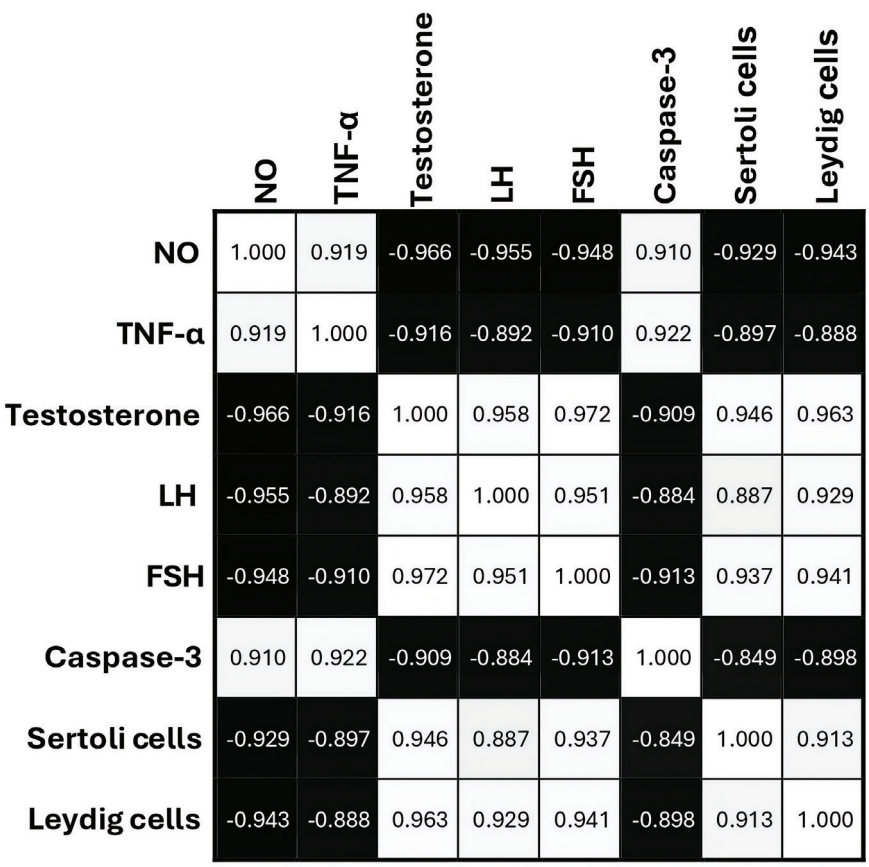


Figure 5. *Pearson correlation heatmap of NO, TNF-α, testosterone, LH, FSH, caspase-3, Sertoli cells, and Leydig cells in diabetic rats. Positive correlations are shown by lighter colors, whereas negative correlations are shown by darker colors.*

Caspase-3 showed strong positive correlations with NO (r=0.91) and TNF-α (r=0.92), but negative correlations with testosterone (r=-0.91), LH (r=-0.88), FSH (r=-0.91), Sertoli cells (r=-0.85), and Leydig cells (r=-0.90). Conversely, testosterone, LH, and FSH showed positive correlations with Sertoli and Leydig cells, with the strongest associations observed between testosterone and Leydig cells (r=0.96), testosterone and FSH (r=0.97), and testosterone and LH (r=0.96). All correlations were statistically significant (p<0.001), indicating strong statistical relationships among the variables included in this study.

Discussion

This research studied how Robusta peaberry coffee influences reproductive hormones and testicular dysfunction in diabetic rats. The findings demonstrated correlations among the variables, supporting the hypothesis that reproductive issues in diabetes stem from oxidative stress that induces inflammation, which activates cell death pathways and hampers steroidogenesis. Diabetes has been reported to lead to an increase in reactive oxygen species (ROS) and an imbalance in antioxidants, thereby exacerbating tissue inflammation, including within the testes.[8] Subsequently, this cascade disrupts steroidogenesis and triggers apoptosis, resulting in decreased testosterone levels and impaired spermatogenesis.[20]

Diabetes mellitus impairs sperm quality and hypothalamic-pituitary function, leading to decreased Leydig cell activity and testosterone production. Increased mitochondrial oxidative stress causes cell death and tissue damage, compromising the blood-testis barrier and leading to lipid peroxidation in germ cell membranes, both of which are crucial for spermatogenesis.[21] Leydig cells are especially vulnerable, affecting steroidogenic proteins and enzyme expression necessary for testosterone synthesis.[18]

Blood glucose levels may be reduced by moderate coffee consumption, and different coffee extract doses can produce different glycemic effects.[17] Robusta coffee has been reported to contain higher CGA than Arabica coffee extract, and aqueous green coffee extracts, particularly Robusta, have shown beneficial effects on pancreatic islet repair in diabetic models.[16,18,22] In diabetes-associated male infertility, testosterone plays an essential role in spermatogenesis and serves as an important marker of testicular function.[23,24] In the present study, diabetic control rats showed decreased LH levels, consistent with previous findings that diabetes is associated with reduced reproductive hormone levels.

The observed increases in LH and FSH following robusta peaberry coffee administration may have

several possible explanations. One interpretation is that the reduction in hyperglycemia-associated oxidative stress and inflammation partially restored hypothalamic pituitary gonadal (HPG) axis function, thereby improving pituitary secretion of gonadotropins.[8,9,28] Alternatively, the elevated LH and FSH concentrations may represent a compensatory endocrine response accompanying the recovery of testicular function rather than a direct stimulatory effect of coffee-derived bioactive compounds on the pituitary gland.[27] Improved Leydig cell integrity may enhance testosterone production, while restoration of Sertoli cell function could normalize feedback mechanisms involved in gonadotropin regulation.[20,27] Because circulating LH and FSH are influenced by multiple physiological factors, including hypothalamic gonadotropin-releasing hormone (GnRH) secretion, pituitary responsiveness, gonadal feedback, metabolic status, and systemic inflammation, the present study cannot determine the primary mechanism responsible for these hormonal changes.[8,9,27,28]

The strong correlations among NO, TNF-α, and IL-1β indicate that these inflammatory mediators increase together during activation of pro-inflammatory pathways. Increased NO can amplify oxidative stress, while TNF-α activates NF-κB and promotes cytokine expression, creating an unfavorable environment for Sertoli and Leydig cells.[25] Robusta peaberry coffee treatment reduced these inflammatory mediators, which may be related to its polyphenol and CGA content that can suppress ROS, inflammatory signaling, and cytokine expression in testicular tissue.[26]

A very strong inverse relationship between testosterone and inflammatory mediators indicates that when the testicular environment is predominantly pro-inflammatory, steroidogenic capacity tends to decline. Physiologically, testosterone is produced by Leydig cells and is highly dependent on mitochondrial integrity, intracellular cholesterol availability, and the expression of cholesterol transport proteins such as StAR and steroidogenic enzymes. Recent evidence suggests that inflammation, particularly through cytokines such as TNF-α, can disrupt androgen production, especially when inflammation occurs alongside oxidative stress.[27] This is in line with a meta-analysis that identifies chronic inflammation as an important mediator of fertility disorders in the diabetic population and highlights that targeting inflammation may be a relevant therapeutic strategy.[28] Thus, the direction of the correlations in this study confirms that testosterone is not merely decreased, but also serves as a marker of testicular function that is sensitive to inflammatory stress. Interestingly, a recent study on CGA in testicular disorder models also reported that CGA can reduce ROS and pro-inflammatory cytokines in the testes, accompanied by histological repair, which is biologically consistent with recovery of the steroidogenesis pathway.[29] If inflammation is suppressed, Leydig cells are more likely to maintain testosterone synthesis, allowing endocrine signaling to Sertoli cells and seminiferous tissue to recover. Activation of caspase-3 generally represents the execution phase of apoptosis, which can be triggered by mitochondrial damage, oxidative stress, and persistent inflammatory signals. Studies on testicular injury models show that when TNF-α, IL-1β, and NF-κB increase, apoptotic indicators, including caspase-3, also tend to rise; conversely, when an intervention suppresses NF-κB, inflammatory markers and caspase-3 also decrease.[30]

This study demonstrates that robusta peaberry coffee may increase reproductive hormone levels and preserve Sertoli and Leydig cells in DM rats after 14 days of administration. This supports the findings of Al-Megrin et al., who reported that green coffee possesses exogenous antioxidant properties.[6] Endogenous antioxidant enzymes such as glutathione peroxidase (GPx), catalase (CAT), and superoxide dismutase (SOD) can be increased by coffee as an exogenous antioxidant source. Coffee flavonoids act as antioxidants by increasing SOD levels and donating electrons and hydrogen ions to superoxide anions, making them more stable and preventing the oxidation of DNA and lipoproteins. The mineral content of coffee can also inactivate reactive components, while its CGA content scavenges free radicals.[31]

Peaberry coffee contains various health-beneficial compounds, including caffeine, flavonoids, tannins, saponins, trigonelline, carbohydrates, fats, amino acids, organic acids, volatile aromatic compounds, and minerals. Chlorogenic acid is the most abundant acid in coffee, accounting for approximately 8% in coffee beans and 4.5% after roasting.[32,33] Chemical compounds such as steroids, alkaloids, and flavonoids may act as aphrodisiac components by increasing male sexual activity through vasodilation, nitric oxide production, and elevated testosterone and gonadotropin levels.[34] Chlorogenic acid lowers blood glucose levels by inhibiting fatty acid production in both in vivo and in vitro settings, a mechanism similar to that of metformin. Both compounds have antioxidant, antimutagenic, and anticancer properties that affect reactive oxygen species (ROS).[19] The mitigation of oxidative stress in the present study was achieved through the administration of coffee-derived extracts, whose antioxidant properties are attributed to their abundant phenolic and flavonoid contents.[35] These bioactive compounds, particularly CGA, are recognized for their ability to enhance insulin sensitivity and improve glucose utilization by peripheral tissues.[6]

Therefore, robusta peaberry coffee may serve as a complementary dietary strategy to modulate glucose metabolism and mitigate metabolic dysfunction associated with diabetes mellitus.[19] Furthermore, bioactive components in coffee, such as CGA and caffeine, exert significant antioxidant potential by scavenging free radicals and inhibiting oxidation

reactions, thereby protecting cellular structures from oxidative damage.

Limitations

This study has several limitations. The relationships among variables were correlational and did not prove direct causality. The sample size was limited, and caspase-3 was measured in serum rather than directly in testicular tissue, so it could not confirm cell-specific apoptosis in Sertoli or Leydig cells. Fertility outcomes, such as sperm quality and reproductive success, were not assessed. In addition, whole robusta peaberry coffee was used, so the specific effects of CGA, caffeine, or other active compounds could not be separated. Detailed phytochemical profiling of each daily extract was also not performed, although preparation consistency was controlled. In addition, the correlation analysis was performed using pooled data from all experimental groups. Therefore, these correlations should be interpreted as exploratory statistical associations rather than evidence of direct mechanistic or causal relationships between variables.

Conclusion

Robusta peaberry coffee reduced inflammatory and apoptosis-related markers and improved reproductive hormone levels, Sertoli cell count, and Leydig cell count in diabetic rats.

These findings indicate that robusta peaberry coffee may support testicular protection under diabetic conditions by modulating inflammation, apoptosis-related responses, and cellular integrity. Further studies are needed to confirm the active compounds, optimal dose, safety profile, and direct fertility outcomes.

Ethical approval

This research has received ethical approval (No. 837/EA/F.XXIII.38/2025) from the Health Ethics Committee of the Polytechnic of Health, Ministry of Health Semarang.

Funding

This research received funding from the DIPA Ministry of Health Polytechnic of Health Semarang.

Conflict of interest

The authors declare no conflicts of interest related to the research, authorship, or publication of this article.

CRediT authorship contribution statement

Mochamad Rizal Maulana: conceptualization, methodology, writing: original draft; **Teguh Budiharjo:** methodology, supervision; **Djoko Priyatno**: investigation, supervision; **Burhannudin**: investigation, discussion of results, resources; **Puji Lestari**: discussion of results, resources; **Sofhia Azqilla Nur**: writing: review and edit; **Yulian Nursaputra Habibina**: writing: review and edit; **Eko Naning Sofyanita**: conceptualization, discussion of results, writing: review and edit, approval of the final text.

Acknowledgements

We appreciate the Ministry of Health Polytechnic of Health for providing the facilities and materials for this research.

References

[1] Setyoadi S, Yusuf A, Kristianingrum ND, Hayati YS, Noviyanti LW, Syafiky NF. The correlation between family support and health status in patients with diabetes mellitus. Healthc Low-Resour Settings. 2023; 11(s1): 11212. doi:10.4081/hls.2023.11212.

[2] Sofyanita EN, Herlina EC, Juniarto AZ. Chlorogenic acid ameliorates testicular damage and hormonal imbalance in diabetic male Wistar rats. Trends Sci. 2026; 23(7): 12673. doi:10.48048/tis.2026.12673.

[3] Rong J, Leng X, Jiang K, Tan J, Dong M, Tan J. Systemic impacts of diabetes on spermatogenesis and intervention strategies: multilayered mechanism analysis and cutting-edge therapeutic approaches. Reprod Biol Endocrinol. 2025; 23(1): 122. doi: 10.1186/s12958-025-01454-4.

[4] Sofyanita EN, Suwondo A, Nugroho H, Afifah DN, Suhartono S, Juniarto AZ. Chlorogenic acid enhanced testosterone-mediated libido through improved glycemic control and redox homeostasis in diabetic male rats. Trends Sci. 2026; 23(8): 13191. doi:10.48048/tis.2026.13191.

[5] Kim H, Kim SR, Jung UJ. Coffee and its major polyphenols in the prevention and management of type 2 diabetes: A comprehensive review. Int J Mol Sci. 2025; 26(12): 5544. doi:10.3390/ijms26125544.

[6] Al-Megrin WAI, El-Khadragy MF, Hussein MH, Mahgoub S, Abdel-Mohsen DM, Taha H, et al. Green Coffea arabica extract ameliorates testicular injury in high-fat diet/streptozotocin-induced diabetes in rats. J Diabetes Res. 2020; 2020: 6762709. doi:10.1155/2020/6762709.

[7] Mohamed AI, Erukainure OL, Salau VF, Islam MS. Impact of coffee and its bioactive compounds on the risks of type 2 diabetes and its complications: A comprehensive review. Diabetes Metab Syndr. 2024; 18(7): 103075. doi:10.1016/j.dsx.2024.103075.

[8] Huang R, Chen J, Guo B, Jiang C, Sun W. Diabetes-induced male infertility: potential mechanisms and treatment options. Mol Med. 2024; 30(1): 11. doi:10.1186/s10020-023-00771-x.

[9] Zhang W, Tong L, Jin B, Sun D. Diabetic testicular dysfunction and spermatogenesis impairment: mechanisms and therapeutic prospects. Front Endocrinol (Lausanne). 2025; 16: 1653975. doi:10.3389/fendo.2025.1653975.

[10] Schwarzmann ET, Washington MP, Rao NZ. Physicochemical analysis of cold brew and hot brew peaberry coffee. Processes. 2022; 10(10):

1989. Doi:10.3390/pr10101989.
[11] Fu X, Li G, Hu F, Huang J, Lou Y, Li Y, et al. Comparative transcriptome analysis in peaberry and regular bean coffee to identify bean quality associated genes. BMC Genomic Data. 2023; 24(1): 12. doi:10.1186/s12863-022-01098-y.
[12] Purwoko T, Suranto, Setyaningsih R, Marliyana SD. Chlorogenic acid and caffeine content of fermented robusta bean. Biodiversitas. 2022; 23(2): 902-6. doi:10.13057/biodiv/d230231.
[13] Awwad S, Issa R, Alnsour L, Albals D, Al-Momani I. Quantification of caffeine and chlorogenic acid in green and roasted coffee samples using HPLC-DAD and evaluation of the effect of degree of roasting on their levels. Molecules. 2021; 26(24): 7502. doi:10.3390/molecules26247502.
[14] Dąbrowska I, Grzędzicka J, Niedzielska A, Witkowska-Piłaszewicz O. Impact of chlorogenic acid on peripheral blood mononuclear cell proliferation, oxidative stress, and inflammatory responses in racehorses during exercise. Antioxidants. 2023; 12(11): 1924. doi:10.3390/antiox12111924.
[15] Nna VU, Abu Bakar AB, Ahmad A, Eleazu CO, Mohamed M. Oxidative stress, NF-κB-mediated inflammation and apoptosis in the testes of streptozotocin-induced diabetic rats: Combined protective effects of Malaysian propolis and metformin. Antioxidants. 2019; 8(10): 465. doi:10.3390/antiox8100465.
[16] Mohamed AI, Olofinsan KA, Salau VF, Erukainure OL, Islam MS. Ethiopian coffee (Coffea arabica) improves glucose uptake and modulates metabolic enzyme activities linked to hyperglycemia-induced infertility in isolated rats' testes. Food Prod Process Nutr. 2024; 6(1): 47. doi:10.1186/s43014-024-00247-3.
[17] Sofyanita E, Juniarto A, Suwondo A, Nugroho H. The effect of chlorogenic acid content in coffee can reduce malondialdehyde (MDA) and increase testosterone hormone levels in DM conditions. Gema Lingkung Kesehat. 2025; 23(1): 20-6. doi:10.36568/gelinkes.v23i1.163.
[18] Hasanuddin R, Jasmiadi J, Abdillah N. The analysis of the chlorogenic acid in the ethanol fraction of Robusta coffee beans and its effect on glucose levels in Wistar rats. Dis Prev Public Health J. 2021; 15(2): 118. doi:10.12928/dpphj.v15i2.4705.
[19] AlAmri OD, Albeltagy RS, Akabawy AMA, Mahgoub S, Abdel-Mohsen DM, Abdel Moneim AE, et al. Investigation of antioxidant and anti-inflammatory activities as well as the renal protective potential of green coffee extract in high fat-diet/streptozotocin-induced diabetes in male albino rats. J Funct Foods. 2020; 71: 103996. doi:10.1016/j.jff.2020.103996.
[20] Soliman GA, Abdel-Rahman RF, Ogaly HA, Althurwi HN, Abd-Elsalam RM, Albaqami FF, et al. Momordica charantia extract protects against diabetes-related spermatogenic dysfunction in male rats: Molecular and biochemical study. Molecules. 2020; 25(22): 5255. doi:10.3390/molecules25225255.
[21] Mou J. The mitochondrial basis of aging and related diseases. E3S Web Conf. 2021; 271: 03033. doi:10.1051/e3sconf/202127103033.
[22] Bisht S, Sisodia SS, Juyal A. Exploring the therapeutic potential of Coffea arabica: Pharmacological evaluation of extracts, fractions, and isolated compounds to treat diabetes-associated hepatic dysfunction and dyslipidemia in rats. Nat Prod Res. 2026 40(10): 2676-2684. doi: 10.1080/14786419.2024.2437026.
[23] Lee R, Lee W, Kim DW, Park HJ. Diazinon induces testicular dysfunction and testicular cell damage through increased reactive oxygen species production in mouse. Cell Death Discov. 2025; 11(1): 113. doi:10.1038/s41420-025-02399-8.
[24] Ahriyasna R, Primal D. The effect of Kawa Daun (Coffea canephora) decoction on blood glucose levels and pancreatic β-cells regeneration in rats with diabetes. Potravinarstvo Slovak J Food Sci. 2023; 17: 122-31. doi: 10.5219/1830.
[25] İrkin LC. Investigation of apoptotic and inflammatory activity in liver tissue of rats fed with clam (Pecten maximus, Linnaeus 1758). Rev Rom Med Lab. 2021; 29: 189-200. doi:10.2478/rrlm-2021-0013.
[26] Zhang W, Lei X, Tu Y, Ma T, Wen T, Yang T, et al. Coffee and the risk of osteoarthritis: A two-sample, two-step multivariable Mendelian randomization study. Front Genet. 2024; 15: 1340044. doi:10.3389/fgene.2024.1340044.
[27] Matsuyama S, DeFalco T. Steroid hormone signaling: multifaceted support of testicular function. Front Cell Dev Biol. 2024; 11: 1339385. doi:10.3389/fcell.2023.1339385.
[28] Hao Y. Inflammation-mediated effects of diabetes mellitus on male fertility: A systematic review and meta-analysis. Front Endocrinol (Lausanne). 2025; 16: 1600565. doi:10.3389/fendo.2025.1600565.
[29] Xing DY, Jin Y, Jin B. A narrative review on inflammaging and late-onset hypogonadism. Front Endocrinol (Lausanne). 2024; 15: 1291389. doi:10.3389/fendo.2024.1291389.
[30] Yalçın T, Kaya S, Yiğin A, Ağca CA, Özdemir D, Kuloğlu T, et al. The effect of thymoquinone on the TNF-α/OTULIN/NF-κB axis against cisplatin-induced testicular tissue damage. Reprod Sci. 2024; 31(8): 2433-46. doi: 10.1007/s43032-024-01567-y.
[31] Fu X, Shen X, Yin X, Zhang YH, Wang X, Han Z, et al. Antioxidant and pro-apoptosis activities of coffee husk (Coffea arabica) anthocyanins. Int Food Res J. 2021; 28: 1187-95. doi:10.47836/ifrj.28.6.10
[32] Sofyanita EN, Maulana MR, Auliya QA, Setyowatiningsih L, Swastyaputri AP. Pengaruh pemberian kopi robusta (Coffea canephora) terhadap penurunan kadar glukosa darah pada

tikus Rattus norvegicus dengan kondisi diabetes melitus. Arter J Ilmu Kesehat. 2024; 5(3): 33-9. doi:10.37148/arteri.v5i3.439.

[33] Maulana MR. Proteksi biji kopi dalam memperbaiki morfologi sperma pada tikus yang diinduksi streptozotocin. J Lab Khatulistiwa. 2024; 7(2): 128-32. doi:10.30602/jlk.v7i2.1421.

[34] Marcolino E, Salavarria D, Silva LGM, Almeida A, Silva FMO, Ribeiro CAAS, et al. Valorization of baobab seeds (Adansonia digitata) as a coffee-like beverage: Evaluation of roasting time on bioactive compounds. J Food Sci Technol. 2023; 61(4): 727-33. doi:10.1007/s13197-023-05873-2.

[35] Saad EA, Hassan HA, Ghoneum M, El-Dein MA. Edible wild plants, chicory and purslane, alleviated diabetic testicular dysfunction and insulin resistance via suppression of 8OHdG and oxidative stress in rats. PLoS One. 2024; 19(4): e0301454. doi:10.1371/journal.pone.0301454.

Journal of Associated Medical Sciences 2026; 59(3): 305-312

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 of self-stigma and psychological distress among parents of children with autism spectrum disorder: A cross-sectional study

Harshavardhana Anandakumar, Raghuram Pattusamy*, Loganathan Shanmugam, Deepika Pakkiri

Faculty of Occupational Therapy, Sri Ramachandra Institute of Higher Education and Research (DU), Chennai, Tamil Nadu, India.

ARTICLE INFO

Article history:
Received 20 February 2026
Accepted as revised 6 July 2026
Available online 22 July 2026

Keywords:
Autism spectrum disorder, mental health, psychological distress, social stigma

ABSTRACT

Background: Autism Spectrum Disorder (ASD) is a lifelong neurodevelopmental disorder characterized by challenges in social communication and restricted behaviors. Parents of children with ASD often face self-stigma-the internalization of negative societal attitudes which contributes to psychological distress. Given the multifaceted caregiving demands associated with ASD, these parents are particularly vulnerable to stress and psychological burden across the developmental course of their child.

Objectives: To investigate the prevalence of, and relationship between, self-stigma and psychological distress among parents of children diagnosed with ASD.

Materials and methods: A cross-sectional study was conducted among 138 parents recruited from special educational centers in Chennai, India. Data were collected using the Parents' Self-Stigma Scale (PSSS) and the Kessler Psychological Distress Scale (K10) and analyzed using IBM SPSS Statistics (Version 26).

Results: Among the parents, 4.3% reported severe self-stigma, 57.2% moderate, and 37.7% mild. Regarding psychological distress, 21% exhibited severe distress, 18.8% moderate, 20.3% mild, and 39.9% were classified as 'well'. A statistically significant positive correlation was observed between self-stigma and psychological distress (r_s=0.311, p<0.001).

Conclusion: Higher levels of self-stigma among parents are significantly associated with elevated psychological distress. Interventions targeting self-stigma reduction, including cognitive reframing, acceptance-based therapeutic strategies, and peer-facilitated support, could mitigate psychological distress and enhance parental well-being.

Introduction

Autism spectrum disorder (ASD) is a lifelong neurodevelopmental disorder, which is characterized by deficits in social communication and the use of language, as well as by restricted interests and repetitive behaviors. According to the Fifth Edition of the Diagnostic and Statistical Manual of Mental Disorders, ASD was defined by integrating distinct subcategories, namely autistic disorder, Asperger's syndrome, and pervasive developmental disorder-not otherwise specified (PDD-NOS).[1] The symptoms observed in ASD vary with age, language capability (with/without speech), IQ, and gender. In infancy and early childhood, common presenting symptoms and signs include delays in language development, regression of language, or a

*Corresponding contributor.
Author's Address: *Faculty of Occupational Therapy, Sri Ramachandra Institute of Higher Education and Research (DU), Chennai, Tamil Nadu, India.*
E-mail address: raghurammot@sriramachandra.edu.in
doi: 10.66285/JAMS.2026.103
E-ISSN: 2539-6056

lack of development of language and communicative gestures.[2-4] Children with ASD may not have many noticeable physical signs but demonstrate behaviors that can be socially misinterpreted, such as flapping, twirling, or having tantrums. Consequently, they may be perceived as 'rebellious' or 'disruptively behaving'.[5]

The global prevalence of ASD has increased substantially over recent decades. Whereas early estimates placed prevalence at approximately 0.2 per 1,000 children, contemporary figures indicate a prevalence of approximately 1 in 100 children worldwide, representing an increase attributable in part to broader diagnostic criteria, improved awareness, and enhanced surveillance methodologies.[6] In India specifically, a systematic review and meta-analysis by Chauhan *et al.* estimated the prevalence of ASD at 0.09% in urban settings and 0.11% in rural settings, with regional and methodological variation warranting continued epidemiological investigation.[7]

Stigma can be defined as a mark of social distinction that is assumed to imply undesirable characteristics or behaviors considered unfavorable in society. Goffman further noted that stigma is situational; the perception of such stigma can produce a deleterious effect on the physical, psycho-social, and communal well-being not only of the person with disability but of his or her family and relatives as well.[8] Parents experience negative attitudes held by society and attribute their child's condition to themselves, which is more evident among parents of children with ASD.[9,10] This internalization only compounds feelings of self-stigma, which stem from societal opinions alongside the notion that it is the parent's responsibility to mold the child into an acceptable social participant.[11-13] Parents and caregivers of a child with ASD may feel that they are inadequate parents, ashamed, guilty, and unable to help the situation. These feelings may be exacerbated by judgments from the outside world and misunderstanding, which further adds to the sense of parental incompetence.[14]

The lived experience of self-stigma among parents of children with ASD may manifest across multiple domains of daily life. For example, a parent who avoids taking their child to a public space out of fear that others will judge their child's atypical behavior or who internalizes such judgment as evidence of their own parental failure exemplifies the behavioral and emotional consequences of internalized stigma.[11] These everyday experiences, repeated across social, educational, and healthcare encounters, compound over time and may translate into measurable psychological burden.

The American Psychological Association defines psychological distress as a set of painful mental and physical symptoms that are associated with normal fluctuations of mood. Parenting children with ASD can provide daily stressors and possible challenges, including the child's behavioral difficulties and impaired communicative functioning.[15] Parents of children with ASD may be more vulnerable to high levels of stress and psychological discomfort throughout the toddler and preschool years because of the special demands and stresses that accompany these developmental stages.[16,17]

The mental health of the parent has an impact on the standard of care and the long-term rehabilitation results of their child. Psychologically distressed caregivers are more likely to engage in less responsive, less consistent, and less stimulating parenting behaviors, which in turn may impede the child's communicative and social development.[18] Moreover, parental stress and depression have been associated with reduced implementation fidelity of home-based intervention programs, diminished advocacy for the child's educational and therapeutic needs, and lower rates of participation in caregiver training initiatives.[19] Supporting parental psychological well-being improves treatment adherence and fosters a stable home environment, leading to better developmental outcomes for children with ASD.[20] These findings highlight the need to address parental mental health as part of comprehensive, family-centered rehabilitation. The stigma stress model, which implies that awareness and agreement with stigmatizing public attitudes, followed by their application to the self, results in diminished self-esteem and increased emotional suffering, can help explain the relationship between self-stigma and psychological distress.[21]

The process known as "courtesy stigma" or "stigma by association", in which caregivers internalize the stigma directed at their child, exacerbates this pathway in parents of children with ASD.[8] Previous studies by Mak and Kwok showed a positive correlation between poor mental health outcomes and parental self-stigma in Hong Kong.[14] However, similar data from the Indian subcontinent are still scarce. In the context of the significant cultural, familial, and systemic disparities that influence caregiving experiences in India, this study was intended to address that gap by exploring the prevalence of self-stigma and psychological distress among parents of children with ASD in Chennai. The results will ultimately help develop more thorough rehabilitation plans for those impacted families.

Materials and methods

Study design and setting

This cross-sectional study was conducted in four specialized educational centers in Chennai, India: Kumaran Special School (Ennore), Sankalp Special School (Kolappancheri), Maithree Special School (Perambur), and B.K.N Opportunity School (Egmore), in August 2024, over a period of 3 weeks. The selection of these sites was intended to collect a representative sample of parents from varying urban/suburban and socioeconomic/educational backgrounds. All eligible parents were provided with a participant information document prior to data collection, which explained the purpose of the study, the voluntary nature of

participation, and anonymity. Written consent was also obtained from each parent. For parents who were more confident communicating in their own language, a validated Tamil version of the consent form and questionnaires was used to ensure that they had fully understood the study's protocols.

Participants

A total of 138 parents of children and adolescents with a diagnosis of ASD aged 4-17 years were recruited. This age group has been chosen to reflect the population enrolled in the participating special educational centers and aligns with the eligibility criteria defined under the Persons with Disabilities Act, India (2016). The World Health Organization states that adolescence lasts until age 19, but the 17-year age limit in this study aligns with the participating schools' policies on admission to their institutions.

Sampling size

The targeted sample size was determined based on the prevalence of internalized stigma identified in previous studies by Mak and Kwok,[14] who investigated self-stigma among parents of children with neuro-developmental impairments. For statistical validity, a power analysis was performed using their reported prevalence rate of 41% as a baseline. The minimum sample size required was determined to be 138 individuals with a confidence level of 95% and a predetermined margin of error of 8%.

Sampling method

A total of 138 parents of children diagnosed with ASD were recruited using a convenience sampling method.

Inclusion criteria

1. Must be the primary parent or legal guardian providing direct care to the child.
2. Parents of children and adolescents aged between 4 and 17 years diagnosed with Autism Spectrum Disorder.
3. Parents able to read and write Tamil or English.

Exclusion criteria

1. Parents who are not the child's primary caregivers.
2. Parents of a child with siblings who also have any other physical and orthopedic conditions.

Data collection tools

Parents' Self-Stigma Scale (PSSS)

The Parents' Self-Stigma Scale (PSSS) is a standardized psychometric instrument developed by Mak and Kwok to measure the extent to which parents of children with neurodevelopmental conditions internalize societal stigma.[14] Self-blame, shame, social disengagement, and perceived inadequacy are the four key psychological characteristics assessed by the self-report instrument. Using a Likert scale with responses ranging from "strongly disagree" to "strongly agree," participants answer a series of items, enabling both a composite self-stigma score as well as specific subscale insights. The validation study reported a Cronbach's alpha of 0.89, indicating excellent internal consistency. Construct validity was demonstrated through confirmatory factor analysis.

Kessler Psychological Distress scale (K10)

This 10-item self-report instrument has been validated to screen for anxiety or depressive disorders and general psychological distress. Respondents are asked to report how often they experience feelings such as anxiety, sadness, restlessness, and exhaustion during the last 30 days using a 5-point Likert scale. Psychometric evaluation showed excellent internal consistency (Cronbach's α=0.92-0.93) and strong criterion validity with sensitivity of 0.87 and specificity of 0.70 for identifying serious mental illness against ICD-10 criteria. Scores are grouped into 4 levels of distress: low (10-19), moderate (20-24), high (25-29), and very high (30-50), and provide an effective means of identifying individuals who may require further clinical assessment.[22]

Procedure

This cross-sectional study was conducted among parents of children with ASD from Special schools in Chennai. The parents were selected based on the selection criteria using convenience sampling. Prior to participation in the study, informed consent was obtained. The PSSS and K10 were administered to all participants. The collected data were analyzed, and results were inferred.

Data analysis

Data were analyzed using IBM SPSS Statistics (Version 26). Descriptive statistics including frequencies, percentages, means, and standard deviations were computed to summarise sociodemographic and outcome variables. Prior to inferential analysis, normality was assessed using the Shapiro-Wilk test, which revealed that both the PSSS and K10 score distributions deviated significantly from normality (p<0.05). Accordingly, Spearman's rank-order correlation coefficient (rs) was used to examine the association between parental self-stigma and psychological distress. Chi-square ($\chi 2$) tests were applied to explore associations between categorical demographic variables and categorical outcome groups.

Results

Sociodemographic characteristics of participants

Of the 138 participants, the majority (N=78,56.5%) were between 36 and 45 years of age. A smaller proportion fell within the 25-35 (N=36,26.1%), 46-55 (N=23,16.7%), and 56-65 (N=1,0.7%) age groups.

Mothers constituted the predominant respondents (N=113, 81.9%), with fathers representing 18.1% (N=25) of the sample (Table 1).

Prevalence of self-stigma and psychological distress

Analysis of the PSSS scores revealed that 57.2% (N=79) of parents experienced moderate self-stigma, 37.7% (N=52) mild, 4.3% (N=6) severe, and only 0.7% (N=1) scored in the normal range. With regard to psychological distress, 39.9% (N=55) of participants scored in the 'well' category, while 20.2% (N=28) reported mild distress, 18.8% (N=26) moderate distress, and 21.1% (N=29) severe distress, indicating that approximately 60% of the sample experienced some degree of psychological distress. These findings are presented in Table 2.

Association between self-stigma and psychological distress

Spearman's rank correlation analysis revealed a statistically significant positive association between parental self-stigma and psychological distress (rs=0.311, p<0.001). Although the magnitude of the correlation is modest, its significance across the full sample indicates that parents reporting higher levels of internalized stigma are consistently more likely to experience elevated psychological distress. (Table 3).

Table 1. Distribution of demographic characteristics of parents

Demographic characteristics of parents	Category	Number of participants
Age range (years)	25-35	36
	36-45	78
	46-55	23
	56-65	1
Gender distribution (N, %)	Mother	113 (81.9%)
	Father	25 (18.1%)

Table 2. Distribution of PSSS and K10 scores.

Measure	Score category	Frequency	Percentage
PSSS score	Normal	1	0.7
	Mild	52	37.7
	Moderate	79	57.2
	Severe	6	4.3
K10 score	Well	55	39.9
	Mild	28	20.2
	Moderate	26	18.8
	Severe	29	21.1

Table 3. Association between self-stigma and psychological distress within participants.

Correlations (Spearman's rho)		PSSS score	K10 score
PSSS score	Correlation coefficient	1.000	0.311***
	Sig. (2-tailed)	-	<0.001
K10 score	Correlation coefficient	0.311***	1.000
	Sig. (2-tailed)	<0.001	-

Discussion

The findings reveal significant insights into how societal perceptions and internalized stigma contribute to the psychological well-being of these parents. The current study found that 4.3% of parents have severe self-stigma; these parents likely face intense societal judgment, lack sufficient support, and may have limited resources or understanding about ASD. This group might also include parents who are more socially or emotionally isolated, making them more vulnerable to severe internalized stigma. Most parents experienced moderate self-stigma 57.2%. This suggests that while they are affected by stereotypical reactions and internalized blame, they are not overwhelmed by these feelings. These parents likely encounter regular challenges related to societal attitudes and may have internal conflicts about their role as a parent; however, they may also possess some coping strategies that prevent these feelings from escalating.

A significant portion of parents falls into the mild self-stigma category, at 37.7%. These parents may experience some societal pressures and occasional feelings of inadequacy or guilt, but they are generally able to manage these emotions. Factors such as moderate social support, education, and understanding of ASD could contribute to a lower level of internalized stigma. Only 0.7% of parents in the study reported a normal level of self-stigma. This could be attributed to strong resilience, a highly supportive environment, or effective coping mechanisms. These parents may have a deep understanding of ASD, access to resources, or a strong social support network that shields them from internalizing stigma.

Mittal *et al*. stated that parents could employ unhealthy coping mechanisms like hiding and social disengagement, which exacerbate the stigma.[23] Additionally, they could adopt adaptive coping mechanisms, including seeking spiritual or religious support, seeking social support, accessing social networks, actively seeking knowledge about ASD, and participating in advocacy and education campaigns. Adaptive coping enables caregivers to avoid being constrained by social stereotypes and judgments about their children's conduct.[18]

Among the parents, 21.0% experienced severe psychological distress, likely due to high caregiving demands, limited support, and persistent stress from managing their child's needs and societal stigma. Moderate distress (18.8%) may reflect partial support from family or community, which helps but does not fully alleviate stress. Mild distress (20.3%) could be attributed to effective coping strategies, such as self-care, professional help, and realistic expectations. 39.9% who were likely to be well appeared to benefit from strong support systems and effective problem-solving skills, helping them maintain better mental health.

The results from the current study show that mothers experience higher levels of self-stigma (severe and extremely severe) compared to fathers, who experience moderate self-stigma. This difference can be attributed to traditional caregiving roles, societal expectations, cultural norms, and the different ways in which mothers and fathers cope with the challenges of raising a child with ASD. Mothers, who constituted 81.9% of respondents, have been shown in multiple studies to assume primary caregiving responsibilities and are therefore disproportionately exposed to stigma-generating social interactions.[24] Cultural norms in India further reinforce maternal caregiving roles, amplifying the social visibility and hence the stigma exposure of mothers in this context.[25]

Association between self-stigma and psychological distress

The analysis reveals a statistically significant, positive correlation between parental self-stigma and psychological distress (rs= 0.311, p<0.001). This means that parents who score higher on the self-stigma scale are significantly more likely to also report higher levels of psychological distress. This finding is consistent with existing evidence by Eaton *et al*., who found a positive association between parental awareness of stigma, self-stigma, and emotional suffering.[26]

The observed Spearman's correlation (rs=0.311, p<0.001) falls within the moderate range according to Cohen's conventional criteria, where values between 0.30-0.49 indicate a moderate association. While this coefficient may appear modest in absolute terms, it is important to contextualize its practical significance. Psychological distress is a complex, multidetermined outcome shaped by factors including social support, coping style, severity of the child's ASD symptoms, socioeconomic status, and marital functioning- none of which were controlled for in the present analysis. Within this context, self-stigma independently accounting for approximately 9.7% of the variance in psychological distress (rs^2=0.097) is both statistically meaningful and clinically relevant. These findings underscore the importance of addressing internalized stigma as one modifiable contributor to parental psychological burden.

The relatively moderate correlation may be explained by several factors. Firstly, psychological distress is shaped by numerous variables beyond self-stigma, including child behavior severity, family support, and financial stress. Secondly, the majority of parents scoring in the moderate self-stigma range results in a restricted range in the predictor variable, which typically attenuates correlation estimates. Thirdly, Indian families may have access to extended family support systems that partially buffer against psychological distress despite high self-stigma levels.[27]

The observed association between self-stigma and psychological distress in this study can be understood through several interrelated mechanisms. According to the cognitive-behavioral model of stigma, when parents internalize negative societal beliefs about their child's

condition, these beliefs activate cognitive schemas of self-blame and inadequacy, which in turn fuel affective states of shame, guilt, and hopelessness.[21] These emotional consequences are consistent with cognitive appraisal theory by Lazarus & Folkman, wherein caregivers who perceive their situation as beyond their coping capacity report greater emotional distress.[28] Furthermore, self-stigma has been shown to reduce help-seeking behavior, thereby limiting access to professional and social support.[26,29] The resulting social isolation creates a self-reinforcing cycle in which stigma leads to withdrawal, which reduces access to corrective social experiences, further entrenching negative self-perceptions and elevating distress.

Hayes *et al.*'s meta-analysis (N=284 studies) confirmed that parents of children with ASD experience significantly higher stress and depression compared to parents of typically developing children.[20] Karst and Van Hecke *et al.* linked parental psychological distress in ASD families to child behavior severity, social isolation, and lack of professional support.[19] Kuhlthau *et al.* documented that unmet service needs independently contribute to parental psychological burden.[30] Notably, studies have reported weaker associations between stigma and distress when robust social support is present, highlighting the moderating role of support systems and explaining the modest rather than strong correlation observed here.[27]

Given the significant association between self-stigma and psychological distress observed in this study, occupational therapy has a specific and valuable role to play in supporting parents of children with ASD. Through role-focused interventions, occupational therapists can help parents reframe their identity beyond the caregiving role, restoring meaningful engagement in occupations such as leisure, social participation, and self-care that may have been relinquished due to the emotional demands of caregiving and the social withdrawal driven by self-stigma.[31] Family-centered practice models, which position the parent as an active collaborator in the therapeutic process rather than a passive recipient of care, may further reduce feelings of inadequacy and enhance parental self-efficacy, both of which are dimensions of self-stigma assessed in the PSSS.[32]

Psychoeducational workshops and individual therapeutic sessions targeting self-stigma reduction have shown promise in analogous populations. Corrigan *et al.* demonstrated that Acceptance and Commitment Therapy (ACT) based approaches significantly reduced internalized stigma in mental health service users.[29] Similarly, self-compassion interventions have been associated with lower shame and improved psychological well-being in caregivers of individuals with developmental disabilities.[33] Encouraging parents to maintain engagement in meaningful occupations unrelated to caregiving such as creative pursuits, social activities, or vocational roles may help rebuild a positive self-concept and reduce the centrality of the stigmatized caregiver identity to overall self-worth.[31,34]

Limitations

Firstly, the cross-sectional design precludes causal inferences regarding the direction of the association between self-stigma and psychological distress. Secondly, convenience sampling may limit the generalisability of findings to parents from rural settings or those whose children are not enrolled in formal educational programs. Thirdly, potential confounders including parental mental health history, child's severity, and socioeconomic status were not assessed. Future research should address these limitations by employing longitudinal designs, probability-based sampling strategies, structured clinical interviews for parental mental health, and the inclusion of child-level variables such as ASD severity and functional profile.

Conclusion

This study provides novel evidence from an underrepresented South Asian context demonstrating that a substantial majority of parents of children with ASD in Chennai experience moderate-to-severe levels of self-stigma, and that higher internalized stigma is significantly associated with greater psychological distress. To the best of our knowledge, this is the first study in the Indian context to employ the PSSS alongside the K10 to systematically characterize this association.

These findings have direct clinical relevance: healthcare professionals and rehabilitation specialists, including occupational therapists, should adopt family-centered approaches that proactively screen for self-stigma and implement targeted interventions such as cognitive reframing, acceptance-based therapeutic strategies, and peer-facilitated support groups.

Ethical approval

The study was approved by the institutional ethics committee for student proposals [CSP/24/JUN/149/259].

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

Harshavardhana Anandakumar: conceptualization, methodology, investigation, data curation, writing: original draft; **Raghuram Pattusamy:** supervision, methodology, writing: review and edit, validation; **Loganathan Shanmugam:** methodology, writing: review and edit, validation; **Deepika Pakkiri:** methodology, supervision, writing: review and edit.

Acknowledgements

The authors would like to thank the management, staff, and students of the participating special schools for their support and cooperation during data collection. The authors also gratefully acknowledge the parents who participated in this study for their time and willingness to share their experiences.

References

[1] American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: APA; 2013. doi: 10.1176/appi.books.9780890425596.

[2] Chawarska K, Shic F, Macari S, Campbell DJ, Brian J, Landa R, et al. 18-month predictors of later outcomes in younger siblings of children with autism spectrum disorder: A baby siblings research consortium study. J Am Acad Child Adolesc Psychiatry. 2014; 53(12): 1317-27. doi: 10.1016/j.jaac.2014.09.015.

[3] Centers for Disease Control and Prevention. Signs and symptoms of autism spectrum disorder. Atlanta, GA: CDC; 2023. Available from: https://www.cdc.gov/autism/signs-symptoms/index.html

[4] Cardon T, editor. Autism navigator: A collection of web-based courses and tools for families and professionals. Tallahassee: Florida State University; 2017. Available from: https://autismnavigator.learnercommunity.com

[5] Botha M, Dibb B, Frost DM. "Autism is me": An investigation of how autistic individuals make sense of autism and stigma. Disabil Soc. 2022; 37(3): 427-53. doi: 10.1080/09687599.2020.1822782.

[6] 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.

[7] Chauhan A, Sahu JK, Jaiswal N, Kumar K, Agarwal A, Kaur J, et al. Prevalence of autism spectrum disorder in Indian children: A systematic review and meta-analysis. Neurol Sci. 2019; 40(7): 1337-47. doi: 10.4103/0028-3886.253970.

[8] Goffman E. Stigma: Notes on the management of spoiled identity. New York: Simon and Schuster Publishers Ltd; 2009.

[9] Farrugia D. Exploring stigma: Medical knowledge and the stigmatisation of parents of children diagnosed with autism spectrum disorder. Sociol Health Illn. 2009; 31(7): 1011-27. doi: 10.1111/j.1467-9566.2009.01174.x.

[10] Gill J, Liamputtong P. Being the mother of a child with Asperger's syndrome: Women's experiences of stigma. Health Care Women Int. 2011; 32(8): 708-22. doi: 10.1080/07399332.2011.555830.

[11] Kinnear SH, Link BG, Ballan MS, Fischbach RL. Understanding the experience of stigma for parents of children with autism spectrum disorder and the role stigma plays in families' lives. J Autism Dev Disord. 2016; 46(3): 942-53. doi: 10.1007/s10803-015-2637-9.

[12] Neely-Barnes SL, Hall HR, Roberts RJ, Graff JC. Parenting a child with an autism spectrum disorder: Public perceptions and parental conceptualizations. J Fam Soc Work. 2011; 14(3): 208-25. doi: 10.1080/10522158.2011.571539.

[13] Woodgate RL, Ateah C, Secco L. Living in a world of our own: The experience of parents who have a child with autism. Qual Health Res. 2008; 18(8): 1075-83. doi: 10.1177/1049732308320112.

[14] Mak WW, Kwok YT. Internalization of stigma for parents of children with autism spectrum disorder in Hong Kong. Soc Sci Med. 2010; 70(12): 2045-51. doi: 10.1016/j.socscimed.2010.02.023.

[15] Bristol AA. A developmental perspective on stress and coping in families of autistic children. Severely Handicapped Young Children and Their Families. 1984: 91-141.

[16] Hall HR, Graff JC. The relationships among adaptive behaviors of children with autism, family support, parenting stress, and coping. Issues Comprehensive Pediatr Nurs. 2011; 34(1): 4-25. doi: 10.3109/01460862.2011.555270.

[17] Behr SK, Murphy DL. Research progress and promise: The role of perceptions in cognitive adaptation to disability. In: Turnbull AP, Patterson JM, Behr SK, Murphy DL, Marquis JG, Blue-Banning MJ, editors. Cognitive coping, families, and disability. Baltimore: Paul H Brookes; 1993. p.151-63.

[18] Osborne LA, McHugh L, Saunders J, Reed P. Parenting stress reduces the effectiveness of early teaching interventions for autistic spectrum disorders. J Autism Dev Disord. 2008; 38(6): 1092-10. doi: 10.1007/s10803-007-0497-7.

[19] Karst JS, Van Hecke AV. Parent and family impact of autism spectrum disorders: A review and proposed model for intervention evaluation. Clin Child Fam Psychol Rev. 2012; 15(3): 247-77. doi: 10.1007/s10567-012-0119-6.

[20] Hayes SA, Watson SL. The impact of parenting stress: A meta-analysis of studies comparing the experience of parenting stress in parents of children with and without autism spectrum disorder. J Autism Dev Disord. 2013; 43(3): 629-42. doi: 10.1007/s10803-012-1604-y.

[21] Corrigan PW, Larson JE, Ruesch N. Self-stigma and the 'why try' effect: Impact on life goals and evidence-based practices. World Psychiatry. 2009; 8(2): 75-81.

[22] Hoffman J, Cossie Q, Ametaj AA, Kim HH, James R, Stroud RE, et al. Construct validity and factor structure of the Kessler-10 in South Africa. BMC Psychol. 2022; 10(1): 177. doi: 10.1186/s40359-022-00883-9.

[23] Mittal D, Sullivan G, Chekuri L, Allee E, Corrigan PW.

Empirical studies of self-stigma reduction strategies: A critical review of the literature. Psychiatr Serv. 2012; 63(10): 974-81. doi: 10.1176/appi.ps.201100459.

[24] Gray DE. Gender and coping: The parents of children with high functioning autism. Soc Sci Med. 2003; 56(3): 631-42. doi: 10.1016/s0277-9536(02)00059-x.

[25] Srinivasan SM, Bhat AN. A review of "music and movement" therapies for children with autism: Embodied interventions for multisystem development. Front Integr Neurosci. 2013; 7: 22. doi: 10.3389/ fnint. 2013.00022.

[26] Eaton K, Ohan JL, Stritzke WG, Corrigan PW. Failing to meet the good parent ideal: Self-stigma in parents of children with mental health disorders. J Child Fam Stud. 2016; 25(10): 3109-23.

[27] Werner S, Shulman C. Subjective well-being among family caregivers of individuals with developmental disabilities: The role of affiliate stigma and psychosocial moderating variables. Res Dev Disabil. 2013; 34(11): 4103-14. doi: 10.1016/j.ridd.2013.08.029.

[28] Lazarus RS, Folkman S. Stress, appraisal, and coping. New York: Springer Publishing Company; 1984.

[29] Corrigan PW, Druss BG, Perlick DA. The impact of mental illness stigma on seeking and participating in mental health care. Psychol Sci Public Interest. 2014; 15(2): 37-70. doi: 10.1177/1529100614531398.

[30] Kuhlthau CC, Heinström J, Todd RJ. The 'information search process' revisited: Is the model still useful. Inf Res. 2008; 13(4): 13-4.

[31] Gutman SA, Schindler VP. The neurological basis of occupation. Occup Ther Int. 2007; 14(2): 71-85. doi: 10.1002/oti.225.

[32] Rodger S, Umaibalan V. The routines and rituals of families of typically developing children compared with families of children with autism spectrum disorder: An exploratory study. Br J Occup Ther. 2011; 74(1): 20-6. doi: 10.4276/030802211X12947686093567.

[33] Neff KD, Faso DJ. Self-compassion and well-being in parents of children with autism. Mindfulness. 2015; 6(4): 938-47. doi: 10.1007/s12671-014-0359-2.

[34] Eaton K, Ohan JL, Stritzke WG, Courtauld HM, Corrigan PW. Mothers' decisions to disclose or conceal their child's mental health disorder. Qual Health Res. 2017; 27(11): 1628-39. doi: 10.1177/1049732317697096.

Journal of Associated Medical Sciences 2026; 59(3): 313-324

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



Reduction of decoder convolutional blocks in Attention Gate U-Net++ to enhance computational efficiency in automatic HR-CTV delineation on CT-based cervical cancer brachytherapy images

Thinnagit Srikhot[1], Chatrawut Pattaweerakul[2], Kamonchanok Nobphuek[2], Titipong Kaewlek[1,3*]

[1]Medical Physics Program, Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Mueang District, Phitsanulok Province, Thailand.
[2]Department of Radiology, Faculty of Medicine, Naresuan University, Mueang District, Phitsanulok Province, Thailand.
[3]Interdisciplinary Health and Data Sciences Research Unit, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.

ARTICLE INFO

Article history:
Received 15 April 2026
Accepted as revised 21 July 2026
Available online 27 July 2026

Keywords:
Cervical cancer, U-Net, U-Net++, attention U-Net++, brachytherapy

ABSTRACT

Background: The contour of the high-risk clinical target volume (HR-CTV) is important in computed tomography (CT)-based cervical cancer brachytherapy to ensure adequate target tumor coverage while sparing radiation exposure to organs at risk (OARs). Although Attention U-Net++ has demonstrated improved segmentation performance, its complex decoder structure substantially increases computational cost and training time.

Objectives: This study aimed to develop a computationally efficient modified Attention U-Net++ architecture by reducing decoder depth while preserving segmentation performance for HR-CTV delineation.

Materials and methods: A retrospective dataset comprising 3,323 CT images from 56 patients with cervical cancer was analyzed. The proposed model eliminates the deepest decoder level to reduce redundant feature propagation and overall model complexity. Segmentation performance was evaluated using the Dice Similarity Coefficient (DSC), Intersection over Union (IoU), 95th percentile Hausdorff Distance (HD95), Accuracy, Precision, and Recall. Computational efficiency was assessed by the number of trainable parameters, training time, and inference time, and was compared with the baseline Attention U-Net++ architecture and related models.

Results: The proposed Attention U-Net++ DL3 reduced the number of trainable parameters from 37.30 million to 8.86 million and shortened training time from 52.27 minutes to 33.11 minutes, while maintaining a comparable inference time (36.90 seconds). Geometric evaluation demonstrated a DSC of 0.9076 and IoU of 0.8376, marginally outperforming the baseline Attention U-Net++ (DSC 0.9057; IoU 0.8338). The HD95 value was 1.3959, which is comparable to existing architectures. Pixel-level classification performance remained robust, with an accuracy of 0.9996, precision of 0.9046, and recall of 0.9298. Overall, the proposed model achieved competitive segmentation performance with substantially reduced model complexity.

Conclusion: Reducing decoder convolutional blocks in Attention U-Net++ significantly improves computational efficiency while preserving segmentation accuracy. The proposed lightweight architecture provides an optimal balance between performance and resource utilization, supporting its practical implementation in CT-based cervical cancer brachytherapy, particularly in resource-limited clinical settings.

* **Corresponding contributor.**
Author's Address: *Medical Physics Program, Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Mueang District, Phitsanulok Province, Thailand.*
E-mail address: titipongk@nu.ac.th
doi: 10.66285/JAMS.2026.104
E-ISSN: 2539-6056

Introduction

Gynecological cancers arise within the female reproductive organs and primarily include cervical, ovarian, and endometrial cancers. Among these, cervical cancer remains one of the leading causes of cancer-related mortality in women worldwide. Data from the Surveillance, Epidemiology, and End Results (SEER) program indicate that cervical cancer accounted for approximately 342,000 deaths between 1975 and 2020, representing the highest mortality rate among gynecological malignancies during this period.[1] Cervical cancer predominantly affects women aged 30-70 years and is strongly associated with persistent infection with high-risk human papillomavirus (HPV), which can induce precancerous cellular abnormalities and ultimately lead to malignant transformation of cervical epithelial cells.[2]

Brachytherapy is an essential treatment modality for cervical cancer, particularly in locally advanced disease. This technique delivers high-dose radiation directly to the tumor using intracavitary applicators inserted into the uterus, enabling precise target coverage while minimizing radiation exposure to surrounding healthy tissues.[3] Current clinical practice follows the recommendations of the GEC-ESTRO Working Group IV, which advocates the use of magnetic resonance imaging (MRI) as the primary imaging modality for brachytherapy planning due to its superior soft-tissue contrast and improved accuracy in delineating target volumes and organs at risk (OARs).[4] However, MRI is not universally available in all clinical settings. In such cases, the IBS-GEC ESTRO-ABS guidelines recommend computed tomography (CT) as an alternative, provided that a pre-treatment MRI is performed to guide CT-based treatment planning throughout the brachytherapy course.[5]

Despite advances in imaging and treatment techniques, CT-based cervical cancer brachytherapy remains challenging. Accurate delineation of the high-risk clinical target volume (HR-CTV) is critical because high radiation doses must be confined to the tumor while sparing adjacent organs, such as the bladder and rectum. Manual contouring is time-consuming and operator-dependent; furthermore, patients are often required to remain immobilized with an intracavitary applicator in place for prolonged periods, which may cause discomfort and reduce treatment tolerance.[6]

To address these challenges, deep learning methods have increasingly been applied to medical image segmentation. Ronneberger *et al.* introduced the U-Net architecture, which demonstrated strong performance in biomedical image segmentation tasks.[7] However, subsequent studies reported that the performance of the conventional U-Net may degrade when segmenting small or low-contrast anatomical structures due to the inclusion of irrelevant background features.[8] To mitigate this limitation, attention mechanisms were introduced to enable selective focus on salient regions.[9] In parallel, Zhou *et al.* proposed U-Net++, which enhances feature representation through densely connected convolutional blocks and deep supervision, thereby improving training stability and segmentation performance.[10] More recently, attention mechanisms have been integrated into U-Net++ skip connections to further refine feature selection.[11] However, these modifications substantially increase model complexity, resulting in higher computational costs and longer training times.

Although Attention U-Net++ improves segmentation accuracy, its high number of trainable parameters may limit practical implementation in clinical settings with constrained computational resources.[12] To date, limited studies have systematically investigated architectural simplification strategies for Attention U-Net++ in CT-based cervical cancer brachytherapy segmentation. Therefore, this study aimed to develop and evaluate a lightweight modified Attention U-Net++ architecture by reducing decoder depth while maintaining segmentation performance.[13]

Materials and methods

This study was conducted to evaluate the proposed segmentation model for CT-based cervical cancer brachytherapy. Ethical approval was obtained from the Institutional Review Board prior to data collection. The study utilized medical imaging data from patients treated between 2022 and 2026. All patient data were anonymized prior to analysis to ensure confidentiality and compliance with ethical standards.

Patient selection

A total of 56 patients diagnosed with cervical cancer, aged between 30-70 years, were included in this study. All patients were clinically staged according to the International Federation of Gynecology and Obstetrics (FIGO) staging system (IA-IIIC). Each patient underwent high-dose-rate (HDR) intracavitary brachytherapy using the Fletcher-Suit applicator system, consisting of tandem and ovoid components (3-4 channels). The radioactive source used for treatment was iridium-192 (^{192}Ir). A prescribed dose of 7 Gy per fraction was delivered for 1-4 fractions depending on clinical indications. Prior to each brachytherapy session, CT simulation was performed with a standardized bladder filling protocol of 50 mL sterile water to maintain consistent organ positioning. CT images were acquired using the Philips Brilliance Big Bore system with a slice thickness of 2 mm. After-loader brachytherapy treatment was delivered using the GammaMed Plus iX system. All imaging data were stored in Digital Imaging and Communications in Medicine (DICOM) format (Table 1). Radiation therapy structures (RT-Structure), including target volumes and organs at risk, were manually delineated by a board-certified radiation oncologist with more than 10 years of clinical experience.

Table 1. Patient characteristics.

Characteristics	
Age (years)	30-70
Retrospective study selection	2022-2026
Radiation source	Iridium-192
Prescription dose per fraction (Gy)	7 Gy x 1-4 fractions
Applicator	Fletcher suit
FIGO clinical stage	IA-IIIC
No. of patients	56
CT slice thickness (mm)	2

Note: *CT: computed tomography, FIGO: International Federation of Gynecology and Obstetrics.*

Patient dataset preparation

A total of 3,323 CT images were collected from 56 patients. On average, each patient underwent approximately four CT simulation sessions (fractions) during brachytherapy treatment, resulting in multiple CT image sets per patient. To prevent data leakage and ensure an unbiased evaluation of model performance, the dataset was partitioned on a patient-wise basis, such that all CT images from the same patient, including all treatment fractions, were assigned exclusively to a single subset.

The dataset was divided into three subsets for deep learning model development: the training set consisted of 2,684 CT images from 37 patients, the validation set comprised 299 CT images from 8 patients, and the testing set included 340 CT images from 11 patients. This patient-wise splitting strategy ensured that no CT images or treatment fractions from the same patient were shared across the training, validation, and testing subsets, thereby minimizing the risk of information leakage and providing a more reliable assessment of the model's generalization performance.

Ground-truth segmentation masks were generated from Digital Imaging and Communications in Medicine (DICOM) datasets comprising CT image files and their corresponding Radiation Therapy Structure Set (RT-STRUCT) files. All datasets were imported into 3D Slicer (version 5.8.1) for visualization and contour processing (Figure 1).

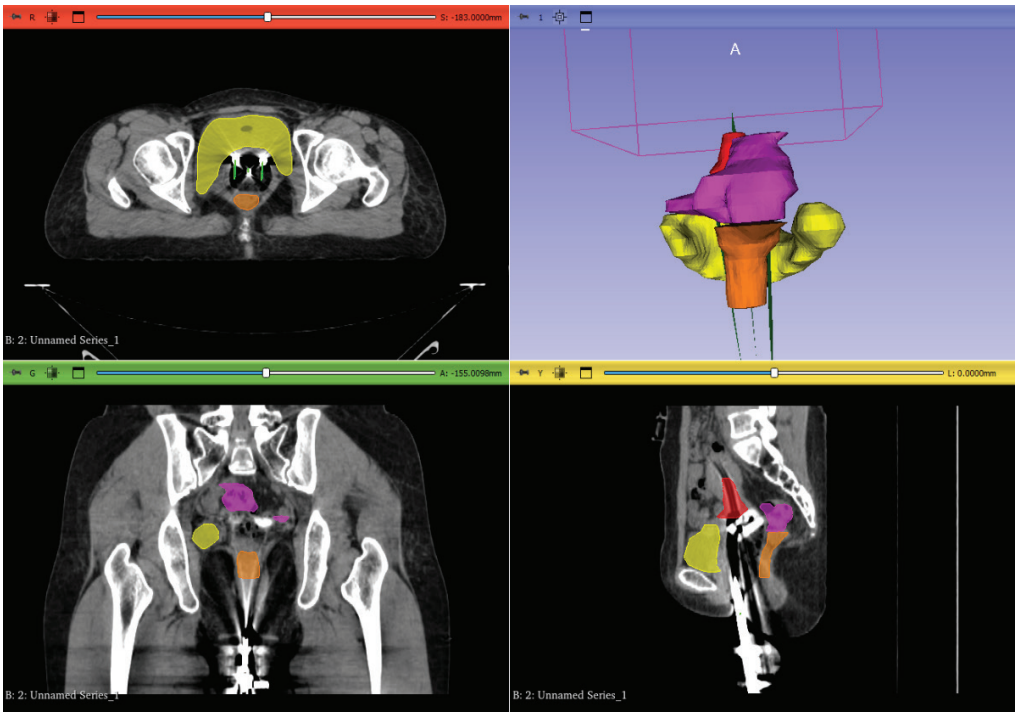


Figure 1. *Visualization of pelvic anatomical structures, highlighting the delineated HR-CTV and organs at risk (OARs) using 3D Slicer software (version 5.8.1) derived from DICOM CT datasets.*

The high-risk clinical target volume (HR-CTV) was derived from RTSTRUCT delineations and processed using 3D Slicer (version 5.8.1) to generate binary segmentation masks. The segmented volumes were exported in Neuroimaging Informatics Technology Initiative (NIfTI; .nii) format and subsequently imported into Google Colab for further processing. Both CT images and corresponding segmentation masks were resized from 512×512 pixels to 256×256 pixels to enhance computational efficiency.[14] The processed images were saved in Portable Network Graphics (PNG) format using automated file naming and indexing procedures to maintain accurate CT-mask pairing. Finally, all images were converted into NumPy array (.npy) format for model training. Data augmentation techniques, including rotation, shearing, horizontal flipping, and vertical flipping, were applied to the training and validation sets to increase data variability and mitigate overfitting.[15] The test set remained unaltered to ensure an unbiased performance evaluation.

The Attention Gate U-Net++ architecture

The integration of attention mechanisms into U-Net++ enhances the network's ability to focus on clinically relevant regions while suppressing irrelevant background information, particularly in low-contrast images and small target structures.[11] In the Attention U-Net++, Attention Gates (AGs) are embedded within skip connections, where encoder feature maps are modulated by gating signals from deeper decoder layers to generate spatial attention coefficients prior to feature fusion. In this study, a modified Attention U-Net++ (AU-Net++) architecture was implemented using TensorFlow. The network accepts two-dimensional CT slices with an input size of 256×256×1 and follows an encoder–decoder structure with nested dense skip pathways. The encoder consists of four downsampling stages and a bottleneck layer, with the number of convolutional filters progressively increasing through 64, 128, 256, 512, and 1,024. Each stage contains two 3×3 convolutional layers with rectified linear unit (ReLU) activation, followed by 2×2 max pooling for spatial downsampling.

The decoder incorporates multi-level nested skip connections, with attention gates applied across multiple levels to refine encoder feature maps before concatenation. Each attention gate employs 1×1 convolutions with ReLU activation, followed by sigmoid activation to generate attention maps. These attention maps are multiplied elementwise with the corresponding encoder features to enhance relevant spatial regions. Finally, the output layer applies a 1×1 convolution with sigmoid activation to generate a binary segmentation probability map (Figure 2).

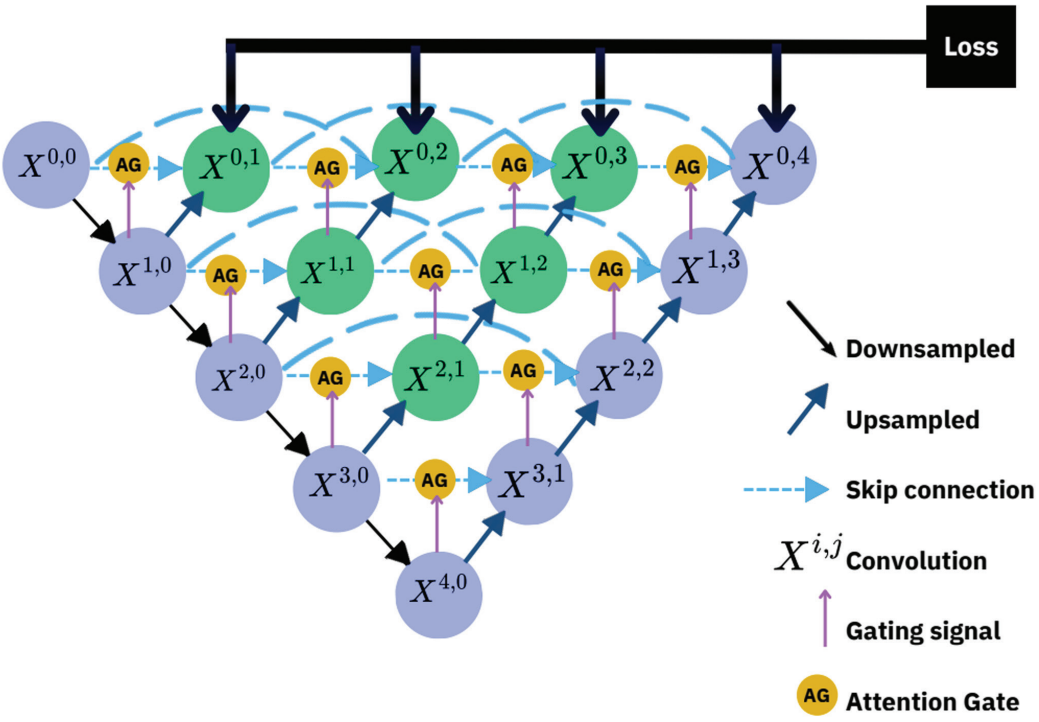


Figure 2. *Architecture of the Attention Gate U-Net++.*

The deep supervision of the Attention Gate U-Net++ architecture

Deep supervision facilitates stable gradient propagation and improves training convergence by applying auxiliary loss functions to intermediate decoder outputs.[11] In the Attention U-Net++ architecture, segmentation outputs are generated from multiple decoder levels (L^1, L^2, and L^3), as illustrated in Figure 3. Each intermediate output is passed through a 1×1 convolution layer followed by a sigmoid activation to produce a probability map. During training, the loss function is computed at each supervised level, and the final loss is defined as the weighted sum of all intermediate losses. This strategy encourages feature refinement at multiple semantic scales and mitigates the vanishing gradient problem in deeper architectures. During inference, only the final decoder output is used to generate the segmentation result, while intermediate outputs are utilized solely to guide the training process.

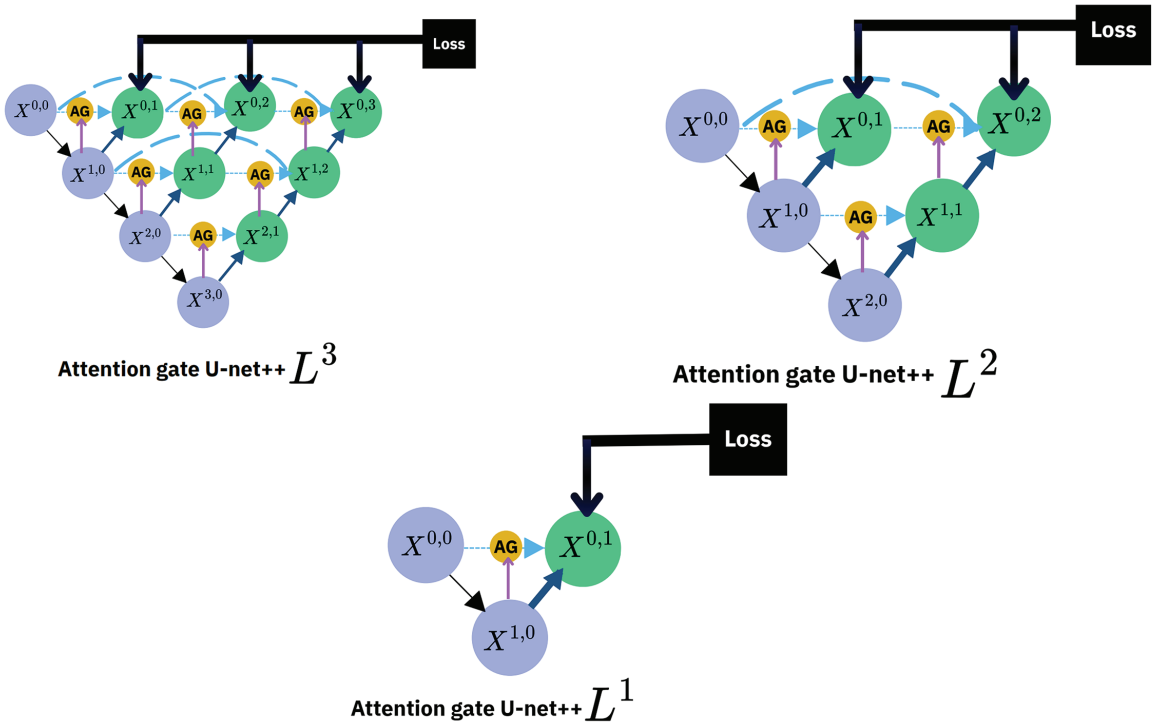


Figure 3. *Deep supervision mechanism in the Attention Gate U-Net++.*

The Attention Gate calculation process

The Attention Gate mechanism operates by jointly processing the upsampled decoder features and the corresponding encoder features from the skip connection.[8-16] Both feature maps are first projected through separate 1×1 convolution layers to reduce and align their channel dimensions. The transformed features are then combined using element-wise summation and passed through a rectified linear unit (ReLU) activation function to introduce non-linearity. Subsequently, a 1×1 convolution followed by a sigmoid activation function is applied to generate the attention coefficient (α), which ranges between 0 and 1. The resulting attention map is upsampled to match the spatial resolution of the encoder feature map and is then multiplied elementwise with the encoder features. This gating operation suppresses irrelevant background responses while enhancing salient features, thereby improving feature representation for subsequent decoding stages (Figure 4).

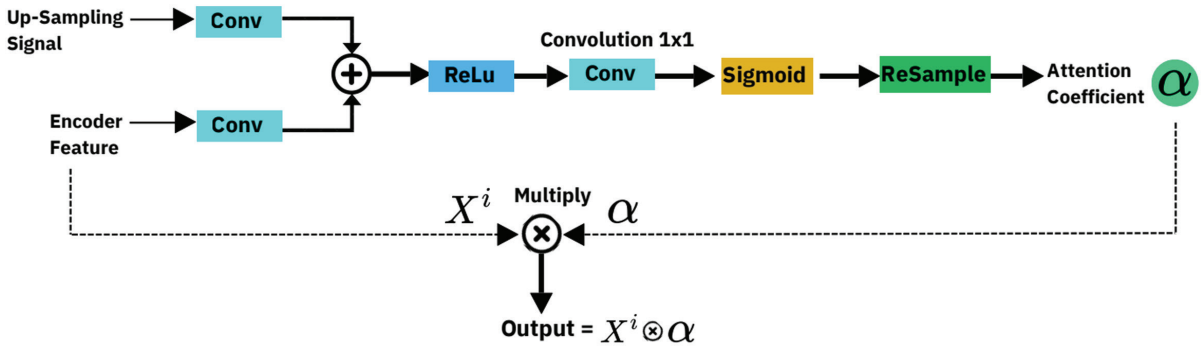


Figure 4. *The Attention Gate calculation process.*

The proposed network architecture

To reduce the computational burden while preserving segmentation performance, a modified Attention U-Net++ (AU-Net++) architecture was developed by progressively reducing decoder depth.[17] As illustrated in Figure 5, three configurations were evaluated: AU-Net++ DL3 (removal of decoder level 4), AU-Net++ DL2 (removal of decoder level 3), and AU-Net++ DL1 (removal of decoder level 2). In deep encoder-decoder architectures, repeated downsampling during feature extraction progressively reduces spatial resolution, leading to the loss of fine-grained boundary information.[18] Consequently, deeper decoder layers may contribute limited spatial detail due to the accumulation of coarse semantic features and boundary blurring. This observation suggests that the deepest decoder levels may introduce redundant computations without substantially improving geometric precision. Based on this rationale, higher decoder levels were systematically removed to reduce model complexity and computational cost. By truncating the deepest decoding stages, the proposed architectures aim to retain sufficient semantic representation while minimizing unnecessary feature propagation.[19] This structural simplification decreases the number of trainable parameters and training time, while potentially improving geometric accuracy by mitigating boundary over-smoothing in the final segmentation output.

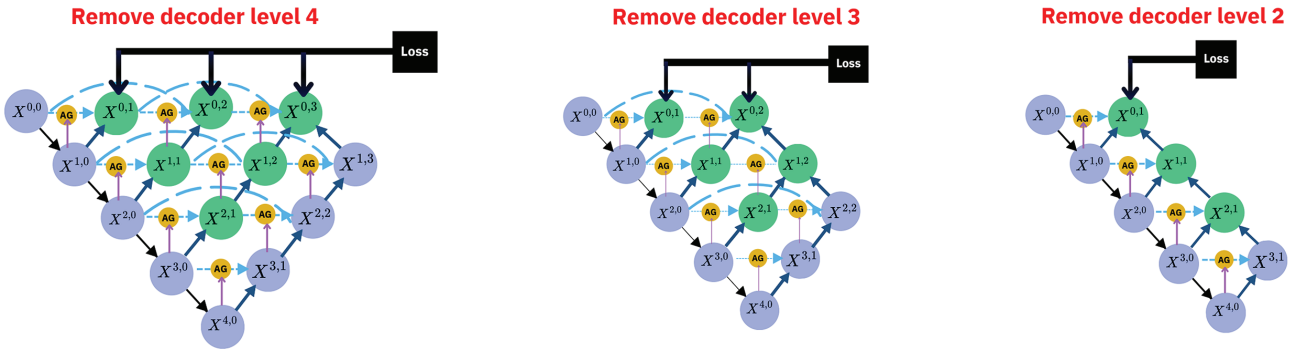


Figure 5. *The AU-Net++ architectures with progressively reduced decoder depths: DL3 (removing Level 4), DL2 (removing Level 3), and DL1 (removing Level 2) to improve computational efficiency.*

Model training

Model development and preliminary experiments were conducted on a laptop (MSI Katana; Intel® Core™ i7 processor; NVIDIA GeForce RTX 4060 GPU). For large-scale training and final model optimization, experiments were implemented in Python using the Keras API integrated within the TensorFlow framework and executed on the Google Colab platform utilizing an NVIDIA A100 GPU runtime. To determine the optimal training configuration, various hyperparameter settings were evaluated. The network was trained using three epoch settings (60, 80, and 100 epochs) and three batch sizes (8, 16, and 32). Dice+Binary Cross-entropy loss is defined in Equation (1):

$$L_{BCE} = -[g\log(p) + (1-g)\log(1-p)] \quad (1)$$

where denotes the predicted probability produced by the model after the sigmoid activation, and represents the corresponding binary ground-truth label. The dice similarity coefficient (DSC) was employed to directly optimize segmentation overlap performance. The Adam optimizer was applied with an initial learning rate of 0.001. Training performance was monitored using

validation loss and DSC metrics to assess convergence behavior. Experimental results indicated that training with 60 epochs and a batch size of 32 achieved stable convergence without observable signs of overfitting, as reflected by consistent validation performance. Consequently, this configuration was selected as the optimal setting, providing stable training behavior, satisfactory segmentation performance, reduced training time, and efficient GPU resource utilization. Training time and inference time were quantitatively evaluated using Python's time library. Furthermore, the total number of trainable parameters was calculated to assess model complexity, particularly in relation to the reduced decoder depth.

Statistical analysis

To assess segmentation performance, both geometric accuracy metrics and pixel-level classification metrics were evaluated. Geometric accuracy was quantified using the Dice Similarity Coefficient (DSC), Intersection over Union (IoU), and the 95th percentile Hausdorff Distance (HD95), (HD95). The DSC and IoU measure the spatial overlap between the predicted segmentation mask and the ground-truth annotation. The HD95 evaluates boundary agreement by computing the 95th percentile of the bidirectional surface distance between predicted and reference contours, thereby reducing sensitivity to extreme outlier points. These metrics are defined in Equations (2)-(4).

$$\mathrm{DSC} = \frac{2|\mathrm{A} \cap \mathrm{B}|}{|\mathrm{A}| + |\mathrm{B}|} \quad (2)$$

$$\mathrm{IoU} = \frac{|\mathrm{A} \cap \mathrm{B}|}{|\mathrm{A} \cup \mathrm{B}|} \quad (3)$$

$$\mathrm{HD}_{95}(\mathrm{A},\mathrm{B}) = \max\left\{\mathrm{P}_{95}\left(\mathrm{mind}_{\mathrm{b}\in\mathrm{B}}(\mathrm{a},\mathrm{b})\right), \mathrm{P}_{95}\left(\mathrm{mind}_{\mathrm{a}\in\mathrm{A}}(\mathrm{b},\mathrm{a})\right)\right\} \quad (4)$$

where *A* and *B* denote the sets of boundary points for the ground-truth and predicted segmentation masks, respectively, *d* represents the Euclidean distance, and indicates the 95th percentile.

The model performance was further evaluated using accuracy, precision, and recall at the pixel level. Accuracy represents the overall proportion of correctly classified pixels. Precision measures the proportion of correctly predicted positive pixels to all predicted positive pixels, while recall quantifies the proportion of true positive pixels correctly identified by the model. These metrics are formally defined in Equations (5)-(7).

$$\mathrm{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \quad (5)$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \quad (6)$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad (7)$$

Specifically, true positive (TP) represents pixels correctly classified as the target region, true negative (TN) denotes pixels correctly classified as background, false positive (FP) signifies background pixels incorrectly classified as target, and false negative (FN) refers to target pixels incorrectly classified as background.

Results

Evaluation of model training

Comparison of the computational efficiency in model evaluation is demonstrated in Table 2. The baseline Attention U-Net++ (AU-Net++) model contains 37.30 million trainable parameters, with a training time of 52.27 minutes and an inference time of 36.97 seconds per test set. Following decoder depth reduction, the AU-Net++ DL3 reduced the number of parameters to 8.86 million and the training time to 33.11 minutes. The AU-Net++ DL2 further decreased the parameter count to 2.10 million, with a corresponding training time of 19.41 minutes. The most compact configuration, AU-Net++ DL1, achieved the smallest model size with only 0.41 million parameters and a training time of 9.26 minutes. A consistent reduction in inference time was also observed across models, with AU-Net++ DL1 demonstrating the fastest performance at 33.29 seconds per test set. These findings indicate that progressive decoder reduction substantially improves computational efficiency while maintaining practical inference speed.

Table 2. Comparison of computational efficiency in model evaluation.

Model	Parameters	Training time (minutes)	Inference time (seconds/test set)
Attention U-net++	37.30 M	52.27	36.9660
AU-Net++ DL3	8.86 M	33.11	36.8953
AU-Net++ DL2	2.10 M	19.41	35.2526
AU-Net++ DL1	0.41 M	09.26	33.2896

Note: *M: millions.*

Evaluation of model stability

Figure 6 illustrates the training and validation Dice loss curves for A: Attention Gate U-Net++, B: AU-Net++ DL3, C: AU-Net++ DL2, and D: AU-Net++ DL1. All models exhibit a rapid reduction in loss during the initial epochs, followed by stable convergence. The training and validation curves remain closely aligned across all configurations, indicating stable learning behavior and the absence of significant overfitting.

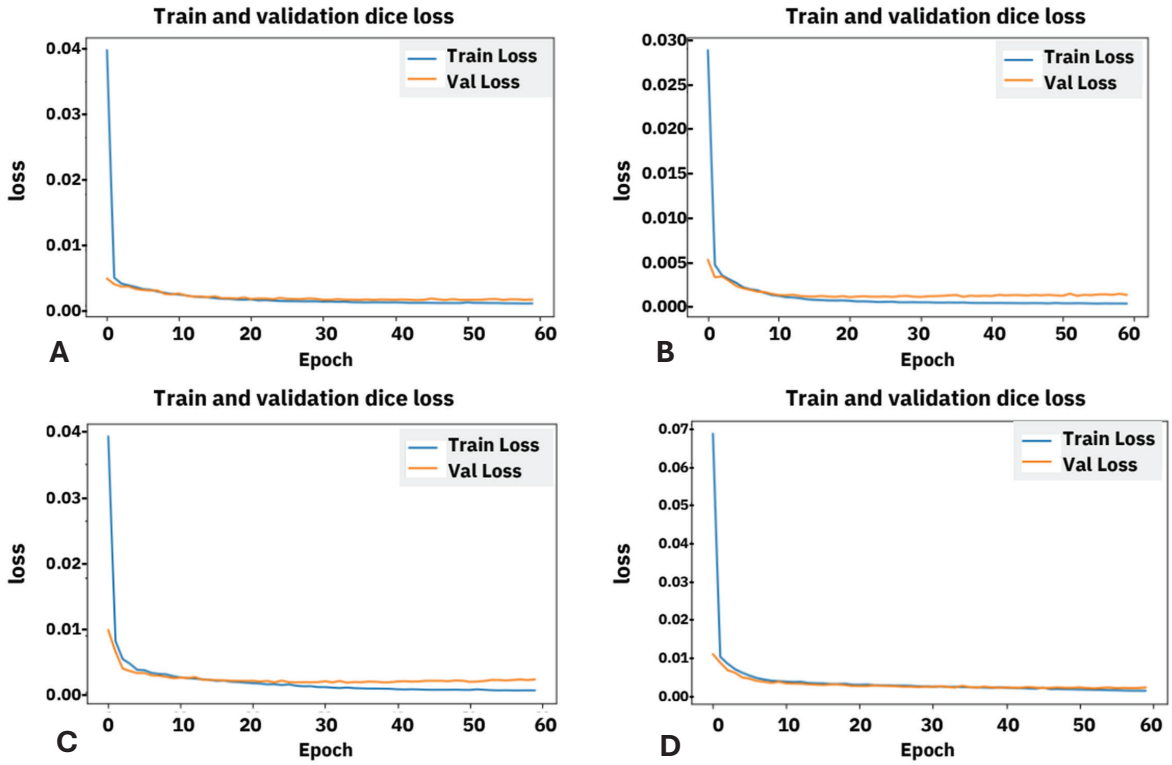


Figure 6. *Training and validation Binary Cross-entropy Dice loss curves. A: Attention Gate U-Net++ model, B: AU-Net++ DL3, C: AU-Net++ DL2, D: AU-Net++ DL1.*

Geometric and performance model evaluation

Table 3 demonstrates that the AU-Net++ DL3 model achieved the best overlap performance, with the highest DSC of 0.9076 and IoU of 0.8376, slightly outperforming Attention U-Net++ (DSC: 0.9057, IoU: 0.8338). Attention U-Net++ obtained the lowest HD95 (1.3279), while DL3 exhibited comparable boundary performance (1.3959). Both models achieved identical accuracy (0.9996), with DL3 providing higher precision (0.9046) and Attention U-Net++ yielding higher recall (0.9382). Performance declined for DL2 and DL1, particularly regarding DSC, IoU, and HD95, although DL1 achieved the highest precision (0.9162). Overall, DL3 offers the optimal balance between accuracy and model simplification.

Table 3. Geometric and performance evaluation.

Model	DSC	IoU	HD95	Accuracy	Precision	Recall
Attention U-net++	0.9057	0.8338	1.3279	0.9996	0.8920	0.9382
AU-Net++ DL3	0.9076	0.8376	1.3959	0.9996	0.9046	0.9298
AU-Net++ DL2	0.8678	0.7783	1.6236	0.9995	0.8729	0.8928
AU-Net++ DL1	0.7514	0.6149	3.3098	0.9991	0.9162	0.6721

The results of boundary overlap analysis

Figure 7 presents the visual comparison between the ground-truth contour (solid red line) and the predicted contour (solid blue line) for Attention Gate U-Net++, AU-Net++ DL3, AU-Net++ DL2, and AU-Net++ DL1. The Attention Gate U-Net++ model demonstrates overall strong boundary agreement, although slight deviations are observed in some peripheral regions. AU-Net++ DL3 exhibits highly precise alignment between the predicted and ground-truth contours, with enhanced consistency along the tumor boundary. While AU-Net++ DL2 maintains reasonable overlap,

it displays minor boundary irregularities compared to DL3. In contrast, AU-Net++ DL1 shows significant boundary mismatches and shape distortions, indicating reduced geometric precision.

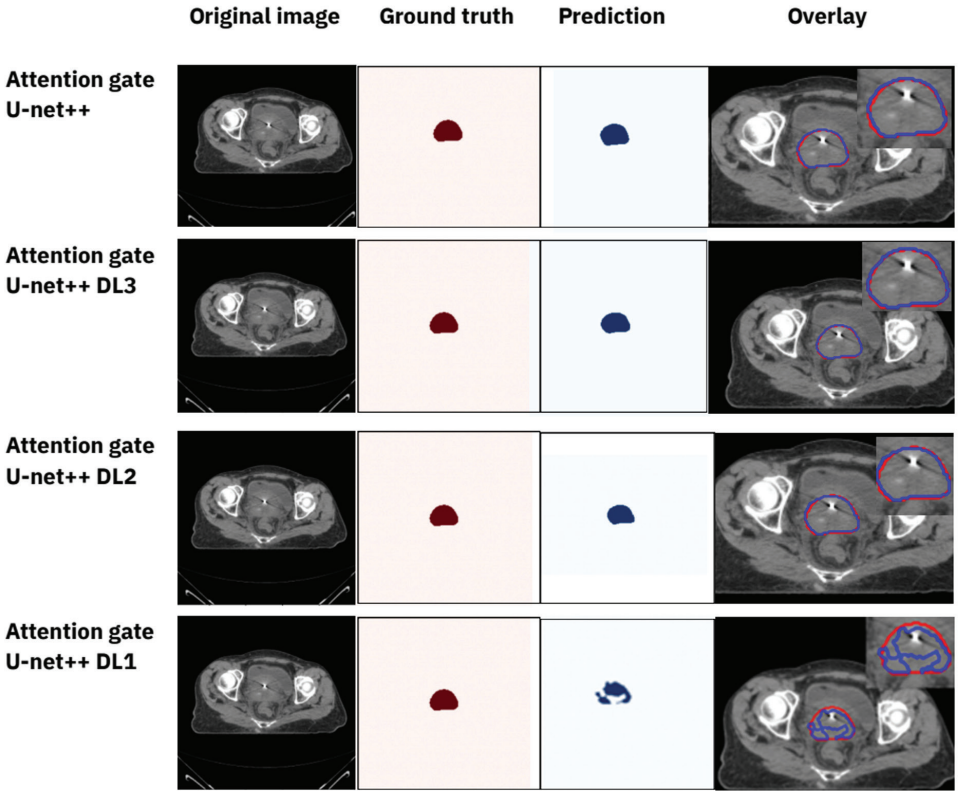


Figure 7. *The boundary overlaps between the solid red contour (ground-truth) and the solid blue contour (predicted segmentation) for the Attention Gate U-Net++, AU-Net++ DL3, AU-Net++ DL2, and AU-Net++ DL1.*

Discussion

This study proposes a modified Attention Gate U-Net++ architecture designed to enhance computational efficiency while maintaining segmentation performance for HR-CTV delineation in CT-based cervical cancer brachytherapy. The architectural modification entailed reducing the decoder depth (DL3) to minimize redundant feature propagation from the deepest decoder level. Tables 4 and 5 compare the proposed model with previously reported architectures, including U-Net,[7] Attention U-Net[8], Residual U-Net,[20] U-Net++,[10] Residual-Attention U-Net++,[16] and Attention U-Net++.[11] Future work should compare the proposed model with recent foundation models, such as MedSAM and other non-U-Net architectures, to further validate its generalizability and clinical applicability.

In terms of computational efficiency (Table 4), AU-Net++ DL3 reduced the number of trainable parameters from 37.30 M in the baseline Attention U-Net++ to 8.86 M, demonstrating approximately a 76% reduction. Training time decreased from 52.27 minutes to 33.11 minutes, while inference time remained comparable (36.89 and 36.97 seconds/test set). These findings indicate that decoder depth predominantly affects backpropagation complexity during training rather than forward-pass inference speed. Although DL2 (2.10 M) and DL1 (0.41 M) further reduced model size, segmentation performance deteriorated significantly, particularly in HD95, suggesting that excessive decoder truncation compromises multi-scale feature reconstruction. Consequently, DL3 represents an optimal structural balance between model compactness and the preservation of spatial detail.

Regarding segmentation performance (Table 5), AU-Net++ DL3 achieved a DSC of 0.9076 and an IoU of 0.8376, which are comparable to Attention U-Net++ (DSC 0.9057) and conventional U-Net (0.9030). Although U-Net++ achieved the highest DSC (0.9114), the performance gain was marginal (<0.004) despite a substantially higher parameter count (36.61 M), indicating diminishing returns with increasing model complexity. Boundary accuracy (HD95=1.3959) remained close to Attention U-Net++ (1.3014) and U-Net (1.2877), suggesting that removing the deepest decoder layer did not significantly compromise contour precision. Pixel-level metrics (accuracy 0.9996, precision 0.9046, recall 0.9298) further confirm stable classification performance.

All experiments were conducted using an NVIDIA A100 GPU runtime on Google Colab, allowing all models to be trained within one hour. However, training the baseline Attention U-Net++ on a local workstation (MSI Katana, Intel Core i7 with NVIDIA RTX 4060) required approximately six hours under identical settings. This highlights the substantial influence of

hardware configuration on deep learning training time and reinforces the practical importance of architectural simplification in resource-limited environments.

Hyperparameter optimization determined that 60 epochs with a batch size of 32 provided stable convergence without overfitting. Increasing the number of epochs extended training time without meaningful improvement in DSC, supporting the study's objective of reducing computational cost while maintaining geometric accuracy.

Overall, the deepest decoder level in Attention U-Net++ contributes limited incremental performance relative to its computational burden. The proposed AU-Net++ DL3 achieves a favorable trade-off between segmentation accuracy and efficiency, supporting its feasibility for clinical implementation, particularly in settings with constrained computational resources.

Table 4. Comparison of execution time and parameter results with previous work.

Author/year	Model	Parameters	Training time (minutes)	Inference time (seconds/test set)
Ronneberger O/ 2015[7]	U-net	31.37 M	16.48	36.1248
Oktay O/2018[8]	Attention U-net	31.90 M	19.43	35.6353
Zhang Z/2018[20]	Residual U-net	33.15 M	22.80	35.7755
Zhou Z/2018[17]	U-net++	36.61 M	41.51	36.1453
Li Z/2022[16]	Residual-Attention U-Net++	39.26 M	60.90	38.1939
Yu Y/2023[11]	Attention U-net++	37.30 M	52.27	36.9660
This study	Attention U-net++ DL3	8.86 M	33.11	36.8953
	Attention U-net++ DL2	2.10 M	19.41	35.2526
	Attention U-net++ DL1	0.41 M	09.26	33.2896

Note: *M: millions.*

Table 5. Comparison of geometric and performance results with previous work.

Author/Year	Model	DSC	IoU	HD95	Accuracy	Precision	Recall
Ronneberger O/ 2015[7]	U-net	0.9030	0.8296	1.2877	0.9996	0.9224	0.9041
Oktay O/2018[8]	Attention U-net	0.9095	0.8394	1.2481	0.9997	0.9298	0.9061
Zhang Z/2018[20]	Residual U-net	0.8967	0.8190	1.3192	0.9996	0.9203	0.8932
Zhou Z/2018[17]	U-net++	0.9114	0.8424	1.4610	0.9997	0.9077	0.9313
Li Z/2022[16]	Residual-Attention U-Net++	0.9024	0.8272	2.6574	0.9996	0.9009	0.9214
Yu Y/2023[11]	Attention U-net++	0.9057	0.8338	1.3014	0.9996	0.8920	0.9382
This study	Attention U-net++ DL3	0.9076	0.8376	1.3959	0.9996	0.9046	0.9298
	Attention U-net++ DL2	0.8678	0.7783	1.6236	0.9995	0.8729	0.8928
	Attention U-net++ DL1	0.7514	0.6149	3.3098	0.9991	0.9162	0.6721

Limitations

This study has several limitations that should be acknowledged. Firstly, the dataset consisted of CT images from only 56 patients collected at a single institution. Although each patient contributed multiple CT simulation sessions during brachytherapy, the relatively small sample size and single-center design may limit the generalizability of the proposed model to other institutions with different patient populations, imaging protocols, scanner types, and clinical workflows. Therefore, external validation using larger multi-center datasets is necessary to confirm the robustness and clinical applicability of the proposed approach.

Secondly, all ground-truth segmentations

were delineated by a single experienced radiation oncologist. Although this ensured contouring consistency throughout the dataset, it did not account for inter-observer variability, which is an important consideration in HR-CTV delineation for cervical cancer brachytherapy. Future studies involving multiple radiation oncologists would enable evaluation of the model against inter-observer variation and improve its clinical reliability.

Thirdly, this study utilized only CT images for automatic HR-CTV segmentation. Although MRI examinations were performed before brachytherapy planning according to clinical guidelines, MRI-guided segmentation was not incorporated for every CT case. Consequently, the proposed model may have reduced soft-tissue discrimination compared with multimodal CT–MRI approaches. Nevertheless, CT-based brachytherapy guided by pre-treatment MRI remains a practical and clinically accepted strategy,[5] particularly in institutions where MRI is not available for every treatment fraction.[21] Future studies integrating multimodal CT–MRI information may further improve segmentation accuracy and anatomical boundary delineation.

Fourthly, the proposed model was developed using a two-dimensional (2D) slice-based segmentation approach. Although 2D segmentation substantially reduces computational complexity and memory requirements, it does not explicitly exploit three-dimensional spatial continuity between adjacent CT slices. As a result, volumetric consistency across slices may be compromised, and no post-processing technique was applied to enforce 3D contour continuity. Future work should investigate three-dimensional or hybrid 2D-3D architectures to better capture volumetric anatomical information while maintaining computational efficiency.

Fifthly, all CT images were acquired with a uniform slice thickness of 2 mm, which provided consistent image quality during model training but may limit generalizability to datasets acquired using different imaging resolutions. A thinner slice thickness (e.g., 1.0-1.3 mm) could improve spatial resolution and enhance geometric accuracy, particularly for boundary-sensitive metrics such as HD95. Therefore, further validation using heterogeneous imaging parameters is warranted.[22]

Lastly, although the proposed decoder-reduced Attention U-Net++ substantially decreased the number of trainable parameters and reduced training time, all experiments were conducted on a high-performance NVIDIA A100 GPU. Consequently, the computational benefits of the proposed architecture have not yet been validated on lower-resource hardware commonly available in routine clinical practice. Future studies should evaluate the model on more accessible GPU platforms and perform additional external benchmarking to further demonstrate its practicality in resource-limited clinical environments.

Conclusion

This study presents a computationally efficient Attention Gate U-Net++ model for the automatic segmentation of high-risk clinical target volumes (HR-CTV) in CT-based cervical cancer brachytherapy. By reducing the decoder depth, the proposed model substantially decreases the number of trainable parameters, thereby reducing memory footprint and training cost while maintaining acceptable geometric accuracy. Despite its lightweight architecture, the proposed approach achieves competitive performance in terms of DSC, IoU, and HD95 compared with more complex models. The reduced computational demand also increases the feasibility of deploying the model on mid-range GPU platforms commonly available in clinical practice, such as the NVIDIA RTX 3060, NVIDIA RTX 4060, and NVIDIA T4, making it a practical alternative for institutions with limited computational resources. Nevertheless, additional benchmarking on these lower-resource hardware platforms is warranted to further validate its real-world computational efficiency and clinical applicability.

Ethical approval

The study was approved by the Ethics Committee of Naresuan University, Thailand (IRB No. P1-0014/2569).

Funding

None

Conflict of interest

None

CRediT authorship contribution statement

Thinnagit Srikhot: data curation, investigation, resources, writing: review and edit; **Chatrawut Pattaweerakul:** clinical validation, investigation, resources, data curation (manual contouring of HR-CTV and OAR on CT images); **Kamonchanok Nobphuek:** resources, data curation (CT image retrieval and clinical data management), validation, writing: review and edit; **Titipong Kaewlek:** 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 Assistant Professor Dr. Titipong Kaewlek for his invaluable guidance and continued support throughout this research. The authors gratefully acknowledge the medical physicist Mrs. Kamonchanok Nobphuek and radiation oncologist Dr. Chatrawut Pattaweerakul from the Department of Radiation Therapy, Naresuan University Hospital, for their kind cooperation, assistance, and facilitation in data collection.

Reference

[1] Priyadarshini S, Swain PK, Agarwal K, Jena D, Padhee S. Trends in gynecological cancer incidence, mortality, and survival among elderly women: a SEER study. Aging Med (Milton). 2024; 7(2): 179-88. doi: 10.1002/agm2.12297.

[2] Burd EM. Human papillomavirus and cervical cancer. Clin Microbiol Rev. 2003; 16(1): 1-17. doi: 10.1128/CMR.16.1.1-17.2003.

[3] Liu T, Wen S, Wang S, Yang Q, Wang X. Artificial intelligence in brachytherapy. J Radiat Res Appl Sci. 2024; 17(2): 100925. doi: 10.1016/j.jrras.2024.100925.

[4] Dimopoulos JCA, Petrow P, Tanderup K, Petric P, Berger D, Kirisits C, et al. Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group (IV): basic principles and parameters for MR imaging within the frame of image based adaptive cervix cancer brachytherapy. Radiother Oncol. 2012; 103(1): 113-22. doi: 10.1016/j.radonc.2011.12.024.

[5] Mahantshetty U, Poetter R, Beriwal S, Grover S, Lavanya G, Rai B, et al. IBS-GEC ESTRO-ABS recommendations for CT based contouring in image guided adaptive brachytherapy for cervical cancer. Radiother Oncol. 2021; 160: 273-84. doi: 10.1016/j.radonc.2021.05.010.

[6] van Vliet-Pérez SM, van Paassen R, Wauben LSGL, Straathof R, van den Berg NJ, Dankelman J, et al. Time-action and patient experience analyses of locally advanced cervical cancer brachytherapy. Brachytherapy. 2024; 23(3): 274-81. doi: 10.1016/j.brachy.2024.01.007.

[7] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer; 2015: pp 234-41. doi: 10.48550/arXiv.1505.04597.

[8] Oktay O, Schlemper J, Folgoc LL, Lee MJ, Heinrich MP, Misawa K, et al. Attention U-Net: learning where to look for the pancreas. arXiv. 2018; 1804.03999. doi: 10.48550/arXiv.1804.03999.

[9] Wang Y, Jian W, Zhu L, Cai C, Zhang B, Wang X. Attention-gated deep-learning-based automatic digitization of interstitial needles in high-dose-rate brachytherapy for cervical cancer. Adv Radiat Oncol. 2024; 9(1): 100941. doi: 10.1016/j.adro.2023.101340.

[10] Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. UNet++: a nested U-Net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer; 2018: pp 3-11. doi: 10.1007/978-3-030-00889-5_1.

[11] Yu Y, Chen S, Wei H. Modified UNet with attention gate and dense skip connection for flow field information prediction with porous media. Flow Meas Instrum. 2023; 89: 102300. doi: 10.1016/j.flowmeasinst.2022.102300.

[12] Zhong X, Ling W, Pan K, Liu C, Wu P, Zhang J, et al. A neural network architecture based on attention gate mechanism for 3D magnetotelluric forward modeling. Comput Geosci. 2025; 207: 106079. doi: 10.48550/arXiv.2503.11408.

[13] Xie Z, Li X, Ma H, Wu S, Cui D. PUNet: a lightweight parallel U-Net architecture integrating Mamba-CNN for high-precision image segmentation. Sci Rep. 2025; 15: 38954. doi: 10.1038/s41598-025-22862-x.

[14] Rajaramana S, Yang F, Zamzmi G, Xue Z, Antani S. Assessing the impact of image resolution on deep learning for TB lesion segmentation on frontal chest X-rays. Diagnostics (Basel). 2023; 13(4): 747. doi: 10.3390/diagnostics13040747.

[15] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data. 2019; 6: 60. doi: 10.1186/s40537-019-0197-0.

[16] Li Z, Zhang H, Li Z, Ren Z. Residual-attention UNet++: a nested residual-attention U-Net for medical image segmentation. Appl Sci. 2022; 12(14): 7149. doi: 10.3390/app12147149.

[17] Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: redesigning skip connections to exploit multiscale features in image segmentation. IEEE Trans Med Imaging. 2019; 39(6): 1856-67. doi: 10.1109/TMI.2019.2959609.

[18] Al Qurri A, Almekkawy M. Improved UNet with attention for medical image segmentation. Sensors. 2023; 23(20): 8589. doi: 10.3390/s23208589.

[19] Zhang L, Li M, Zhang P, Liu P. EfficientSegNet: lightweight semantic segmentation with multi-scale feature fusion and boundary enhancement. Sensors. 2025; 25: 5934. doi: 10.3390/s25195934.

[20] Zhang Z, Liu Q, Wang Y. Road extraction by deep residual U-Net. IEEE Geosci Remote Sens Lett. 2018; 15(5): 749-53. doi: 10.1109/LGRS.2018.2802944.

[21] Dankulchai P, Prasartseree T, Sittiwong W, Chansilpa Y, Apiwarodom N, Petsuksiri J, et al. Clinical outcome comparison between CT-guided versus all MRI-guided scenarios in brachytherapy for cervical cancer: a single-institute experience. Clin Oncol (R Coll Radiol). 2024; 36(11): e456-67. doi: 10.1016/j.clon.2024.08.003.

[22] Qing K, Yue NJ, Hathout L, Ma C, Reyhan M, Zhu J, et al. The combined use of 2D scout and 3D axial CT images to accurately determine the catheter tips for high-dose-rate brachytherapy plans. J Appl Clin Med Phys. 2021; 22(3): 273-8. doi: 10.1002/acm2.13184.

Journal of Associated Medical Sciences 2026; 59(3): 325-333

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 a tele-pulmonary rehabilitation program in individuals at risk of chronic obstructive pulmonary disease from artificial intelligence predictions

Busaba Chuatrakoon[1*], Siriphun Kongsawasdi[1], Kittichai Wantanajittikul[2], Sompong Sriburee[2], Supatcha Konghakote[1]

[1]*Integrated Neuro-Musculoskeletal, Chronic Disease, and Aging Research Engagement Center (ICARE Center), Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Thailand.*
[2]*Department of Radiology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand.*

ARTICLE INFO

Article history:
Received 6 May 2025
Accepted as revised 21 July 2026
Available online 31 July 2026

Keywords:
COPD, tele-rehabilitation, exercise capacity, pulmonary rehabilitation, exercise

ABSTRACT

Background: Chronic Obstructive Pulmonary Disease (COPD) is a major global health burden, and early intervention in at-risk individuals may improve outcomes. Emerging technologies such as artificial intelligence (AI) offer novel means of identifying individuals at risk of COPD, enabling targeted preventive strategies. Tele-pulmonary rehabilitation (Tele-PR) has shown promise in managing chronic respiratory conditions, but its effectiveness in AI-identified at-risk individuals remains underexplored.

Objectives: To evaluate the effectiveness of a 12-week Tele-PR program in improving disease-related symptoms, muscular strength, and exercise capacity in individuals identified as at risk of COPD by AI prediction.

Materials and methods: Out of 286 screened participants, 38 were identified by AI as being at risk of COPD and enrolled in the study. They were allocated by village-level cluster assignment (quasi-experimental design) to either an intervention group (N=20) or a control group (N=18). The intervention group participated in a Tele-PR program conducted three sessions per week for 12 weeks. The control group received only verbal information on COPD and preventive self-management. Outcome measures included the modified Medical Research Council (mMRC) dyspnea scale, COPD Assessment Test (CAT), back and leg muscle strength using a dynamometer, and six-minute walk distance (6MWD). Assessments were conducted at baseline and post-intervention.

Results: All participants completed the study without adverse events. Following the intervention, the Tele-PR group showed a statistically significant improvement in 6MWD compared to the control group (p=0.011). However, no significant differences were observed between groups in mMRC scores, CAT scores, or muscle strength measurements.

Conclusion: A 12-week Tele-PR program significantly enhances exercise capacity, as measured by 6MWD, in individuals identified by AI as being at risk for COPD. These findings support the use of Tele-PR as an early intervention strategy in high-risk populations.

* **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, Thailand.*
E-mail address: busaba.c@cmu.ac.th
doi: 10.66285/JAMS.2026.105
E-ISSN: 2539-6056

Introduction

Chronic obstructive pulmonary disease (COPD) is a significant public health problem and the fourth leading cause of death worldwide.[1] It is characterized by persistent respiratory symptoms and airflow limitations that are progressive and not fully reversible.[2] Not only impaired respiratory function, but also muscle dysfunction, balance impairment, decreased exercise capacity, and

poor quality of life (QoL) are well recognized in individuals with COPD.[3,4] Diagnosis of COPD is based on respiratory symptoms and confirmation of airflow limitation using spirometry. An assessment to diagnose COPD involves consideration of personal history and risk factors, along with additional investigations such as chest radiographs, CT scans, and arterial blood gas analyses.[2]

During the COVID-19 pandemic, spirometry was often restricted due to infection control concerns. Consequently, alternative screening approaches have gained attention. Recently, intelligent systems applying machine learning techniques to chest radiographs have been developed to estimate the likelihood of obstructive pulmonary disease. Such systems may facilitate early identification of individuals at increased risk of COPD in community settings, particularly where access to spirometry is limited.

COPD imposes a substantial burden on patients, families, and healthcare systems due to its chronic and progressive nature. Most previous studies have focused on rehabilitation strategies in patients with clinically established COPD.[5-8] In contrast, little attention has been given to preventive interventions targeting individuals identified as being at increased predicted risk before clinically overt airflow limitation develops.

Pulmonary rehabilitation (PR) is strongly recommended by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) as an essential component of COPD management.[2] Evidence demonstrates that PR improves exercise capacity, physical fitness, and health-related quality of life.[6,7] However, access to center-based PR programs may be limited. Tele-pulmonary rehabilitation has emerged as an alternative approach and has been shown to provide benefits comparable to hospital-based programs in patients with COPD.[8]

Unlike previous studies evaluating telerehabilitation in confirmed COPD populations, the present study integrates an AI-assisted radiographic screening system with a preventive tele-pulmonary rehabilitation intervention in community-based adults identified as being at increased predicted risk of COPD. This study centers on two key components: (1) AI-assisted risk stratification to identify individuals prior to clinically established disease, and (2) implementation of a structured tele-pulmonary rehabilitation program in this at-risk population. By combining early risk identification with remote functional training, this approach shifts the focus from disease management to early functional preservation.

Therefore, this study aimed to examine the effects of a 12-week tele-pulmonary rehabilitation program (Tele-PR) on disease-related symptoms, muscle strength, and exercise capacity in individuals at increased predicted risk of COPD.

Materials and methods

Participants

A total of 286 community-based adults who had previously undergone chest radiography as part of their annual health check-up were screened using the AI-based model. Among them, 38 individuals with predicted COPD risk scores ≥0.75 were identified. These individuals were subsequently assessed against the predefined inclusion and exclusion criteria. All 38 participants met the eligibility criteria and were enrolled in the study. The inclusion criteria were a minimum age of 40, the ability to comprehend the training program, and the capacity to communicate well with telemonitoring. Exclusion criteria were individuals with a history of home mechanical ventilator use, neurological and musculoskeletal conditions affecting postural control, symptomatic cardiovascular diseases, uncorrected visual problems, and non-smartphone users. The study protocol was approved by the Human Ethical Review Board of the primary investigator's institution (COAAMSEC-65EX-019). All eligible participants signed an informed consent form prior to enrollment in the study.

Study design

This study was a community-based quasi-experimental controlled study with cluster allocation at the village level. Four villages were included; two villages were assigned to the Tele-PR intervention group and two villages to the control group to minimize contamination between participants. No individual randomization was performed.

Sample size calculation

The sample size calculation in this study was based on a previous study that investigated the effect of tele-pulmonary rehabilitation in patients with COPD on their exercise capacity as measured by a 6-minute walk test (6MWT).[9] The mean and standard deviation of 6MWT between the intervention and control groups were 420.2±74.9 and 339.9±110.1, respectively. This yielded an effect size of 0.85. With 80% power and a 5% type-I error, a minimum sample size of 36 participants was required to detect a between-group difference in 6MWD. Considering a possible dropout rate of 10%, 40 participants (20 per group) were enrolled in the study.

Instrument

AI-assisted screening tool for individuals at risk of COPD

A deep learning-based screening model was developed using a transfer learning approach with a pretrained ResNet-50 convolutional neural network architecture.[10] The model was implemented in MATLAB (Version 2023b, The MathWorks, Inc., Natick, MA, USA).

For model development, 540 labeled digital chest radiographs from normal individuals and patients across different COPD severity stages were used for supervised training, with an additional 60 radiographs used for internal validation. These datasets were independent of the participants included in the present intervention study. The trained model generated a probability score representing the likelihood of radiographic patterns

associated with airflow limitation. Feature extraction was performed automatically through convolutional layers without predefined radiographic variable selection.

In the present study, a separate cohort of 286 community-based adults who had undergone annual chest radiography was screened using the trained model. From this independent screening cohort, 38 individuals with probability scores≥0.75 were enrolled. The probability score was applied as a risk stratification tool to identify individuals with relatively high predicted risk. A relatively high probability threshold (0.75) was selected to reduce potential false-positive classification and to focus on individuals with more pronounced radiographic patterns suggestive of increased risk. This threshold was pragmatically defined for inclusion in this exploratory preventive study and does not represent a clinically validated diagnostic cut-off. External validation in independent populations is warranted to further establish the model's generalizability (Figure 1).

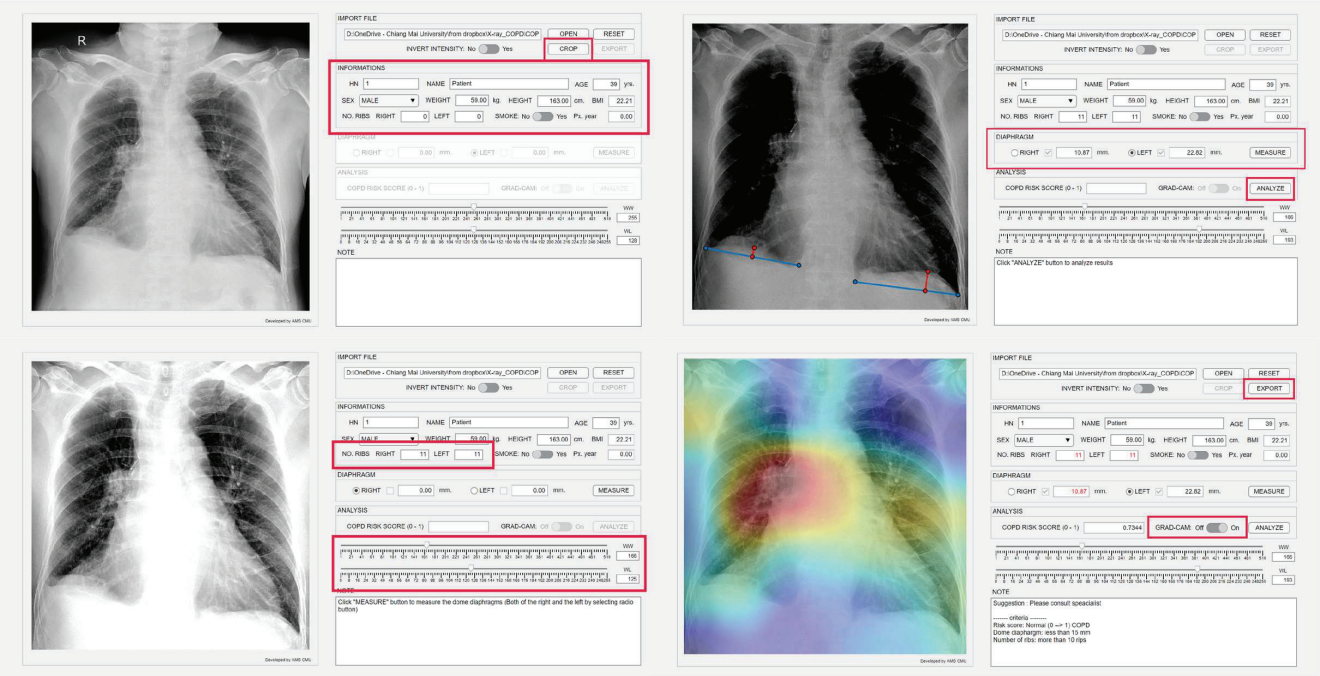


Figure 1. *A deep learning-based screening system for the risk of COPD.*

Tele-pulmonary rehabilitation (Tele-PR) program

In this study, a home-based Tele-PR program was delivered via the Line application on a smartphone platform. Participants in the intervention group performed the program three times per week for 12 weeks. Each session lasted approximately 45-60 minutes and consisted of 1) breathing and chest mobilization exercises (approximately 10 minutes) to enhance ventilatory mechanics, 2) stretching exercises (approximately 10 minutes) to promote flexibility, 3) resistance exercises for upper and lower limbs using elastic bands (15-20 minutes) to improve muscle strength, and 4) aerobic walking training (15-20 minutes) to enhance cardiopulmonary endurance. Each session consisted of structured exercise components, supplemented by self-management education and adherence support delivered through instructional materials and weekly telemonitoring via the Line application.

At the initial visit, participants received face-to-face instruction on proper exercise techniques and application use. Written and video instructional materials were provided through the Line application. The program consisted of three progression levels, which were advanced every four weeks according to participant tolerance. Exercise intensity was monitored using the Borg Rating of Perceived Exertion scale, targeting levels 3-5 for strengthening exercises and 3-6 for aerobic training (Table 1).[11]

Table 1. Details of the tele-pulmonary rehabilitation program (12 weeks, 3 sessions/week).

Intervention components	**Level 1** (Weeks 1-4)	**Level 2** (Weeks 5-8)	**Level 3** (Weeks 9-12)
Breathing exercise	6 breaths/set, 5 sets/session	Same as Level 1	Same as Level 1
Stretching exercise (8 muscle groups)	10 second hold × 5 repetitions	Same as Level 1	Same as Level 1
Strengthening exercise (6 muscle groups, elastic band, Borg 3-5)	10 reps × 2 sets	10 reps × 3 sets (increased resistance)	12 reps × 3 sets (increased resistance)
Endurance exercise (walking)	Borg 3-4, 15 minutes	Borg 4-5, 20 minutes	Borg 5-6, 25 minutes
Self-management and education	Initial instruction + weekly reinforcement	Ongoing reinforcement	Ongoing reinforcement
Estimated session duration	~45-60 minutes	~45-60 minutes	~45-60 minutes

Note: *reps: repetitions.*

Resistance was progressed by selecting elastic bands with greater tension when participants were able to complete the prescribed repetitions within the target Borg RPE range (3-5) without excessive fatigue. If participants reported RPE below the target range, resistance was increased in the subsequent session; if RPE exceeded the target range, resistance was adjusted downward. Stretching exercises were performed with a 10-second hold per repetition, repeated five times per muscle group.

Weekly follow-up messages were sent to reinforce adherence, and participants were required to report exercise completion through the application. Participants were instructed to discontinue exercise and contact the research team if they experienced chest pain, dizziness, severe dyspnea, or other abnormal symptoms. No adverse events were reported during the intervention period (Figure 2).

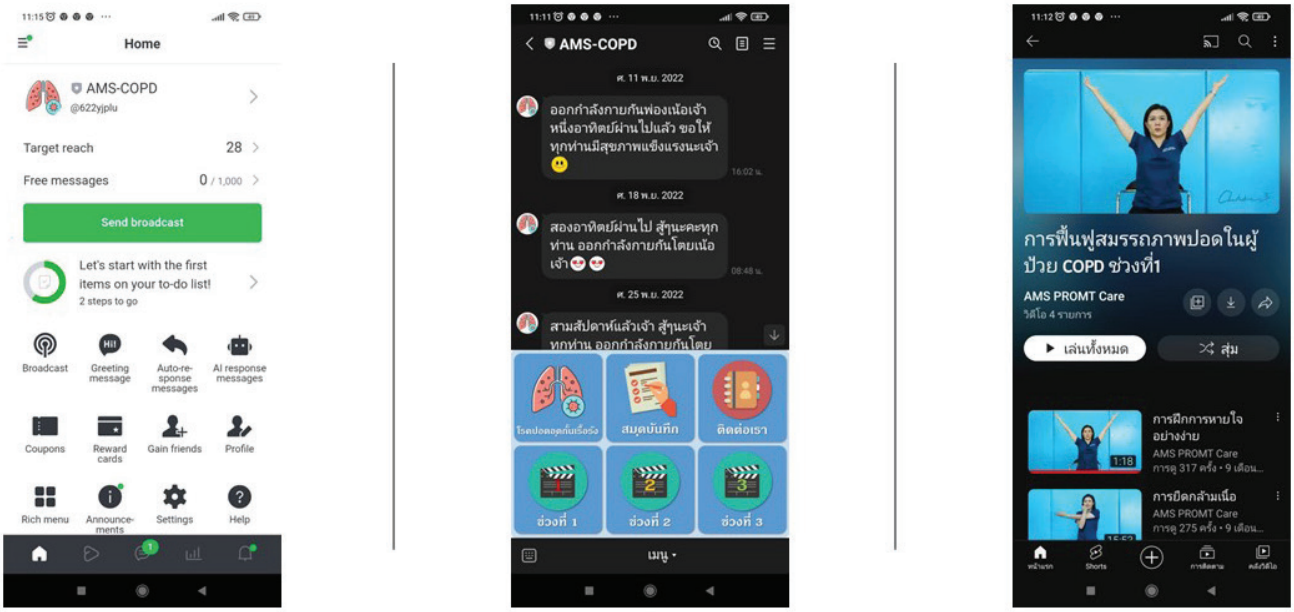


Figure 2. *The Tele-PR via Line application.*

Baseline characteristics and outcome measures

Participant characteristics, spirometry, and all outcome measures were assessed using standard, structured protocols at baseline and post-12-week intervention by trained assessors. Airflow obstruction was defined as FEV_1/FVC <0.70.

Outcome measures

The primary outcome of this study was exercise capacity measured by the Six-Minute Walk Distance (6MWD).

The Six-Minute Walk Test (6MWT) was conducted according to standard guidelines.[12] Participants walked at their own pace along a 30-meter corridor for six minutes, and the total walking distance (meters) was recorded for analysis. The 6MWD was selected as the primary outcome due to its established validity and clinical relevance in evaluating functional exercise capacity in respiratory populations.

Secondary outcomes included symptom burden, health-related quality of life, and muscle strength:

The COPD assessment test-Thai version (CAT-TH), a short, simple, valid, and reliable self-administered questionnaire, was used to determine health-related quality of life. High CAT scores indicate low QoL.[13]

The modified Medical Research Council dyspnea scale (mMRC) was used to identify disability levels related to dyspnea. It is a five-point Likert scale (0-4), with a higher score indicating greater breathlessness. The mMRC has good test-retest reliability.[14]

A back-leg dynamometer was used to measure the back and leg muscles' strength. The resisted force was recorded in kilograms.[15]

Study procedure

Participants were recruited through verbal announcements and poster advertisements within the community. After providing written informed consent, their existing chest radiographs were analyzed using the trained AI-based screening model. Individuals with AI-predicted COPD risk scores ≥ 0.75 were contacted by telephone to confirm eligibility.

Thirty-eight eligible participants were scheduled for baseline assessment at the community health center prior to group allocation. The intervention group (N=20) received the Tele-PR program delivered through the Line application with remote monitoring for 12 weeks. The control group (N=18) received general primary healthcare education delivered by community healthcare personnel. The educational content included advice on smoking cessation, physical activity promotion, respiratory hygiene, and chronic disease prevention. Education was provided in a standard informational format without individualized exercise prescription or supervised training. Participants in the control group were advised to continue their usual daily activities during the study period. All participants were assessed on outcome measures at baseline and after the 12-week intervention period.

Data analysis

Normality of data distribution was assessed using the Shapiro–Wilk test. Baseline demographic characteristics between groups were compared using independent t-tests for continuous variables and chi-square tests for categorical variables. Between-group differences in changes from baseline were analyzed using independent t-tests. Continuous variables are presented as means with 95% confidence intervals. A two-tailed significance level of $p<0.05$ was considered statistically significant. Data were analyzed using SPSS software (version 21.0, IBM Corporation, Chicago, IL, USA).

Results

Participants' characteristics

Of the 286 participants in the communities, 38 were eligible and allocated either to intervention or control groups. No differences in demographic characteristics between the two groups were observed at baseline (Table 2). Baseline pulmonary function, characterized by FEV_1/FVC ratio and % predicted FEV_1, did not differ significantly between groups. Raw FVC values were not available from the archived dataset. The term "spirometry-defined airflow obstruction" refers to participants with FEV_1/FVC <0.70 and does not indicate a clinical diagnosis of COPD. All participants completed the training program with no reported adverse events. The average exercise attendance rate was 92%.

Table 2. Demographic characteristics of the participants (N=38)

Variables	Tele-PR (N=20)	Control (N=18)	*p* value
Gender (male, %)	7 (35)	4 (22.2)	0.485
Age (year)	58.3±9.5	52.7±8.1	0.059
AI-risk score (%)	0.9±0.1	0.9±0.1	0.644
Height (m)	1.6±0.1	1.6±0.1	0.533
Weight (kg)	55.8±8.9	54.9±8.4	0.762
BMI (kg/m^2)	22.7±3.3	21.8±2.2	0.336
Smoking (N, %)	3 (15)	2 (11.1)	1.000
Occupational pollution time (hour)	2.1±2.0	1.4±1.3	0.202
FEV_1/FVC	78.8±7.0	83.5±7.7	0.054
Predicted FEV_1 (%)	88.5±10.9	90.0±16.4	0.740
Spirometry-defined airflow obstruction (FEV_1/FVC<0.70, N (%)	3 (15)	2(11.1)	1.000

Note: *AI: artificial intelligence, BMI: body mass index, FEV_1: forced expiratory volume in the first second, Tele-PR: tele-pulmonary rehabilitation program.*

Effectiveness of the Tele-PR program

Descriptive data of all outcome variables from baseline to the post-intervention period are presented in Table 3. Baseline comparisons of all outcome variables between groups revealed no statistically significant differences (all $p>0.05$). Baseline back and leg strength values appeared to be within expected ranges for community-based adults of similar age, suggesting relatively preserved musculoskeletal function. A significant improvement in 6MWD was observed in the intervention group compared to the control group after the 12-week training period ($p=0.011$; effect size =0.87).

Table 3. Outcome measures for the intervention and control groups.

Variables	Tele-PR (N=20)			Control (N=18)			95% Confidence Interval	*p* value	Effect size
	Pretest	Posttest	Change	Pretest	Posttest	Change			
CAT score (point)	4.1±4.1	2.2±4.9	-1.9±6.6	3.3±3.3	2.8±3.5	-0.6±2.5	-2.01 to 4.70	0.422	0.26
mMRC (level)	0(0.75)	0(0)	0(0)	0(0.25)	0(0)	0(0)	-0.35 to 0.49	0.747	-
Back strength (kg)	65.5±28.8	70.8±31.8	5.3±11.0	54.9±22.7	59.1±20.1	4.2±5.2	-6.91 to 4.65	0.693	0.13
Leg strength (kg)	72.8±34.3	75.6±34.0	2.9±11.1	67.8±29.0	71.4±22.8	3.6±11.9	-6.86 to 8.28	0.850	0.06
6MWD (m)	428.5±57.7	466.9±70.3	38.4±40.9	430.9±34.1	441.2±37.7	10.3±19.8	-49.16 to -7.08	0.011*	0.87

Notes: **p<0.05 difference between the change score of tele-PR and control groups. CAT: COPD assessment test, mMRC: modified medical research council dyspnea scale, Tele-PR: tele-pulmonary rehabilitation, 6MWD: 6-minute walk distance.*

Discussion

Pulmonary rehabilitation (PR) has been widely recommended as an effective strategy for individuals with COPD. In addition to center-based programs, growing evidence supports the effectiveness of home-based and tele-pulmonary rehabilitation interventions.[5,16-18] However, most previous studies have focused on patients with established COPD across varying severity stages.[5,6,16-18] The present study differs conceptually in that it targeted community-based adults identified as having an increased predicted risk of COPD using an AI-assisted radiographic screening system. Rather than focusing on established diseases, this study explored a preventive approach aimed at promoting functional capacity before substantial physiological decline occurs. Although baseline spirometry parameters and symptom scores were generally within normal ranges, indicating that most participants did not have clinically established COPD, the AI-derived risk score may capture subtle structural radiographic changes that precede clinically detectable airflow limitation. Therefore, the intervention targeted individuals with increased predicted risk rather than confirmed COPD, emphasizing early health promotion.

This study is novel in that it integrates AI-assisted radiographic risk stratification with a structured tele-pulmonary rehabilitation program in a preventive community-based setting. Unlike previous studies that focused on patients with established COPD, this approach applies artificial intelligence to identify individuals at increased predicted risk and implements early functional intervention before clinically overt disease develops. To our knowledge, this combined screening-intervention framework has not been previously reported in community-based populations.

The main finding demonstrated that the 12-week Tele-PR program significantly improved exercise capacity compared to education alone. The mean improvement in 6MWD (38.4 meters) was not only statistically significant with a large effect size but also exceeded commonly reported minimal clinically important difference (MCID) thresholds of approximately 25-35 meters in respiratory populations.[19] Although MCID values have primarily been established in patients with chronic respiratory disease, exceeding these thresholds suggests that the improvement is likely to be perceptible and functionally meaningful in daily activities, even in individuals without established COPD. In a preventive population, such enhancement in walking capacity may reflect improved physical resilience and reduced risk of future functional decline. The improvement in exercise capacity may be attributed to enhanced cardiovascular endurance, peripheral conditioning, and improved exercise adherence supported by structured telemonitoring.

Given the relatively preserved baseline lung function and low symptom burden, the observed functional gains likely reflect improved submaximal exercise efficiency rather than measurable changes in pulmonary mechanics. This may explain why significant changes were observed only in exercise capacity, whereas symptom scores and muscle strength demonstrated limited improvement. Exercise intensity during training was monitored using the Borg scale; however, RPE and SpO_2 were not systematically recorded during the 6MWT, which limits further physiological interpretation.

In contrast, disease-related symptoms, muscle strength, and health-related quality of life showed only slight improvements in both groups without significant differences. Baseline CAT and mMRC scores were low, indicating minimal symptom burden at study entry. Although these measures were included to explore potential subclinical symptom changes in individuals at increased predicted risk of COPD, the limited improvement observed is consistent with the preventive stage of this population. A potential ceiling effect may therefore explain these findings, as participants demonstrated relatively preserved baseline lung function and low symptom severity. Baseline back and leg strength values appeared to be within expected ranges for community-based adults of similar age, suggesting relatively preserved musculoskeletal function at baseline. Furthermore, the strengthening component of the Tele-PR program was adapted from protocols originally designed for symptomatic COPD populations and may not have provided sufficient overload stimulus for this relatively healthy group.

From a public health perspective, the findings suggest that tele-pulmonary rehabilitation may serve not only as a therapeutic strategy for established COPD but also as a preventive health promotion tool in individuals at increased predicted risk. Early functional enhancement may potentially contribute to long-term reductions in healthcare utilization and associated costs, although longitudinal studies are required to confirm this hypothesis. While improvements in 6MWD may be expected with regular exercise, pulmonary rehabilitation is a structured, progression-based intervention that integrates respiratory techniques, aerobic conditioning, resistance training, and adherence support. In individuals at increased predicted risk of COPD, early implementation of such a structured program may enhance functional reserve and promote long-term physical resilience before clinically significant physiological decline occurs. This preventive framework distinguishes the present study from therapeutic rehabilitation in established COPD populations. Regarding the AI-assisted screening model, it should be emphasized that the model was developed and internally validated during the training phase and was applied in this study as a risk stratification tool rather than a diagnostic instrument. Although promising as a screening support system, external validation in independent cohorts is required before broader clinical implementation.

Limitations

This study has several limitations. Firstly, the absence of a follow-up period limits conclusions regarding the long-term sustainability of the intervention

effects. Secondly, spirometry was not reassessed post-intervention, as the primary focus was functional exercise capacity rather than short-term pulmonary function change. Thirdly, baseline habitual physical activity levels were not objectively assessed, and differences in daily activity patterns between groups cannot be completely excluded. Although baseline functional capacity and muscle strength did not differ significantly, unmeasured variability in habitual activity may have influenced the magnitude of response to the intervention. Lastly, the strengthening intensity may not have been sufficient for participants with relatively preserved baseline musculoskeletal function. Future studies should consider individualized progression strategies, objective monitoring of physical activity levels, longitudinal follow-up, and external validation of the AI-based screening model.

Conclusion

The 12-week tele-pulmonary rehabilitation program significantly improved exercise capacity in community-based adults identified as having increased predicted risk of COPD through an AI-assisted radiographic screening approach. Although participants did not have established COPD, the findings suggest that early, structured tele-rehabilitation may enhance functional capacity in individuals with elevated predicted risk. This preventive strategy may contribute to promoting long-term respiratory health in community settings. Further research is warranted to optimize intervention intensity, evaluate long-term outcomes, and externally validate the AI-based risk stratification model.

Ethical approval

The study was performed in compliance with relevant laws and institutional guidelines, and the appropriate institutional committees approved it (COAAMSEC-65EX-019); participants also signed a written consent form.

Funding

This research received no external funding.

Conflict of interest

The authors report there are no conflicts of interest in this work.

Data availability statement

The data presented in this study are available upon request from the corresponding authors.

CRediT authorship contribution statement

Busaba Chuatrakoon: conceptualization, methodology, investigation, data collection, data curation, writing original draft, review and edit, supervision; **Siriphun Kongsawasdi:** methodology, data curation, review and edit; **Kittichai Wantanajittikul:** data curation, review and edit; **Sompong Sriburee:** data collection, review and edit; **Supatcha Konghakote:** data collection, review and edit.

Acknowledgments

The authors would like to thank all the participants who participated in this study. Thanks also go to the Faculty of Associated Medical Sciences, Chiang Mai University, for supporting this research project.

References

[1] Kyu HH, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018; 392: 1859-922. doi: 10.1016/S0140-6736(18)32335-3.

[2] Global Initiative for Chronic Obstructive Lung Disease [Internet]. USA: GOLD; [cited 2025 Apr]. Available from: http://goldcopd.org/2022-gold-reports-2/.

[3] Roig M, Eng JJ, MacIntyre DL, Road JD, Reid WD. Deficits in muscle strength, mass, quality, and mobility in people with chronic obstructive pulmonary disease. J Cardiopulm Rehabil Prev. 2011; 31: 120-4. doi: 10.1097/HCR.0b013e3181f68ae2.

[4] Chuatrakoon B, Ngai SPC, Sungkarat S, Uthaikhup S. Balance impairment and effectiveness of exercise intervention in chronic obstructive pulmonary disease-a systematic review. Arch Phys Med Rehabil. 2020; 101: 1590-602. doi: 10.1016/j.apmr.2020.02.007.

[5] Zhang L, Maitinuer A, Lian Z, Li Z, Ding W, Wang W, et al. Home based pulmonary telerehabilitation under telemedicine system for COPD: a cohort study. BMC Pulm Med. 2022; 22(1): 284. doi: 10.1186/s12890-022-02077-w.

[6] Li LSK, Butler S, Goldstein R, Brooks D. Comparing the impact of different exercise interventions on fatigue in individuals with COPD: a systematic review and meta-analysis. Chronic Respir Dis. 2019. doi: 10.1177/1479973119894855.

[7] Moore E, Palmer T, Newson R, Majeed A, Quint JK, Soljak MA. Pulmonary rehabilitation as a mechanism to reduce hospitalizations for acute exacerbations of COPD: a systematic review and meta-analysis. Chest. 2016; 150(4): 837-59. doi: 10.1016/j.chest.2016.05.038.

[8] Ora J, Prendi E, Attinà ML, Cazzola M, Calzetta L, Rogliani P. Efficacy of respiratory telerehabilitation in COPD patients: systematic review and meta-analysis. Monaldi Arch Chest Dis. 2022. doi: 10.4081/monaldi.2022.2105.

[9] Vasilopoulou M, Papaioannou AI, Kaltsakas G, Louvaris Z, Chynkiamis N, Spetsioti S, et al. Home-based maintenance tele-rehabilitation reduces the

risk for acute exacerbations of COPD, hospitalisations and emergency department visits. Eur Respir J. 2017. doi: 10.1183/13993003.02129-2016.

[10] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway (NJ): IEEE; 2016. p. 770-8. doi: 10.1109/CVPR.2016.90.

[11] Chuatrakoon B, Uthaikhup S, Ngai SPC, Liwsrisakun C, Pothirat C, Sungkarat S. The effectiveness of home-based balance and pulmonary rehabilitation program in individuals with chronic obstructive pulmonary disease: a randomized controlled trial. Eur J Phys Rehabil Med. 2022; 58(3): 478-86. doi: 10.23736/S1973-9087.22.06866-1.

[12] ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med. 2002; 166: 111-7. doi: 10.1164/ajrccm.166.1.at1102.

[13] Pothirat C, Kiatboonsrian S, Chuchottaworn C. Validation of the new COPD assessment test translated into Thai in patients with chronic obstructive pulmonary disease. BMC Pulm Med. 2014. doi: 10.1186/1471-2466-14-193.

[14] Mahler DA, Ward J, Waterman LA, McCusker C, ZuWallack R, Baird JC. Patient-reported dyspnea in COPD reliability and association with stage of disease. Chest. 2009; 136: 1473-9. doi: 10.1378/chest.09-0260.

[15] Coldwells A, Atkinson G, Reilly T. Sources of variation in back and leg dynamometry. Ergonomics. 1994; 37(1): 79-86. doi: 10.1080/00140139408963625.

[16] Tsutsui M, Gerayeli F, Sin DD. Pulmonary rehabilitation in a post-COVID-19 world: telerehabilitation as a new standard in patients with COPD. Int J Chron Obstruct Pulmon Dis. 2021; 16: 379-91. doi: 10.2147/COPD.S263031.

[17] Paneroni M, Colombo F, Papalia A, Colitta A, Borghi G, Saleri M, et al. Is telerehabilitation a safe and viable option for patients with COPD? A feasibility study. COPD. 2015; 12(2): 217-25. doi: 10.3109/15412555.2014.933794.

[18] Tsai LL, McNamara RJ, Moddel C, Alison JA, McKenzie D, McKeough ZJ. Home-based telerehabilitation via real-time videoconferencing improves endurance exercise capacity in patients with COPD: the randomized controlled TeleR study. Respirology. 2017; 22(4): 699-707. doi: 10.1111/resp.12966.

[19] Holland AE, Nici L. The return of the minimum clinically important difference for 6-minute-walk distance in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2013; 187(4): 335-6. doi: 10.1164/rccm.201212-2272ED.

Journal of Associated Medical Sciences 2026; 59(3): 334-343

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



Prediction of coronavirus disease severity based on increased levels of microvesicles: A systematic review and meta-analysis

Nastasya Nunki[1,2], Yetti Hernaningsih[3,4,5*], Puspa Wardhani[3,4,6], Citrawati Dyah Kencono Wungu[6,7], Agil Saputra[8], Narazah Mohd Yusoff[3,9]

[1]Laboratory Medicine, Master Program of Basic Medical Science, Faculty of Medicine, Universitas Airlangga, Surabaya, East Java, Indonesia.
[2]Department of Medical Laboratory Technology, Faculty of Health, Universitas Nahdlatul Ulama Surabaya, Surabaya, East Java, Indonesia.
[3]Department of Clinical Pathology, Faculty of Medicine, Universitas Airlangga, Surabaya, East Java, Indonesia.
[4]Dr. Soetomo General Academic Hospital, Surabaya, East Java, Indonesia.
[5]Airlangga University Hospital, Surabaya, East Java, Indonesia.
[6]Institute of Tropical Disease, Universitas Airlangga, Surabaya, East Java, Indonesia.
[7]Department of Physiology and Medical Biochemistry, Faculty of Medicine, Universitas Airlangga, Surabaya, East Java, Indonesia.
[8]Department of Medical Laboratory Technology, Faculty of Medicine, Universitas Palangka Raya, Palangka Raya, Central Borneo, Indonesia.
[9]Genetics Unit, Clinical Diagnostics Lab, Advanced Medical and Dental Institute, Universiti Sains Malaysia, Penang, Malaysia.

ARTICLE INFO

Article history:
Received 6 March 2026
Accepted as revised 10 July 2026
Available online 31 July 2026

Keywords:
Biomarker, COVID-19, microvesicles, procoagulant, severity

ABSTRACT

Background: Elevated microvesicle (MV) levels are potential biomarkers of coronavirus disease 2019 (COVID-19). We compared different types of microvesicles (total microvesicles, TMVs), platelet-derived microvesicles (PMVs), endothelium-derived microvesicles (EMVs), and tissue factor-bearing microvesicles (TF-MVs).

Objectives: We investigated the correlations between microvesicles and D-dimers, as well as between microvesicles and COVID-19 severity.

Materials and methods: A systematic literature search was conducted in PubMed, Web of Science, Scopus, and ScienceDirect to identify relevant studies comparing microvesicle types between patients with COVID-19 and healthy controls. Pooled estimates of the standardized mean difference (SMD) for each MV type were calculated using RevMan 5.4, along with the corresponding 95% confidence intervals (CIs).

Results: This meta-analysis included six studies comprising 183 COVID-19 patients and 135 healthy controls for quantitative analysis. Compared with controls, COVID-19 patients presented significantly higher levels of circulating TMVs, PMVs, EMVs, and TF-MVs (TMVs: SMD=4.27, 95% CI 2.18-6.36, p<0.0001; PMVs: SMD=2.07, 95% CI 0.99-3.15, p=0.0002; EMVs: SMD=1.20, 95% CI 0.56-1.84, p=0.0002; and TF-MVs: SMD=2.56, 95% CI 1.98-3.15, p<0.00001). D-dimer levels were correlated with MV levels (N=82, r=0.430, 95% CI 0.227-0.592, p<0.0001). In general, MVs were correlated with COVID-19 severity (N=168; r=0.942; 95% CI, 0.558-0.993; p=0.002).

Conclusion: Patients with severe COVID-19 had elevated levels of TMVs, PMVs, EMVs, and TF-MVs. Therefore, MVs may contribute to the procoagulant effects associated with prognostic and severity biomarkers of COVID-19.

Corresponding contributor.
Author's Address: *Department of Clinical Pathology, Faculty of Medicine, Universitas Airlangga, Surabaya, East Java, Indonesia.*
E-mail address: yetti-h@fk.unair.ac.id
doi: 10.66285/JAMS.2026.106
E-ISSN: 2539-6056

Introduction

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2

(SARS-CoV-2), affects many body systems, including the hematological system, and has garnered global attention since its initial discovery.[1] The invasion of host cells by SARS-CoV-2 can lead to thrombosis,[2] for which hypercoagulopathy is a contributory factor and potential cause of severe COVID-19.[3] Recently, extracellular vesicles have been identified as novel diagnostic biomarkers of triggers of hypercoagulation and subsequent thrombosis in COVID-19 patients [4] and have been implicated in various human diseases.[5]

Microvesicles (MVs), a type of extracellular vesicle,[6] are released from various cells following cellular hyperactivation, including platelets, leukocytes, and endothelial cells.[7] MVs possess intact membrane structures and express membrane glycoproteins specific to their origin cells.[8] Thus, MVs can be isolated from platelet-free plasma (PFP) by ultracentrifugation and exhibit diverse immunophenotypes.[8] Upon cellular activation, MVs adhere to the activated endothelium along with neutrophil extracellular traps (NETs), modulate the coagulation cascade, and induce excessive thrombin production, which ultimately results in a hypercoagulable state.[9]

As MVs are released under pathophysiological conditions, their clinical relevance as disease markers has increasingly generated interest.[10] Furthermore, MVs may influence COVID-19 severity. MVs contribute to blood clotting by exposing negatively charged phospholipids and tissue factors (TFs), which initiate the coagulation pathway *in vivo*.[11] MV formation disrupts cellular systems and leads to asymmetry of the phospholipid bilayer in the cell membrane, resulting in the externalization of phosphatidylserine (PS). A notable feature of MVs is their involvement in blood coagulation as a source of TF and as a negatively charged surface on which clotting factor complexes can accumulate.[11] Thus, MVs are known for their procoagulant properties.[12,13] Moreover, D-dimer is a part of the routine screening protocol for determining the severity of COVID-19 and is associated with MV-related tissue factor-dependent activation of blood cells. Thus, the presence of MVs supports fibrin formation, with consequent D-dimer formation and fibrin degradation by plasmin.

This review compared platelet-derived, endothelium-derived, and TF-bearing MV levels in COVID-19 patients with those in healthy controls. Moreover, the correlations between MVs and D-dimer, an essential marker of thrombus formation, and between MVs and COVID-19 severity were assessed. To strengthen these findings, a meta-analysis explored the potential application of MVs as novel biomarkers of thrombus occurrence and disease severity in COVID-19 patients.

Materials and methods

Ethical consideration

The protocol for this systematic review has been registered in the International Prospective Registry of Systematic Reviews (PROSPERO CRD42022362375). This review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines. Only publicly available data from published studies were analyzed, and no direct patient interaction or identifiable data were used.

Type of sampling and reasons for selection

Purposive sampling of studies was used for a systematic review and meta-analysis. The selected studies were chosen based on specific inclusion criteria (human studies on COVID-19 and microvesicles, comparison with healthy controls, quantification of MV levels, and correlation with D-dimer levels and disease severity) to ensure the relevance and robustness of the findings. A systematic literature search was conducted in PubMed, Web of Science, Scopus, and ScienceDirect.

Search strategy

A systematic search was conducted using PubMed, Web of Science, Scopus, and ScienceDirect for articles published through May 10, 2023. The search was not limited by region or release type, and a focused strategy was employed to identify human studies that examined MVs from COVID-19 patients. The search strategy involved a combination of improvisation keywords, MeSH terms (COVID-19, coronavirus, SARS-CoV-2, microvesicles, microparticles, procoagulant), and abbreviations (MVs and MPs).

Inclusion criteria

The eligibility criteria for study inclusion were based on five parameters: population, intervention, comparator, outcome, and study design (PICOS). This meta-research methodology ensures an unbiased search, inclusion, and assessment of articles.

Population

Only studies involving human participants were included.

Intervention

The included studies followed MV isolation protocols for blood samples obtained from human participants; patients who were diagnosed using RT-PCR, patients who were in the post-COVID-19 convalescent stage after up to four weeks of being PCR negative for the virus; patients who were subjected to D-dimer examination; and patients with comorbidities, such as diabetes, hypertension, coronary heart disease, hyperlipidemia, and chronic renal diseases.

Comparator

Blood samples from healthy controls were used as comparator groups. The severity of COVID-19 patients and patients whose D-dimer levels were also measured were used as the assessment groups.

Outcome

MVs were quantified in patients who were clinically evaluated for COVID-19, and disease severity was graded as mild to severe according to World Health Organization guidelines. Compared with those in the control group, the specific MVs of interest in COVID-19 patients included total microvesicles (TMVs), platelet-derived microvesicles (PMVs), endothelium-derived microparticles (EMVs), and tissue factor-bearing microvesicles (TF-MVs). The correlations of MVs with disease severity and D-dimer levels were investigated.

Study design

This systematic review included observational, analytical, cohort, and cross-sectional studies that measured MV levels in cohorts of COVID-19 patients. The studies were published in 2020.

Exclusion criteria

Studies involving nonhuman subjects were excluded. We excluded studies of patients who followed a home isolation protocol, had bacterial inflammation (non-COVID-19) or malignancies, or preclinical animal models. Review articles, case reports, case series, meta-analyses, duplicate studies, and non-English articles were excluded.

Study and data collection

Titles and abstracts that met the inclusion criteria were screened for full-text review, and all reviewers participated. Any discrepancies were resolved through discussions among the reviewers. Figure 1 depicts the Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart of the study selection process. The extracted data were subjected to a double-blinded process. The studies were assessed based on various characteristics, including specimen type, MV isolation technique, instruments used for detection, and other relevant aspects, all of which were recorded and registered in the PROSPERO protocol. To ensure specificity and consistency, the terms “microvesicles” and “microparticles” were used to identify studies that defined MVs as particles in the 100-1,000 nm size range. The characterization of MVs by researchers includes the quantification of MVs, analysis of MV size, and positive labeling of MV markers.

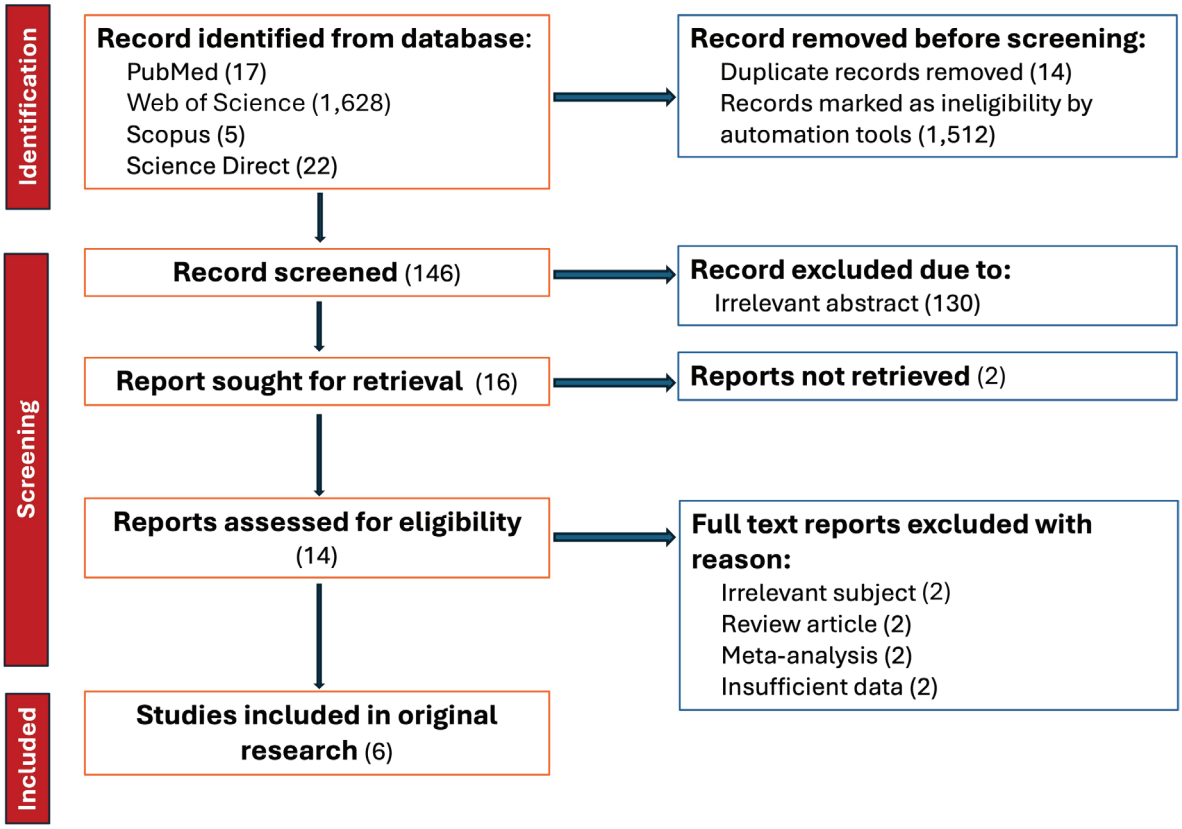


Figure 1. *PRISMA flowchart of the literature search and study selection*

Quality assessment of studies

The methodological quality of the included studies was evaluated using the Newcastle-Ottawa Quality Scale (NOS), which comprises three components: participant selection, comparability of the study groups, and exposure. Each study was assigned a score from 0 to 10.

Statistical analysis

Heterogeneity between studies was assessed using the Q test and I^2 test. A random or fixed-effects model was used to calculate the pooled standardized mean difference (SMD). When significant heterogeneity ($p<0.05$) was indicated across studies, a random effects model was used for the analysis. Otherwise, a fixed-effects model was used ($p>0.05$). The 95% confidence intervals (CIs) of the SMDs were calculated. Begg's funnel plot and Egger's test for publication bias were recommended, with 10 studies included to assess the evidence of publication bias. If publication bias was present, the funnel plot was asymmetric, and Egger's test was significant ($p<0.05$). Review Manager meta-analysis version 5.4 (The Cochrane Collaboration, UK) was used to calculate the SMD and CI.

Standardized mean difference (SMD) was used to compare microvesicle levels between COVID-19 patients and control groups because the included studies utilized different measurement scales and reporting units. SMD values were interpreted using conventional thresholds: 0.2 (small effect), 0.5 (moderate effect), and ≥0.8 (large effect). Positive SMD values indicated higher microvesicle levels in COVID-19 patients than in controls, whereas negative values indicated lower levels.

For studies reporting correlations between microvesicle and D-dimer levels or between microvesicle levels and COVID-19 severity, correlation coefficients were Fisher-transformed prior to pooled analysis.

Results

Study characteristics

In this systematic review, the initial search retrieved 1,672 studies. After removing duplicate records and records deemed ineligible based on automated tools (e.g., publication year and language), 146 studies remained. The abstracts of these studies were reviewed, and 130 articles were excluded as irrelevant. Two studies were subsequently retrieved for review, whereas two were not, resulting in the eligibility of 14 studies. Finally, based on a thorough examination of the full texts, only six studies met the inclusion criteria and were included in the analysis of MVs/microparticles in COVID-19 patients.

The essential characteristics of the included studies are presented in Table 1. Most of the included studies were conducted in Egypt (N=2), followed by the United States (N=1), Saudi Arabia (N=1), Germany (N=1), and China (N=1). The first study to discuss MVs in COVID-19 patients was published in 2020. Several studies have examined MVs in COVID-19 patients via blood samples. However, only four studies (N=4) conducted original research on MVs in humans. Zahran *et al*. and Guo *et al*. reported data on MVs in patients with varying severities of COVID-19.[14-15] Sun *et al.,* Hamali *et al*., Rausch *et al*., and Abdelmaksoud *et al.* reported the presence of MVs in patients with COVID-19.[16-19] Four of the six included studies employed a cross-sectional design, whereas the others used cohort and case-control designs. The quality of the studies was assessed using the NOS, and all studies deemed high quality had scores higher than 7.

Table 1. Characteristics of the systematic studies.

First Author	Study location, Year	MVs sources	MVs isolation, size/ morphology analysis	Surface marker	Sample size		Study design
					COVID-19 patient	Healthy control	
Zahran AM *et al*.[14]	Egypt, 2021	PFP	Ultracentrifugation, specific monoclonal antibodies with fluorochrome; calibration beads (1.0 µm), flow cytometry	Endothelial MVs ($CD146^+$ $CD45^-$ and platelet MVs are $CD41^+$	42	20	Cross-sectional
Sun Haoyu *et al*.[16]	USA, 2022	PFP	Ultracentrifugation, specific monoclonal antibodies with fluorochrome, calibration beads (100, 200, 380, 490, and 1,000 nm), TEM, flow cytometry	MVs from white blood cells ($CD45^+$), platelets ($CD61^+$), and endothelial cells ($CD144^+$),	5	5	Cross-sectional
Hamali HA *et al*.[17]	Saudi Arabia, 2022	PFP	Ultracentrifugation, MV-Activity Kit, ELISA	MVs and TF-bearing MVs	40	37	Case control
Rausch L *et al*.[18]	Germany, 2021	Peripheral blood mononuclear cells	Ultracentrifugation, TEM, flow cytometry	PS+ EVs and PMVs-marker+ (CD41/ CD63/CD274/ CD62P) cells.	54	35	Cohort
Guo Da-Zhi *et al*.[15]	China, 2020	PFP	Ultracentrifugation, flow cytometry	Annexin V+ MVs+	71	30	Cross-sectional
Abdelmaksoud MF *et al*.[19]	Egypt, 2023	Not described	ELISA Kit	PDMVs (anti-CD42a/GP-IX and anti-CD42b/GPIb antibodies	25	32	Cross-sectional

Note: *PFP: platelet-free plasma; all samples used in all studies were from humans.*

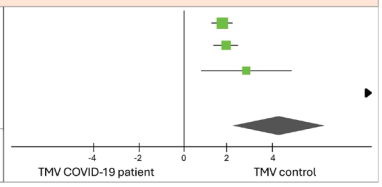
Meta-analysis of MVs in COVID-19 patients

This meta-analysis examined the differences in MV levels between patients with COVID-19 and healthy controls. The analysis revealed significantly higher TMV levels in COVID-19 patients than in controls (SMD=4.27, 95% CI=2.18–6.36, p<0.0001; Figure 2A). Similar significant differences were observed for PMVs (SMD=2.07, 95% CI=0.99-3.15, p=0.0002), EMVs (SMD =1.20, 95% CI=0.56-1.84, p=0.0002), and TF-MVs (SMD=2.56, 95% CI=1.98-3.15, p<0.00001) between COVID-19 patients and controls (Figure 2B-2D). Notably, PMV levels were higher than those of EMVs; however, TF-MVs presented the highest levels, followed by PMVs and EMVs (Figure 2B-2D). Furthermore, this study investigated the correlation between MVs and routine COVID-19 test parameters, such as D-dimer levels, which indicate thrombus formation due to increased coagulation activity in COVID-19 patients. The analysis revealed a significant correlation between MVs and D-dimer levels (N=82, r=0.430, 95% CI=0.227-0.592, p<0.0001) (Figure 2E). Additionally, MVs increased with disease severity (N=168, r=0.942, 95% CI=0.558-0.993, p=0.002) (Figure 2F).

A

Study or subgroup	TMV COVID-19 patient Mean	SD	Total	TMV control Mean	SD	Total	Weight (%)	Std. mean difference IV, random, 95% CI
Guo 2020	2,286.47	299	71	1,851.53	30.13	30	31.9	1.71 [1.22, 2.20]
Hamali 2022	21.7	12.5	40	3.9	1.8	37	31.8	1.94 [1.39, 2.48]
Sun 2022	70.14	29.28	5	5.45	0.49	5	24.9	2.82 [0.81, 4.84]
Zahran 2021	67.9	2.5	19	26	1.2	20	11.4	21.11 [16.13, 26.09]
Total (95% CI)			**135**			**92**	**100**	**4.27 [2.18, 6.36]**

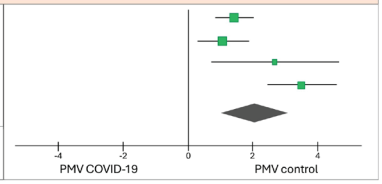
Heterogeneity: Tau^2=3.50; Chi^2=58.46, df=3 (p<0.00001); I^2=95%
Test for overall effect: Z=4.01 (p<0.0001)



B

Study or subgroup	PMV COVID-19 Mean	SD	Total	PMV control Mean	SD	Total	Weight (%)	Std. mean difference IV, random, 95% CI
Abdelmaksoud 2023	38.7	10.6	25	18.9	15.3	32	30.2	1.45 [0.86, 2.04]
Rausch 2021	24.5	18.6	23	7.1	4.4	11	28.4	1.09 [0.32, 1.86]
Sun 2022	36.71	16.02	5	2.94	0.38	5	15.9	2.69 [0.73, 4.65]
Zahran 2021	62.01	2.6	19	53.38	2.2	20	25.5	3.52 [2.48, 4.55]
Total (95% CI)			**72**			**68**	**100**	**2.07 [0.99, 3.15]**

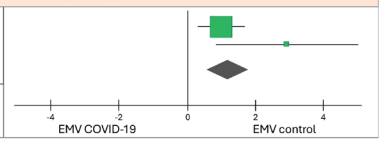
Heterogeneity: Tau^2=0.92; Chi^2=15.91, df=3 (p=0.001); I^2=81%
Test for overall effect: Z=3.75 (p=0.0002)



C

Study or subgroup	EMV COVID-19 Mean	SD	Total	EMV control Mean	SD	Total	Weight (%)	Std. mean difference IV, fixed, 95% CI
Zahrah 2021	11.94	0.83	19	11.26	0.43	20	90.5	1.02 [0.34, 1.69]
Sun 2022	1.73	0.7	5	0.12	0.01	5	9.5	2.94 [0.87, 5.00]
Total (95% CI)			**24**			**25**	**100**	**1.20 [0.56, 1.84]**

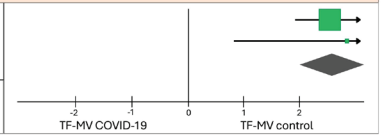
Heterogeneity: Chi^2=3.00, df=1 (p=0.08); I^2=67%
Test for overall effect: Z=3.68 (p=0.0002)



D

Study or subgroup	TF-MV COVID-19 Mean	SD	Total	TF-MV control Mean	SD	Total	Weight (%)	Std. mean difference IV, fixed, 95% CI
Hamali 2022	1.9	0.7	40	0.5	0.3	37	91.8	2.54 [1.93, 3.15]
Sun 2022	59.8	24.91	5	4.18	0.34	5	8.2	2.85 [0.82, 4.88]
Total (95% CI)			**45**			**42**	**100**	**2.56 [1.98, 3.15]**

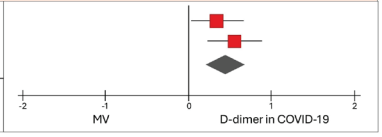
Heterogeneity: Chi^2=0.08, df=1 (p=0.77); I^2=0%
Test for overall effect: Z=8.63 (p<0.00001)



E

Study or subgroup	Fisher's Z	SE	Weight (%)	Fisher's Z IV, fixed, 95% CI
Zahran 2021	0.34844977	0.160128	51.3	0.35 [0.03, 0.66]
Hamali 2022	0.57087283	0.164399	48.7	0.57 [0.25, 0.89]
Total (95% CI)			**100**	**0.46 [0.23, 0.68]**

Heterogeneity: Chi^2=0.94, df=1 (p=0.33); I^2=0%
Test for overall effect: Z=3.98 (p<0.0001)



F

Study or subgroup	Fisher's Z	SE	Weight (%)	Fisher's Z IV, random, 95% CI
Guo 2022	2.32345903	0.094916	51.7	2.32 [2.14, 2.51]
Rausch 2021	1.16604452	0.235702	48.3	1.17 [0.70, 1.63]
Total (95% CI)			**100**	**1.76 [0.63, 2.90]**

Heterogeneity: Tau^2=0.64; Chi^2=20.75, df=1 (p<0.00001); I^2=95%
Test for overall effect: Z=3.05 (p=0.002)

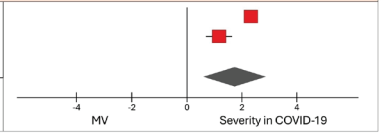


Figure 2. *Forest plots and meta-analyses comparing microvesicle types between COVID-19 patients and controls. A: Forest plot and meta-analysis of TMVs, B: Forest plot and meta-analysis of PMVs, C: Forest plot and meta-analysis of EMVs, D: Forest plot and meta-analysis of TF-MVs, E: Forest plot and meta-analysis of MVs correlated with D-dimer levels in COVID-19 patients, F: Forest plot and meta-analysis of MVs correlated with disease severity in COVID-19 patients.*

Assessment of studies

Three independent reviewers assessed the quality of the studies to evaluate the risk of bias (Supplementary data). The included cohort studies did not indicate publication bias, and none were deemed of low quality. The NOS assessment tool incorporates assessment parameters specific to each study design. All studies measured TMVs in COVID-19 patients, except the study by Rausch *et al.,* which precisely measured CD41 PMVs.[18] However, the quality assessment results did not differ significantly between the studies (Supplementary data). Specifically, this analysis focused on TMVs, PMVs, EMVs, and TF-MVs expressed in COVID-19 patients and healthy controls. All the results collected within the groups were consistent with the overall analysis. To understand the role of MVs in COVID-19, this study analyzed the association between MVs and COVID-19 severity, as well as the correlation between MVs and D-dimer, a coagulation parameter. All the findings are consistent with the original analysis.

Publication bias and sensitivity analysis

Begg's funnel plot or Egger's test was not performed to assess evidence of publication bias because there were fewer than 10 studies; according to the Cochrane guidelines, funnel plots are recommended for studies exceeding this threshold.

Discussion

Regardless of the type of control or comorbidity, our systematic review revealed a significant increase in annexin V-labeled TMVs, PMVs, and EMVs in patients with COVID-19. This is the first systematic review to evaluate MVs in COVID-19 patients and compare them with MV levels in healthy controls. Despite the potential heterogeneity among studies, the results of the subgroup and sensitivity analyses were robust and indicated a significant increase in MVs in COVID-19 patients. Many severe cases of COVID-19 are associated with hemostatic markers such as D-dimer.[20] The increased coagulation and thrombosis observed in COVID-19 are triggered by cellular hyperactivation, which is initiated when the spike surface receptor of SARS-CoV-2 binds to ACE-2 receptors on host cells through the transmembrane protease serine-type 2 (TMPRSS2), facilitating fusion with host cells.[21]

SARS-CoV-2 infection involves the endothelium, and entry of the virus into ACE-2-expressing endothelial cells leads to endothelial dysfunction and loss of endothelial integrity.[22] Consequently, uncontrolled viral invasion triggers endothelial dysfunction, activating the immune system and hemostatic responses. Endothelial dysfunction, in turn, contributes to the release of MVs from the endothelium, further increasing the production of proinflammatory cytokines, such as IL-6, IL-8, IL-10, and TNF-α, and ultimately exacerbates the severity of COVID-19.[23] Additionally, during acute inflammation, various cells, including neutrophils, monocytes, lymphocytes, and platelets, undergo activation and release MVs.[24]

According to Balbi *et al.*, many EVs carrying surface antigens found in COVID-19 patients originate from platelets and the endothelium.[8] Consistent with this, the studies included in our meta-analysis predominantly employed markers derived from platelets and the endothelium, with more recent studies investigating tissue factor expression. Kaur *et al.* reported the upregulation of IL-6, IL-8, IL-10, and TNF-α, along with increased prothrombotic and coagulation factors, in patients with severe COVID-19.[25] MVs can form coagulation complexes and support thrombin formation, contributing to the prothrombotic changes observed in COVID-19 patients and playing a significant role in the development of a hypercoagulable state.[17] MVs of COVID-19 patients with comorbidities are more procoagulant. In addition to adhering to the endothelium, it enables the assembly of enzyme complexes and thrombin formation, thus contributing to hypercoagulation.[16] In line with these findings, Guo et al. confirmed that increased production of circulating MVs is associated with COVID-19 progression and that the number of MVs may indicate disease severity.[15]

Previous studies have consistently demonstrated an increase in the number of microvesicles, and our meta-analysis further confirmed that MV levels are elevated not only in COVID-19 patients but also in healthy controls and are positively correlated with D-dimer levels. Furthermore, the expression of TFs on MV surfaces was significantly higher in COVID-19 patients than in healthy controls. Notably, TF-bearing MVs appeared to be the most abundant, followed by MVs originating from platelets and the endothelium.

TFs are expressed on the surface of MVs when cells are activated, causing cell membrane translocation and facilitating TF-initiated coagulation.[26] Several tissue factors are released in response to inflammation.[22] Generally, TF is found in the subendothelial matrix. In previous studies, TF-MVs were reported to play a role in thrombosis, indicating their severity and death, and these TF-bearing MVs are procoagulants that trigger the start of the coagulation cascade.[27,28,12]

The presence of hyperactive platelets in patients with COVID-19 increases the number of PMVs in circulation.[18] MV derivatives, such as PMVs, contain cellular receptors that may trigger an inflammatory storm. The ability of PMVs to mediate inflammatory cytokines can exacerbate the inflammatory response. In COVID-19, there is an increase in PMVs due to secondary inflammatory conditions when the virus is active.[19] PMVs can change the function of cells, including neutrophils, endothelial cells, and megakaryocytes, as platelet precursors, thereby enabling communication between cells during infection.[11]

PMVs and EMVs influence the expression of cell

adhesion molecules, enabling interactions with the endothelium and the recruitment of inflammatory cells.[29] Garnier *et al*. reported that the concentration of EMVs was higher in intubated COVID-19 patients than in uninfected controls.[30] Subgroup analysis of molecular cluster of differentiation (CD) markers was not performed because of insufficient data. Increased D-dimer levels in COVID-19 patients are associated with cell activation and coagulation due to the release of MVs.[30,31,14] D-dimer is often found in patients with COVID-19, is a risk factor that correlates with disease severity and can predict hypercoagulability status due to the risk of thrombosis and mortality.[32-34] Thus, when D-dimer levels correlate with MV release, MV may indirectly trigger the presence of cross-linked fibrin breakdown products.

Microvesicles have been shown to increase in number and correlate with D-dimer, which predicts hypercoagulability. MVs can be identified early in the diagnosis of thrombosis and have potential as diagnostic biomarkers. In addition, as part of the EV subtype, microvesicles play an essential role in cellular communication, making MVs promising candidates for therapeutic agents and drug delivery systems. MV may be a promising noninvasive biomarker for monitoring biological processes in the diagnosis and treatment of thrombosis.[35,36]

Microvesicles may also serve as promising therapeutic candidates due to their natural role in intercellular communication, ability to carry bioactive cargo (including proteins, lipids, and nucleic acids), biocompatibility, and low immunogenicity. These characteristics make them potential targeted drug-delivery systems for improving therapeutic efficacy while minimizing adverse effects.[35]

The results of this meta-analysis were consistent with those of previous studies. This meta-analysis used the standard deviation of the mean to validate studies that presented only numerical data. This study was conducted at the end of the COVID-19 pandemic. For this reason, studies have discussed potential markers that are not widely known for the treatment of COVID-19, such as PMV, EMV, and TF-bearing MVs, which can be used as new reliable markers of the response to pathophysiological severity in patients with COVID-19.

The potential role of microvesicles in predicting COVID-19 severity may be reflected by the magnitude of the pooled SMD, where larger SMD values indicate greater differences in microvesicle levels between severe COVID-19 patients and comparison groups.[37] Hypercoagulation and thrombosis in COVID-19 patients are attributable to MVs from platelets and endothelium derivatives or TF-bearing MVs, which increase over time, along with the severity of the effects. However, there is no consensus regarding the association of MVs with COVID-19. MV quantification in COVID-19 patients should be developed to initiate management to minimize disease worsening and high mortality.

These studies are limited in that few studies discuss microvesicles in COVID-19 patients, and not all have examined the associations between MVs and other laboratory parameters. Investigating this association in future studies could provide insights into the complex pathophysiology resulting from increased MVs in COVID-19 patients, particularly hypercoagulation and thrombosis. Another limitation of this study is the heterogeneity of microvesicle isolation, characterization, and detection methods across included studies, which may influence measurement consistency and comparability of reported microvesicle levels. This meta-analysis included a limited number of studies with relatively small sample sizes, which may affect the generalizability of the findings. Therefore, the results should be interpreted cautiously, and larger multicenter studies are needed to validate the role of microvesicles as predictive biomarkers in COVID-19.

Conclusion

In conclusion, the number of MVs detected in COVID-19 patients is significantly higher than that detected in healthy controls. The presence of MVs in COVID-19 patients is positively associated with disease severity. Specifically, TF-MVs are present at higher levels in patients with COVID-19 and are more abundant than PMVs and EMVs. MVs also demonstrate a positive correlation with D-dimer, a crucial parameter in hemostasis. The predictive value of MVs in assessing COVID-19 severity has been shown consistently across numerous studies, with high statistical significance (p<0.05). Therefore, MVs hold promise as valuable markers for predicting the prognosis of patients with COVID-19.

Ethical approval

None

Funding

None

Conflict of Interest

The Authors declared no conflict of interest.

CRediT authorship contribution statement

Nastasya Nunki: conceptualization, methodology, formal analysis, investigation, resources, data curation, writing: review and edit, project administration; **Yetti Hernaningsih**: conceptualization, validation, writing: original draft, supervision, funding acquisition; **Puspa Wardhani**: conceptualization, supervision; **Citrawati Dyah Kencono Wungu**: investigation, data curation, visualization; **Agil Saputra**: methodology, formal analysis, resources; **Narazah Mohd Yusoff**: supervision. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

The authors would like to acknowledge the professional editor at Editage for assisting in editing the manuscript.

References

[1] Li X, Zhong X, Wang Y, Zeng X, Luo T, Liu Q. Clinical determinants of the severity of COVID-19: A systematic review and meta-analysis. PLoS One. 2021; 16(5): e0250602. doi: 10.1371/journal.pone.0250602.

[2] Wool GD, Miller JL. The impact of COVID-19 disease on platelets and coagulation. Pathobiology. 2021; 88(1): 15-27. doi: 10.1159/000512007.

[3] Rostami M, Mansouritorghabeh H. D-dimer level in COVID-19 infection: a systematic review. Expert Rev Hematol. 2020; 13(11): 1265-75. doi: 10.1080/17474086.2020.1813023.

[4] Georgescu A, Simionescu M. Extracellular vesicles: versatile nanomediators, potential biomarkers and therapeutic agents in atherosclerosis and COVID-19-related thrombosis. Int J Mol Sci. 2021; 22(11): 5967. doi: 10.3390/ijms22115967.

[5] Barberis E, Vanella VV, Falasca M, Caneapero V, Cappellano G, Raineri D, et al. Circulating exosomes are strongly involved in SARS-CoV-2 infection. Front Mol Biosci. 2021; 8: 632290. doi: 10.3389/fmolb.2021.632290.

[6] Cappellano G, Raineri D, Rolla R, Giordano M, Puricelli C, Vilardo B, et al. Circulating platelet-derived extracellular vesicles are a hallmark of SARS-CoV-2 infection. Cells. 2021; 10(1): 85. doi: 10.3390/cells10010085.

[7] O'Kennedy N, Duttaroy AK. Platelet hyperactivity in COVID-19: Can the tomato extract Fruitflow® be used as an antiplatelet regime. Med Hypotheses. 2021; 147: 110480. doi: 10.1016/j.mehy.2021.110480.

[8] Balbi C, Burrello J, Bolis S, Lazzarini E, Biemmi V, Pianezzi E, et al. Circulating extracellular vesicles are endowed with enhanced procoagulant activity in SARS-CoV-2 infection. EBioMedicine. 2021; 67: 103369. doi: 10.1016/j.ebiom.2021.103369.

[9] Wijaya T, Hsieh PP, Melinda F, Suarca IK. Hypercoagulation state in infant with COVID-19: a case report. Indones J Biomed Sci. 2022; 16(1): 66-9. doi: 10.15562/ijbs.v16i1.354.

[10] Koupenova M, Corkrey HA, Vitseva O, Tanriveardi K, Somasundaran M, Liu P, et al. SARS-CoV-2 initiates programmed cell death in platelets. Circ Res. 2021; 129(6): 631-46. doi: 10.1161/CIRCRESAHA.121.319195.

[11] Whiteside TL. Procoagulant activity of extracellular vesicles in plasma of patients with SARS-CoV-2 infection. EBioMedicine. 2021; 68: 103411. doi: 10.1016/j.ebiom.2021.103411.

[12] Taus F, Salvagno G, Canè S, Fava C, Mazzaferri F, Carrara E, et al. Platelets promote thromboinflammation in SARS-CoV-2 pneumonia. Arterioscler Thromb Vasc Biol. 2020; 40(12): 2975-89. doi: 10.1161/ATVBAHA.120.315175.

[13] Althaus K, Marini I, Zlamal J, Pelzl L, Singh A, Häberle H, et al. Antibody-induced procoagulant platelets in severe COVID-19 infection. Blood. 2021; 137(8): 1061-71. doi: 10.1182/blood.2020008083.

[14] Zahran AM, El-Badawy O, Ali WA, Mahran ZG, Mahran EEMO, Rayan A. Circulating microparticles and activated platelets as novel prognostic biomarkers in COVID-19; relation to cancer. PLoS One. 2021; 16(2): e0246806. doi: 10.1371/journal.pone.0246806.

[15] Guo D, Lv Y, Qi Y, Pan S. Increased circulating microparticles and inflammatory factors aggravate coronavirus disease 2019 (COVID-19). Research Square [Preprint]. 2020. doi: 10.21203/rs.3.rs-53108/v1.

[16] Sun H, Du Y, Kumar R, Buchkovich N, He P. Increased circulating microparticles contribute to severe infection and adverse outcomes of COVID-19 in patients with diabetes. Am J Physiol Heart Circ Physiol. 2022; 323(5): H1176-H1193. doi: 10.1152/ajpheart.00293.2022.

[17] Hamali HA, Saboor M, Dobie G, Madkhali AM, Akhter MS, Hakamy A, et al. Procoagulant microvesicles in COVID-19 patients: possible modulators of inflammation and prothrombotic tendency. Infect Drug Resist. 2022; 15: 2359-68. doi: 10.2147/IDR.S361405.

[18] Rausch L, Lutz K, Schifferer M, Winheim E, Gruber R, Oesterhaus EF, et al. Binding of phosphatidylserine-positive microparticles by PBMCs classifies disease severity in COVID-19 patients. J Extracell Vesicles. 2021; 10(14): e12173. doi: 10.1002/jev2.12173.

[19] Abdelmaksoud MF, Abdelmaksoud SS, Abdelsamee HF, Ezzelregal HG, Alfeky MA. Platelets derived microparticles in COVID-19: correlation to inflammatory and coagulation state. J Appl Hematol. 2021; 12(4): 195-202. doi: 10.4103/joah.joah_73_21.

[20] Novena OD. Skin manifestations in coronavirus disease 2019 (COVID-19): A literature review. Indones J Biomed Sci. 2021; 15(2): 113-6. doi: 10.15562/ijbs.v15i2.327.

[21] Hassanpour M, Rezaie J, Nouri M, Panahi Y. The role of extracellular vesicles in COVID-19 virus infection. Infect Genet Evol. 2020; 85: 104422. doi: 10.1016/j.meegid.2020.104422.

[22] Zaid Y, Puhm F, Allaeys I, Naya A, Oudghiri M, Khalki L, et al. Platelets can associate with SARS-CoV-2 RNA and are hyperactivated in COVID-19. Circ Res. 2020; 127(11): 1404-18. doi: 10.1161/CIRCRESAHA.120.317703.

[23] Che Mohd Nassir CMN, Hashim S, Wong KK, Abdul Halim S, Idris NS, Jayabalan N, et al. COVID-19 infection and circulating microparticles-reviewing

evidence as microthrombogenic risk factor for cerebral small vessel disease. Mol Neurobiol. 2021; 58(8): 4188-215. doi: 10.1007/s12035-021-02422-1.

[24] Xiang M, Wu X, Jing H, Liu L, Wang C, Wang Y, et al. The impact of platelets on pulmonary microcirculation throughout COVID-19 and its persistent activating factors. Front Immunol. 2022; 13: 955654. doi: 10.3389/fimmu.2022.955654.

[25] Kaur S, Singh A, Kaur J, Verma N, Pandey AK, Das S, et al. Upregulation of cytokine signaling in platelets increases risk of thrombophilia in severe COVID-19 patients. Blood Cells Mol Dis. 2022; 94: 102653. doi: 10.1016/j.bcmd.2022.102653.

[26] Wang J, Pendurthi UR, Yi G, Rao LVM. SARS-CoV-2 infection induces the activation of tissue factor-mediated coagulation via activation of acid sphingomyelinase. Blood. 2021; 138(4): 344-9. doi: 10.1182/blood.2021010915.

[27] Franco C, Lacroix R, Vallier L, Judicone C, Bouriche T, Laroumagne S, et al. A new hybrid immunocapture bioassay with improved reproducibility to measure tissue factor-dependent procoagulant activity of microvesicles from body fluids. Thromb Res. 2020; 196: 414-24. doi: 10.1016/j.thromres.2020.10.003.

[28] Liu G, Yin XM. The role of extracellular vesicles in liver pathogenesis. Am J Pathol. 2022; 192(10): 1358-67. doi: 10.1016/j.ajpath.2022.06.009.

[29] Rodrigues KF, Pietrani NT, Fernandes AP, Bosco AA, de Sousa MCR, de Fátima Oliveira Silva I, et al. Circulating microparticles levels are increased in patients with diabetic kidney disease: A case-control research. Clin Chim Acta. 2018; 479: 48-55. doi: 10.1016/j.cca.2018.01.011.

[30] Garnier Y, Claude L, Hermand P, Sachou E, Claes A, Desplan K, et al. Plasma microparticles of intubated COVID-19 patients cause endothelial cell death, neutrophil adhesion and netosis, in a phosphatidylserine-dependent manner. Br J Haematol. 2022; 196(5): 1159-69. doi: 10.1111/bjh.18001.

[31] Canzano P, Brambilla M, Porro B, Cosentino N, Tortorici E, Vicini S, et al. Platelet and endothelial activation as potential mechanisms behind the thrombotic complications of COVID-19 patients. JACC Basic Transl Sci. 2021; 6(3): 202-18. doi: 10.1016/j.jacbts.2020.12.009.

[32] Yusuf F, Abubakar A, Maghfirah D, Baswin A. Relationship red distribution width to platelet ratio with fibrosis degrees based on transient elastography in chronic hepatitis B patients. Bali Med J. 2021; 10(2): 793-7. doi: 10.15562/bmj.v10i2.2530.

[33] Hernaningsih Y, Aryati, Notopuro PB, Hajat A, et al. Laboratory Aspects of COVID-19. Surabaya: Airlangga University Press; 2021.

[34] Wiyono N, Hernaningsih Y, Hajat A, Notopuro PB, Yusoff NM, Moses EJ. Thromboelastography parameter and its association with survival of COVID-19 patients: A retrospective cross-sectional study. Indones J Biomed Sci. 2023; 19(1): 124-9. doi: 10.15562/ijbs.v19i1.472 .

[35] Kumar MA, Baba SK, Sadida HQ, et al. Extracellular vesicles as tools and targets in therapy for diseases. Signal Transduct Target Ther. 2024; 9(1): 105. doi: 10.1038/s41392-024-01815-3.

[36] Nunki N, Hernaningsih Y, Wardhani P, et al. Platelet and Monocyte Microvesicles as Potential Biomarkers of COVID-19 Severity: A Cross-Sectional Analysis. Ann Lab Med. 2024; 44(5): 392-400. doi: 10.3343/alm.2023.0450.

[37] Wungu CDK, Kurniawan RB, Oktafia P. Tinjauan sistematis dan metaanalisis untuk kedokteran klinis [Systematic review and meta-analysis for clinical medicine]. Surabaya: Airlangga University Press; 2025.

Journal of Associated Medical Sciences 2026; 59(3): 344-356

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 PETG-based 3D-printed lumbar spine phantom for bone mineral density assessment

Thunyarat Chusin, Pakaphorn Kongfun, Suchanya Chanapim, Suphichaya Thanashappokin, Thanyawee Pengpan, Kingkarn Aphiwatthanasumet*

Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.

ARTICLE INFO

Article history:
Received 27 April 2026
Accepted as revised 31 July 2026
Available online 5 August 2026

Keywords:
Computed tomography, bone mineral density, dual-energy X-ray absorptiometry, phantom

ABSTRACT

Background: Recent advances in three-dimensional printing technology have enabled the development of customized phantoms for quantitative imaging applications, including bone mineral density assessment.

Objectives: This study aimed to develop and validate a 3D-printed lumbar spine phantom by comparing its radiological characteristics with those of the commercial anthropomorphic RSD RS-113T pelvic phantom and evaluating its performance against the GE Lunar iDXA quality control (QC) phantom.

Materials and methods: A commercial RSD RS-113T pelvic phantom was scanned using CT and DXA to obtain reference CT numbers and BMD values. Lumbar spine structures were segmented from CT images to design a 3D-printed lumbar spine phantom using fused deposition modeling (FDM). Polyethylene terephthalate glycol (PETG; densities 1.24 and 1.27 gm/cm^3), combined with calcium carbonate powder, was used to simulate bone mineral content. CT number and BMD measurements of the developed phantoms were compared with those of the RSD RS-113T pelvic phantom using Bland-Altman agreement analysis. BMD repeatability of the developed phantoms was subsequently validated using the GE Lunar iDXA QC phantom.

Results: The PETG black phantom (1.27 gm/cm^3) showed closer agreement with the RSD RS-113T pelvic phantom than the PETG white phantom (1.24 gm/cm^3), with smaller Bland–Altman mean biases for CT number (8.65 vs -297.10 HU) and BMD (0.22 vs 0.84 gm/cm^2). The RSD RS-113T pelvic phantom exhibited BMD CV values of less than 2%. The PETG black phantom also demonstrated better BMD repeatability (CV: 0.58%-2.92%) than the PETG white phantom, although some measurements exceeded the 2% precision criterion.

Conclusion: The PETG black 3D-printed lumbar spine phantom demonstrated superior BMD repeatability compared with the PETG white phantom. It also showed good agreement with the RSD RS-113T pelvic phantom and satisfactory BMD performance when validated using the GE Lunar iDXA QC phantom.

Introduction

Recent advances in three-dimensional (3D) printing have created new opportunities for developing customized phantoms in medical imaging and radiological physics.[1] Additive manufacturing techniques enable precise control over anatomical geometry, internal structure, and material composition, allowing the fabrication of phantoms with tunable radiological properties.[2,3] In particular, fused deposition modeling (FDM) combined

* **Corresponding contributor.**
Author's Address: *Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*
E-mail address: kingkarna@nu.ac.th
doi: 10.66285/JAMS.2026.107
E-ISSN: 2539-6056

with polymer-based materials[4] and mineral additives offers a promising approach for simulating bone-equivalent attenuation and density characteristics relevant to both dual-energy X-ray absorptiometry (DXA) and computed tomography (CT) imaging.[5-7]

DXA is a widely used quantitative imaging modality for assessing bone mineral density (BMD) based on the differential attenuation of X-ray photons at two distinct energy spectra. BMD values expressed in grams per square centimeter (gm/cm^2) serve as a fundamental parameter for the diagnosis, classification, and longitudinal monitoring of osteoporosis and other metabolic bone disorders.[8] Owing to its high precision, low radiation dose, and short acquisition time, DXA is regarded as the reference standard for clinical BMD evaluation.[9] However, the accuracy and reproducibility of DXA-derived BMD measurements depend on rigorous quality assurance procedures, which rely heavily on standardized phantoms to ensure measurement consistency across time and imaging systems.[10,11] Variations in geometry and material composition across manufacturers restrict cross-system comparability.[12] In computed tomography (CT), phantoms are similarly essential for evaluating system performance, including image quality, CT number accuracy, and measurement consistency. They provide standardized references for calibration and validation of quantitative imaging metrics under controlled conditions.[13,14]

In this context, 3D printing offers a flexible platform for developing customized phantoms with controlled structural and material properties. Nevertheless, the extent to which 3D-printed phantoms can accurately replicate the densitometric characteristics of established reference phantoms remains insufficiently explored. We hypothesize that optimizing filament density and material composition will yield a 3D-printed phantom with CT number and BMD values comparable to those of a reference phantom.

Accordingly, this study aimed to develop and quantitatively validate a 3D-printed lumbar spine phantom by comparing its radiological characteristics with those of the anthropomorphic RSD RS-113T pelvic phantom and evaluating its BMD performance using the GE Lunar iDXA QC phantom as the reference.

Materials and methods

The RSD RS-113T pelvic phantom was initially scanned using both a CT system and a DXA system to obtain baseline imaging and bone density data. For CT analysis, regions of interest (ROIs) with an area of 7.34 mm^2 were placed at three specific locations (one central and two lateral) on axial CT images of the lumbar vertebral bodies of L1-L5 to measure CT number values. Subsequently, bone mineral density (BMD) measurements of the same vertebrae were performed using the DXA system to provide corresponding reference values. Following this, the lumbar spine phantom was fabricated using FDM with filaments (eSUN, China; IEMAI, China) and Creality 3D printers (Shenzhen Creality 3D Technology Co., Ltd. China), specifically the Creality K1 Max and Creality Ender-3 V3 models. The printing materials included polyethylene terephthalate glycol (PETG) filaments with densities of 1.24 and 1.27 gm/cm^3 and polylactic acid (PLA) filament with a density of 1.23 gm/cm^3. BMD measurements were conducted using a GE Lunar iDXA Prodigy system (GE Healthcare, USA) operated at 76 kV, 3 mA.

CT imaging was performed with a PHILIPS Ingenuity scanner (Philips Healthcare, The Netherlands) following a standardized lumbar spine protocol set at 120 kV, 30 mA, and 1 mm slice thickness, balancing adequate X-ray penetration and spatial resolution to ensure precise bone structure assessment. This methodology enabled a comprehensive comparison of the imaging and bone density characteristics of the developed 3D-printed lumbar spine phantom with those of the reference phantom.

For radiological characterization, the CT number and BMD measurements of the developed 3D-printed phantom were compared with those of the commercial anthropomorphic RSD RS-113T pelvic phantom using Bland-Altman agreement analysis to assess measurement agreement. Subsequently, the BMD performance of the developed 3D-printed phantom was validated using the GE Lunar iDXA QC phantom.

3D block sample design

Firstly, a rectangular prism (cuboid) model was designed in Autodesk Fusion 360 with dimensions of 15 mm×15 mm in width and length, and a thickness of 21 mm, corresponding to the L3 vertebral body. Block samples were then printed using two commonly used filament materials: PETG and PLA. For each material, samples were fabricated at infill densities of 80% and 100%, utilizing two infill patterns: Gyroid, selected for its flexibility and mechanical strength, and Crosshatch, selected for its printing efficiency and economical material usage. To evaluate their radiological properties, CT scans were performed on the samples, with a single ROI of 7.34 mm^2 placed at the center of each block sample to measure CT numbers. Additionally, BMD values were obtained via DXA. Based on clinically reported DXA ranges, a BMD value of 0.50 gm/cm^2 was defined as low-density, 1.00 gm/cm^2 as intermediate-density, and 1.50 gm/cm^2 as nominal density target, and were selected for this study. The resulting CT number and BMD values from the block samples were then compared with those of the RSD RS-113T pelvic phantom to assess similarity.

In the second phase, the block samples were redesigned as box-shaped phantoms with lids, maintaining overall dimensions consistent with the L3 vertebral body (15x15x21 mm). The lids and edges were printed at 100% infill density to enhance structural integrity. Samples were fabricated using PETG and PLA at infill densities of 0%, 20%, and 40%, and printed

with three infill patterns: Grid, Crosshatch, and Gyroid (Figure 1). To evaluate the incorporation of bone-mimicking material, samples were weighed before and after filling with calcium carbonate ($CaCO_3$) powder, allowing assessment of filling efficiency. Subsequently, CT and DXA measurements were conducted, and the data were compared with those from the RSD RS-113T pelvic phantom. Through this process, the optimal combination of material, infill density, and pattern, yielding CT number and BMD values close to those of the pelvic phantom, was identified and selected for final phantom fabrication.

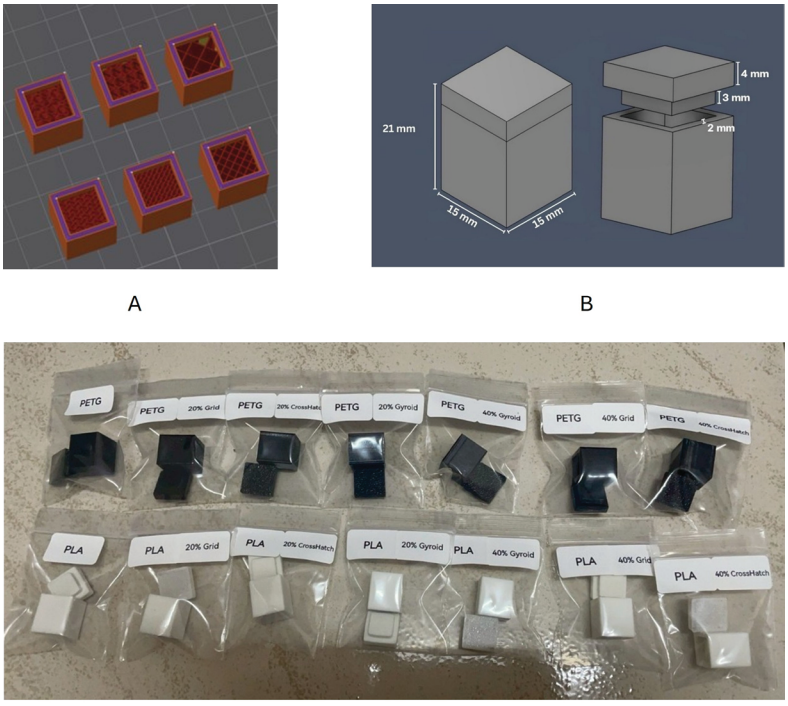


Figure 1. *Design and appearance of 3D-printed block samples. A: Block samples printed at infill densities of 20% and 40% using Grid, Cross Hatch, and Gyroid patterns. B: Dimensions of the block sample: 15 mm×15 mm×21 mm. C: PETG and PLA samples printed at infill densities of 0%, 20%, and 40%, with Grid, Cross Hatch, and Gyroid patterns.*

Fabrication, characterization, and validation of the lumbar spine phantom

The final lumbar spine phantom was generated by applying threshold-based segmentation to the CT images of the RSD RS-113T pelvic phantom using 3D Slicer software, thereby creating a 3D model that accurately replicated the size and shape of the original phantom. The model was then 3D-printed using the optimized parameters determined from the block sample tests, including the printing material, infill density, and printing pattern. Following fabrication, measurements were performed using both CT (Figure 2) and DXA. To validate the radiological characteristics of the printed phantom, CT number measurements were taken from three ROIs, one central and two laterals, on CT images of the printed phantom (Figure 3). To evaluate its suitability for bone density quality control applications, the repeatability of five BMD measurements was assessed and compared with that of the GE Lunar iDXA QC phantom (Figure 4). Finally, the measured parameters were statistically compared with those of the reference phantoms to confirm the reliability of the developed 3D-printed phantom.



Figure 2. *CT imaging of the developed 3D-printed lumbar spine phantom fabricated from PETG with calcium carbonate ($CaCO_3$) powder with various material densities. Left: 1.24 gm/cm^3, Right: 1.27 gm/cm^3.*

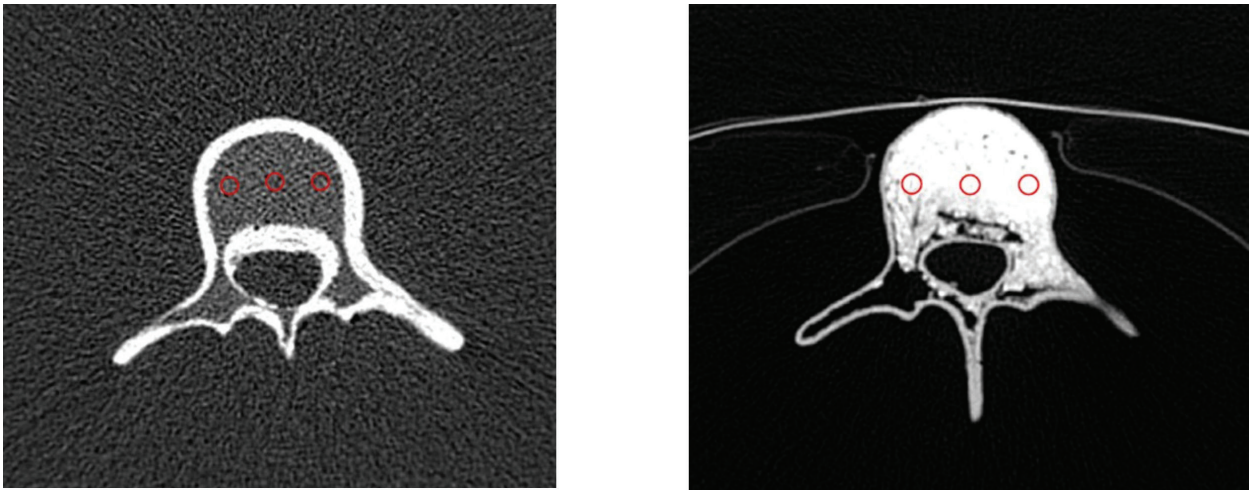


Figure 3. *Example of three ROIs, each measuring 7.34 mm^2, placed at different locations. Left: within vertebral bodies of the RSD RS-113T pelvic phantom, Right: 3D-printed lumbar spine phantom.*



Figure 4. *GE Lunar iDXA QC phantom used as the reference phantom for BMD validation.*

Data analysis

To evaluate the radiological characteristics of the developed 3D-printed lumbar spine phantom in comparison with the RSD RS-113T pelvic phantom, descriptive statistics were calculated and are presented as mean±standard deviation (SD). Agreement between the two phantoms was assessed using Bland-Altman analysis to evaluate systematic bias and the limits of agreement for the measured parameters. All statistical analyses were performed using JASP software (version 0.19.6; JASP Team, Amsterdam, The Netherlands), an open-source statistical analysis package.

To validate the BMD performance of the developed 3D-printed phantom, the coefficient of variation (CV) was calculated from five repeated BMD measurements and compared with the manufacturer's recommended precision criteria (≤2%).

Results

Table 1 summarizes the repeatability of CT number and BMD measurements obtained from five repeated scans of the RSD RS-113T pelvic phantom. The phantom exhibited mean CT number and BMD values of 752.94 HU and 1.80 gm/cm^2 (L1-L5), respectively. These values served as the reference for subsequent Bland-Altman agreement analysis with the developed 3D-printed phantoms. In the first phase of block sample testing, cuboid samples measuring 15×15×21 mm were printed using PETG and PLA filaments with infill densities of 80% and 100%, utilizing Gyroid and Crosshatch infill patterns. DXA measurements revealed that these samples did not generate sufficient signal intensity for reliable BMD calculation, necessitating redesign (Supplementary Table 1).

Table 1. CT number and BMD results obtained from five repeated scans of the RSD RS-113T pelvic phantom.

RSD RS-113T pelvic phantom (Scan number)	**CT number** (HU)				
	L1	**L2**	**L3**	**L4**	**L5**
1	733.10	787.40	755.30	793.80	695.50
2	737.92	789.52	761.97	795.41	702.30
3	740.20	785.60	750.80	788.47	703.60
4	736.05	792.44	745.32	789.67	696.30
5	730.60	784.70	751.19	780.50	695.85
Min	730.60	784.70	745.32	780.50	695.50
Max	740.20	792.44	761.97	795.41	703.60
Average	735.57	787.93	752.92	789.57	698.71
SD	3.41	2.79	5.53	5.21	3.50
%CV	0.46	0.35	0.73	0.66	0.50
RSD RS-113T pelvic phantom (Scan number)	**BMD** (gm/cm^2)				
	L1	**L2**	**L3**	**L4**	**L5**
1	1.74	1.94	1.88	1.74	1.81
2	1.73	1.91	1.87	1.72	1.80
3	1.75	1.92	1.84	1.73	1.77
4	1.71	1.90	1.82	1.69	1.75
5	1.74	1.93	1.85	1.71	1.80
Min	1.71	1.90	1.82	1.69	1.75
Max	1.75	1.94	1.88	1.74	1.81
Average	1.73	1.92	1.85	1.72	1.79
SD	0.01	0.01	0.02	0.02	0.02
%CV	0.58	0.52	1.08	1.16	1.12

As a result of the second phase, block samples were redesigned as box-shaped phantoms with lids, maintaining dimensions consistent with the L3 vertebral body. The lids and edges were printed at 100% infill density. Samples were fabricated using PETG and PLA with infill densities of 0%, 20%, and 40%, printed using Grid, Crosshatch, and Gyroid patterns. Samples were weighed before and after filling with $CaCO_3$ powder to assess filling efficiency. Weight measurements (Supplementary Table 2) confirmed a significant mass increase following $CaCO_3$ addition, which correlated with enhanced CT number and BMD signals. In Table 2, CT and DXA measurements showed comparable average CT numbers for PETG and PLA samples containing $CaCO_3$. However, the PETG sample printed at 0% infill demonstrated an average BMD of 1.69 gm/cm^2, closely approximating the RSD RS-113T pelvic phantom's BMD of 1.80 gm/cm^2. Therefore, PETG was selected as the optimal material for fabricating the lumbar spine phantom.

Table 2. CT number and BMD results from block sample tests conducted in the second phase.

3D printed block % Infill & Pattern	**PETG**		**PLA**	
	CT number (HU)	**BMD** (gm/cm^2)	**CT number** (HU)	**BMD** (gm/cm^2)
20% Grid	999.00	1.22	1182.00	0.96
20% Crosshatch	1231.20	1.00	1041.50	0.74
20% Gyroid	1094.90	0.86	955.30	0.72
40% Grid	748.60	0.80	839.50	0.56
40% Crosshatch	961.10	0.75	985.30	0.66
40% Gyroid	860.50	0.67	974.80	0.59
0%	1691.40	1.69	1603.70	0.97

Testing of the 3D-printed lumbar spine phantoms using PETG filaments with densities of 1.24 and 1.27 gm/cm^3 showed that the sample with 1.24 gm/cm^3 density had mean CT number and BMD values of 1061.70 HU and 0.96 gm/cm^2 (L1-L5), respectively. While structurally accurate, this phantom did not adequately replicate X-ray attenuation compared to the pelvic phantom. In contrast, the phantom printed at 1.27 gm/cm^3 exhibited mean CT number and BMD values of 744.26 HU and 1.58 gm/cm^2 (L1-L5), closely matching the pelvic phantom in both structural and radiological properties (Figure 5).

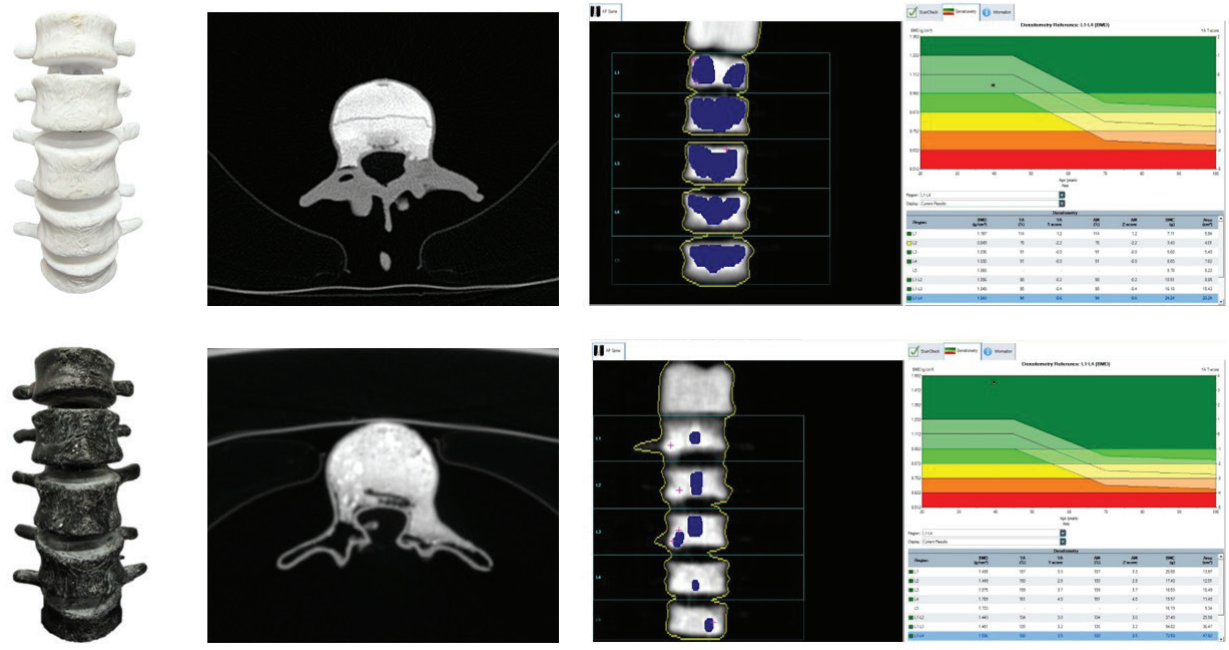


Figure 5. *The 3D-printed lumbar spine phantom fabricated using PETG material with calcium carbonate ($CaCO_3$) powder added to simulate bone mineral content, showing CT and BMD measurements.* Top: white indicates a density of 1.24 gm/cm^3, Bottom: black indicates a density of 1.27 gm/cm^3.

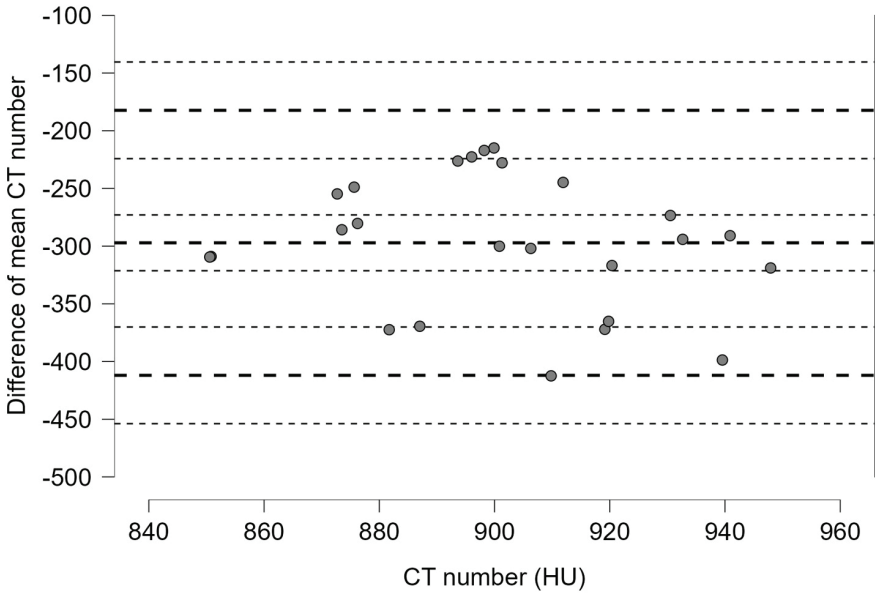
The repeatability of CT number and BMD measurements, derived from five repeated scans of the 3D-printed lumbar spine phantom fabricated from PETG with $CaCO_3$ powder added to simulate bone mineral content, is summarized in Tables 3 and 4. Bland-Altman analysis was performed to evaluate the agreement of CT number and BMD measurements between the developed 3D-printed phantom and the RSD RS-113T pelvic phantom (Figure 6 and Figure 7).

Table 3. Repeatability of CT number measurements obtained from five repeated scans of the 3D-printed lumbar spine phantom fabricated using PETG (white, 1.24 gm/cm^3, IEMAI, China) and PETG (black, 1.27 gm/cm^3, eSUN, China) containing calcium carbonate powder to simulate bone mineral content.

3D-printed PETG "White" filament density = 1.24 gm/cm^3 (Scan Number)	**CT number** (HU)				
	L1	**L2**	**L3**	**L4**	**L5**
1	1105.20	1102.40	1138.90	1016.40	1083.00
2	1015.22	1034.29	1079.73	1007.37	1015.86
3	1057.29	1078.80	1050.90	1000.07	1090.36
4	1067.25	1086.37	1107.39	1006.75	1058.12
5	1067.93	1071.75	1116.04	1005.29	1079.83
Min	1015.22	1034.29	1050.90	1000.07	1015.86
Max	1105.20	1102.40	1138.90	1016.40	1090.36
Average	1062.58	1074.72	1098.59	1007.18	1065.43
SD	28.77	22.64	30.45	5.28	27.01
%CV	2.71	2.11	2.77	0.52	2.54
3D-printed PETG "Black" filament density = 1.27 gm/cm^3 (Scan number)	**CT number** (HU)				
	L1	**L2**	**L3**	**L4**	**L5**
1	767.25	793.42	724.45	758.70	784.44
2	721.07	722.92	706.43	746.79	771.20
3	725.70	748.54	734.28	750.35	776.95
4	763.93	733.84	728.79	754.53	725.05
5	727.01	752.51	719.06	735.04	734.22
Min	721.07	722.92	706.43	735.04	725.05
Max	767.25	793.42	734.28	758.70	784.44
Average	740.99	750.25	722.60	749.08	758.37
SD	20.21	24.03	9.51	8.08	24.01
%CV	2.73	3.20	1.32	1.08	3.17

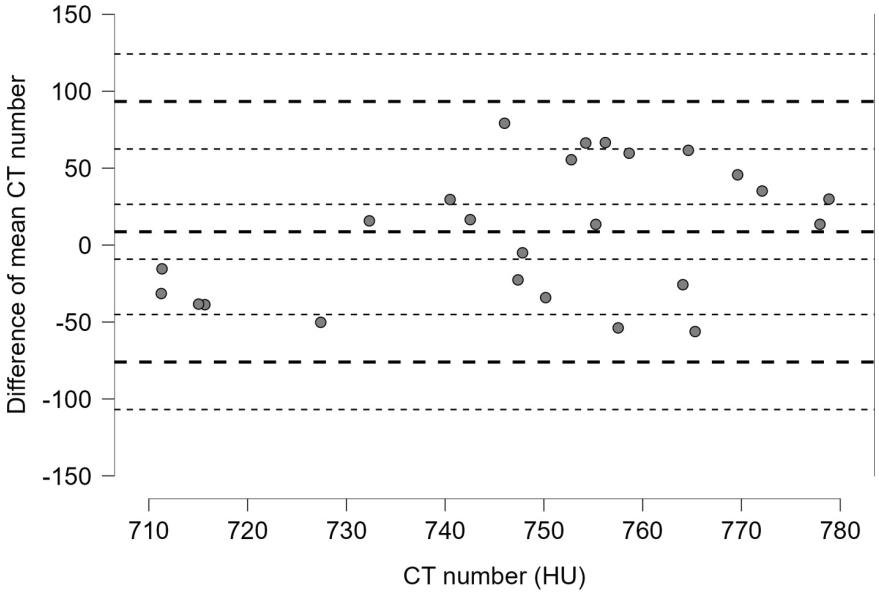
Table 4. Repeatability of BMD measurements obtained from five repeated scans of the 3D-printed lumbar spine phantom fabricated using PETG (white, 1.24 gm/cm^3, IEMAI, China) and PETG (black, 1.27 gm/cm^3, eSUN, China) containing calcium carbonate powder to simulate bone mineral content.

3D-printed PETG "White" filament density = 1.24 gm/cm^3 (Scan Number)	**BMD** (gm/cm^2)				
	L1	**L2**	**L3**	**L4**	**L5**
1	1.20	0.85	1.04	1.03	1.07
2	1.16	0.69	1.01	0.96	0.87
3	1.13	0.68	0.93	1.02	0.83
4	1.14	0.75	1.03	0.93	0.95
5	1.15	0.75	0.97	0.98	0.89
Min	1.13	0.68	0.93	0.93	0.83
Max	1.20	0.85	1.04	1.03	1.07
Average	1.16	0.74	1.00	0.98	0.92
SD	0.02	0.06	0.04	0.04	0.08
%CV	1.72	8.06	4.00	4.08	8.70
3D-printed PETG "Black" filament density = 1.27 gm/cm^3 (Scan Number)	**BMD** (gm/cm^2)				
	L1	**L2**	**L3**	**L4**	**L5**
1	1.45	1.45	1.60	1.71	1.71
2	1.39	1.49	1.62	1.67	1.72
3	1.44	1.41	1.63	1.72	1.72
4	1.49	1.46	1.53	1.64	1.75
5	1.44	1.42	1.55	1.79	1.74
Min	1.39	1.41	1.53	1.64	1.71
Max	1.49	1.49	1.63	1.79	1.75
Average	1.44	1.45	1.59	1.71	1.73
SD	0.03	0.03	0.04	0.05	0.01
%CV	2.08	2.07	2.52	2.92	0.58



3D-printed PETG (white) phantom and RSD RS-113T pelvic phantom

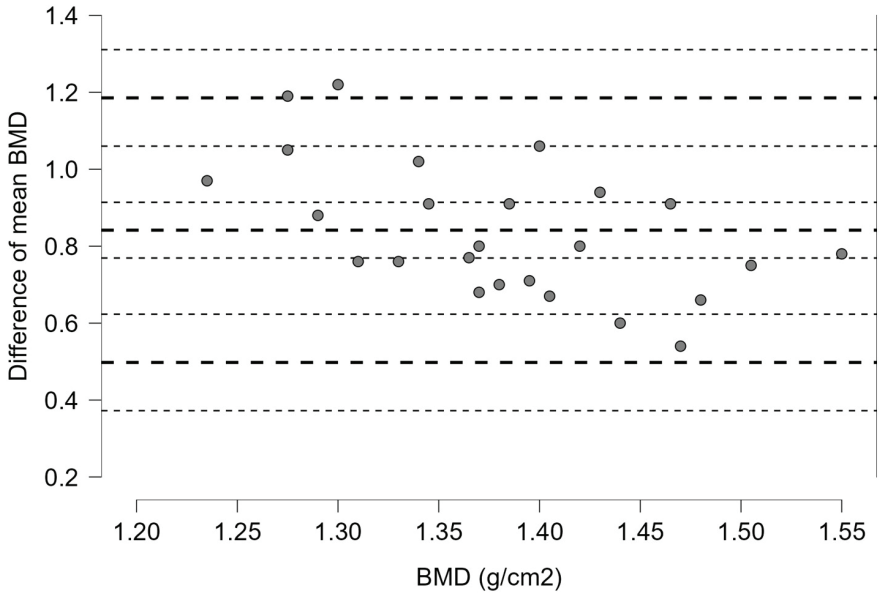
Bias & Limits	Point Value	Lower 95% CI	Upper 95% CI
Mean difference + 1.96 SD	-182.30	-224.20	-140.40
Mean difference	-297.10	-321.30	-273.00
Mean difference - 1.96 SD	-412.00	-453.90	-370.10



3D-printed PETG (black) phantom and RSD RS-113T pelvic phantom

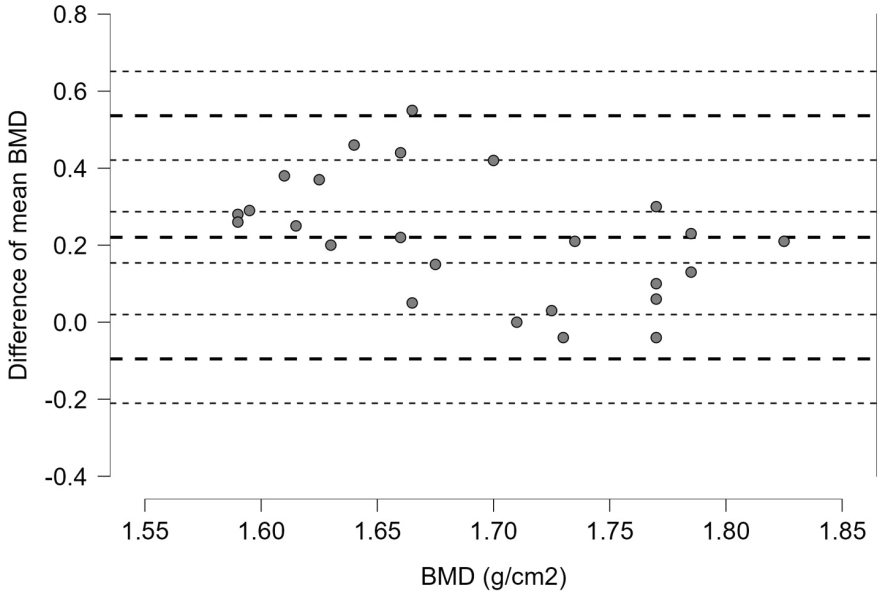
Bias & Limits	Point Value	Lower 95% CI	Upper 95% CI
Mean difference + 1.96 SD	93.32	62.43	124.20
Mean difference	8.65	-9.18	26.48
Mean difference - 1.96 SD	-76.01	-106.90	-45.13

Figure 6. *Bland-Altman plots showing the agreement in mean CT numbers. Top: 3D-printed PETG "white" versus RSD RS-113T phantom, Bottom: 3D-printed PETG "black" versus RSD RS-113T phantom.*



3D-printed PETG (white) phantom and RSD RS-113T pelvic phantom

Bias & Limits	Point Value	Lower 95% CI	Upper 95% CI
Mean difference + 1.96 SD	1.185	1.060	1.311
Mean difference	0.842	0.769	0.914
Mean difference - 1.96 SD	0.498	0.372	0.623



3D-printed PETG (black) phantom and RSD RS-113T pelvic phantom

Bias & Limits	Point Value	Lower 95% CI	Upper 95% CI
Mean difference + 1.96 SD	0.536	0.421	0.651
Mean difference	0.220	0.154	0.287
Mean difference - 1.96 SD	-0.095	-0.210	0.020

Figure 7. Bland-Altman plots showing the agreement in mean BMD values. Top: 3D-printed PETG "white" versus RSD RS-113T phantom, Bottom: 3D-printed PETG "black" versus RSD RS-113T phantom.

For CT number measurements (Figure 6), the developed 3D-printed PETG "white" phantom showed a mean difference of -297.10 HU, with 95% limits of agreement ranging from -412.00 HU to -182.30 HU. In contrast, the developed 3D-printed PETG "black" phantom showed a substantially smaller mean difference of 8.65 HU, with 95% limits of agreement ranging from -76.01 HU to 93.32 HU. Similarly, Bland-Altman analysis of BMD measurements (Figure 7) demonstrated greater agreement between the developed 3D-printed PETG "black" phantom and the standard RSD phantom than between the developed 3D-printed PETG "white" phantom and the standard RSD phantom. The developed 3D-printed PETG "white" phantom showed a mean difference of 0.84 gm/cm^2, with 95% limits of agreement ranging from 0.50 gm/cm^2 to 1.18 gm/cm^2. By comparison, the developed 3D-printed PETG "black" phantom showed a mean difference of 0.22 gm/cm^2, with 95% limits of agreement ranging from -0.09 gm/cm^2 to 0.54 gm/cm^2. Overall, the developed 3D-printed PETG "black" phantom demonstrated closer agreement with the standard RSD RS-113T pelvic phantom than the "white" phantom, demonstrating smaller systematic bias and narrower limits of agreement for both CT number and BMD measurements.

Finally, developed PETG 3D-printed phantoms were validated using the GE Lunar iDXA QC phantom. In terms of repeatability, the GE Lunar iDXA QC phantom exhibited a coefficient of variation (CV) of less than 2% for BMD measurements (Table 5). The PETG black phantom demonstrated comparable repeatability, with BMD CV values ranging from 0.58% to 2.92%, whereas the PETG white phantom showed consistently higher variability (CV: 1.72%-8.70%).

Table 5. Repeatability of BMD measurements obtained from five repeated scans of GE Lunar iDXA QC phantom.

GE Lunar iDXA QC phantom (Scan Number)	**BMD** (gm/cm^2)
1	1.501
2	1.501
3	1.500
4	1.501
5	1.502
Min	1.500
Max	1.502
Average	1.501
SD	0.001
%CV	0.047

Discussion

This study focused on developing a 3D-printed lumbar spine phantom for X-ray-based imaging applications. Using FDM 3D printing technology, the phantom required approximately 6 to 9 hours. The average CT number and BMD values of the 3D-printed phantom closely corresponded to those of the RSD RS-113T pelvic phantom. Notably, PETG filament with a density of 1.27 gm/cm^3 produced the developed phantoms with mean CT number and BMD values of 744.26 HU and 1.58 gm/cm^2, respectively, closely matching the pelvic phantom's values of 752.94 HU and 1.80 gm/cm^2. This demonstrates the effectiveness of the developed phantoms in simulating both structural and radiological properties. These findings support the feasibility of tuning phantom radiological characteristics through material selection and density control, thereby enabling the fabrication of 3D-printed phantoms that more closely replicate human bone for CT number and DXA measurements.

3D printing offers several advantages, including the ability to fabricate complex anatomical geometries, tune internal density, and produce customized phantoms through a relatively straightforward manufacturing process. These capabilities make 3D printing particularly well suited for designing imaging phantoms for radiological testing. Supporting this approach, previous studies have demonstrated the feasibility of using 3D-printed phantoms for quality assurance and performance evaluation in CT,[15] and X-ray radiography by carefully tailoring material composition16 and internal structure to achieve clinically relevant attenuation, CT number values,[17] and BMD ranges.[18,19]

Based on the findings from the first phase of block sample testing, cuboid samples (15×15×21 mm) were fabricated using PETG and PLA filaments at infill densities of 80% and 100%, employing Gyroid and Crosshatch infill patterns. DXA measurements demonstrated that these samples produced insufficient signal intensity for reliable BMD quantification. This was attributed to the relatively low density of the filament materials, which limited DXA's ability to detect and measure bone-equivalent signals. Therefore, the sample designs were modified to include $CaCO_3$

powder with an infill density of 0% to increase X-ray absorption and improve BMD detectability. The results showed that PETG- and PLA-based samples containing $CaCO_3$ exhibited comparable mean CT numbers; however, the PETG sample fabricated with 0% infill achieved an average BMD of 1.69 gm/cm^2 (Table 2), closely approximating the pelvic phantom's BMD of 1.80 gm/cm^2. These findings are consistent with previous research by Badiuk *et al.*,[20] who reported that adding $CaCO_3$ and strontium carbonate into resin increased CT numbers up to approximately 1000-1500 HU for bone modeling. While Badiuk's study employed resin-based printing, this work utilized PETG filament. Although PETG is more expensive and more challenging to print than PLA, its superior mechanical strength and flexibility justified its selection, with a 0% infill density chosen for the fabrication of the 3D-printed phantom.

Limitations

This study has several limitations. Firstly, the amount of $CaCO_3$ powder incorporated during the 3D printing process could not be precisely controlled, which may have introduced variability in phantom composition and imaging properties. Secondly, the absence of surrounding soft tissue-equivalent materials may have affected CT attenuation and BMD accuracy. In addition, the two phantoms were scanned under different conditions; the RSD RS-113T phantom includes a soft tissue-equivalent surround, whereas the 3D-printed lumbar spine phantom was scanned in air. Given the influence of beam hardening, scatter, and attenuation from surrounding materials, these differences may have introduced systematic bias. Future work should incorporate soft tissue-equivalent materials or housing to better simulate clinical conditions and enable more direct comparison with anthropomorphic phantoms. Further development should also explore composite bone-mimicking filaments and alternative printing technologies to improve spatial resolution, material uniformity, and phantom homogeneity. Finally, further optimization and cross-vendor evaluation across different DXA systems are warranted to establish its broader applicability.

Conclusion

This study developed and validated a 3D-printed lumbar spine phantom using PETG filament (density 1.27 gm/cm^3) containing $CaCO_3$ powder. The phantom demonstrated radiological characteristics comparable to those of the commercial RSD RS-113T pelvic phantom and provided reliable BMD measurements when evaluated using the GE Lunar iDXA QC phantom, supporting its potential as an alternative phantom for DXA quality assurance and research applications. Further optimization and multicenter, cross-vendor validation across different DXA systems are warranted to establish its broader applicability.

Ethic approval

Ethical approval was not required for this study, as no human subjects were involved and no human data were collected.

Funding

This work was supported in part by the Faculty of Allied Health Sciences, Naresuan University in Thailand.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Availability of data

Data required for this study may be made available by the author(s) upon reasonable request.

CRediT authorship contribution statement

Thunyarat Chusin: conceptualization, review and edit, supervision; **Pakaphorn Kongfun**: methodology, data collection, data analysis; **Suchanya Chanapim**: methodology, data collection, data analysis; **Suphichaya Thanashappokin**: methodology, data collection, data analysis; **Thanyawee Pengpan**: conceptualization, review and edit, supervision; **Kingkarn Aphiwatthanasumet**: conceptualization, methodology, original draft, review and edit, supervision. All authors jointly discussed the results and approved the final version of the manuscript.

Generative AI and AI-assisted use

Not applicable

Acknowledgements

The authors would like to thank the Department of Radiology, PRINC Pitsanuvej Hospital, Phitsanulok, for providing CT imaging facilities, and the radiographers for their assistance with data collection.

References

[1] Filippou V, Tsoumpas C. Recent advances on the development of phantoms using 3D printing for imaging with CT, MRI, PET, SPECT, and ultrasound. Med Phys. 2018; 45(9): e740-60. doi: 10.1002/mp.13058.

[2] Okkalidis F, Chatzigeorgiou C, Okkalidis N, Dukov N, Milev M, Bliznakov Z, et al. Characterization of commercial and custom-made printing filament materials for computed tomography imaging of radiological phantoms. Technologies. 2024; 12(8): 139. doi: 10.3390/technologies12080139.

[3] Pullen MW, Pooley RA, Kofler JM, Valero-Moreno F, Ramos-Fresnedo A, Domingo RA, et al. A radiographic analysis of common 3D print materials and assessment of their fidelity within vertebral models. Ann 3D Print Med. 2022; 8: 100080. doi: 10.1016/j.stlm.2022.100080.

[4] Kim SY, Park JW, Park J, Yea JW, Oh SA. Fabrication of 3D printed head phantom using plaster mixed with polylactic acid powder for patient-specific QA in intensity-modulated radiotherapy. Sci Rep. 2022; 12: 17500. doi: 10.1038/s41598-022-22520-6.

[5] Hatamikia S, Kronreif G, Unger A, Oberoi G, Jaksa L, Unger E, et al. 3D printed patient-specific thorax phantom with realistic heterogenous bone radiopacity using filament printer technology. Z Med Phys. 2022; 32(4): 438-52. doi: 10.1016/j.zemedi.2022.02.001.

[6] Cochran BJ, Trajanovska S, Sunn N, Rye KA, Ryder WJ, editors. Development of 3D printing filaments for bespoke X-Ray phantoms. In: IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC); 2018 Nov 10-17; Sydney, NSW, Australia. IEEE; 2018.

[7] Kozee M, Weygand J, Andreozzi JM, Hunt D, Perez BA, Graham JA, et al. Methodology for computed tomography characterization of commercially available 3D printing materials for use in radiology/radiation oncology. J Appl Clin Med Phys. 2023; 24(6): e13999. doi: 10.1002/acm2.13999.

[8] Sangondimath G, Sen RK, T FR. DEXA and imaging in osteoporosis. Indian J Orthop. 2023; 57(Suppl 1): 82-93. doi: 10.1007/s43465-023-01059-2.

[9] Slart RHJA, Punda M, Ali DS, Bazzocchi A, Bock O, Camacho P, et al. Updated practice guideline for dual-energy X-ray absorptiometry (DXA). Eur J Nucl Med Mol Imaging. 2024; 52: 539-63. doi: 10.1007/s00259-024-06912-6.

[10] Frimeth J, Galiano E, Webster D. Some physical and clinical factors influencing the measurement of precision error, least significant change, and bone mineral density in dual-energy x-ray absorptiometry. J Clin Densitom. 2010; 13(1): 29-35. doi: 10.1016/j.jocd.2009.09.004.

[11] Lewiecki EM, Binkley N, Morgan SL, Shuhart CR, Camargos BM, Carey JJ, et al. Best practices for dual-energy x-ray absorptiometry measurement and reporting: international society for clinical densitometry guidance. J Clin Densitom. 2016; 19(2): 127-40. doi: 10.1016/j.jocd.2016.03.003.

[12] Kim HS, Yang SO. Quality control of DXA system and precision test of radio-technologists. J Bone Metab. 2014; 21(1): 2-7. doi: 10.11005/jbm.2014.21.1.2.

[13] Mansour Z, Mokhtar A, Sarhan A, Ahmed MT, El-Diasty T. Quality control of CT image using American College of Radiology (ACR) phantom. Egypt J Radiol Nucl Med. 2016; 47(4): 1665-71. doi: 10.1016/j.ejrnm.2016.08.016.

[14] Li Y, Jiang Y, Liu H, Yu X, Chen S, Ma D, et al. A phantom study comparing low-dose CT physical image quality from five different CT scanners. Quant Imaging Med Surg. 2022; 12(1): 766-80. doi: 10.21037/qims-21-245.

[15] Hofmann T, Buschmann M, Homolka P. X-ray attenuation properties of additive manufacturing and 3d printing materials for mimicking tissues in radiographic phantoms measured by CT from 70 to 140 kV: 2025 update. Biomimetics (Basel). 2026; 11(3): 202. doi: 10.3390/biomimetics11030202.

[16] Ahmed AM, Buschmann M, Breyer L, Kuntner C, Homolka P. Tailoring the mass density of 3D printing materials for accurate X-ray imaging simulation by controlled underfilling for radiographic phantoms. Polymers. 2024; 16(8): 1116. doi: 10.3390/polym16081116.

[17] Jusufbegović M, Pandžić A, Šehić A, Jašić R, Julardžija F, Vegar-Zubović S, et al. Computed tomography tissue equivalence of 3D printing materials. Radiography (Lond). 2022; 28(3): 788-92. doi: 10.1016/j.radi.2022.02.008.

[18] Mei K, Pasyar P, Geagan M, Liu LP, Shapira N, Gang GJ, et al. Design and fabrication of 3D-printed patient-specific soft tissue and bone phantoms for CT imaging. Sci Rep. 2023; 13(1): 17495. doi: 10.1038/s41598-023-44602-9.

[19] Petraikin AV, Mikhailova AM, Kudryavtsev ND, Cherkasskaya MV, Yastrebova VO, Omelyanskaya OV, et al. Development of L1 vertebral anthropomorphic model for densitometric phantom improvement. Sovrem Tekhnologii Med. 2025; 17(4): 52-7. doi: 10.17691/stm2025.17.4.05 (in Russian).

[20] Badiuk SR, Sasaki DK, Rickey DW. An anthropomorphic maxillofacial phantom using 3-dimensional printing, polyurethane rubber and epoxy resin for dental imaging and dosimetry. Dentomaxillofac Radiol. 2022; 51(1): 20200323. doi: 10.1259/dmfr.20200323.

Journal of Associated Medical Sciences 2026; 59(3): 357-363

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



A review article: DIRFloortime® on improving social communication skills in Thai children with autism spectrum disorder

Natwipa Wanicharoen

Communication Sciences and Disorders Division, Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 17 February 2026
Accepted as revised 27 July 2026
Available online 7 August 2026

Keywords:
DIRFloortime®, social communication, autism spectrum disorder

ABSTRACT

Background: Autism Spectrum Disorder (ASD) is characterized by persistent deficits in social communication and interaction, necessitating early, developmentally appropriate intervention. The Developmental, Individual-differences, Relationship-based model (DIRFloortime®) is a relationship-centered, child-directed approach designed to strengthen foundational capacities such as engagement, reciprocal communication, and symbolic thinking. Although increasingly implemented in Thailand, its empirical evidence base remains limited.

Objectives: This review aimed to systematically synthesize existing Thai research examining the effects of DIRFloortime® on social communication outcomes in children with ASD and to explore its practical implications within the Thai context.

Materials and methods: A structured literature search was conducted across Scopus, Thai Journal Online (ThaiJO), Google Scholar, and reference lists of relevant publications. Empirical studies published in English or Thai that evaluated DIRFloortime® interventions in Thai children with ASD were included.

Results: Across studies, DIRFloortime® demonstrated consistent positive effects on functional emotional development, engagement, reciprocal interaction, and foundational social communication skills. Reductions in autism symptom severity and improvements in language-related outcomes were observed particularly in longer-duration interventions. Parental fidelity and caregiver involvement emerged as significant predictors of treatment outcomes.

Conclusion: Current evidence suggests that DIRFloortime® is a culturally congruent and developmentally meaningful intervention for Thai children with ASD. While methodological limitations remain, findings support its potential as a caregiver-mediated approach to enhancing social communication. Further large-scale, high-quality studies are warranted to strengthen the evidence base and inform national implementation strategies.

* **Corresponding contributor.**
Author's Address: *Communication Sciences and Disorders Division, Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: natwipa.w@cmu.ac.th
doi: 10.66285/JAMS.2026.108
E-ISSN: 2539-6056

Introduction

Autism spectrum disorder (ASD) is characterized by persistent deficits in social communication and interaction, along with restricted and repetitive behaviors and atypical sensory responses.[1] Impairments in social communication are particularly significant, as they are associated with academic difficulties, behavioral challenges, and long-term relational problems.[2-4] Early intervention is therefore essential to improve developmental outcomes.[5]

Developmental interventions, grounded in cognitive[6,7] and social[8] constructivist theories, emphasize child-directed learning through meaningful caregiver engagement and target foundational skills such as shared attention, imitation, and reciprocal interaction.[9] The Developmental, Individual-differences, Relationship-based (DIR) model is a prominent relationship-based framework that promotes engagement, two-way communication, and higher-order thinking.[10] However, effective implementation requires skilled and responsive caregivers or professionals.[5]

In Thailand, evidence on DIRFloortime® remains limited.[11-14] This review synthesizes the available Thai research and examines its implications for speech-language pathologists (SLPs) and parent-mediated intervention.

Autism spectrum disorder (ASD)

Autism spectrum disorder is a neurodevelopmental condition characterized by persistent deficits in social communication and social interaction, accompanied by restricted and repetitive patterns of behavior, circumscribed interests, and/or atypical sensory responses. These features emerge in early childhood and persist across development, presenting with varying degrees of severity and functional impact among individuals.[1]

Although the presence of impairments in both core domains-social communication deficits and restricted, repetitive behaviors-is required for a diagnosis of ASD, difficulties in social communication warrant particular attention. Social communication encompasses the sharing of information, thoughts, feelings, and intentions through verbal and nonverbal means within socially meaningful contexts. It involves not only the transmission of messages but also the reciprocal exchange and mutual understanding that underpin effective interpersonal interaction.[2] Deficits in this domain are strongly associated with a wide range of adverse developmental outcomes, including diminished academic achievement, increased incidence of problem behaviors, and persistent challenges in establishing and maintaining meaningful interpersonal relationships.[2-4]

A recent meta-analysis by Issac *et al*.[15] identified 66 epidemiological studies examining the prevalence of ASD, encompassing a total sample of 21,313,061 children worldwide. Of these, 25 studies were conducted in Europe, 22 in Asia, and 13 in the Americas, with three studies each originating from Africa and Australia. The pooled global prevalence of ASD was estimated at 0.77% among children, with a higher prevalence observed in boys (1.14%). Australia demonstrated the highest pooled prevalence estimate (effect size = 2.18), underscoring the need for sustained public health attention and resource allocation in that region.[15] In Thailand, the prevalence of ASD among children aged 2-5 years in 2022 was estimated at 6 per 1,000, corresponding to 16,892 cases among 2,814,823 children nationwide.[16]

Early intervention is essential for children with ASD, as timely and appropriate support can significantly improve developmental trajectories. Numerous intervention programs have been developed for children with autism, each grounded in distinct theoretical and philosophical frameworks. These approaches broadly encompass behavioral interventions, cognitive-behavioral models, and developmental paradigms.[5] In the present review, we focus exclusively on a developmental intervention approach-specifically, the Developmental, Individual-differences, Relationship-based (DIR) model-as a framework for supporting children with ASD.

Developmental interventions

Developmental interventions are principally grounded in cognitive[6,7] and social[8] constructivist theories, which hypothesize that children actively construct knowledge and skills through exploration, interaction, and personal discovery. A defining feature of these interventions is their child-directed orientation.[9] Rather than imposing externally structured tasks, developmental approaches prioritize meaningful social engagement-particularly interactions with primary caregivers-as the primary context for learning. These interventions seek to strengthen foundational capacities that are often delayed or underdeveloped in children with ASD but are essential for broader developmental progress. Targeted competencies typically include initiating social interactions, sharing attention and interest (e.g., through gestures and joint attention), observing and imitating others' behaviors, and participating in reciprocal exchanges such as turn-taking during play and early conversational routines.[9]

Developmental interventions are frequently described as child-led, as they capitalize on the child's intrinsic motivation to communicate, and naturalistic, reflecting their implementation within everyday environments and routines. Common strategies include modeling language and actions, imitating the child's behavior to attribute communicative intent, and employing interactional techniques-such as strategic pausing and expectant waiting-as well as environmental arrangements, such as placing preferred objects slightly out of reach, to elicit communicative attempts.[9] Within these frameworks, caregivers are systematically trained to integrate such strategies into daily activities, thereby enhancing the frequency, duration, synchrony, and reciprocity of parent–child interactions.[9]

Although no universally endorsed set of principles has been formally delineated for developmental interventions, certain programs-such as Pediatric Autism Communication Therapy-illustrate foundational elements that are widely shared across developmental models. These typically include: 1) a developmental orientation, whereby intervention goals are organized according to a hierarchical progression of developmental stages; 2) a naturalistic mode of delivery, emphasizing authentic interactions within everyday caregiver-child contexts; and 3) a caregiver-mediated model, in which caregivers act as collaborative partners and are supported in building the skills, confidence, and autonomy necessary to

facilitate their child's ongoing learning and development.[17]

Among the most widely recognized developmental approaches is the Greenspan model, formally known as the Developmental, Individual-differences, Relationship-based (DIR) framework.[10]

Developmental, Individual-differences, Relationship-based (DIR)/Floortime (DIRFloortime®)

The DIRFloortime® approach was developed in 1989 by child psychiatrist Stanley Greenspan, MD, and child psychologist Serena Wieder, Ph.D. The Greenspan model is conceptualized as a relationship-based approach in which the primary objective is to cultivate the child's capacity to form meaningful interpersonal connections that serve as the foundation for cognitive and developmental growth.[10] The DIR model is organized around three core components: 1) functional emotional developmental capacities, 2) individual differences in sensory processing and biological functioning, and 3) relationships and interactions. Together, these elements provide a comprehensive framework for understanding and supporting the developmental trajectories of children with ASD.[18]

Within this framework, relational engagement is regarded as the driving force of development. By strengthening emotionally meaningful interactions, the model seeks to facilitate the acquisition of essential developmental capacities, including attention and self-regulation, engagement and relating, purposeful nonverbal communication, affective signaling, complex social problem-solving, symbolic communication, and ultimately abstract and logical thinking.[10] Collectively, these developmental milestones organize emerging mental capacities and underscore the central role of affect and emotional experience in shaping cognitive, social, and communicative advancement.[19]

Floortime represents the practical application of the DIR framework. It is a play-based, interactive intervention approach that emphasizes individual differences, child-centered interests, and affective engagement between the child and caregiver. Grounded in Greenspan's developmental theory, floortime involves shared play and learning experiences-often conducted on the floor-to promote reciprocal and emotionally attuned interactions. The approach posits that previously unmet developmental milestones may be achieved through intensive, child-directed play and sustained relational engagement with caregivers or other significant adults.[20]

The floortime process comprises five sequential steps: 1) observation, 2) opening circles of communication, 3) following the child's lead, 4) extending and expanding play, and 5) supporting the child in closing circles of communication. The process begins with careful observation to discern the child's affect, communicative style, and interests. The adult then initiates interaction in a manner that is attuned to the child's emotional state, thereby opening a communicative exchange. By following the child's lead, the caregiver fosters warmth, shared engagement, and relational connection. Play is subsequently expanded to stimulate creative thinking and emotional expression. Finally, the child completes the communicative loop through purposeful responses, enabling repeated reciprocal exchanges that reinforce understanding of two-way communication and promote socio-emotional development.[20]

The DIR model is highly dependent on the competence and responsiveness of the parent or professional implementing the intervention. Effective delivery requires accurate interpretation of the child's cues, timely recognition of opportunities for engagement, and contingent responses that extend and deepen interaction. Such demands call for advanced observational, reflective, and relational skills and may present challenges for maintaining consistent, high-fidelity implementation in community-based contexts.[5] Moreover, the DIR framework advocates for a comprehensive, multidisciplinary approach that integrates family support, school and home programs, biomedical considerations, and related therapies.[5] Coordinated intervention across diverse settings is intended to address the child's holistic needs and optimize developmental progress.[21]

Over the past decades, a growing body of research has examined the effects of the DIRFloortime® model on social communication outcomes among young and school-aged children with ASD in both clinical and home-based settings.[22-26] For example, a recent systematic review by Divya *et al*.[22] reported significant improvements across multiple functional domains in children who participated in Floortime interventions. Home-based Floortime was associated with gains in emotional functioning, communication abilities, and daily living skills, as well as enhanced parent-child interactions reported by mothers. The review further identified certain parental demographic factors as significant moderators of intervention outcomes and reported no adverse events for children or caregivers. Overall, these findings suggest that Floortime is a cost-effective, child-led approach that can be implemented early in development and may substantially support social and emotional growth when initiated promptly by healthcare professionals.[22]

In Thailand, empirical investigations examining the effects of DIRFloortime® on social communication outcomes among children with ASD remain limited.[11-14] In the present review, we synthesize the available Thai evidence on the impact of DIRFloortime® on social communication development. Additionally, we examine the practical implications of implementing this approach within the Thai sociocultural context, with particular emphasis on its relevance for Thai speech-language pathologists and the active involvement of parents in the intervention process.

DIRFloortime® effects on social communication skills in Thai children with ASD

A comprehensive literature search was conducted across electronic databases, including Scopus and

Thai Journal Online (ThaiJO). Additional searches were performed through Google Scholar and manual screening of reference lists from relevant articles. The search included studies published from 2010 through December 2025. Articles published in English or Thai were included. Search terms were combined using Boolean operators and included the following keywords: "DIR" OR "developmental individual-differences Relationship-based" OR "Floortime" AND "Autism" OR "autism spectrum disorder" OR "ASD" AND "Thailand" OR "Thai children" AND "social communication" OR "communication" OR "functional emotional development".

Four empirical studies conducted in Thailand collectively provide emerging evidence regarding the effects of the DIRFloortime® model on social communication outcomes in children with ASD.[11-14] Overall, the findings suggest that DIRFloortime® yields positive effects on functional emotional development and foundational social communication skills. These include a pilot randomized controlled trial (RCT),[11] a one-year prospective parent-training follow-up study,[12] a quasi-experimental hospital-based longitudinal study with 6-18 months of follow-up,[14] and a community-based pilot intervention.[13] Although these studies differ in methodological rigor, duration, and service context, they demonstrate convergent findings indicating improvements in social communication and functional emotional development.

The pilot RCT conducted by Pajareya and Nopmaneejumruslers[11] provides the strongest level of methodological evidence. Preschool children with ASD who received home-based DIRFloortime® (mean 15.2 hours per week for three months) in addition to routine care demonstrated significantly greater improvements in the Functional Emotional Assessment Scale (FEAS), Childhood Autism Rating Scale (CARS), and Functional Emotional Questionnaires compared with those receiving routine care alone. These findings indicate that DIRFloortime® confers additive benefits beyond standard behavioral services. Specifically, the study reported statistically significant gains in FEAS scores, reflecting advancement in engagement, purposeful emotional interaction, and social problem-solving capacities. These domains correspond to Greenspan's Functional Developmental Levels (FDL) 1-4, encompassing self-regulation, engagement, and reciprocal communication.[11]

The subsequent one-year prospective follow-up study by Pajareya and Nopmaneejumruslers[12] extended these findings. Sustained home-based implementation (mean 14.2 hours per week) resulted in statistically significant increases in FEAS and Functional Emotional Developmental Questionnaire (FEDQ) scores, as well as significant reductions in CARS scores ($p \leq 0.001$). Nearly half of the participants achieved gains of at least 1.5 FDL levels. Improvements in FEAS and FEDQ scores suggest enhanced shared attention, increased frequency of communicative exchanges ("circles of communication"), and improved affective reciprocity. Importantly, parental fidelity and baseline symptom severity were significant predictors of outcome, underscoring the central role of caregiver engagement.

The quasi-experimental hospital-based longitudinal study by Singsathorn *et al*.[14] demonstrated statistically significant improvements in speech, language, communication, sociability, and sensory-motor perception domains at 6-, 12-, and 18-month follow-up. The progressive nature of these improvements suggests a cumulative or dose-response effect and highlights the importance of sustained intervention. These findings indicate that early relational gains may consolidate into more stable communicative behaviors over time.

Similarly, the community-based pilot study by Yuwapoom *et al*.[13] implemented a holistic DIR-informed model over three months in a district hospital setting. Children with ASD demonstrated significant improvements on the Greenspan Social-Emotional Growth Chart (SEGC), particularly in two-way communication and emotional engagement. Although changes in standardized developmental skill measures (DSI) were less pronounced in the short term, the study also reported significant reductions in parental stress as measured by the Parenting Stress Index (PSI). These findings underscore both the relational impact of the intervention and its feasibility in resource-limited community settings.

Reductions in global autism symptom severity were most evident in studies employing longer durations and controlled designs. Pajareya and Nopmaneejumruslers[11] reported significant reductions in CARS scores in the intervention group compared with routine care, suggesting that enhancing emotional reciprocity and social engagement may attenuate observable autistic symptomatology. Their one-year follow-up study[12] further demonstrated statistically significant decreases in CARS scores ($p \leq 0.001$), with greater improvements associated with higher parental fidelity, supporting a dose-response relationship between consistent relational intervention and symptom reduction. Singsathorn *et al*.[14] likewise observed continued improvements in sociability and communication domains over time. Although Yuwapoom *et al*.[13] did not report CARS outcomes, improvements in SEGC scores and parent-reported interaction quality indirectly suggest reduced symptom severity in relational domains.

Language-related outcomes were less uniformly reported but reveal important trends. Singsathorn *et al*.[14] demonstrated significant improvements in speech and language domains across follow-up periods, suggesting that relational gains may scaffold structural language development over time. Pajareya and Nopmaneejumruslers reported functional emotional improvements associated with enhanced gestural and symbolic communication, indicating progression from affective reciprocity toward representational communication.[12] In contrast, Yuwapoom *et al*.[13] found limited short-term gains in standardized developmental skills (DSI) among children with ASD, despite improvements in emotional engagement. This discrepancy suggests that structural language development may require

longer intervention duration, greater intensity, or both.

Across all four studies, caregiver involvement emerged as a central mechanism of change. Both Pajareya and Nopmaneejumruslers emphasized structured parent training and home-based implementation, with parental fidelity significantly predicting outcomes.[11,12] Yuwapoom *et al*.[13] documented significant reductions in parental stress, suggesting that improved relational attunement benefits not only children but also caregiver well-being. Singsathorn *et al*.[14] similarly incorporated sustained caregiver participation within both clinic-based and home routines. These findings reinforce the theoretical premise of DIRFloortime®: relational attunement and emotionally meaningful interaction serve as primary drivers of social communication development, while reduced parental stress may further enhance interaction quality through a positive feedback mechanism.

However, it is important to note that, across all four studies, the intervention providers were not SLPs. In three studies, the DIRFloortime® intervention was implemented by physicians who had received certified training in DIRFloortime® approaches.[11-13] In the remaining study, the intervention was delivered by an occupational therapist; however, the authors did not specify whether formal certification in DIRFloortime® had been obtained.[14] This observation highlights the interdisciplinary nature of DIRFloortime® implementation within the Thai context and underscores the need for further professional training and competency standards across related disciplines.

In summary, evidence from the four Thai studies suggests that DIRFloortime® primarily enhances the relational and affective foundations of social communication in children with ASD, with reductions in symptom severity and language gains emerging as secondary outcomes following sustained intervention. Stronger methodological designs demonstrated more robust evidence of efficacy, while community-based studies supported the feasibility and cultural relevance of the approach in Thailand. Notably, none of the reviewed studies implemented the intervention through SLPs; instead, interventionists primarily consisted of physicians with DIRFloortime® certification and an occupational therapist. These findings highlight both the interdisciplinary nature of DIRFloortime® implementation and the potential role of Thai SLPs in expanding relationship-based, communication-focused services for children with ASD.

Practical implications for Thai speech-language pathologists

Although none of the four reviewed studies implemented the DIRFloortime® intervention through SLPs,[11-14] the DIRFloortime® framework nevertheless holds substantial implications for the evolving role of Thai SLPs. Its relationship-based and communication-centered principles align closely with the core competencies of the speech-language pathology profession, particularly in facilitating social communication, emotional engagement, and caregiver-mediated interaction in children with ASD. Rather than positioning the clinician solely as a direct service provider, this approach advances a coaching-oriented model in which SLPs support caregivers in cultivating emotionally attuned and reciprocal interactions with their children. In this capacity, SLPs enhance caregiver sensitivity to the child's communicative cues and facilitate affective scaffolding that promotes engagement and shared attention-core foundations of social communication.

Moreover, DIRFloortime® emphasizes integrating pragmatic and social-emotional communication goals within naturally occurring routines across home and community contexts. This orientation positions Thai SLPs to collaborate effectively with multidisciplinary teams in both hospital and community-based settings, thereby ensuring continuity of care and contextual relevance. Ongoing monitoring of intervention fidelity and developmental progression is essential to maintain implementation quality and optimize child outcomes. Importantly, this coaching-based model expands service accessibility, particularly in regions where intensive therapist-delivered programs are constrained by limited professional resources.

Parental involvement represents a central mechanism of change within the DIRFloortime® framework. Evidence demonstrating the predictive value of caregiver fidelity highlights the importance of structured parent education, systematic supervision, and regular follow-up. The establishment of peer-support networks may further strengthen caregiver confidence, mitigate stress, and promote sustained engagement. Critically, reframing play not merely as recreational activity but as purposeful therapeutic interaction can enhance parental commitment and support the consistent implementation necessary to achieve meaningful developmental gains.

Future directions

Findings from the four Thai studies collectively indicate that DIRFloortime® represents a feasible, culturally congruent, and developmentally meaningful intervention for children with ASD.[11-14] To translate these research findings into sustainable impact, future efforts must extend beyond efficacy studies toward policy advancement and systemic implementation.

Firstly, integration into the national early intervention framework is warranted. Given consistent evidence demonstrating improvements in functional emotional development and social communication, DIRFloortime® should be considered for inclusion within Thailand's national early childhood intervention policies. Policymakers may explore incorporating caregiver-mediated, relationship-based approaches alongside established behavioral models within services administered by the Ministry of Public Health and the Ministry of Education. Developing

national clinical guidelines that recognize DIRFloortime® as an evidence-informed intervention would expand access and diversify available treatment pathways.

Secondly, workforce development and capacity building are essential. A coordinated national training initiative targeting speech-language pathologists, developmental pediatricians, nurses, and community health workers should be established. As caregiver coaching and implementation fidelity are critical determinants of treatment outcomes, structured certification pathways and continuing professional development programs in DIRFloortime® are recommended. Integrating relevant training modules into university curricula for SLPs and allied health professionals would further ensure long-term sustainability.

Thirdly, expansion into primary and community healthcare systems is imperative. Evidence from community-based implementation demonstrates feasibility within district hospitals and primary care settings. Accordingly, policy initiatives should prioritize scaling DIR-informed parent coaching programs across community health centers, particularly in rural and underserved regions. Embedding such interventions within Thailand's primary healthcare infrastructure would help mitigate disparities in service access.

Fourthly, standardized parent training and support systems should be institutionalized. At the practice level, structured parent education programs, systematic follow-up mechanisms, and peer-support networks are critical to maintaining caregiver engagement. Developing nationally standardized manuals and culturally adapted Thai-language materials would promote consistency in implementation. Public awareness initiatives that frame play as purposeful therapeutic engagement may further enhance parental participation and reduce stigma.

Fifthly, sustainable funding mechanisms must be established. Caregiver-mediated models such as DIRFloortime® offer potential cost advantages in resource-constrained settings. Policymakers should consider reimbursement structures that support parent coaching sessions and home-based intervention, thereby incentivizing family-centered care. Economic evaluations are necessary to inform resource allocation within Thailand's universal health coverage system.

Finally, effective implementation will require intersectoral collaboration. Coordinated efforts across healthcare, education, and social welfare sectors are essential to ensure continuity of care. Integrated service models linking hospitals, early childhood centers, and special education services can promote consistent developmental support and strengthen long-term outcomes for children with ASD.

Conclusion

Although available evidence from Thai studies examining the DIRFloortime® model and social communication outcomes remains limited, the DIRFloortime® model holds promise as a culturally congruent and developmentally grounded intervention for enhancing social communication in children with ASD. Across diverse service settings, DIRFloortime® consistently demonstrates positive effects on functional emotional development, reciprocal engagement, and foundational communicative capacities, with downstream benefits observed in symptom severity and language outcomes when intervention is sustained. The model's emphasis on caregiver-mediated, emotionally attuned interaction aligns well with Thailand's family-centered cultural context and supports the evolving role of speech-language pathologists as relational coaches. Although the current evidence base remains limited in scope and methodological rigor, findings collectively suggest that DIRFloortime® represents a feasible and meaningful framework for early intervention. Continued rigorous research and systematic implementation efforts are warranted to strengthen its empirical foundation and guide national policy and practice in Thailand.

Ethical approval

This is a review article, which did not directly involve human subjects. Therefore, ethical approval was not required for this paper.

Funding

There was no funding for this work.

Conflict of interest

The author declares no conflict of interest.

CRediT authorship contribution statement

Natwipa Wanicharoen: conceptualization, investigation, project administration, supervisor, writing: draft, review and edit, correspondence.

References

[1] American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. 5th ed. Washington (DC): American Psychiatric Association; 2013.

[2] Mundy P, Sigman M, Ungerer J, Sherman T. Defining the social deficits of autism: the contribution of non-verbal communication measures. J Child Psychol Psychiatry. 1986; 27(5): 657-69. doi: 10.1111/j.1469-7610.1986.tb00190.x.

[3] Koegel RL, Koegel LK, Surratt A. Language intervention and disruptive behavior in preschool children with autism. J Autism Dev Disord. 1992; 22(2): 141-53. doi: 10.1007/bf01058147.

[4] Bauminger N, Kasari C. Loneliness and friendship in high-functioning children with autism. Child Dev. 2000; 71(2): 447-56. doi: 10.1111/1467-8624.00156.

[5] Corsello CM. Early intervention in autism. Infants Young Child. 2005; 18(2): 74-85. doi: 10.1097/00001163-200504000-00002.

[6] Piaget J. Play, dreams and imitation in childhood. London: Routledge; 2013.
[7] Piaget J, Inhelder B. The psychology of the child. New York: Basic Books; 2008.
[8] Vygotsky LS. Mind in society: the development of higher psychological processes. Cambridge (MA): Harvard University Press; 1978.
[9] Sandbank M, Bottema-Beutel K, Crowley S, Cassidy M, Dunham K, Feldman JI, et al. Project AIM: autism intervention meta-analysis for studies of young children. Psychol Bull. 2020; 146(1): 1-29. doi: 10.1037/bul0000215.
[10] Greenspan SI, Wieder S. Developmental patterns and outcomes in infants and children with disorders in relating and communicating. J Dev Learn Disord. 1997; 1: 87-142.
[11] Pajareya K, Nopmaneejumruslers K. A pilot randomized controlled trial of DIR/Floortime™ parent training intervention for pre-school children with autistic spectrum disorders. Autism. 2011; 15(5): 563-77. doi: 10.1177/1362361310386502.
[12] Pajareya K, Nopmaneejumruslers K. A one-year prospective follow-up study of a DIR/Floortime™ parent training intervention for preschool children with autistic spectrum disorders. J Med Assoc Thai. 2012; 95(9): 1184-93.
[13] Yuwapoom W, Nopmaneejumruslers K, Jiawiwatkul A, Kleebpung N. A study of the holistic approach to child development according to the DIRFloortime concept: Banlat Hospital, Phetchaburi (in Thai). Int J Child Dev Ment Health. 2017; 5(1): 43-54.
[14] Singsathorn D, Suksong W, Permsirivanich W. Effects of development promotion based on DIR-Floortime techniques in autistic children, Queen Savang Vadhana Memorial Hospital (in Thai). J Phrapokklao Nurs Coll. 2016; 27(1): 104-22.
[15] Issac A, Halemani K, Shetty A, Thimmappa L, Vijay VR, Koni K, et al. The global prevalence of autism spectrum disorder in children: a systematic review and meta-analysis. Osong Public Health Res Perspect. 2025; 16(1): 3-15. doi: 10.24171/j.phrp.2024.0286.
[16] Health Data Center, Ministry of Public Health (TH). Percentage of autistic patients accessing services. 2023 [cited 2026 Feb 13]. Available from: hdcservice.moph.go.th.
[17] Aldred C, Green J, Adams C. A new social communication intervention for children with autism: pilot randomised controlled treatment study suggesting effectiveness. J Child Psychol Psychiatry. 2004; 45(8): 1420-30. doi: 10.1111/j.1469-7610.2004.00848.x.
[18] Wieder S, Greenspan SI. Climbing the symbolic ladder in the DIR model through floor time/interactive play. Autism. 2003; 7(4): 425-35. doi: 10.1177/1362361303007004008.
[19] Greenspan SI. DIRFloortime model. 2008. Available from: www.stanleygreenspan.com.
[20] de Boer SR, Griswold DE, Myles BS, Byrd SE, Ganz JB, Cook KT, et al. Autism spectrum disorders: interventions and treatments for children and youth. Thousand Oaks (CA): Corwin Press; 2004.
[21] Greenspan SI, Wieder S. Engaging autism: using the floortime approach to help children relate, communicate, and think. Cambridge (MA): Da Capo Lifelong Books; 2006.
[22] Dwidjantoro MI, Sukinah S. Effectiveness of DIR-Floortime intervention on communication skills development in elementary school students with autism spectrum disorder. J Innov Res Prim Educ. 2025; 4(4): 2547-58. doi: 10.56916/jirpe.v4i4. 2365.
[23] Divya KY, Begum F, John SE, Francis F. DIR Floortime in engaging autism: a systematic review. Iran J Nurs Midwifery Res. 2023; 28(2): 132-8. doi: 10.4103/ijnmr.ijnmr_272_21.
[24] Pacheco P, Pacheco M, Molini-Avejonas D. Study of 18 months of follow-up DIRFloortime intervention in preschool children with autism spectrum disorder. Eur Psychiatry. 2021; 64(S1): S503. doi: 10.1192/j.eurpsy.2021.1346.
[25] Liao ST, Hwang YS, Chen YJ, Lee P, Chen SJ, Lin LY. Home-based DIR/Floortime™ intervention program for preschool children with autism spectrum disorders: preliminary findings. Phys Occup Ther Pediatr. 2014; 34(4): 356-67. doi: 10.3109/01942638.2014.918074.
[26] Birkland MF. The effect of the DIRFloortime model on communication in children with autism [dissertation]. Pullman (WA): Washington State University; 2010.

Journal of Associated Medical Sciences 2026; 59(3): 364-376

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



Influence of Eu_2O_3 incorporation on the gamma-ray shielding performance of germanoborate glasses

Natthakridta Chanthima[1,2], Supakit Yonphan[1,2], Chalermpon Mutuwong[1,2], Wuttichai Chaiphaksa[1,2*]

[1]Physics program, Faculty of Science and Technology, Nakhon Pathom Rajabhat University, Nakhon Pathom Province, Thailand.
[2]Center of Excellence in Glass Technology and Materials Science (CEGM), Nakhon Pathom Rajabhat University, Nakhon Pathom Province, Thailand.

ARTICLE INFO

Article history:
Received 28 May 2026
Accepted as revised 7 August 2026
Available online 18 August 2026

Keywords:
Lead-free, europium oxide, radiation shielding

ABSTRACT

Background: The widespread use of ionizing radiation in medical and industrial applications has increased the demand for effective and lead-free radiation shielding materials. Rare earth doped glass systems have attracted attention due to their optical transparency and favorable attenuation properties.

Objectives: This study investigates the physical and gamma-ray shielding properties of Eu_2O_3-doped germanoborate glasses.

Materials and methods: Glasses with the composition $40GeO_2$: $25Gd_2O_3$: $(35-x)B_2O_3$: xEu_2O_3 (x=0, 1, 3, 5, 7 and 9 mol%) were prepared using the melt-quenching method. Density and molar volume were evaluated experimentally, while shielding parameters, including mass attenuation coefficients (μ_m), the effective atomic number (Z_{eff}), the effective electron density (N_{eff}), and half-value layers (HVL) and Pb equivalent thickness, were calculated using the WinXCom program.

Results: The density of the glasses increased with increasing Eu_2O_3 concentration, while the shielding parameters exhibited strong energy dependence according to dominant photon interaction mechanisms. Distinct K-absorption edges of Eu and Gd were observed near 0.04852 MeV and 0.05024 MeV, respectively. The incorporation of Eu_2O_3 slightly enhanced the attenuation parameters and reduced the HVL values, indicating improved gamma-ray shielding performance.

Conclusion: The developed Eu_2O_3-doped germanoborate glasses demonstrate promising physical and radiation shielding properties, suggesting their potential as transparent and lead-free shielding materials for medical and industrial applications.

Introduction

Ionizing radiation plays an essential role in modern medical applications, particularly in diagnostic imaging and radiation therapy. Nevertheless, continuous exposure to ionizing radiation may cause harmful biological effects in both patients and medical personnel, thereby increasing the demand for effective radiation shielding materials in healthcare environments.[1] The attenuation capability of a shielding material strongly depends on its density, elemental composition, and atomic number. Materials containing high atomic number elements generally exhibit superior photon attenuation

* **Corresponding contributor.**
Author's Address: *Physics program, Faculty of Science and Technology, Nakhon Pathom Rajabhat University, Nakhon Pathom Province, Thailand.*
E-mail address: wuttichai@webmail.npru.ac.th
doi: 10.66285/JAMS.2026.109
E-ISSN: 2539-6056

performance because they enhance the probability of photon interaction processes such as the photoelectric effect and Compton scattering.[2] Lead has long been utilized as a conventional radiation shielding material owing to its high density and excellent attenuation efficiency. However, the toxicity, heavy weight and environmental concerns associated with lead-based materials have motivated researchers to explore alternative eco-friendly shielding materials.[1]

Concrete is another widely used shielding material because of its low cost and good mechanical strength. Despite its effectiveness in radiation protection, the opaque nature of concrete limits its application in environments where optical transparency is required, particularly in medical and diagnostic facilities. Consequently, transparent glass materials have attracted considerable attention as promising candidates for radiation shielding applications due to their optical transparency, compositional flexibility, chemical durability, and ease of fabrication.[3] Significant attention has therefore been directed toward the development of lead-free materials for gamma-ray shielding applications.[4] Among various glass systems, borate glasses are recognized as excellent host materials because of their superior glass-forming ability, thermal stability, and capability to accommodate different modifier oxides without significant crystallization. In addition, borate glasses exhibit good transparency and can be tailored to achieve desirable physical and shielding properties through compositional modification.[5] The incorporation of GeO_2 into borate glass systems is known to improve structural connectivity, chemical stability, refractive index, and optical transparency. Owing to its relatively high density and low phonon energy, GeO_2 also enhances the physical and optical performance of germanoborate glasses, making them suitable for photonic and radiation shielding applications.[6]

For gamma-ray shielding applications, Gd_2O_3 is considered one of the most attractive heavy metal oxides because gadolinium possesses a high atomic number (Z=64) and high density, which significantly improve photon attenuation efficiency. Moreover, Gd_2O_3 incorporation can enhance radiation shielding performance while maintaining the optical transparency of the glass matrix.[7] In addition, Eu_2O_3 has recently attracted interest as a rare earth dopant due to its unique optical and luminescent properties. Europium ions can modify the structural arrangement of glass networks and contribute to photon interaction behavior owing to their relatively high atomic number (Z=63). Furthermore, Eu^{3+} doped glasses are well known for their strong luminescence characteristics, narrow emission bands, and potential applications in photonic and scintillation materials.

Therefore, the incorporation of Eu_2O_3 into germanoborate glass systems may provide additional functionality while simultaneously contributing to radiation attenuation performance.[8] Glasses doped with rare earth and heavy metal oxides have been widely investigated as lead-free gamma-ray shielding materials. However, most reported systems are based on borate, phosphate, tellurite or silicate hosts and employ only a small amount of a single high-Z rare-earth modifier.

For example, the Eu-, Te- and Pb-containing telluroborate glasses[9] and the Ce- and Gd-reinforced borosilicate glasses reported previously derive most of their attenuation from the heavy-metal or lead component of the host and incorporate the rare-earth oxide only as a minor dopant.[10] In contrast, the present work combines a germanoborate host with a dual rare-earth design, in which a fixed 25 mol% Gd_2O_3 is reinforced with variable Eu_2O_3 (0-9 mol%), so that the rare-earth oxides act as the principal attenuating constituents rather than as minor dopants. This configuration also introduces two closely spaced K-absorption edges (Eu≈48.52 keV and Gd≈50.24 keV) that extend the high-attenuation region within the diagnostic energy range, while keeping the glass lead-free and optically active via Eu^{3+} luminescence.

In the present study, Eu_2O_3-doped germanoborate glasses with the composition $40GeO_2$: $25Gd_2O_3$: $(35\text{-}x)B_2O_3$: xEu_2O_3 were prepared using the conventional melt-quenching technique. The physical and gamma-ray shielding properties of the prepared glasses were systematically investigated through density, molar volume, mass attenuation coefficient, effective atomic number, effective electron density, half-value layer, and lead equivalent thickness analyses. The photon interaction parameters were theoretically evaluated using the WinXCom program based on the NIST photon cross section database.[11]

Materials and methods

Glass sample preparation

The germanoborate glass system containing Eu_2O_3 was synthesized according to the composition $40GeO_2$: $25Gd_2O_3$: $(35\text{-}x)B_2O_3$: xEu_2O_3, where x=0, 1, 3, 5, 7, and 9 mol%, using the conventional melt-quenching method. High purity chemical reagents were employed throughout the preparation process. Each batch was prepared with a constant total weight of 20 gm and measured using a high-precision analytical balance with an accuracy of four decimal places. The prepared mixtures were transferred into covered alumina crucibles and heated in an electric furnace at 1,400 °C for 3 hours until complete melting and homogenization were achieved. Subsequently, the molten glasses were cast into graphite molds at ambient conditions and annealed at 500 °C for 3 hours to eliminate residual thermal stress before cooling gradually to room temperature. The fabricated glass samples were then shaped and polished into dimensions of approximately 1.0×1.5×0.3 cm for subsequent analyses. The measured physical parameters, including density and molar volume, are illustrated in Figure 1.

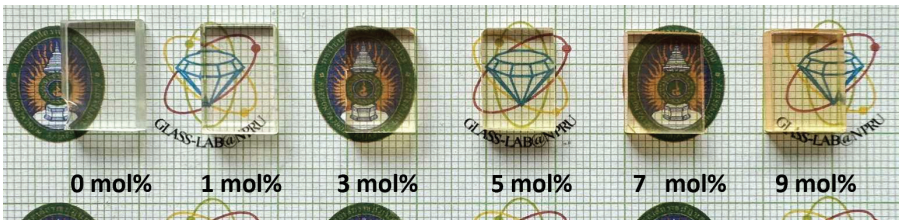


Figure 1. *The glass samples of the Eu_2O_3-doped germanoborate glasses.*

Measurement of physical properties

The physical characteristics of the prepared Eu_2O_3-doped germanoborate glasses were investigated through density and molar volume measurements. Density is considered a fundamental physical parameter in glass materials, as it provides important information regarding the structural compactness and atomic packing of the glass network. Furthermore, density plays a crucial role in determining the interaction probability between the material and incident gamma photons. In addition, the measured density values are required to determine several radiation shielding parameters, including the linear attenuation coefficient, effective atomic number, and electron density. The density (ρ) of the fabricated glass samples was experimentally determined at room temperature using the Archimedes method with distilled water as the immersion medium. All measurements were carried out using a high-resolution analytical balance (AND, HR-200). The density values were calculated using the following equation:

$$\rho = \left(\frac{W_{air}}{W_{air} - W_{water}}\right) \times \rho_w \quad (1)$$

Where and correspond to the weights of the sample measured in air and the immersion liquid, respectively, while represents the density of the buoyant medium.

The molar volume was determined from the experimentally measured density values and the corresponding molecular weights of the glass compositions. This parameter provides useful information regarding the structural arrangement, packing efficiency, and spatial distribution within the glass network. Variations in molar volume are commonly associated with changes in bond configuration, network connectivity, and the formation of bridging oxygen (BO) and non-bridging oxygen (NBO) species in the glass structure. Consequently, molar volume is recognized as an important parameter for correlating the structural characteristics of glass systems with their physical and radiation shielding properties.[12] The molar volume (V_m) of the prepared glasses was calculated using the following equation:

$$V_m = \frac{M}{\rho} \quad (2)$$

Where denotes the molecular weight of the glass composition and ρ represents the density of the corresponding glass sample.

Theoretical background

The attenuation of gamma rays passing through a material is governed by the exponential Beer-Lambert law, which describes the reduction in photon intensity as a function of material thickness and attenuation capability:

$$I = I_0 e^{-\mu x} \quad (3)$$

where I and I_0 represent the transmitted and incident photon intensities, respectively, μ is the linear attenuation coefficient (cm^{-1}), and x denotes the thickness of the absorber material (cm). The mass attenuation coefficient (μ_m) is obtained by normalizing μ to the density of the material, thereby providing a density-independent parameter for evaluating the gamma-ray attenuation capability of the investigated glasses:

$$\mu_m = \sum_i w_i \left(\frac{\mu}{\rho}\right)_i \quad (4)$$

When w_i represents the weight fraction of each constituent element, and μ_m corresponds to the mass attenuation coefficient of the individual elements within the glass composition. The theoretical attenuation parameters were calculated using the WinXCom program, which is based on the NIST photon cross-section database and provides photon interaction coefficients as a function of elemental composition and photon energy.[12]

The average atomic cross-section, (σ_a), can be evaluated from the mass attenuation coefficient using the following relation:

$$\sigma_a = \frac{\mu}{N_A \sum_i \frac{w_i}{A_i}} \quad (5)$$

Similarly, the average electronic cross-section (σ_e), can be determined using the following equation:

$$\sigma_e = \frac{1}{N_A} \sum_i \left(\sum_i \frac{f_i A_i}{Z_i} \right) w_i \quad (6)$$

where f_i and Z_i denote the fractional abundance and atomic number of the constituent element, respectively.

For composite glass materials, the effective atomic number (Z_{eff}) represents an equivalent atomic number that characterizes the material's photon interaction capability of the material.

$$Z_{eff} = \frac{\sigma_a}{\sigma_{el}} \quad (7)$$

The effective electron density (N_{eff}) defines the number of electrons per unit mass, which is related to the Z_{eff}, and is formulated as follows.[13]

$$N_{eff} = N_A \frac{Z_{eff}}{\sum_i \frac{w_i}{A_i}} \quad (8)$$

The half-value layer (HVL) is defined as the thickness of a shielding material required to reduce the incident gamma-ray intensity to one-half of its original value, and it can be calculated using the following equation[14]:

$$HVL = \frac{0.693}{\mu} \quad (9)$$

where μ is the linear attenuation coefficient. HVL is a practical parameter for assessing and comparing the shielding effectiveness of different materials.

The lead equivalent thickness represents the thickness of lead required to provide the same radiation attenuation performance as the investigated shielding material under identical exposure conditions. This parameter is widely used to compare the shielding effectiveness of alternative materials with conventional lead-based shields. The equivalent lead thickness can be estimated from the relationship between the linear attenuation coefficient and the thickness of the material for a specific photon energy, and is expressed by the following equation[15]:

$$d_{pb} = \frac{\mu_{material}}{\mu_{pb}} \times d_{material} \quad (10)$$

where μ_{pb} is the linear attenuation coefficient of lead, d_{pb} denotes the calculated lead equivalent thickness, $\mu_{material}$ represents the linear attenuation coefficient of the investigated material, and $d_{material}$ corresponds to the thickness of the shielding sample.

WinXCom program

WinXCom is a computational program developed based on the NIST XCOM photon cross-section database and is widely used to evaluate photon interaction parameters in shielding materials. The software provides theoretical mass attenuation coefficients together with partial photon interaction cross sections, including photoelectric absorption, Compton scattering, and pair production, over a broad photon energy range from approximately 1 keV to 100 GeV, as illustrated in Figure 2. By specifying the elemental composition and photon energy range, the program can generate attenuation-related parameters for compounds, mixtures, and glass systems with high computational efficiency.

In addition to mass attenuation coefficients, WinXCom can also be used to derive important shielding parameters such as the effective atomic number (Z_{eff}) and effective electron density (N_{eff}), which are commonly utilized to evaluate gamma-ray shielding performance. Owing to its reliability, rapid calculation capability, and broad applicability, WinXCom has become a standard tool extensively used in radiation shielding investigations.[16] It should be noted that the gamma-ray shielding parameters presented in this study were obtained from WinXCom calculations, without experimental validation through gamma-ray attenuation measurements.

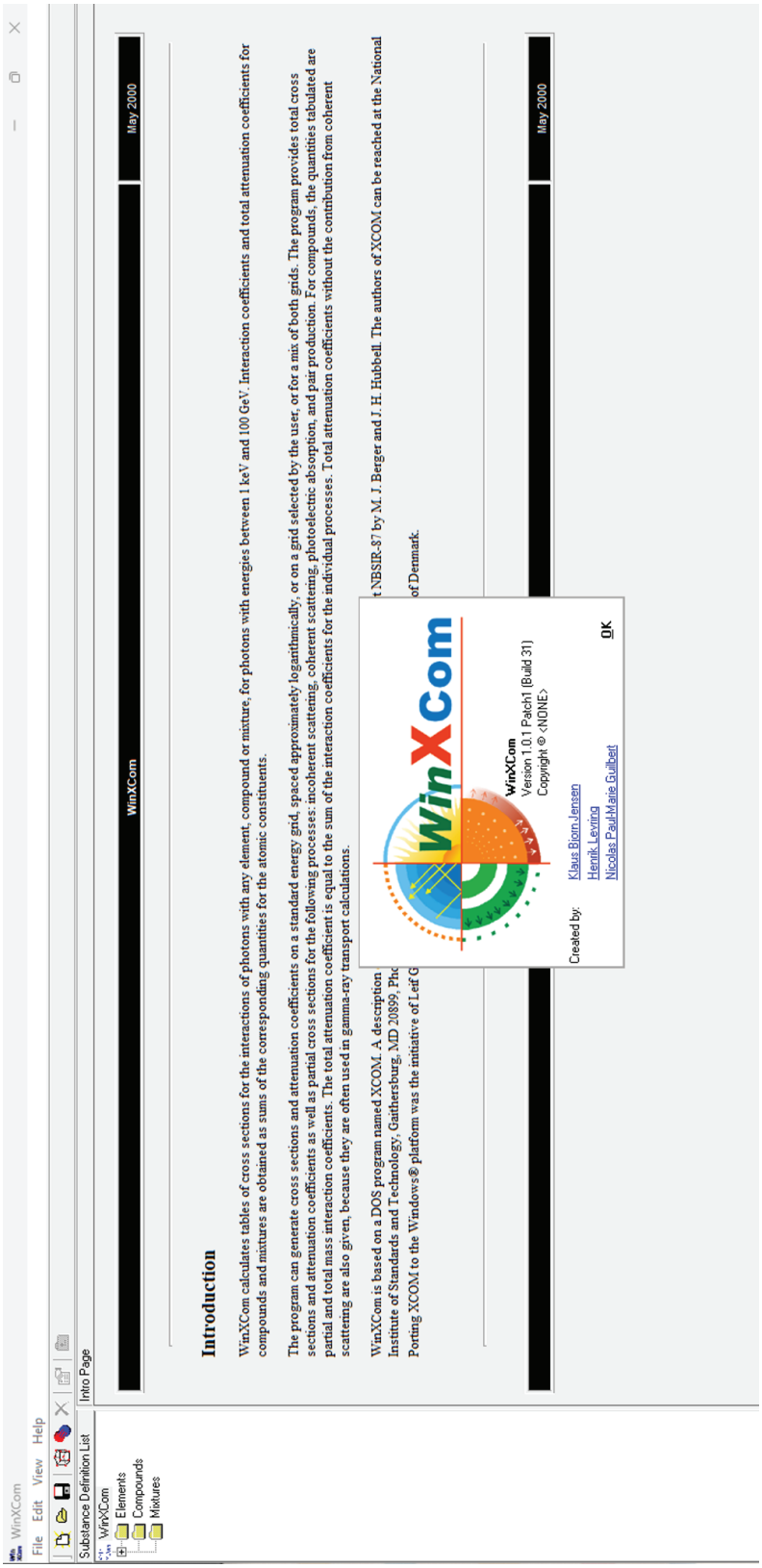


Figure 2. *WinXCom program*[16]

Results

Physical properties

The density and molar volume of the Eu_2O_3-doped germanoborate glasses with the composition $40GeO_2$: $25Gd_2O_3$: $(35-x)B_2O_3$: xEu_2O_3 were systematically investigated, as illustrated in Figure 3. The results reveal that the density of the prepared glasses gradually increased with increasing Eu_2O_3 concentration, rising from approximately 4.58 g/cm^3 for the undoped sample to 5.14 g/cm^3 for the glass containing 9 mol% Eu_2O_3. This behavior can be attributed to the substitution of B_2O_3, which possesses relatively low molecular weight and density, by Eu_2O_3, which contains heavier europium ions with a high atomic mass. The incorporation of Eu^{3+} ions therefore enhances the overall density and compactness of the glass network. In contrast, the molar volume exhibits a slightly different trend with increasing Eu_2O_3 content. Initially, the molar volume decreased from approximately 34.7 cm^3/mol to 34.5 cm^3/mol with increasing Eu_2O_3 concentration up to 3 mol%, suggesting structural compaction and improved atomic packing within the glass matrix. This behavior may result from the occupation of interstitial positions by Eu^{3+} ions, reducing free volume and enhancing network connectivity. However, at higher Eu_2O_3 concentrations, the molar volume gradually increased and reached approximately 35.8 cm^3/mol for the glass containing 9 mol% Eu_2O_3. The increase in molar volume may be associated with the formation of non-bridging oxygen (NBO) sites and partial disruption of the borate glass network caused by the modifier role of Eu^{3+} ions. Such structural modifications can expand the glass network and increase interatomic spacing within the system.[17,18]

The observed variations in density and molar volume confirm that Eu_2O_3 incorporation significantly influences the structural arrangement, packing efficiency, and network connectivity of the germanoborate glass system.

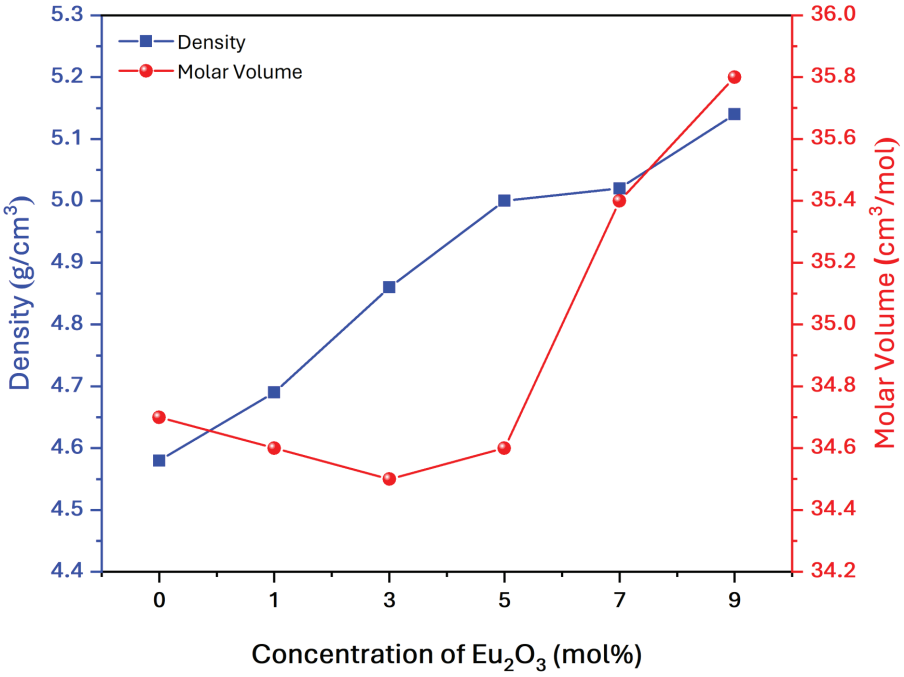


Figure 3. *The density and molar volume of Eu_2O_3-doped germanoborate glasses.*

Radiations properties

Figure 4 illustrates the dominant photon interaction mechanisms of the Eu_2O_3-doped germanoborate glass containing 9 mol% Eu_2O_3 over the photon energy range of 0.015-15 MeV. The attenuation processes can be classified into three principal interaction mechanisms: the photoelectric effect, Compton scattering, and pair production. In the low-energy region below approximately 0.1 MeV, the photoelectric effect is the dominant interaction process. This behavior is strongly associated with the presence of high atomic number elements such as gadolinium (Z=64) and europium (Z=63), which significantly enhance the probability of photon absorption. In addition, two distinct absorption edges can be observed near 0.04852 MeV and 0.05024 MeV, corresponding to the K-absorption edges of Eu and Gd, respectively. These absorption discontinuities originate from the interaction between incident photons and K-shell electrons, resulting in a sudden increase in attenuation efficiency within this energy region. As the photon energy increases into the intermediate energy range (approximately 0.1-1.0 MeV), Compton scattering becomes the predominant interaction mechanism. This interaction mainly depends

on the electron density of the material and involves the inelastic scattering of photons by loosely bound electrons. Consequently, Compton scattering plays a major role in gamma-ray attenuation for medical diagnostic energies and industrial radiographic applications. At higher photon energies exceeding 1.02 MeV, pair production gradually becomes more significant. In this process, incident photons interact with the nuclear field and are converted into electron-positron pairs. The increasing contribution of pair production at high energies demonstrates the capability of the investigated glass system to attenuate high-energy gamma photons, which is particularly important for shielding applications involving high-energy radiation facilities and linear accelerators.

The total attenuation behavior represents the combined contribution of all photon interaction mechanisms and exhibits strong energy dependence across the investigated energy range. The incorporation of Eu_2O_3 together with Gd_2O_3 contributes to enhanced attenuation performance, particularly in the low-energy region where the photoelectric effect and the characteristic K-absorption edges of Eu and Gd strongly influence photon interaction probability.

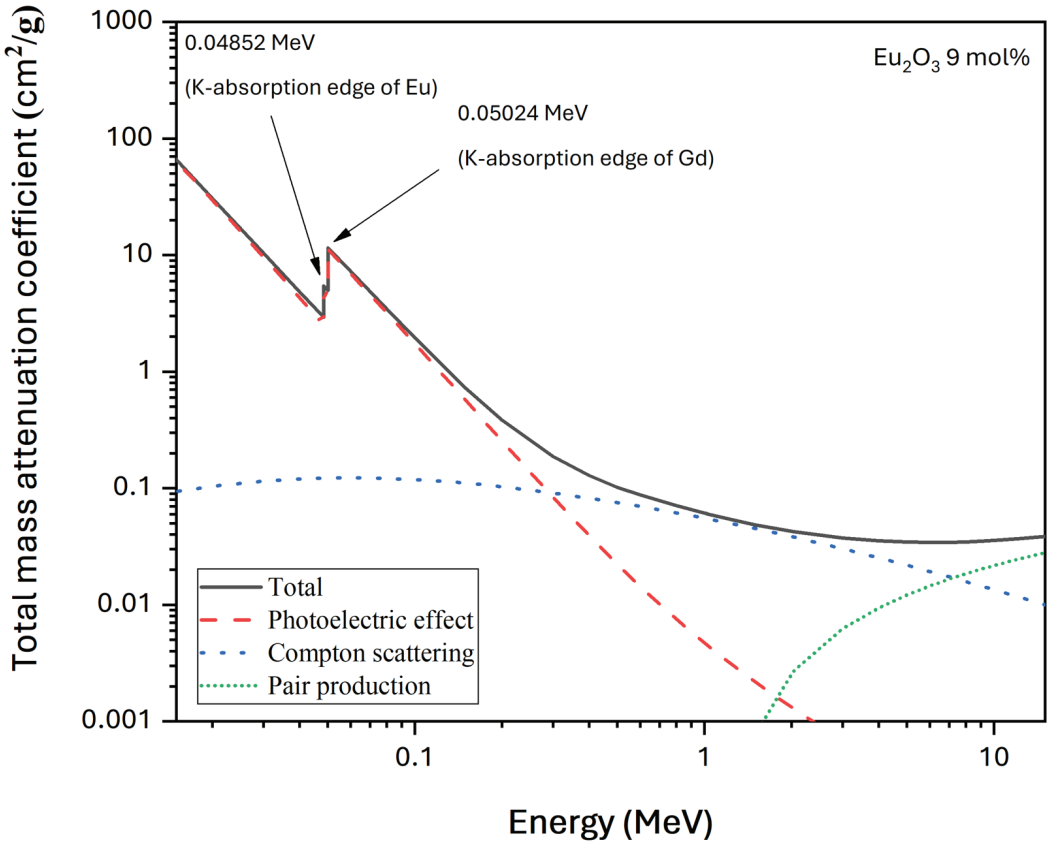


Figure 4. *The partial interaction of Eu_2O_3-doped germanoborate glasses.*

Figure 5 presents the variation of the mass attenuation coefficient (μ_m) for Eu_2O_3-doped germano borate glasses with the composition $40GeO_2$: $25Gd_2O_3$: $(35\text{-}x)B_2O_3$: xEu_2O_3 (where x=0, 1, 3, 5, 7, and 9 mol%) over the photon energy range of 0.015-15 MeV using the WinXCom program. The results demonstrate that the mass attenuation coefficient decreases continuously with increasing photon energy, which is consistent with the energy dependence of the dominant photon interaction mechanisms, including the photoelectric effect, Compton scattering, and pair production. In the low-energy region below approximately 0.2 MeV, the photoelectric effect is the predominant attenuation process, resulting in relatively high μ_m values. In this region, the attenuation coefficients decrease rapidly as photon energy increases due to the strong energy dependence of the photoelectric absorption process. In other words, because the photoelectric interaction becomes far more probable as the photon energy decreases, μm attains its highest values in this low-energy region. Furthermore, two noticeable discontinuities can be observed near 0.04852 MeV and 0.05024 MeV, corresponding to the K-absorption edges of europium (Eu) and gadolinium (Gd), respectively. At these energies, incident photons possess sufficient energy to ionize K-shell electrons, leading to a sudden enhancement in photon absorption probability and attenuation efficiency. These jumps occur exactly where the photon energy just exceeds the K-shell binding energy of Eu and Gd, opening an additional photoelectric

absorption channel. As the photon energy increases into the intermediate region (approximately 0.2-1.0 MeV), the mass attenuation coefficient gradually decreases, and the dominant interaction mechanism shifts toward Compton scattering. In this energy range, photon interactions are mainly governed by electron density rather than atomic number, resulting in relatively small differences among the investigated glass compositions.

Hence, the μm curves of all compositions nearly coincide here, because Compton scattering depends on electron density rather than on atomic number.[19] At higher photon energies exceeding 1.02 MeV, pair production gradually becomes the dominant interaction mechanism. This process involves the conversion of high-energy photons into electron-positron pairs in the vicinity of the atomic nucleus. Consequently, the attenuation behavior at high energies becomes less dependent on elemental composition compared with the low-energy region. Comparative analysis of the investigated glass systems indicates that increasing Eu_2O_3 concentration slightly enhances the mass attenuation coefficient, particularly at low photon energies. This improvement can be attributed to the relatively high atomic number and density of europium ions, which increase photon interaction probability through the photoelectric effect. However, the influence of Eu_2O_3 concentration becomes less pronounced at intermediate and high photon energies because Compton scattering and pair production are less sensitive to atomic number. Overall, the obtained results confirm that Eu_2O_3 incorporation contributes positively to the gamma-ray attenuation capability of germanoborate glasses, demonstrating their potential for radiation shielding applications.

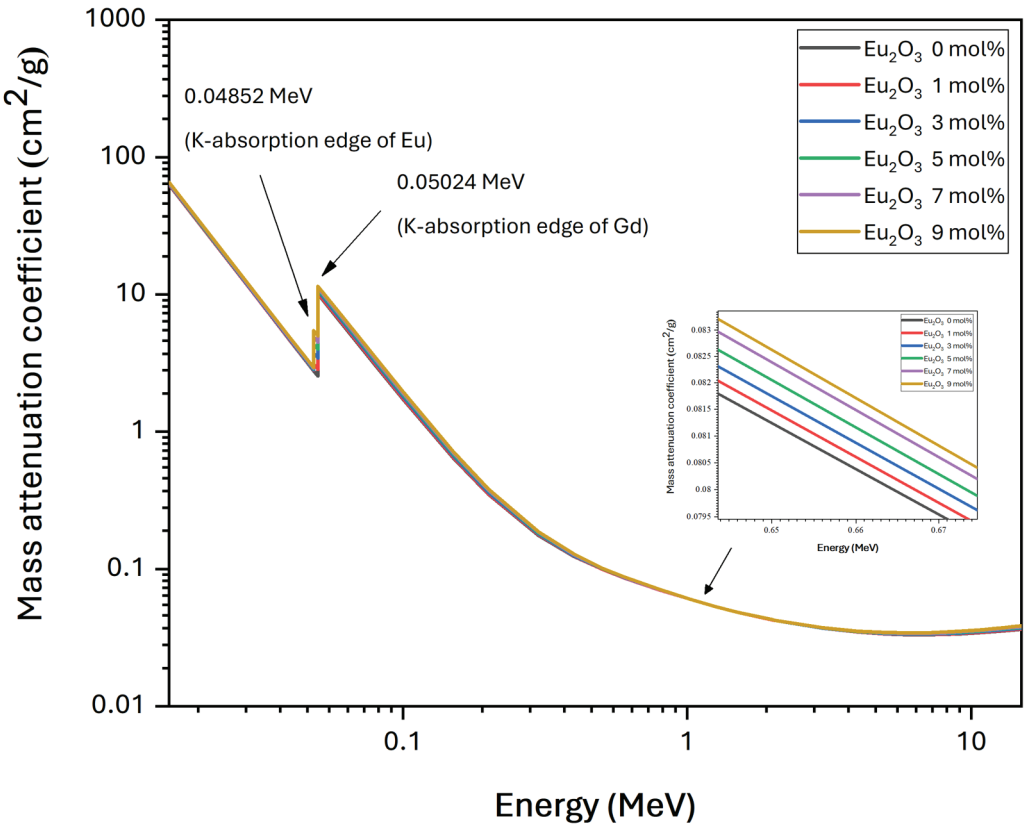


Figure 5. *The mass attenuation coefficient of Eu_2O_3-doped germanoborate glasses.*

Figure 6 illustrates the variation of the effective atomic number (Z_{eff}) of Eu_2O_3-doped germanoborate glasses with different Eu_2O_3 concentrations (0-9 mol%) over the photon energy range of 0.015-15 MeV. The obtained results demonstrate that the Z_{eff} values strongly depend on photon energy due to the dominant photon interaction mechanisms occurring within different energy regions. In the low-energy region below approximately 0.1 MeV, relatively high Z_{eff} values are observed for all investigated glass compositions. This behavior is mainly attributed to the dominance of the photoelectric effect, which strongly depends on the atomic number of the constituent elements. Therefore, the incorporation of high atomic number oxides such as Gd (Z=64) and Eu (Z=63) significantly enhances the photon interaction probability and increases the effective atomic number of the glass system. The high atomic numbers of Gd and Eu enhance the photoelectric effect, resulting in the maximum Z_{eff} in the low-energy region. A distinct discontinuity in

the Z_{eff} curves can be observed near 0.04852 MeV and 0.05024 MeV, corresponding to the K-absorption edges of Eu and Gd, respectively. At these characteristic energies, incident photons possess sufficient energy to ionize K-shell electrons, resulting in a sudden increase in the atomic interaction cross-section and consequently producing sharp enhancements in the effective atomic number.

The presence of these absorption edges confirms the strong contribution of Eu and Gd ions to photon attenuation in the low-energy region. As photon energy increases beyond the absorption edge region, the Z_{eff} values gradually decrease and reach minimum values within the intermediate energy range of approximately 0.2-2.0 MeV. This trend is associated with the predominance of Compton scattering, where photon interactions depend mainly on electron density rather than atomic number. Consequently, the differences in Z_{eff} among the investigated glass compositions become less pronounced within this energy region. At higher photon energies exceeding 1.02 MeV, the Z_{eff} values exhibit a slight increasing trend due to the increasing contribution of pair production. Since pair production probability is influenced by the atomic number of the absorber material, glasses containing higher Eu_2O_3 concentrations demonstrate slightly higher effective atomic numbers at high photon energies. Z_{eff} rises again in this region because pair production, like the photoelectric effect, is enhanced by the high-Z Eu and Gd atoms.[20]

The energy-dependent variation of Z_{eff} closely follows the behavior observed for the mass attenuation coefficient, particularly within the photoelectric absorption region. The incorporation of Eu_2O_3 together with Gd_2O_3 therefore enhances the photon interaction capability of the germanoborate glass system across a broad photon energy range, indicating its potential applicability for gamma ray shielding applications.

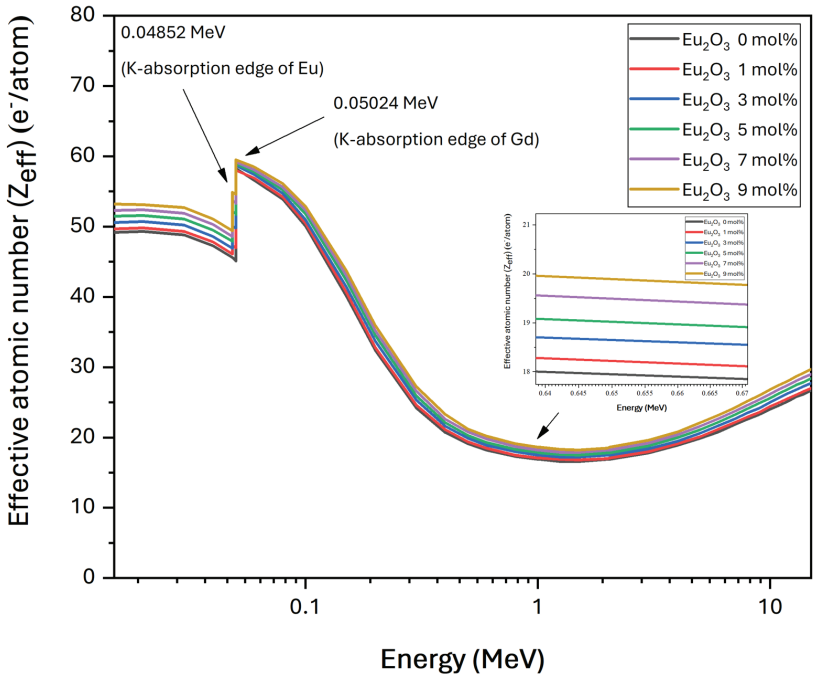


Figure 6. *The effective atomic number of Eu_2O_3 doped germanoborate glasses.*

Figure 7 illustrates the variation of the effective electron density N_{eff} of Eu_2O_3 doped germanoborate glasses over the photon energy range of 0.015-15 MeV. In the low-energy region, relatively high N_{eff} values are observed together with noticeable discontinuities near 0.04852 MeV and 0.05024 MeV, corresponding to the K-absorption edges of Eu and Gd, respectively. A slight inversion trend can also be observed below the absorption edge region, where glasses containing lower Eu_2O_3 concentrations exhibit slightly higher N_{eff} values compared with highly doped compositions. This behavior may be associated with the influence of the Eu and Gd absorption edges, which produce abrupt changes in photon interaction probability and consequently affect the effective electron density response of the investigated glasses. As photon energy further increases, the N_{eff} values gradually rise again due to the increasing contribution of pair production interactions.

Glasses containing higher Eu_2O_3 concentrations exhibit slightly enhanced N_{eff} values at medium and high photon energies owing to the additional electron contribution introduced by europium ions into the glass network.[21]

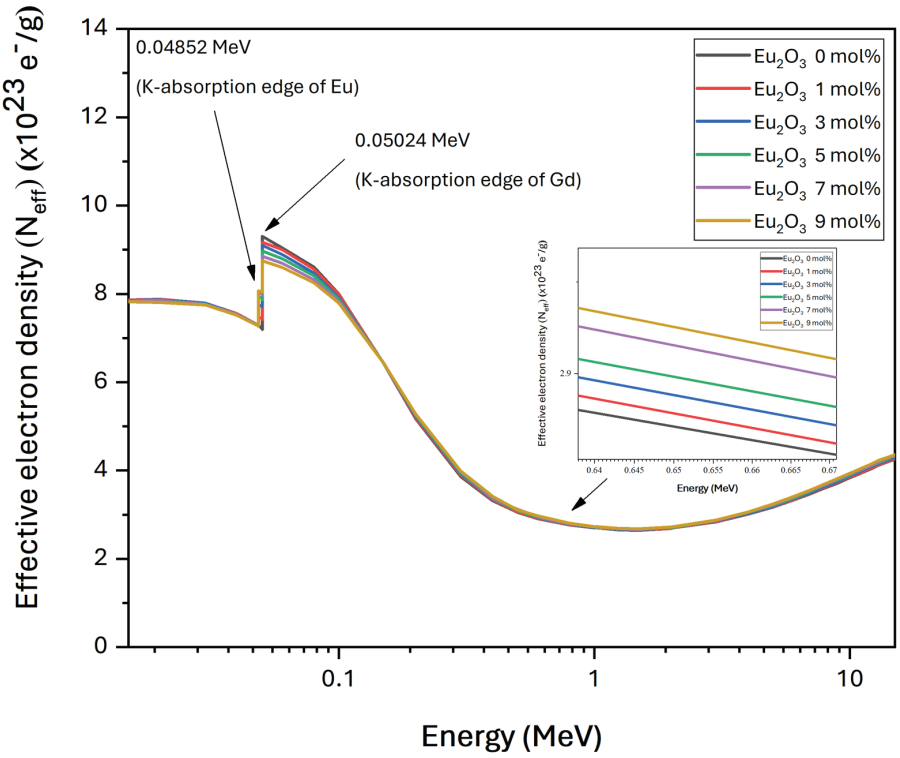


Figure 7. *The effective electron density of Eu_2O_3-doped germanoborate glasses.*

Figure 8 presents the variation of the half-value layer (HVL) for Eu_2O_3-doped germanoborate glasses at a photon energy of 0.662 MeV. It can be observed that the HVL values slightly decrease as the Eu_2O_3 concentration increases, reflecting the improved ability of the glasses to attenuate incident gamma photons. The reduction in HVL indicates that a thinner glass thickness is required to reduce the radiation intensity to half its initial value, confirming the enhanced shielding performance after Eu_2O_3 incorporation. To evaluate the shielding effectiveness, the obtained HVL values were compared with several conventional shielding materials, including commercial window glass,[22] ordinary concrete,[23] and PbO-based glass (50% PbO).[24] The investigated glasses exhibit noticeably lower HVL values than those of commercial window glass and ordinary concrete, indicating superior gamma-ray attenuation capability. Nevertheless, the HVL values remain moderately higher than those of PbO-based glass, which can be attributed to the superior attenuation efficiency of lead-containing materials resulting from the higher atomic number and density of Pb atoms. Among all investigated compositions, the sample containing 9 mol% Eu_2O_3 demonstrates the minimum HVL value, suggesting the highest shielding efficiency within the studied glass system.

These results indicate that the incorporation of Eu_2O_3 positively contributes to the radiation shielding characteristics of germanoborate glasses while maintaining the advantages of transparency and lead-free composition.

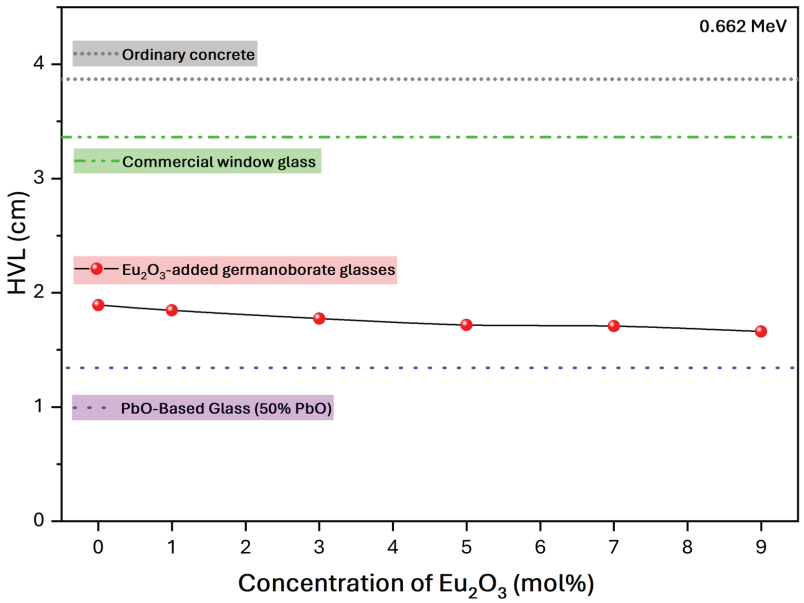


Figure 8. *The half-value layer (HVL) values of Eu2O3-doped germanoborate glasses.*

Figure 9 presents the Pb equivalent thickness of Eu_2O_3-doped germanoborate glasses at 0.662 MeV as a function of Eu_2O_3 concentration. The results indicate that the Pb equivalent values gradually increase with increasing Eu_2O_3 content, rising from approximately 0.094 cm for the undoped glass to about 0.107 cm for the glass containing 9 mol% Eu_2O_3. This behavior suggests that the incorporation of Eu_2O_3 enhances the gamma-ray attenuation capability of the investigated glass system due to the contribution of high atomic number europium ions to photon interaction probability. In comparison with reference shielding materials, the Pb equivalent values of the investigated glasses are significantly higher than those of commercial window glass,[23] demonstrating their superior radiation shielding performance. However, the obtained values remain slightly lower than those reported for PbO-based shielding glasses,[24] which is expected because lead possesses a higher atomic number and density. Nevertheless, the developed Eu_2O_3-doped germanoborate glasses provide the additional advantages of lead-free composition, optical transparency, and improved environmental safety, making them promising candidates for practical radiation shielding applications.

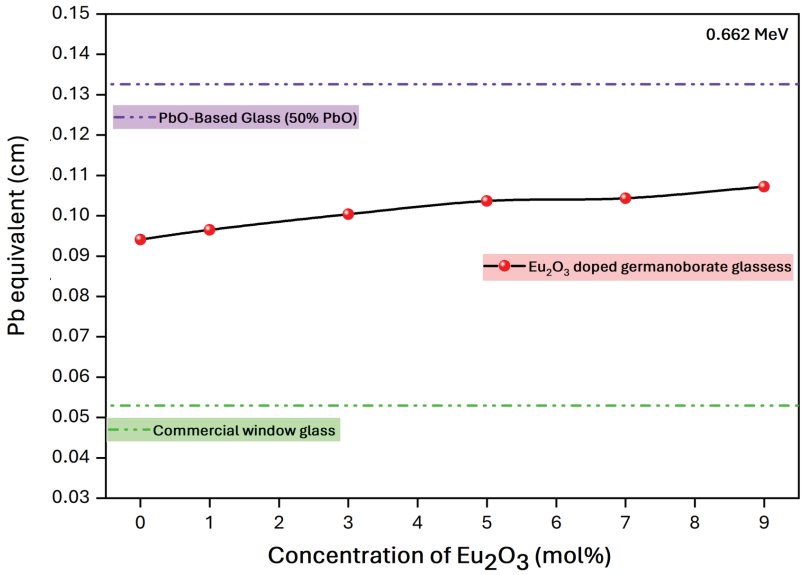


Figure 9. *Pb equivalent of Eu_2O_3-doped germanoborate glasses.*

Limitation

This study has several limitations. The glass samples were successfully fabricated using the conventional melt-quenching technique, and their physical properties, including density and molar volume, were experimentally measured. However, all gamma-ray shielding parameters were theoretically calculated using the WinXCom program and were not validated by experimental attenuation measurements. Therefore, the reported shielding performance should be interpreted as theoretical predictions rather than direct experimental results.

Conclusion

In this work, the physical and gamma-ray shielding properties of Eu_2O_3-doped germanoborate glasses with the composition $40GeO_2$: $25Gd_2O_3$: $(35\text{-}x)B_2O_3$: xEu_2O_3 (x=0, 1, 3, 5, 7 and 9 mol%) were systematically investigated. The glass samples were successfully fabricated using the conventional melt-quenching technique. The obtained results revealed that the density of the glasses gradually increased with increasing Eu_2O_3 concentration, while the molar volume exhibited slight structural variations associated with modifications in glass network packing and connectivity. The radiation shielding parameters, including the μ_m, Z_{eff}, N_{eff}, HVL and Pb equivalent thickness, were evaluated using the WinXCom program. The investigated parameters exhibited strong energy dependence corresponding to the dominant photon interaction mechanisms, namely the photoelectric effect at low energies, Compton scattering at intermediate energies, and pair production at high energies. Distinct discontinuities associated with the K-absorption edges of Eu and Gd were also observed near 0.04852 MeV and 0.05024 MeV, respectively. The incorporation of Eu_2O_3 slightly enhanced the attenuation capability of the investigated glasses, particularly in the low-energy region, due to the contribution of europium ions with relatively high atomic number. In addition, the Pb equivalent values increased with increasing Eu_2O_3 concentration, indicating improved shielding effectiveness compared with commercial window glass, although the performance remained slightly lower than that of PbO-based shielding glasses.

The developed Eu_2O_3-doped germanoborate glasses exhibit favorable physical characteristics together with promising gamma-ray shielding performance. The combination of lead-free composition, optical transparency, and enhanced radiation attenuation capability suggests that these glasses are potential candidates for transparent radiation shielding applications in medical, industrial, and nuclear radiation environments.

Ethical approval

Not applicable. This study did not involve human participants, animals, or identifiable personal data.

Funding

National Research Council of Thailand for supporting this research (project number TSRI_69_4.1).

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Natthakridta Chanthima: writing: original draft, validation, investigation, funding acquisition, conceptualization; **Supakit Yonphan:** writing: review and edit, resources, investigation; **Chalermpon Mutuwong**: methodology, investigation, software; **Wuttichai Chaiphaksa:** writing: review and edit, validation, conceptualization, supervision.

Generative AI and AI-assisted use

Not applicable

Acknowledgements

The authors would like to thank Thailand Science Research and Innovation (TSRI), Center of Excellence in Glass Technology and Materials Science (CEGM), Nakhon Pathom Rajabhat University, and National Research Council of Thailand for supporting this research (project number TSRI_69_4.1).

References

[1] Safari A, Rafie P, Taeb S, Najafi M, Mortazavi SMJ. Development of lead-free materials for radiation shielding in medical settings: a review. J Biomed Phys Eng. 2024; 14(3): 229-44. doi.org/10.31661/jbpe.v0i0.2404-1742.

[2] Juhim F, Chee FP, Awang A, Moh PY, Mohd Salleh KA, Ibrahim S, et al. Study of gamma radiation shielding on tellurite glass containing TiO2 and Al2O3 nanoparticles. Heliyon. 2023; 9(11): e22529. doi.org/10.1016/j.heliyon.2023.e22529.

[3] Geist WH, Santi PA, Swinhoe MT, editors. Non-destructive assay of nuclear materials for safeguards and security. Cham: Springer; 2024. doi.org/10.1007/978-3-031-58277-6.

[4] Uosif MAM, Issa SAM, Ene A, Ivanov V, Mostafa AMA, Atta A, et al. Optimal composition for radiation shielding in BTCu-x glass systems as determined by FLUKA simulation. J Mater Res Technol. 2023; 25: 2088-96. doi.org/10.1016/j.jmrt.2023.06.107.

[5] Yasmin S, Kamislioglu M, Sayyed MI. Assessment of radiation shielding performance of Li2O-BaO-Bi2O3-P2O5 glass systems within the energy range from 0.081 MeV to 1.332 MeV via MCNP6 code. Optik. 2023; 274: 170529. doi.org/10.1016/j.ijleo.2023.170529.

[6] Al-Buriahi MS, Kurtulus R, Eke C, Alomairy S, Olarinoye IO. An insight into advanced glass systems for radiation shielding applications: a review on different modifiers and heavy metal

oxides-based glasses. Heliyon. 2024; 10(22): e40249. doi.org/10.1016/j.heliyon.2024.e40249.

[7] Hongtong W, Kaewjaeng S, Kothan S, Meejitpaisan P, Cheewasukhanont W, Limkitjaroenporn P, et al. Development of gadolinium doped calcium phosphate oxyfluoride glasses for X-ray shielding materials. Mater Today Proc. 2018; 5(6): 14063-8. doi.org/10.1016/j.matpr.2018.02.062.

[8] Felix AT, Ton ND, Luan NT, Kim HJ. Luminescence and scintillation properties of Eu3+ doped lithium borate glass. New Phys Sae Mulli. 2025; 75(8): 595-601. doi.org/10.3938/NPSM.75.595.

[9] ALMisned G, Tekin HO, Ene A, Issa SAM, Kilic G, Zakaly HMH. A closer look on nuclear radiation shielding properties of Eu3+ doped heavy metal oxide glasses: impact of Al2O3/PbO substitution. Materials (Basel). 2021; 14(18): 5334. doi.org/10.3390/ma14185334.

[10] Abou Hussein EM, Shaban SE, Rammah YS, Misbah MH, Marzouk MA. Newly developed CeO2 and Gd2O3-reinforced borosilicate glasses from municipal waste ash and their optical, structural, and gamma-ray shielding properties. Sci Rep. 2024; 14(1): 13673. doi.org/10.1038/s41598-024-63207-4.

[11] Gerward L, Guilbert N, Jensen KB, Levring H. WinXCom-a program for calculating X-ray attenuation coefficients. Radiat Phys Chem. 2004; 71(3-4): 653-4. doi.org/10.1016/j.radphyschem.2004.04.040.

[12] Humaid M, Asad J, Aboalatta A, Shaat SKK, Musleh H, Ramadan K, et al. Gamma and neutron shielding properties of lead-borosilicate shielded glass; novel technique of solid waste recycling. Constr Build Mater. 2023; 375: 130896. doi.org/10.1016/j.conbuildmat.2023.130896.

[13] Elmahroug Y, Tellili B, Souga C. Determination of total mass attenuation coefficients, effective atomic numbers and electron densities for different shielding materials. Ann Nucl Energy. 2015; 75: 268-74. doi.org/10.1016/j.anucene.2014.08.015.

[14] Kolavekar SB, Hiremath GB, Patil PN, Badiger NM, Ayachit NH. Investigation of gamma-ray shielding parameters of bismuth phospho-tellurite glasses doped with varying Sm2O3. Heliyon. 2022; 8(11): e11788. doi.org/10.1016/j.heliyon.2022.e11788.

[15] Mukamil S, Ullah I, Sarumaha C, Wabaidur SM, Islam MA, Khattak SA, et al. Lead-borate glass system doped with Sm3+ ions for the X-ray shielding applications. Results Phys. 2022; 43: 106121. doi.org/10.1016/j.rinp.2022.106121.

[16] Gerward L, Guilbert N, Jensen KB, Levring H. X-ray absorption in matter. Reengineering XCOM. Radiat Phys Chem. 2001; 60(1-2): 23-4. doi.org/10.1016/S0969-806X(00)00324-8.

[17] Madshal MA, El-Damrawi G, Abdelghany AM, Abdelghany MI. Structural studies and physical properties of Gd2O3-doped borate glass. J Mater Sci Mater Electron. 2021; 32(11): 14642-53. doi.org/10.1007/s10854-021-06022-1.

[18] Chandra Shekhar Reddy M, Prabhakar Reddy A, Krishnamurthy Goud K, Appa Rao B. Physical and optical properties of PbO-Sb2O3-B2O3 glasses doped with Gd2O3. Mater Today Proc. 2016; 3(10): 3970-5. doi.org/10.1016/j.matpr.2016.11.058.

[19] El-Khayatt AM, Saudi HA, AlRowis NH. Synthesis and characterization of zinc-lead-phosphate glasses doped with europium for radiation shielding. Sustainability. 2023; 15(12): 9245. doi.org/10.3390/su15129245.

[20] Hussein KI, Alqahtani MS, Meshawi AA, Alzahrani KJ, Zahran HY, Alshehri AM, et al. Evaluation of the radiation shielding properties of a tellurite glass system modified with sodium oxide. Materials (Basel). 2022; 15(9): 3172. doi.org/10.3390/ma15093172.

[21] Al-Buriahi MS, Alajerami YSM, Abouhaswa AS, Alalawi A, Nutaro T, Tonguc B. Effect of chromium oxide on the physical, optical, and radiation shielding properties of lead sodium borate glasses. J Non Cryst Solids. 2020; 544: 120171. doi.org/10.1016/j.jnoncrysol.2020.120171.

[22] Madbouly AM, Alazab HA, Borham E, Ezz-ElDin FM. Study of gamma radiation dosimeter and radiation shielding parameters of commercial window glass. Appl Phys A. 2021; 127(10): 761. doi.org/10.1007/s00339-021-04889-9.

[23] Singh K, Singh H, Sharma V, Nathuram R, Khanna A, Kumar R, et al. Gamma-ray attenuation coefficients in bismuth borate glasses. Nucl Instrum Methods Phys Res B. 2002; 194(1): 1-6. doi.org/10.1016/S0168-583X(02)00498-6.

[24] Hussein EMA, Madbouly AM. Chemical and radiation shielding effectiveness of some heavy metal oxide glasses for immobilizing radioactive wastes. J Aust Ceram Soc. 2024; 60(1): 127-42. doi.org/10.1007/s41779-023-00951-2.

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