JOURNAL OF ASSOCIATED MEDICAL SCIENCES

JAMS

JOURNAL OF ASSOCIATED MEDICAL SCIENCES

THE OFFICIAL PEER-REVIEWED ONLINE JOURNAL

Volume 59 Number 1 January -April 2026 E-ISSN: 2539-6056

Journal of Associated Medical Sciences

Aims and scope

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

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

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

Objectives

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

Peer review process

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

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

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

Publication schedule

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

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

Access Policy and Archiving

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

Types of Manuscripts

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

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

Ownership and Revenue Sources

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

Journal History

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

Language

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

Publication Fee

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

Informed Consent Policy

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

Advertising and Direct Marketing

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

Editorial Board

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

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

Journal of Associated Medical Sciences

Vol. 59 No. 1 : January - April 2026

Content

Content

Content

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

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Phrenic nerve and cervical functional interdependence and its sensorimotor anomalies: A musculoskeletal pain consideration review

Shaheem Rasheed[1], Mashael Alsobhi[2*], Mohamed Faisal Chevidikunnan[2], Fayaz Khan[2]

[1]Zayed Military Hospital, Al Batayih, Sharjah, United Arab Emirates.
[2]Department of Physical Therapy, Faculty of Medical Rehabilitation Sciences, King Abdulaziz University, Jeddah, Saudi Arabia.

ARTICLE INFO

Article history:
Received 8 October 2024
Accepted as revised 28 August 2025
Available online 5 September 2025

Keywords:
Musculoskeletal, interdependence, breathing, diaphragm, motor control.

ABSTRACT

Background: Optimal breathing occurs when using the diaphragm muscle, which is known as diaphragmatic breathing. Any disruption or faulty habit of this breathing pattern would result in breathing pattern disorders (BPDs).

Objectives: The aim of this review is to shed light on the importance of detecting the presence and extent of asymptomatic functional anomalies of cervical segments among individuals with breathing pattern disorders.

Materials and methods: This narrative review was performed by searching in MEDLINE using terms related to phrenic nerve and cervical functional, interdependence, and sensorimotor anomalies. Additionally, articles regarding this topic were extracted in full for a comprehensive and critical analysis. Any study design, such as observational studies, systematic reviews, or narrative reviews, was included. The results were narratively synthesized and described.

Results: Breathing is a physiological process that occurs without any conscious effort. Optimal breathing utilizes the deep diaphragmatic muscles, which lead to filling the lungs with air more efficiently. However, breathing pattern disorders may result from inadequate diaphragm use. BPDs can lead to structural musculoskeletal pathologies, including neck, low back, temporomandibular joint pain, and/or scapular dyskinesis. BPDs cause motor reflexive changes in the cervical spine, which may contribute to pain, restriction, and functional impairments.

Conclusion: Early identification of this faulty habitual suboptimal breathing is required to minimize its adverse effect, which can be disruptive to the motor control mechanism.

Introduction

Breathing is an essential and constant series of oxygen and carbon dioxide gas exchanging cyclic activity humans do, yet most individuals are unaware of how breathing persistently happens, and little attention is often paid to it throughout the day. Optimal breathing is referred to diaphragmatic breathing, which is three-dimensional in nature, and it utilizes the diaphragm muscle to perform adequate functionality[1] (Figure 1). Also, it involves synchronous movement of the abdomen and lower rib cage.[2] Normal eupneic breathing can be affected by multiple factors, including biomechanical, biochemical, and psychological, etc..[3-6]

However, faulty/suboptimal breathing patterns occur when the diaphragm is not adequately used for

∗ **Corresponding contributor.**
Author's Address: *Department of Physical Therapy, Faculty of Medical Rehabilitation Sciences, King Abdulaziz University, Jeddah, Saudi Arabia.*
E-mail address: mgalsobhi@kau.edu.sa
doi: 10.12982/JAMS.2026.001
E-ISSN: 2539-6056

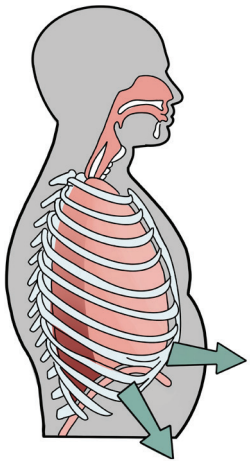


Figure 1. *Diaphragmatic Breathing/Belly Breathing.*

breathing, and it results in breathing pattern disorders (BPDs) or hyperventilation syndrome (HVS). BPD is defined as inappropriate breathing that is persistent enough to cause symptoms with no apparent organic cause.[7] However, a paradoxical breathing pattern, which is considered the most extreme form of BPDs, is often accompanied with muscle imbalance and/or dysfunction and cervical spine pain.[8,9] It is postulated that BPDs have a detrimental effect on functional movements by affecting the motor control of cervical, scapular, and lumbar segments.[10] As like poor posture, a poor breathing pattern is not a disease but a disorder. BPDs, which are habitual, can lead to various faulty functional adaptations, which in turn may cause structural musculoskeletal pathologies, such as neck,[11] low back,[12] temporomandibular joint pain,[13] and scapular dyskinesis.[3] BPDs functional anomalies cause reflexive changes in cervical structure, which are insidious and without pain. Moreover, these faulty habitual patterns act physically and physiologically against biologically sustained patterns and, in a vicious cycle, promote abnormal function and alter structure, which then prevents the return to normal function.

The adaptation reflex continues to progress up to a point where the adaptive capacity is exhausted and symptoms inevitably emerge involving pain, restriction, impairments, and structural breakdown.[14] Early identification of these reflexive motor control changes in the cervical region due to BPD is imperative to preclude the incidence of further debilitating cervical structural anomalies. Up to our knowledge, no studies have been conducted to find out the presence and extent of asymptomatic functional anomalies of cervical segments among individuals with BPDs. Also, its extended adverse effects have not been further studied in this population.

Diaphragm and breathing pattern disorder

Literature explained breathing as a constant, fundamental, and automatic act of supplying oxygen through physiological respiration and mechanical ventilation. Respiration is the exchange of gas between lungs-alveoli and capillaries, whereas ventilation is the process of inhalation and exhalation, i.e., air goes in and out of the pulmonary system. On the contrary, breathing is not a mere oxygen supply process only, but it is also associated with the well-being of the entire body system, maintaining allostasis[8] along with providing biomechanical stability and mobility of the trunk and spine.[2] The diaphragm, which is the primary respiratory muscle, contributes to 70-80% of respiratory force, and it has also been evaluated for its role as postural, sphincteric, visceral, and metabolic functions.[8,14] This horizontal dome-shaped muscle is a repertoire having a primary role in respiration,[8] and a secondary role in the trunk and postural stability,[10] spinal decompression,[15] emotional regulation,[16] visceral mobility,[17] and fluid dynamics.[18]

Among the general population, BPDs are more common than reported.[8] It is estimated that 10% of the population is affected by BPD, which escalates to 30% with asthma sufferers and 83% in individuals with anxiety.[14] An average of 10% of patients in a population are diagnosed with HVS.[19] However, suboptimal breathing is more prevalent in women (14%) than in men (2%).[20] Many individuals breathe inefficiently by having shallow breathing or using upper chest muscles instead of utilizing diaphragm muscles, which do not meet the physiological needs of humans (Figure 2). These breathing pattern changes are unnoticed as it is subtle, yet clinically significant.[20] BPDs may range from simple upper chest breathing to hyperventilation to the most extreme form of paradoxical breathing (Figures 3 & 4).[21] In paradoxical breathing, the chest and abdominal muscle functions are in an opposite pattern, i.e., the patient inhales using the chest muscles.[22] When breathing becomes dysfunctional, a general sequence of progressive, adaptive changes results from BPDs.[14] The sensory-motor system often utilizes the secondary respiratory muscles, which are over-activated. These accessory and superficial muscles (Sternocleidomastoid, Scalenus, Upper

trapezius, Pectoralis Minor and Major, and Levator scapulae) become hypertonic,[13] which can lead to the development of trigger points[24] that become fibrotic, and will also be palpable and observable. As a result, the head is protruded forward, and the shoulders are hunched anteriorly and pulled forward.[24,25] Progressively, the lower cervical spine becomes rigid with fixed lordosis and decreased mobility of the second cervical segments and second rib.[26] Figure 5 shows the etiopathogenesis of PBDs affecting the cervical motor control mechanism.[25]

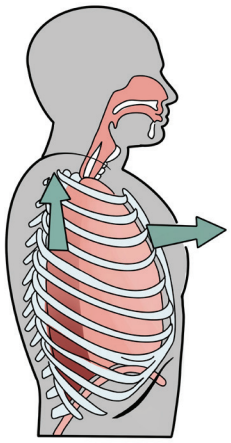


Figure 2. *Chest breathing.*

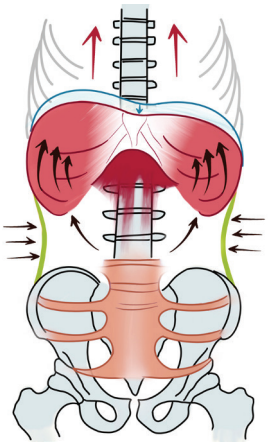


Figure 3. *Paradoxical inspiration.*

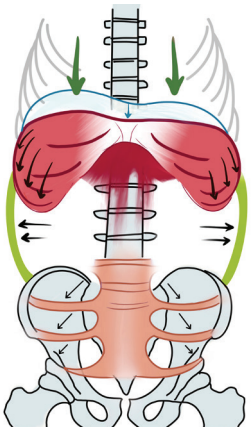


Figure 4. *Paradoxical expiration.*

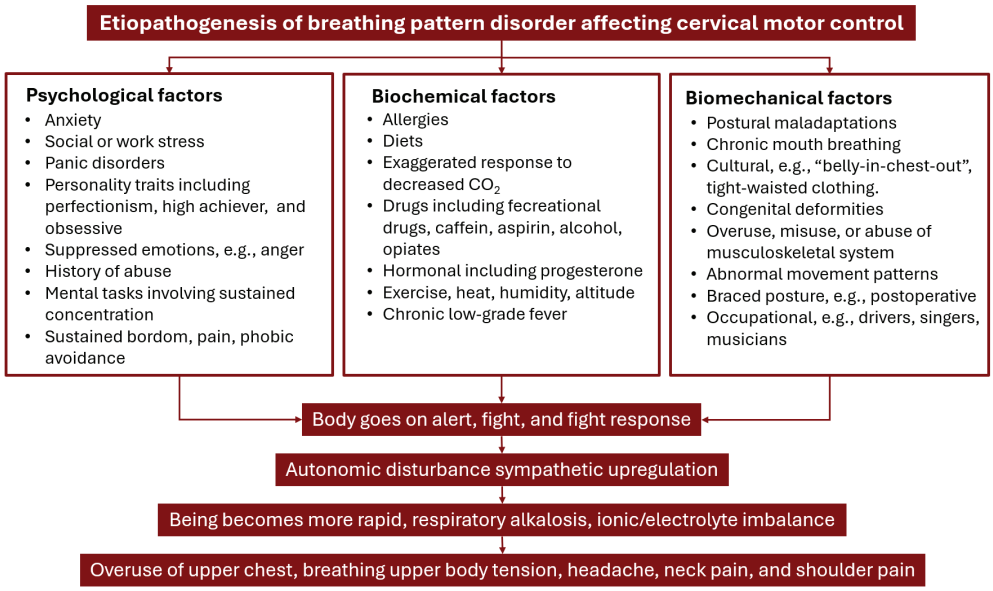


Figure 5. *Flow chart showing the etiopathogenesis of breathing pattern disorders affecting cervical motor control.*

Cervical disorders

The cervical spine is a mobile segment providing a wide range of movement. Usually, activities of daily living (ADL) utilize 30% to 50% of cervical ROM.[27] In regard to neck stability, muscle strength is critical to maintain proper alignment, smooth movement, and stabilization. Nevertheless, cervical structural and functional pathologies exhibit an increasing prevalence in the present-day population, and it is associated with a significant morbidity rate. As a great number of individuals are becoming accustomed to electronic gadgets and indulging in tabletop jobs, neck pain and related symptoms are reported to be the most common disorders which affect approximately 66% of the population at some stage of their lives.[28,29] Even though a large population suffers from neck pain, people may not experience a complete recovery of symptoms as it is difficult to be diagnosed and managed.[30] Secondary muscle activation in BPD is linked to neck pain,[11] trigger point formation,[31] and scapular dyskinesis.[32] Chronic neck pain is usually diagnosed without a history of trauma, structural involvement, and poor posture, which is referred as mechanical neck pain.

Functional neurology delineates the phrenic nerve and cervical neural interconnectivity through the cervical from C_1 to C_4 and brachial from C_5 to T_1 plexus, sharing its roots for the phrenic nerve from C_3 to C_6, which supplies the diaphragm. BPDs induce aberrant neural drive that can cause symptoms like motor control dysfunction in the neck, shoulder, and arm, along with brachialgia and thoracic outlet syndrome (TOS), etc. (Figures 6 and Figure 7).[33,34] Furthermore, the central or rural portion of the diaphragm is innervated by vagus nerve, thus diaphragm dysfunction will affect the cervical base, floor of the mouth, eyes, and the dura matter via vagal connection to the trigeminal nucleus through the medial longitudinal fasciculus.[35,36] The trigeminal nucleus is connected to C_2-C_4 nerve roots,[35,37,38] rectus capitis posterior minor,[39] and nuchal ligament, which is directly connected to dura.[40]

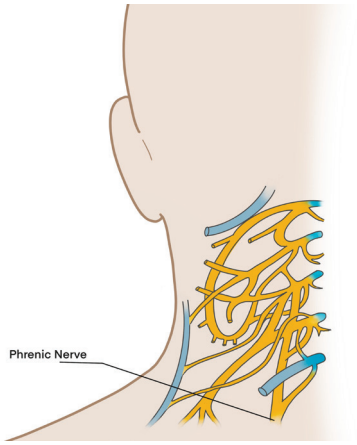


Figure 6. Phrenic nerve and its relation to the brachial plexus.

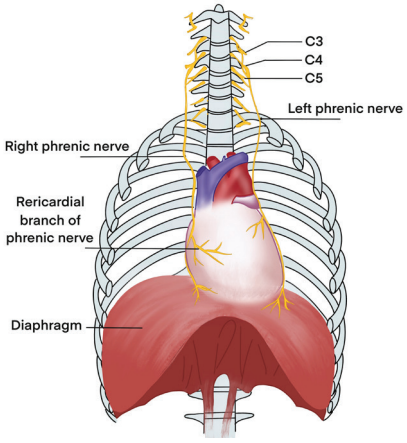


Figure 7. Phrenic Nerve and its branches.

Somatosensory alteration

Humans are born with a basic movement control system called primitive reflexes, balance, and righting reactions. Through both asymmetric tonic neck reflex (ATNR) and symmetric tonic neck reflex (STNR), the cervical spine facet proprioceptors/mechanoreceptors contribute to the function and protection of the cervical spine,[41] postural stability,[42] and they play a significant role in cervical pain.[43] Small and slender suboccipital muscles have more proprioceptive receptors than those of lower extremity large muscles, such as the gluteus maximus.[44,45] Excessive neural drive due to BPD primarily affects the lower threshold tonic motor neurons supplying deep cervical muscles, which are crucial in maintaining stability, motor control, and the optimal instantaneous axis of rotation (OIAR). BPD induces neural dysfunction, which causes anomalies in suboccipital muscles, triggers neck pain, and radiates pain in the relevant dermatome or myotome distribution pattern.[46]

These breathing disruptions induce cervical, neural, and muscular mechanisms that remain sub-clinical as functional instability for months and years until symptoms eventually emerge. Functional instability is defined as the sensation of instability or recurrent sprain (or both) due to proprioceptive neuromuscular deficit,[47] which is clinically exhibited as recurrent neck knots (wry neck), neck sprain, chronic neck pain, cervicogenic headaches, suboccipital muscular atrophy,[48] and standing balance deterioration.[49] The movement control system is functioning by controlled and coordinated activity of three subsystems, i.e., the motor control system (CNS), the passive support system (Skeletal System), and the active movement system (Muscular System). Dysfunction of any of these subsystems debilitates stability and causes pain.[50] This functional instability of any of the subsystems is compensated by the higher motor system using one or more subsystems, which is achieved by altering muscle activation and movement patterns. Eventually, it becomes ingrained in the motor cortex, and later individuals may find it difficult to retrain to normal, even with conventional rehabilitation approaches. The long-term utilization of these adaptive mechanisms leads to the inception of irreversible structural pathologies. Thus, BPDs can be disruptive to the motor control mechanism.

Anomalies of the autonomic nervous system

Human body allostasis is effectively regulated by the autonomic nervous system (ANS) through various involuntary systems in the body, like the respiratory, muscular, circulatory, lymphatic, and endocrine systems.[51,52] ANS with its branches, sympathetic nervous system (SNS) and parasympathetic nervous system (PNS), responds to different perceptions such as pain, fear, anxiety, and other emotional stressors by adjusting breathing, muscle tone, heart rate, blood pressure, and hormones.[53,54] Generally, the ANS passes the diaphragm through the central crural region.[55,56] This suggests that any anomalies in the diaphragm, such as spasm or atrophy, will have deleterious effects on ANS in both up and downward directions, affecting the visceral organs as well as the metamers innervated by the ANS plexus such as muscles and vertebral facets leading to chronic pain.[57-59] "Up-regulation" is referred to as the heightened and aroused response of the sympathetic nervous system.[60] Continuous "Up-regulation" of SNS predisposes to alteration in breathing pattern by changing the motor control pattern and also the respiratory muscle recruitment pattern[61] leading to BPD, acute or chronic musculoskeletal pain.

Persistent body function with an "up regulated" ANS leads to hypersensitivity to touch, hyper-reflexia, and exaggerated pain perception. Normal palpation or stimulus causes increased pain perception and withdrawal or removal of the body part in a pattern for protection.[60] These hyperactive startle responses theoretically delineate the cause and perpetuation of musculoskeletal pain without an organic cause in patients with BPDs.[62] This stage is very challenging for the therapist to approach patients, as any attempt will be excruciating for the subject.[63]

Regional interdependence

There is a growing awareness about the concept of regional interdependency. Human body systems are integrated and interconnected as fascio-neuro-skeletally to depend upon each other locally, regionally as well as globally. The concept of regional interdependence began with the underlying thought process, which describes the human body as a kinetic chain.[64] This theory suggests that if one part of the kinetic chain is unable to perform any motor pattern sufficiently, another part of the body will compensate for the deficiency.[64]

Though muscles function individually, they are functionally interconnected across the body with fascial webbing traced as "myofascial meridians". These fascial meridians distribute and transmit tension, strain, fixation, stability, resilience, and postural compensation to other remote portions of the body. The deep front line (DFL) myofascial meridian better explains the relationship between the diaphragm and cranio-cervical region. DFL through its upper posterior line, upper middle line, and upper anterior line connects the diaphragm to the cranio-cervical region.[65] The upper posterior line connects the posterior diaphragm, anterior longitudinal ligament, and longus coli to longus capitis while the upper middle line connects the central tendon of the diaphragm, pericardium, mediastinum, parietal pleura, fascia prevertebralis to scalenes. However, the upper anterior line connects the anterior diaphragm, transversus thoracis, infrahyoid, and suprahyoid to the jaw muscles.

The dysfunctional breathing patterns might create a variety of negative physiological and biomechanical adverse effects on the entire myofascial line. Decreased diaphragm mobility causes cervical erector spinae hypertonicity, ribcage hypomobility, and thoracic prevertebral muscle tightness, which may lead to exaggerated thoracic kyphosis, cervical hyperlordosis, and capital extension.

Conclusion

It is axiomatic that in chronic neck conditions, pain and movement dysfunctions are not always due to the structural theory of pain-spasm-pain; on the contrary, pain results from muscular adaptation due to breathing pattern disorder mediated by the CNS. Hypertonicity of accessory muscles of respiration due to constant activation, eliciting a muscle imbalance syndrome (not a symptom) because of functional pathology leading to cervical motor control disorder. A breathing pattern assessment in patients with neck pain can attribute to a differential diagnosis between structural and functional cervical disorders. Pain is a protective response to muscle overuse or trauma, which has to be addressed and corrected promptly. Future research should investigate the effect of breathing pattern retraining on musculoskeletal pain syndrome to provide clinicians with evidence for integrating breathing retraining into their clinical practices.

Ethical approval

This is a review study for already existing data, and it 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 authors declare no conflict of interest.

CRediT authorship contribution statement

Shaheem Rasheed: conceptualization, investigation, methodology, writing original draft; **Mashael Alsobhi**: conceptualization, investigation, resources, writing, review and editing; **Mohamed Faisal Chevidikunnan**: investigation, conceptualization, methodology, visualization, validation, reviewing and editing; **Fayaz Khan**: conceptualization, visualization, validation, project administration, reviewing and editing.

Acknowledgements

The authors would like to thank Ms. Shatha Mukhtar for her help in editing the figures used in this review.

References

[1] Pryor JA, Prasad SA, editors. Physiotherapy for respiratory and cardiac problems-adults and paediatrics. Edinburgh: Churchill Livingstone; 2002.

[2] Kaminoff L. What yoga therapists should know about the anatomy of breathing. Int J Yoga Ther. 2006; 16: 67-77. doi:10.17761/ijyt.16.1.64603tq46513j242.

[3] CliftonSmith T, Rowley J. Breathing pattern disorders and physiotherapy: inspiration for our profession. Phys Ther Rev. 2011; 16(1): 75-86. doi:10.1179/1743288X10Y.0000000025.

[4] Chaitow L. Breathing pattern disorders, motor control, and low back pain. J Osteopath Med. 2004; 7(1): 33-40. doi:10.1016/S1443-8461(04)80007-8.

[5] Courtney R, Cohen M, van Dixhoorn J. Relationship between dysfunctional breathing patterns and ability to achieve target heart rate variability with features of “coherence” during biofeedback. Altern Ther Health Med. 2011; 17: 38-44.

[6] Perri M, Halford E. Pain and faulty breathing: a pilot study. J Bodyw Mov Ther. 2004; 8(4): 297-306. doi:10.1016/S1360-8592(03)00085-8.

[7] Vickery R. The effect of breathing pattern retraining on performance in competitive cyclists [Internet]. 2008 [cited 2021 Mar 7]. Available from: /paper/The-effect-of-breathing-pattern-retraining-on-in-Vickery/80d5467cb0fb59f8e9777c943a8688ea6e-cd8135.

[8] Chaitow L, Gilbert C, Bradley D. Recognizing and treating breathing disorders. 2nd Ed. [Internet]. Elsevier; [cited 2021 Mar 2]. Avai lable from: https://www.elsevier.com/books/recognizing-and-treating-breathing-disorders/chaitow/978-0-7020-4980-4.

[9] Perri M. Rehabilitation of breathing pattern disorders. In: Liebenson C, editor. Rehabilitation of the spine: A practitioner’s manual. 2nd Ed. Philadelphia: Lippincott Williams and Wilkins; 2007: pp 93-109. J Can Chiropr Assoc. 2007; 51(1): 62.

[10] Hodges PW, Heijnen I, Gandevia SC. Postural activity of the diaphragm is reduced in humans when respiratory demand increases. J Physiol. 2001; 537(Pt 3): 999-1008. doi:10.1111/j.1469-7793.2001.00999.

[11] Kapreli E, Vourazanis E, Billis E, Oldham JA, Strimpakos N. Respiratory dysfunction in chronic neck painpatients. A pilot study. Cephalalgia. 2009; 29(7): 701-10. doi:10.1111/j.1468-2982.2008.01787.

[12] Roussel NA, Nijs J, Truijen S, Smeuninx L, Stassijns G. Low back pain: clinimetric properties of the Trendelenburg test, active straight leg raise test, and breathing pattern during active straight leg raising. J Manipulative Physiol Ther. 2007; 30(4): 270-8. doi:10.1016/j.jmpt.2007.03.001.

[13] Hruska RJ. Influences of dysfunctional respiratory mechanics on orofacial pain. Dent Clin North Am. 1997; 41(2): 211-27.

[14] Courtney R. The functions of breathing and its dysfunctions and their relationship to breathing

therapy. Int J Osteopath Med. 2009; 12(3): 78-85. doi: 10.1016/j.ijosm.2009.04.002.
[15] Kolar P, Sulc J, Kyncl M, et al. Postural function of the diaphragm in persons with and without chronic low back pain. J Orthop Sports Phys Ther. 2012; 42(4): 352-62. doi:10.2519/jospt.2012.3830.
[16] Keleman S. Emotional anatomy: the structure of experience. Berkeley: Center Press; 1985.
[17] Yu X, Jiang HY, Zhang CX, Jin ZH, Gao L, Wang RD, et al. The role of the diaphragm in postural stability and visceral function in Parkinson's disease. Front Aging Neurosci. 2021; 13: 785020. doi:10.3389/fnagi.2021.785020.
[18] Salah HM, Goldberg LR, Molinger J, Felker GM, Applefeld W, Rassaf T, et al. Diaphragmatic function in cardiovascular disease: JACC Review Topic of the Week. J Am Coll Cardiol. 2022; 80(17): 1647-59. doi:10.1016/j.jacc.2022.08.760.
[19] Rice RL. Symptom patterns of the hyperventilation syndrome. Am J Med. 1950; 8(6): 691-700. doi:10.1016/0002-9343(50)90093-3.
[20] Thomas M, McKinley RK, Freeman E, Foy C, Price D. The prevalence of dysfunctional breathing in adults in the community with and without asthma. Prim Care Respir J. 2005; 14(2): 78-82. doi:10.1016/j.pcrj.2004.10.007.
[21] Lum LC. Hyperventilation syndromes in medicine and psychiatry: a review. J R Soc Med. 1987; 80(4): 229-31. doi:10.1177/014107688708000413.
[22] Chapman EB, Hansen-Honeycutt J, Nasypany A, Baker RT, May J. A clinical guide to the assessment and treatment of breathing pattern disorders in the physically active: Part 1. Int J Sports Phys Ther. 2016; 11(5): 803-9.
[23] Lucas KR, Polus BI, Rich PA. Latent myofascial trigger points: their effects on muscle activation and movement efficiency. J Bodyw Mov Ther. 2004; 8(3): 160-6. doi: 10.1016/j.jbmt.2003.12.002.
[24] Chapman EB, Hansen Honeycutt J, Nasypany A, Baker RT, May J. A clinical guide to the assessment and treatment of breathing pattern disorders in the physically active: part 1, a case series. Int J Sports Phys Ther. 2016; 11(6): 971-9.
[25] Kapreli E, Vourazanis E, Billis E, Oldham JA, Strimpakos N. Respiratory dysfunction in chronic neck pain patients. A pilot study. Cephalalgia. 2009;29(7):701–10. doi:10.1111/j.14682982.2008.01787.
[26] Kolnes LJ. Embodying the body in anorexia nervosa--a physiotherapeuti approach. J Bodyw Mov Ther. 2012; 16(3): 281-8. doi:10.1016/j.jbmt.2011.12.005.
[27] Bennett SE, Schenk RJ, Simmons ED. Active range of motion utilized in the cervical spine to perform daily functional tasks. J Spinal Disord Tech. 2002; 15(4): 307-11. doi:10.1097/00024720-200208000-00008.
[28] Binder AI. Cervical spondylosis and neck pain. BMJ. 2007; 334(7592): 527-31. doi:10.1136/bmj.39130.530182.80.
[29] Fejer R, Hartvigsen J. Neck pain and disability due to neck pain: what is the relation?. Eur Spine J. 2008; 17(1): 80-8. doi:10.1007/s00586-007-0521-9.
[30] Abhishek A, Doherty M. Diagnosis and clinical presentation of osteoarthritis. Rheum Dis Clin North Am. 2013; 39(1): 45-66. doi:10.1016/j.rdc.2012.10.007.
[31] Travell JG, Simons DG. Myofascial Pain and Dysfunction: The Trigger Point Manual, Vol. 2, The Lower Extremities. 1st Ed. Baltimore: Williams & Wilkins; 1993; pp 628.
[32] Bradley H, Esformes J. Breathing pattern disorders and functional movement. Int J Sports Phys Ther. 2014; 9(1): 28-39.
[33] Barraclough A, Triplett J, Tuch P. Brachial neuritis with phrenic nerve involvement. J Clin Neurosci. 2012; 19(9): 1301-2. doi:10.1016/j.jocn.2011.12.009.
[34] Prakash, Prabhu LV, Madhyastha S, Singh G. A variation of the phrenic nerve: case report and review. Singapore Med J. 2007; 48(12): 1156-7.
[35] Messlinger K, Fischer MJ, Lennerz JK. Neuropeptide effects in the trigeminal system: pathophysiology and clinical relevance in migraine. Keio J Med. 2011; 60(3): 82-9. doi:10.2302/kjm.60.82.
[36] Kemp WJ 3rd, Tubbs RS, Cohen-Gadol AA. The innervation of the cranial dura mater: neurosurgical case correlates and a review of the literature. World Neurosurg. 2012; 78(5): 505-10. doi:10.1016/j.wneu.2011.10.045.
[37] Adeeb N, Mortazavi MM, Tubbs RS, Cohen-Gadol AA. The cranial dura mater: a review of its history, embryology, and anatomy. Childs Nerv Syst. 2012; 28(6): 827-37. doi:10.1007/s00381-012-1744-6.
[38] Upadhyay J, Knudsen J, Anderson J, Becerra L, Borsook D. Noninvasive mapping of human trigeminal brainstem pathways. Magn Reson Med. 2008; 60(5): 1037-46. doi:10.1002/mrm.21682.
[39] Kahkeshani K, Ward PJ. Connection between the spinal dura mater and suboccipital musculature: evidence for the myodural bridge and a route for its dissection--a review. Clin Anat. 2012; 25(4): 415-22. doi:10.1002/ca.21261.
[40] Dean NA, Mitchell BS. Anatomic relation between the nuchal ligament (ligamentum nuchae) and the spinal dura mater in the craniocervical region. Clin Anat. 2002; 15(3): 182-5. doi:10.1002/ca.10001.
[41] Cavanaugh JM, Lu Y, Chen C, Kallakuri S. Pain generation in lumbar and cervical facet joints. J Bone Joint Surg Am. 2006; 88(Suppl2): 63-7. doi:10.2106/JBJS.E.01411.
[42] Abrahams VC. The physiology of neck muscles; their role in head movement and maintenance of posture. Can J Physiol Pharmacol. 1977; 55(3): 332-8. doi:10.1139/y77-047.
[43] McLain RF. Mechanoreceptor endings in human

cervical facet joints. Spine (Phila Pa 1976). 1994; 19(5): 495-501. doi: 10.1097/00007632-199403000-00001.

[44] Kulkarni V, Chandy MJ, Babu KS. Quantitative study of muscle spindles in suboccipital muscles of human foetuses. Neurol India. 2001; 49(4): 355-9.

[45] Proske U, Gandevia SC. The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. Physiol Rev. 2012; 92(4): 1651-97. doi:10.1152/physrev.00048.2011.

[46] Grgić V. Cervicogenic headache: etiopatho genesis, characteristics, diagnosis, differential diagnosis and therapy. Lijec Vjesn. 2007; 129(6-7): 230-6.

[47] Tropp H. Commentary: Functional ankle instability revisited. J Athl Train. 2002; 37(4): 512-5.

[48] McPartland JM, Brodeur RR, Hallgren RC. Chronic neck pain, standing balance, and suboccipital muscle atrophy-a pilot study. J Manipulative Physiol Ther. 1997; 20(1): 24-9.

[49] Madeleine P, Prietzel H, Svarrer H, ArendtNielsen L. Quantitative posturography in altered sensory conditions: a way to assess balance instability in patients with chronic whiplash injury. Arch Phys Med Rehabil. 2004; 85(3): 432-8. doi:10.1016/j.apmr.2003.03.003.

[50] Panjabi MM. The stabilizing system of the spine. Part I. Function, dysfunction, adaptation, and enhancement. J Spinal Disord. 1992; 5(4): 383-97. doi:10.1097/00002517-199212000-00001.

[51] Check JH, Cohen R, Katsoff B, Check D. Hypofunction of the sympatheticnervous system is an etiologic factor for a wide variety of chronic treatment-refractory pathologic disorders which all respond to therapy with sympathomimetic amines. Med Hypotheses. 2011; 77(5): 717-25. doi:10.1016/j.mehy.2011.08.024.

[52] Cohen H, Neumann L, Shore M, Amir M, Cassuto Y, Buskila D. Autonomic dysfunction in patients with fibromyalgia: application of power spectral analysis of heart rate variability. Semin Arthritis Rheum. 2000; 29(4): 217-27. doi:10.1053/sarh.2000.3621.

[53] Hallman DM, Lyskov E. Autonomic regulation, physicalactivity and perceived stress in subjects with musculoskeletal pain: 24-hour ambulatory monitoring. Int J Psychophysiol. 2012; 86(3): 276-82. doi:10.1016/j.ijpsycho.2012.09.004.

[54] Fonseca D, Beda A, Sá A, Simpson D. Gain and coherence estimates between respiration and heart-rate: Differences between inspiration and expiration. Auton Neurosci. 2013; 178: doi:10.1016/j.autneu.2013.02.013.

[55] Restrepo CS, Eraso A, Ocazionez D, Lemos J, Martinez S, Lemos DF. The diaphragmatic crura and retrocrural space: normal imaging appearance, variants, and pathologic conditions. Radiographics. 2008; 28(5): 1289-305. doi:10.1148/rg.285085022.

[56] Gest TR, Hildebrandt S. The pattern of the thoracic splanchnic nerves as they pass through the diaphragm. Clin Anat. 2009; 22(7): 809-14. doi:10.1002/ca.20867.

[57] Zou Y, Zhao X, Hou YY, Liu T, Wu Q, Huang YH, et al. Meta-analysis of effects of voluntary slow breathing exercises for control of heart rate and blood pressure in patients with cardiovascular diseases. Am J Cardiol. 2017; 120(1): 148-53. doi:10.1016/j.amjcard.2017.03.009.

[58] Loukas M, Klaassen Z, Merbs W, Tubbs RS, Gielecki J, Zurada A. A review of the thoracic splanchnic nerves and celiac ganglia. Clin Anat. 2010; 23(5): 512-22. doi:10.1002/ca.21018.

[59] Loukas M, Du Plessis M, Louis RG, Tubbs RS, Wartmann CT, Apaydin N. The subdiaphragmatic part of the phrenic nerve - morphometry and connections to autonomic ganglia. Clin Anat. 2016; 29(1): 120-8. doi:10.1002/ca.22577.

[60] Iams J. When reflex rule: A new paradigm in understanding why some people don't get well [Internet]. Primal Reflex Release Technique; [cited 2021 Mar 29]. Available from: https://theprrt.com/pain-relief-articles/.

[61] Fantazzi F. Primal Reflex Release Technique Part 1 - PRRT Physical Therapy - CyberPT [Internet]. [cited 2021 Mar 22]. Available from: http://www.cyberpt.com/prrt.asp.

[62] Javanshir K, Ortega-Santiago R, Mohseni-Bandpei MA, Miangolarra-PageJC, Fernández-de-Las-Peñas C. Exploration of somatosensory impairments in subjects with mechanical idiopathic neck pain: a preliminary study. J Manipulative Physiol Ther. 2010; 33(7): 493-9. doi:10.1016/j.jmpt.2010.07.005.

[63] Wehrwein EA, Joyner MJ. Chapter 8 - Regulation of blood pressure by the arterial baroreflex and autonomic nervous system. In: Buijs RM, Swaab DF, Editors. Handbook of Clinical Neurology [Internet]. Elsevier; 2013 [cited 2025 Feb 16]. pp 89-102. (Autonomic Nervous System; vol. 117). Available from: https://www.sciencedirect.com/science/article/pii/B9780444534910000080.

[64] Sueki DG, Cleland JA, Wainner RS. A regional interdependence model of musculoskeletal dysfunction: research, mechanisms, and clinical implications. J Man Manip Ther. 2013; 21(2): 90-102. doi:10.1179/2042618613Y.0000000059.

[65] Myers TW. Anatomy Trains: Myofascial Meridians for Manual and Movement Therapists. Elsevier Health Sciences; 2009: pp 308.

Journal of Associated Medical Sciences 2026; 59(1): 9-17

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

Single time-point kidney dosimetry in ^{177}Lu-PSMA therapy: A comparison between AI-based and manual segmentation approaches

Thitaya Chaiwongsa[1], Putthiporn Charoenphun[1,2], Wichana Chamroonrat[2], Krisanat Chuamsaamarkkee[1,2*]

[1]Master of Science Program in Medical Physics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.
[2]Division of Nuclear Medicine, Department of Diagnostic and Therapeutic Radiology, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 15 May 2025
Accepted as revised 21 August 2025
Available online 9 September 2025

Keywords:
Single time point dosimetry, ^{177}Lu-PSMA, kidneys absorbed dose, kidneys segmentation.

ABSTRACT

Background: Single time-point (STP) dosimetry has become a practical and efficient approach for personalised radioligand therapy (RLT), with 48-hours post-injection identified as optimal for kidney dose estimation in ^{177}Lu-PSMA therapy for prostate cancer. However, segmentation accuracy remains a critical factor affecting dosimetry reliability. AI-based segmentation has recently been integrated into commercial software to improve efficiency and reduce variability.

Objectives: This study aims to quantify kidney absorbed doses in patients receiving ^{177}Lu-PSMA therapy using STP dosimetry and to compare the accuracy and consistency of AI-based segmentation versus manual segmentation techniques.

Materials and methods: Eight treatment cycles from 5 patients of ^{177}Lu-PSMA were retrospectively analysed. In this work, whole-body SPECT/CT imaging was performed approximately 48 hours post-injection. Then, kidney dosimetry was calculated using voxel-based STP (Hänscheid method) within MIM SurePlan™ MRT software. Kidney volumes of interest (VOIs) were segmented using three approaches: 1) AI-based automatic segmentation, 2) AI-based with manual refinement, and 3) fully manual segmentation. Mean absorbed doses and VOI volumes were compared across methods. Statistical analyses included ANOVA, Dice Similarity Coefficient (DSC), and Jaccard Similarity Coefficient (JSC).

Results: No significant differences in mean kidney absorbed doses were found across segmentation methods (p=0.964), while kidney VOI volumes showed significant variation (p<0.05). AI-based segmentation achieved high concordance with manual delineation (DSC: 0.898±0.019; JSC: 0.816±0.031).

Conclusion: AI-based segmentation provides comparable absorbed dose results to manual segmentation, with reduced time and inter-observer variability.

Introduction

While the Medical Internal Radiation Dose (MIRD) Committee recommends multi-time-point (MTP) imaging to determine time-integrated activity (TIA) and absorbed dose accurately, MTP protocols are often impractical in routine clinical settings due to time and resource constraints.[8] To address this, Hänscheid et al. proposed a simplified single time-point (STP) method that estimates absorbed dose using one single quantitative SPECT/CT image and an assumed effective half-life to address.[9] Recent studies, including those by Brosch-Lenz et al.,[10] have validated

* **Corresponding contributor.**
Author's Address: *Master of Science Program in Medical Physics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.*
E-mail address: Krisanat.chu@mahidol.ac.th
doi: 10.12982/JAMS.2026.002
E-ISSN: 2539-6056

the accuracy of kidney absorbed dose estimation using STP at approximately 48 hours post-injection, showing minimal deviation compared to MTP-derived references for ^{177}Lu-PSMA in prostate cancer patients, which was in the same way as other related STP dosimetry studies.[11-13]

One of the primary sources of uncertainty in image-based dosimetry is organ segmentation. Accurate delineation of the kidneys directly impacts the volume and activity quantification, and thus the calculated absorbed dose.[14,15] Although widely used, manual segmentation is time-consuming and operator-dependent, leading to inter-observer variability and potential inconsistencies.[16-18] In contrast, artificial intelligence (AI)-based segmentation, particularly using deep learning (DL), offers the potential for rapid, reproducible, and standardized delineation, and has been increasingly integrated into commercial dosimetry platforms such as MIM SurePlan™ MRT (version 7.3.6) and Hermes Voxel-based Dosimetry (version 3.0).[18-20]

Therefore, this study aims to quantify kidney absorbed doses in patients receiving ^{177}Lu-PSMA therapy using STP dosimetry and to compare the accuracy and consistency of AI-based segmentation versus manual segmentation techniques.

Materials and methods

Patient selection

This study was conducted on eight treatment cycles in five ^{177}Lu-PSMA patients treated at the Department of Nuclear Medicine, Faculty of Medicine, Ramathibodi Hospital, Bangkok, Thailand, from January 2024 to December 2024. This retrospective study was approved by the local ethics committee of the Faculty of Medicine Ramathibodi Hospital, Mahidol University. Patients were aged 58-88 years (mean: 67.49±11.56 years). The mean administration activity was 6731.57±1026.49 MBq. Whole-body single time-point SPECT/CT scans were performed approximately 48 hours post-injection (mean: 49.18±2.61 hours post administration) as shown in Table 1. Patient data were tabulated.

Table 1. Patient demographic Information.

Overall (N=8)	Age (years)	Administration activity (MBq)	Acquisition time point (h p.i.)
Mean±SD	67.49±11.56	6731.57±1026.49	49.18±2.61
Median	65.00	6710.15	50.22
Range	58.00-88.00	5607.70-8103.00	43.77-51.95

Image acquisition protocol

A dual-head hybrid SPECT/CT scanner (GE Discovery 870 DR, GE Healthcare, MI, USA) equipped with a Medium Energy General Purpose (MEGP) collimator was used for image acquisition. Each bed position was acquired using 60 frames per head at 5 seconds per frame, resulting in a total acquisition time of approximately 7 minutes.

Quantitative image reconstruction was performed using an Ordered Subset Expectation Maximisation (OSEM) algorithm in Hermes Hybrid Recon 3.0 (Hermes Medical Solutions, Sweden), with 16 iterations, 9 subsets, and a matrix size of 128×128.

To correct for collimator-detector response, a Gaussian model was applied during reconstruction, and the specific parameters for the MEGP collimator of the GE Discovery 870 DR were entered into the software. CT images for attenuation correction were acquired using a low-dose protocol to minimise additional radiation exposure.

Voxel-based single time-point dosimetry

Voxel S-value dosimetry for the kidneys was performed using MIM SurePlan™ MRT dosimetry software (version 7.3.6; GE Healthcare, MI, USA). Single time-point dosimetry, which estimates the time-integrated activity (TIA) based on monoexponential decay fitting using the effective half-life (T_{eff}), was applied as proposed by Hänscheid et al.[10] as shown in Equation 1.

$$TIA \approx \frac{A(t)\cdot 2^{\frac{t}{T_{eff}}}\cdot T_{eff}}{\ln(2)} \qquad (1)$$

If the imaging time point was interval 0.75 to 2.5, STP TIA had less than 10% error compared with the MTP TIA.[10,11] STP TIAs were calculated using the measured activity at approximately 48 hours post-injection for ^{177}Lu-PSMA.

For kidneys STP dosimetry, the volume of interest (VOI) was contoured, limited to the renal cortex on each axial slice as illustrated in Figure 1. This VOI was used to determine both the organ volume (in mL) and the mean absorbed dose (in Gy).

To evaluate the effect of segmentation methods on dosimetric outcomes, three different approaches were applied for kidney delineation:

(a) AI-based automatic segmentation

This method utilized the deep learning algorithm Contour ProtégéAI®, integrated within MIM SurePlan™ MRT software (version 7.3.6). The segmentation process was fully automated and initiated upon image import. The AI algorithm was trained to segment normal organs (including kidneys) based on CT datasets, providing a consistent and time-efficient alternative to manual contouring.

(b) AI-based segmentation with manual modification

In this approach, the automatically generated VOIs from Contour ProtégéAI® were reviewed and manually adjusted by a trained operator to improve anatomical accuracy, particularly in regions where the AI model may have under- or over-segmented the renal cortex. This method represents a hybrid of AI support with human expert refinement.

(c) *Manual segmentation (reference standard)*

Full manual segmentation was performed on CT images by a medical physicist with one year of experience in dosimetry. The segmentation was conducted slice-by-slice, with careful delineation of the renal cortex. This method served as the reference standard for comparison. Importantly, all manual segmentations were performed independently and blinded to the AI-generated results, ensuring unbiased evaluation.

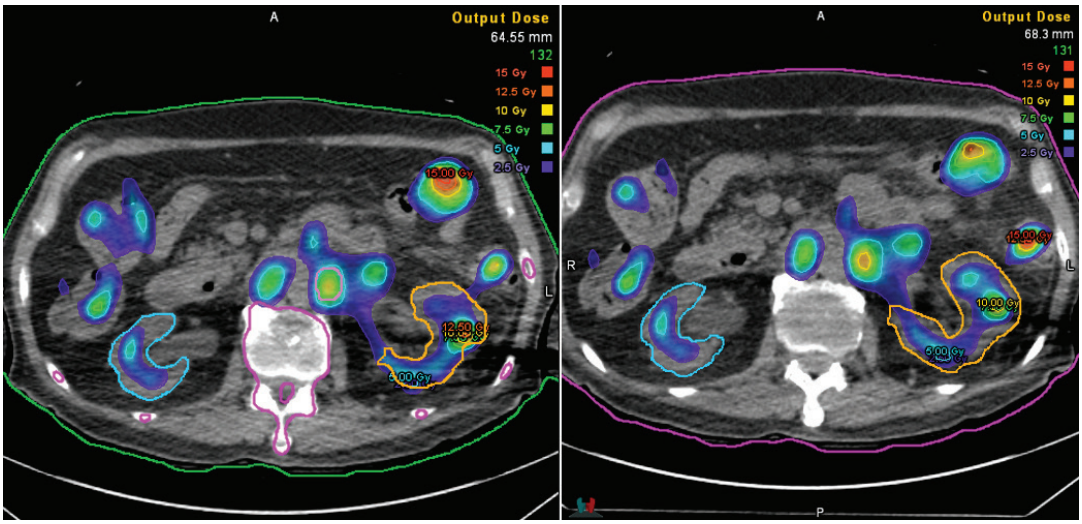


Figure 1. *Kidneys contouring. Left: AI-based segmentation, and right: manual segmentation.*

Each segmentation method was subsequently used to calculate the kidney absorbed dose using the MIM SurePlan MRT dosimetry workflow, which incorporates quantitative SPECT/CT data and voxel-based dose computation. The comparison among segmentation methods aimed to assess the consistency in dosimetric outcomes and segmentation accuracy. Specifically, the mean absorbed dose and contouring volume of the kidneys were compared across the three segmentation techniques using percentage difference. For evaluating segmentation accuracy, Dice Similarity Coefficient (DSC) and Jaccard Similarity Coefficient (JSC) were calculated to quantify the overlap between segmented volumes. These metrics were computed using Python (Version 3.13.2).

Statistical analysis

Following kidneys dosimetry, dosimetric parameters from each segmentation method, including absorbed dose and contouring volume, were reported. Comparisons of kidney absorbed doses and contouring volume across segmentation methods were conducted using One-way analysis of variance (ANOVA) in IBM SPSS Statistics. The $p<0.05$ was considered statistically significant.

Results

Kidney dosimetry and contouring volume

The kidney mean-absorbed dose, estimated using the single time-point (STP) method, along with corresponding contouring volumes, was evaluated using three segmentation approaches: manual, AI-assisted, and AI-assisted with manual adjustment. Details of these measurements, including percentage differences compared to manual segmentation (used as the reference), are presented in Table 2. Additionally, segmentation performance was assessed using Dice Similarity Coefficients (DSC) and Jaccard Similarity Coefficients (JSC), as shown in Table 3.

Table 2. Mean absorbed dose (Gy) and kidney volume (mL), including percentage differences (%), derived from manual segmentation and two alternative methods: AI-based segmentation and AI-based segmentation with manual modification, for each patient.

Study No.	Mean absorbed dose (Gy) (Percentage difference)*			Contouring volume (mL) (Percentage difference)*		
	AI-based with			AI-based with		
	Manual segmentation	Manual modification segmentation	AI-based segmentation	Manual segmentation	Manual modification segmentation	AI-based segmentation
1	1.82	2.00 (9.42)	2.04 (11.40)	446.13	366.85 (-19.50)	332.13 (-29.30)
2	1.84	1.93 (4.77)	1.95 (5.80)	376.59	337.62 (-10.91)	321.49 (-15.79)
3	1.61	1.70 (5.44)	1.72 (6.61)	404.33	363.63 (-10.60)	347.34 (-15.16)
4	2.57	2.71 (5.30)	2.76 (7.13)	290.12	254.82 (-12.96)	247.96 (-15.67)
5	5.17	5.37 (3.80)	5.37 (3.80)	342.36	267.60 (-21.25)	276.60 (-24.51)
6	2.89	3.05 (5.39)	3.11 (7.33)	303.30	271.01 (-11.24)	254.67 (-17.43)
7	2.47	2.49 (0.81)	2.50 (1.21)	354.79	351.57 (-0.91)	346.92 (-2.24)
8	2.09	2.21 (5.58)	2.22 (6.03)	367.13	337.98 (-8.27)	329.44 (-10.82)
Mean±SD	2.39±1.41	2.52±1.17	2.54±1.16	357.46±50.91	315.79±46.47	309.94±39.08

Note: The number in brackets quantifies the relative change as a percentage derived from manual segmentation.

Table 3. Dice similarity coefficient and Jaccard similarity coefficient comparing manual segmentation with two alternative methods—AI-based segmentation and AI-based segmentation with manual modification—for kidney volumes in each patient.

Study No.	Dice similarity coefficient (DSC)		Jaccard similarity coefficient (JSC)	
	AI-based segmentation	AI-based with manual modification segmentation	AI-based segmentation	AI-based with manual modification segmentation
1	0.867	0.881	0.765	0.788
2	0.913	0.928	0.840	0.866
3	0.920	0.933	0.849	0.875
4	0.908	0.915	0.831	0.844
5	0.905	0.926	0.826	0.862
6	0.913	0.930	0.839	0.869
7	0.873	0.890	0.774	0.802
8	0.893	0.903	0.807	0.824
Mean±SD	0.898±0.019	0.9131±0.020	0.816±0.031	0.840±0.033

Comparison of absorbed dose and contouring volume

Figures 2 and 3 show the distribution of kidney absorbed dose across the three segmentation techniques, both overall and for individual cases. The AI-assisted method yielded the highest mean absorbed dose (2.54±1.16 Gy), followed closely by the AI-assisted with manual adjustment (2.52±1.17 Gy), while the manual method resulted in a slightly lower mean value (2.39±1.41 Gy).

A similar trend was observed in kidney contouring volumes (Figure 4). The manual approach produced the largest volume (357.46±50.91 mL), followed by AI with manual adjustment (315.79±46.47 mL), and then AI-only segmentation (309.94±39.08 mL).

DSC and JSC values, which reflect how well the automated contours matched the manually segmented ones, are summarised in Table 3. Overall, segmentation with manual adjustment outperformed AI-only: the average DSC improved from 0.898±0.019 to 0.913±0.020, and the JSC from 0.816±0.031 to 0.840±0.033. These differences are visualised as box plots in Figures 5 and 6.

Statistical comparison

A one-way ANOVA was conducted to examine differences in absorbed dose and kidney volume among the three segmentation methods, as shown in Table 4.

While the absorbed dose did not significantly differ between methods ($p>0.05$), kidney volumes did show statistically significant differences ($p<0.05$), particularly between manual and AI-only segmentation. This suggests that while absorbed dose estimates remained consistent, the accuracy of kidney contouring varied depending on the segmentation approach used.

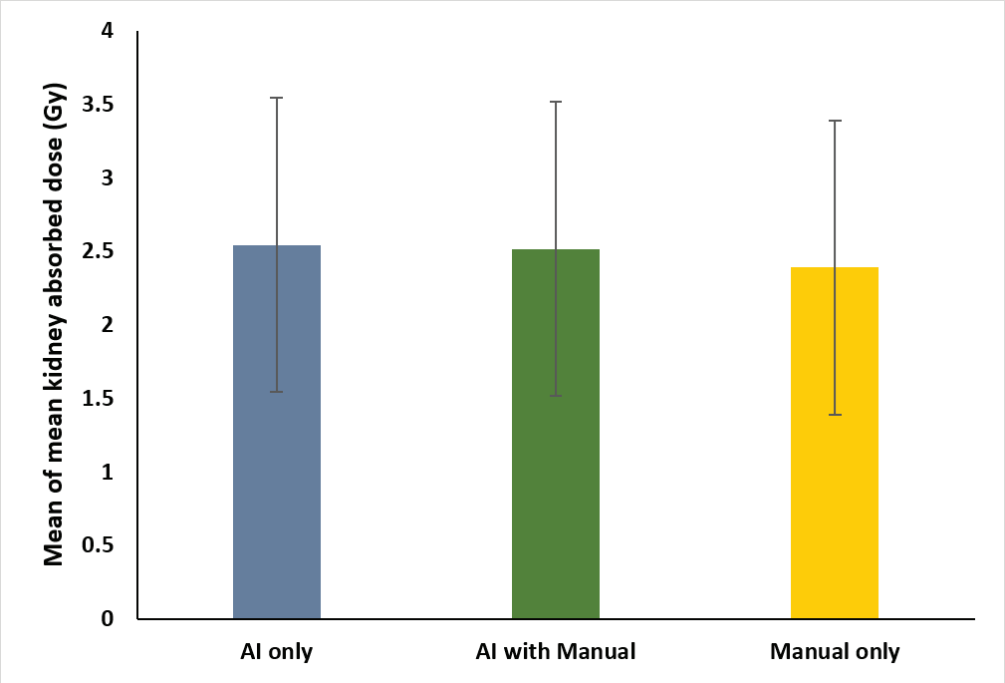


Figure 2. *Comparison of the mean kidney absorbed dose (±SD) obtained using three different segmentation methods: AI-based segmentation, AI-based segmentation with manual modification, and manual segmentation.*

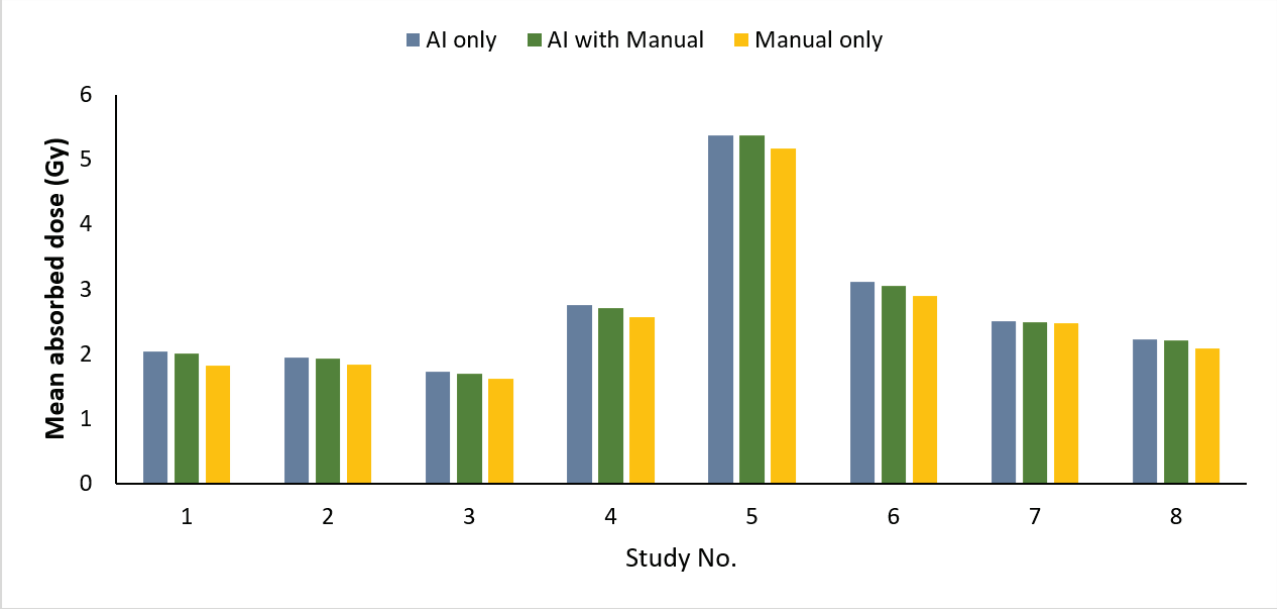


Figure 3. *Comparison of the mean kidney absorbed dose obtained using three segmentation methods: AI-based segmentation, AI-based segmentation with manual modification, and manual segmentation across individual studies.*

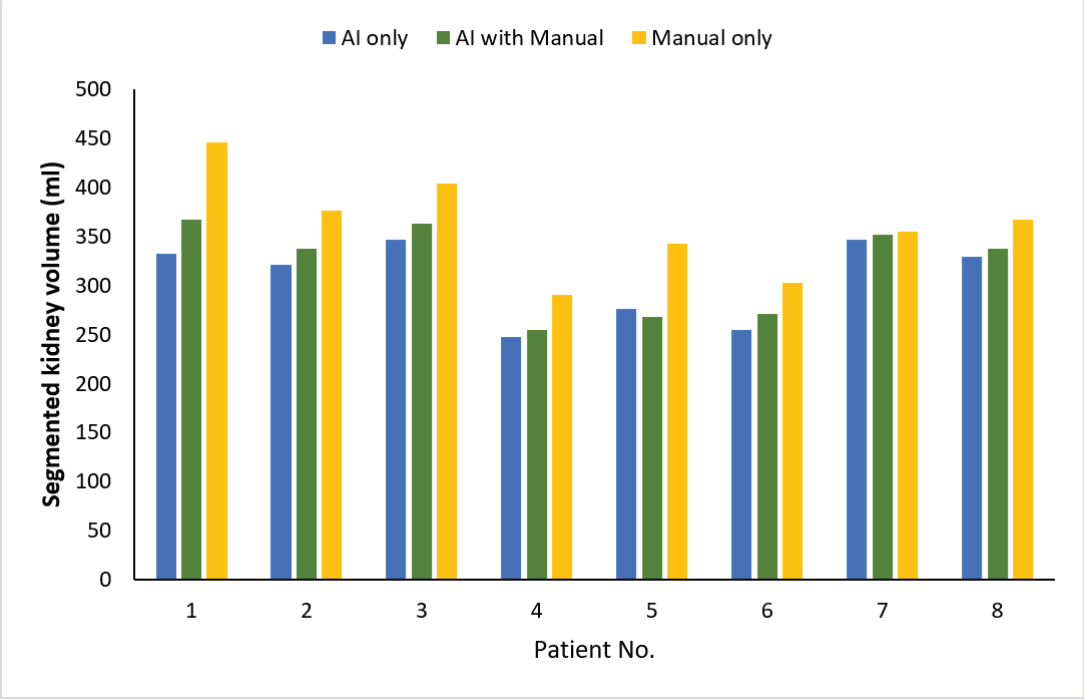


Figure 4. *Comparison of the mean kidney contouring volume (±SD) obtained using three segmentation methods: AI-based segmentation, AI-based segmentation with manual modification, and manual segmentation.*

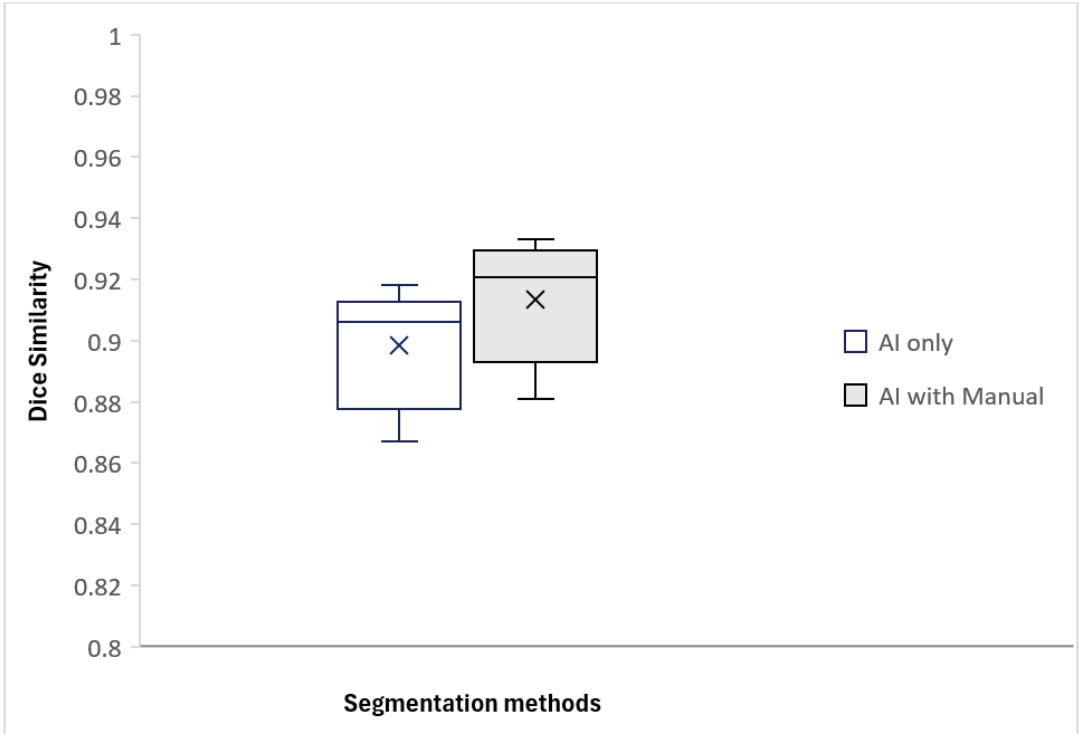


Figure 5. *Box plot of dice similarity coefficient (DSC) for kidney contouring volumes obtained from AI-based segmentation and AI-based segmentation with manual modification.*

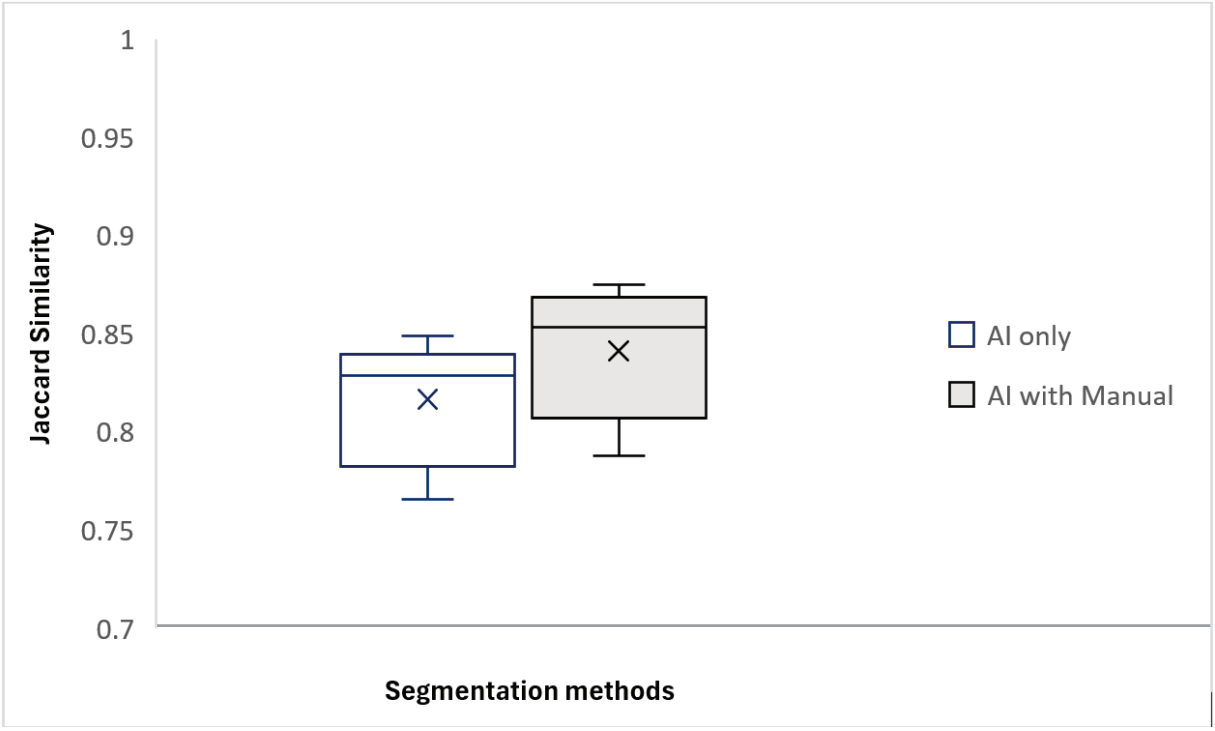


Figure 6. *Box plot of Jaccard Similarity Coefficient (JSC) for kidney contouring volumes obtained from AI-based segmentation and AI-based segmentation with manual modification.*

Statistical comparison

A one-way ANOVA was conducted to examine differences in absorbed dose and kidney volume among the three segmentation methods, as shown in Table 4. While the absorbed dose did not significantly differ between methods ($p>0.05$), kidney volumes did show statistically significant differences ($p<0.05$), particularly between manual and AI-only segmentation. This suggests that while absorbed dose estimates remained consistent, the accuracy of kidney contouring varied depending on the segmentation approach used.

Table 4. The p values from one-way ANOVA comparing kidney mean absorbed dose and contouring volume obtained using manual segmentation, AI-based segmentation, and AI-based segmentation with manual modification.

	Manual vs AI-based	Manual vs AI with manual modification	AI-based vs AI with manual modification	All segmentation methods
Mean absorbed dose	0.813	0.821	0.992	0.964
Contouring volume	$p<0.05$	0.074	0.477	$p<0.05$

Discussion

In this study, three different kidney segmentation approaches for single time-point (STP) dosimetry in patients treated with ^{177}Lu-PSMA was investigated: manual segmentation by a medical physicist, fully automated segmentation, and automated segmentation with manual refinement.

When comparing these methods, the mean absorbed dose to the kidneys did not show a statistically significant difference ($p>0.05$). The percentage difference in absorbed dose between AI-based segmentation and manual segmentation was approximately 6% (range: 0%-11%), while the difference between AI-only and AI with manual modification was about 5% (range: 0%-9%).

These findings align with previous work by Dewaraja et al. who evaluated deep-learning-based kidney segmentation against manual contours in a multiple time-point (MTP) SPECT/CT study following ^{177}Lu-PRRT.[21] They reported mean absorbed dose differences of 3% for CNN-only segmentation and 2% (range: 0%-4%) with manual segmentation. The slightly higher variability observed in our study may be attributed to the use of STP dosimetry based on time-integrated activity (TIA) estimation from a STP approach, while the previous study used MTP. Nonetheless, our results

demonstrated that STP dosimetry performed around 48 hours post-injection can achieve a mean error of less than 10%, consistent with the recommendation by Hänscheid et al.,[9] and supported by Resch et al.,[22] as a practical alternative to MTP approach.

In terms of segmentation accuracy, our study reported average DSC values of 0.898 for AI-only segmentation and 0.913 for AI with manual modification when compared with manual segmentation. These results are comparable to those reported by Dewaraja et al. who found a DSC of 0.91 for deep-learning-only segmentation and 0.93 when manual refinement was included.[21]

Despite the similar absorbed dose estimates, a statistically significant difference was found in the kidney contouring volumes, particularly between manual and AI-only methods. Manual contours tended to include larger volumes, likely due to added margins to account for spill-out activity, limited spatial resolution (partial volume effects), and CT-related artefacts. These factors may lead to underestimation of the absorbed dose when smaller volumes are contoured, potentially impacting dose-response evaluations. This observation is supported by Nazari et al. who reported interobserver variability in manual segmentation due to inconsistent inclusion of spill-out activity beyond CT-defined anatomical boundaries in kidney and liver VOIs during deep-learning dosimetry evaluation. [23]

Limitations

This study has several limitations. First, the patient cohort was relatively small, which may limit the generalisability of the findings. Second, single time-point (STP) dosimetry was not validated against multi-time-point (MTP) dosimetry, as MTP was not routinely performed at our institution as a baseline for voxel-based dosimetry. Lastly, manual segmentation was conducted by a single observer, which may introduce observer bias and limit assessment of interobserver variability.

Conclusion

This study demonstrated that single time-point (STP) dosimetry is a promising approach for voxel-based dosimetry in ^{177}Lu-PSMA therapy and can be feasibly implemented in routine clinical practice. While different segmentation methods resulted in variations, particularly in the segmented kidney volume, the absorbed dose estimates remained comparable across methods. AI-based segmentation produced results consistent with manual segmentation, while offering significant advantages in terms of time efficiency, reliability, accuracy, and consistency, making it a robust alternative for clinical dosimetry workflows.

Ethical approval

This study was approved by the Human Ethics Committee of the Faculty of Medicine Ramathibodi Hospital, Mahidol University (COA No. MURA2025/422).

Funding

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

Conflict of interest

There is no conflict of interest.

CRediT authorship contribution statement

Thitaya Chaiwongsa: conducted the experiments, analyzed the data, and drafted the manuscript; **Putthiporn Charoenphun:** co-supervised the study; **Wichana Chamroonrat:** supervised the clinical part and verified the VOIs in the study; **Krisanat Chuamsaamarkkee:** drafted the manuscript, prepared the figures and tables, and served as the corresponding author.

Acknowledgements

We would like to express our sincere thanks to Transmedic (Thailand) for providing the MIMS demo workstation to the Division of Nuclear Medicine, Department of Diagnostic and Therapeutic Radiology, Faculty of Medicine Ramathibodi Hospital.

References

[1] Meyer C, Szidonya L, Winters C, Mench A, Mallak N, Mittra E. Quantitative imaging for 177Lu-PSMA treatment response monitoring and dosimetry. Front Nucl Med. 2023; 3: 1291253. doi: 10.3389/fnume.2023.1291253.

[2] Violet J, Jackson P, Ferdinandus J, Sandhu S, Akhurst T, Iravani A, et al. Dosimetry of 177Lu-PSMA-617 in metastatic castration-resistant prostate cancer: Correlations between pretherapeutic imaging and whole-body tumor dosimetry with treatment outcomes. J Nucl Med. 2019; 60(4): 517-23. doi:10.2967/jnumed.118.219352.

[3] Juzeniene A, Stenberg VY, Bruland ØS, Larsen RH. Preclinical and clinical status of PSMA-targeted alpha therapy for metastatic castration-resistant prostate cancer. Cancers (Basel). 2021; 13(4): 779. doi:10.3390/cancers13040779.

[4] Alsadi R, Djekidel M, Bouhali O, Doherty JO. Towards routine clinical use of dosimetry in [177Lu] Lu-PSMA prostate cancer radionuclide therapy: Current efforts and future perspectives. Front Phys. 2022; 10: 1-13 doi:10.3389/fphy.2022.940677.

[5] Ramonaheng K, Qebetu M, Ndlovu H, Swanepoel C, Smith L, Mdanda S, et al. Activity quantification and dosimetry in radiopharmaceutical therapy with reference to 177Lutetium. Front Nucl Med. 2024; 4: 1355912. doi:10.3389/fnume.2024.1355912.

[6] Brosch-Lenz J, Uribe C, Gosewisch A, Kaiser L, Todica A, Ilhan H, et al. Influence of dosimetry method on bone lesion absorbed dose estimates

in PSMA therapy: application to mCRPC patients receiving Lu-177-PSMA-I&T. EJNMMI Phys. 2021; 8(1): 26. doi:10.1186/s40658-021-00369-4.

[7] Gustafsson J, Sundlöv A, Sjögreen Gleisner K. SPECT image segmentation for estimation of tumour volume and activity concentration in 177Lu-DOTATATE radionuclide therapy. EJNMMI Res. 2017; 7(1): 18. doi: 10.1186/s13550-017-0262-7.

[8] Siegel JA, Thomas SR, Stubbs JB, Stabin MG, Hays MT, Koral KF, et al. MIRD pamphlet no. 16: Techniques for quantitative radiopharmaceutical biodistribution data acquisition and analysis for use in human radiation dose estimates. J Nucl Med. 1999;40(2): 37S-61S.

[9] Hänscheid H, Lapa C, Buck AK, Lassmann M, Werner RA. Dose mapping after endoradiotherapy with 177Lu-DOTATATE/DTATOC by a single measurement after 4 days. J Nucl Med. 2018; 59(1): 75-81. doi:10.2967/jnumed.117.193706.

[10] Brosch-Lenz J, Delker A, Völter F, Unterrainer LM, Kaiser L, Bartenstein P, et al. Toward single-time-point image-based dosimetry of 177Lu-PSMA-617 therapy. J Nucl Med. 2023; 64(5): 767-74. doi:10.2967/jnumed.122.264594.

[11] Kurth J, Heuschkel M, Tonn A, Schildt A, Hakenberg OW, Krause BJ, et al. Streamlined schemes for dosimetry of 177Lu-labeled PSMA targeting radioligands in therapy of prostate cancer. Cancers (Basel). 2021; 13(15): 3884. doi:10.3390/cancers13153884.

[12] Jackson PA, Hofman MS, Hicks RJ, Scalzo M, Violet J. Radiation dosimetry in 177Lu-PSMA-617 therapy using a single treatment SPECT/CT scan: A novel methodology to generate time- and tissue-specific dose factors. J Nucl Med. 2020; 61(7): 1030-6. doi:10.2967/jnumed.119.233411.

[13] Peters SMB, Mink MCT, Privé BM, de Bakker M, de Lange F, Muselaers CHJ, et al. Optimization of the radiation dosimetry protocol in Lutetium-177-PSMA therapy: toward clinical implementation. EJNMMI Res. 2023; 13(1): 6. doi:10.1186/s13550-023-00952-z.

[14] Gawel J, Rogulski Z. The challenge of single-photon emission computed tomography image segmentation in the internal dosimetry of 177Lu molecular therapies. J Imaging. 2024; 10(1): 27-40. doi: 10.3390/jimaging10010027.

[15] Makris NE, van Velden FH, Huisman MC,Menke CW, Lammertsma AA, Boellaard R. Validation of simplified dosimetry approaches in 89 Zr-PET/CT: The use of manual versus semi-automatic delineation methods to estimate organ absorbed doses. Medical Physics. 2014; 41(10): 102503. Doi: 10.1118/1.4895973.

[16] Vomacka L, Gosewisch A, Delker A, Fendler W, Bartenstein P, Boening G. Automatic image segmentation for 3D dosimetry in Lu-177 DKFZ-PSMA-617 therapy of castrate-resistant metastatic prostate cancer using a robust cluster algorithm on 4D SPECT data. J Nucl Med. 2016; 57(suppl2): 497.

[17] Hesamian MH, Jia W, He X, Kennedy P. Deep learning techniques for medical image segmentation: achievements and challenges. J Digit Imaging. 2019; 32(4): 582-96. doi: 10.1007/s10278-019-00227-x.

[18] Danieli R, Milano A, Gallo S, Veronese I, Lascialfari A, Indovina L, et al. Personalized dosimetry in targeted radiation therapy: A look to methods, tools and critical aspects. J Pers Med. 2022; 12(2): 205. doi: 10.3390/jpm12020205.

[19] Wendler T, Kreissl MC, Schemmer B, Rogasch JMM, De Benetti F. Artificial Intelligence-powered automatic volume calculation in medical images-available tools, performance and challenges for nuclear medicine. Nuklearmedizin. 2023; 62(6): 343-53. doi: 10.1055/a-2200-2145.

[20] MIM software Inc. MIM SurePlan™ MRT for Molecular Radiotherapy Dosimetry User Guide.

[21] Dewaraja YK, Mirando DM, Peterson AB, Niedbala J, Millet JD, Mikell JK, et al. A pipeline for automated voxel dosimetry: Application in patients with multi-SPECT/CT imaging after 177Lu-peptide receptor radionuclide therapy. J Nucl Med. 2022; 63(11): 1665-72. doi: 10.2967/jnumed.121.263738.

[22] Resch S, Ziegler SI, Sheikh G, Unterrainer LM, Zacherl MJ, Bartenstein P, et al. Impact of the reference multiple-time-point dosimetry protocol on the validity of single-time-point dosimetry for [177Lu]Lu-PSMA-I&T therapy. J Nucl Med. 2024; 65(8): 1272-8. doi: 10.2967/jnumed.123.266871.

[23] Nazari M, Jiménez-Franco LD, Schroeder M, Kluge A, Bronzel M, Kimiaei S. Automated and robust organ segmentation for 3D-based internal dose calculation. EJNMMI Res. 2021; 11(1): 53. doi: 10.1186/s13550-021-00796-5.

Journal of Associated Medical Sciences 2026; 59(1): 18-23

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Prevalence and hematological characteristics of bacterial vaginosis in postmenopausal women

Yenti Purnamasari[1,2*], Firdaus Hamid[3], Juminten Saimin[4], Agussalim Bukhari[5]

[1]*Doctoral Medical Science, Faculty of Medicine, Hasanuddin University, Makassar, Indonesia.*
[2]*Microbiology Department, Faculty of Medicine, Halu Oleo University, Kendari, Indonesia.*
[3]*Microbiology Department Faculty of Medicine, Hasanuddin University, Makassar, Indonesia.*
[4]*Obstetric and Gynecology Department, Halu Oleo University, Kendari, Indonesia.*
[5]*Nutrition Department Faculty of Medicine, Hasanuddin University, Makassar, Indonesia.*

ARTICLE INFO

Article history:
Received 21 May 2025
Accepted as revised 3 September 2025
Available online 9 September 2025

Keywords:
Bacterial vaginosis, menopause, haematology examination, inflammation, platelet-to-lymphocyte ratio.

ABSTRACT

Background: Bacterial vaginosis (BV) is a common vaginal dysbiosis associated with systemic inflammatory responses. However, its hematological impact in postmenopausal women remains unclear. This study investigates hematological parameters in postmenopausal women with BV to assess potential systemic inflammatory alterations.

Objectives: This study aimed to investigate various hematological parameters across different BV conditions to better understand their potential role in BV diagnosis and pathophysiology.

Materials and methods: A total of twenty-five postmenopausal women were recruited and categorized into three groups: bacterial vaginosis (BV), intermediate vaginal microbiota, and healthy vaginal microbiota based on Nugent scoring. Vaginal samples were collected aseptically with the participants in the lithotomy position by swabbing the vaginal walls circumferentially near the cervical fornix. Nugent scoring was performed on Gram-stained smears to classify the subjects. Subsequently, Verify® urinalysis reagent strips were directly applied to the vaginal wall to measure pH, protein, and glucose levels. Additionally, hematological parameters including leukocyte count, lymphocyte, eosinophil, neutrophil, monocyte, basophil, platelet counts, and platelet-to-lymphocyte ratio (PLR) were assessed from peripheral blood samples. All data were statistically analysed and compared among the three groups using ANOVA and Kruskal–Wallis tests, with a significance level set at $p<0.05$.

Results: The prevalence of BV in postmenopausal women was 72%. Hematology parameters did not show significant differences across BV, intermediate, and healthy groups ($p>0.05$). Leukocyte, neutrophil, and PLR values were slightly higher in BV cases, but not to a statistically significant degree. These findings suggest that BV in postmenopausal women may not elicit strong systemic inflammatory responses compared to premenopausal populations. Additionally, pH, glucose, and protein levels did not differ significantly among the groups, highlighting the need to reconsider standard vaginal health biomarkers in postmenopausal women due to physiological changes induced by menopause.

Conclusion: The prevalence of bacterial vaginosis (BV) in this study was higher than that reported in previous studies; however, it was not associated with significant hematological alterations. Further research is needed to identify reliable systemic and vaginal biomarkers for BV diagnosis in this population, considering the potential influence of hormonal and immunological factors.

* **Corresponding contributor.**
Author's Address: *Doctoral Medical Science, Faculty of Medicine, Hasanuddin University, Makassar, Indonesia.*
E-mail address: yenti.purnamasari@uho.ac.id
doi: 10.12982/JAMS.2026.003
E-ISSN: 2539-6056

Introduction

In postmenopausal women, the reduction in estrogen levels contributes to vaginal atrophy, decreased lactobacilli colonization, and a higher susceptibility to bacterial vaginosis (BV). It is characterized by an imbalance in the vaginal microbiota, where the dominance of *Lactobacillus* species is replaced by anaerobic bacteria such as *Gardnerella vaginalis*, *Atopobium vaginae*, and *Mobiluncus* species. This dysbiosis leads to an increased vaginal pH and a reduction in protective lactic acid production, which predisposes women to secondary infections and inflammatory conditions.[1,2] Although we understand the burden of BV in women of reproductive age, much less is known about the burden of BV in postmenopausal women.

The prevalence of BV varies globally, with estimates ranging from 2% to 57% among postmenopausal women, depending on the diagnostic criteria used.[3,4] While BV is often asymptomatic, it has been associated with serious health consequences, including an increased risk of sexually transmitted infections (STIs), pelvic inflammatory disease, and adverse pregnancy outcomes. Despite its clinical relevance, BV remains underdiagnosed in menopausal women due to the overlap of symptoms with other vaginal conditions such as atrophic vaginitis.[4]

Gold standard diagnostic methods for BV, such as Amsel's criteria and Nugent scoring, rely on clinical symptoms and microscopic evaluation.[5] However, these methods may be challenging to apply in postmenopausal women due to the discomfort associated with swab collection, which is exacerbated by vaginal atrophy.[6,7] Additionally, some postmenopausal women may find it difficult to assume the lithotomy position required for swab collection due to age-related musculoskeletal limitations. BV not only disrupts the vaginal microbiota and compromises vaginal health but may also contribute to systemic inflammatory responses, as indicated by alterations in hematological parameter. Although BV is primarily considered a localized condition of the lower genital tract, emerging evidence suggests that it may also elicit systemic inflammatory responses. Several studies have reported elevated serum levels of inflammatory markers such as C-reactive protein (CRP) and pro-inflammatory cytokines (e.g., IL-6, TNF-α) in women with BV, indicating a potential systemic immune activation beyond the vaginal environment.[8] However, studies explain hematological parameters in BV remain limited and inconclusive. In this study, we aim to investigate various hematological parameters across different BV conditions to better understand their potential role in BV diagnosis and pathophysiology.

Materials and methods

This cross-sectional study aims to evaluate the hematological profile as an indicator of BV status. The study population recruited menopausal women who attended the obstetrics and gynecology polyclinic of an independent clinic in Kendari City, Southeast Sulawesi, Indonesia, between April and June 2024 and consented to participate as research subjects. During the study period, 78 postmenopausal women attended the sampling site. Of these, 25 met the inclusion and exclusion criteria and were defined as the accessible population. All eligible subjects were recruited using a total sampling. We collected 25 samples with inclusion criteria consisted of menopausal women without a history of systemic disease and without symptoms indicative of reproductive tract infection, with menopause defined as the absence of menstruation for a minimum of 12 consecutive months.

Subject information and sample collection

Vaginal fluid specimens were obtained from participants using sterile cotton-tipped swabs using ESwab (COPAN Diagnostics, Murrieta, CA, USA) during a pelvic examination. Participants were placed in the lithotomy position to facilitate sample collection and direct application Vaginal swab at the vaginal wall with approximately 2 inches depth. The Verify® Urinalysis Reagent Strips (Verify®, Indonesia) were employed to evaluate vaginal secretions for pH, glucose, and protein levels.

BV assessment

Obstetricians and gynaecologists collected vaginal swabs according to established protocols. We performed Gram staining on all vaginal swabs to determine the Nugent score. The gram-stained smears were heat-fixed, and sequentially stained with crystal violet, iodine solution, decolorized with alcohol, and counterstained with safranin. The stained slides were examined under a microscope at 1000x magnification to assess the presence of *Lactobacillus* morphotypes and other which were then scored based on the Nugent criteria.[9]

pH, glucose and protein levels

A single reagent strip was directly applied to the lateral vaginal wall, ensuring contact between the reagent pads and the vaginal secretions for approximately 1-2 seconds. Care was taken to fully moisten the pads corresponding to pH, glucose, and protein parameters. The reagent strip was then withdrawn and held horizontally to prevent mixing of the reagents. Colour changes on the pads were compared to the manufacturer's reference colour chart at 60 seconds. The resulting colour intensities were recorded and categorized based on the manufacturer's semi-quantitative scale.

Hematological assessment

We collected venous blood samples to evaluate hematological parameters, including leukocytes, eosinophils, basophils, neutrophils, lymphocytes, monocytes, and platelets. Blood samples were drawn

from the antecubital vein using a sterile technique and collected into ethylenediaminetetraacetic acid (EDTA) tubes to prevent coagulation. Routine hematological analysis was performed using an automated haematology analyser to determine complete blood counts (CBC). The PLR was calculated by dividing the absolute platelet count by the absolute lymphocyte count. We need no specific patient preparation, such as fasting, was required prior to blood collection. All samples were processed promptly to ensure accuracy and reliability of the results.

Data analysis

Statistical analysis was performed to compare biochemical and hematological parameters among the BV, intermediate, and healthy groups. The associations between BV status and pH, protein, and glucose levels were analysed using the Chi-square test or Fisher's exact test, as appropriate. Hematological variables were compared using one-way ANOVA or the Kruskal–Wallis test, depending on data distribution. A $p<0.05$ was considered statistically significant.

Results

This study involved 25 postmenopausal women who consented to participate. The baseline characteristics of the participants are summarized in Table 1. We performed the BV diagnosis using the Nugent score, a well-established scoring system for classifying vaginal microbiota status based on Gram-stain microscopy. Bacterial vaginosis assessment according to Nugent score microscopy.[9] Health (score 0-3) defined as Gram-positive rods (lactobacilli) predominate representing normal vaginal flora; Intermediate (score 4-6) when decreased lactobacilli with increased presence of small Gram-variable rods (*Gardnerella vaginalis*) and Gram-negative curved rods; BV (score 7-10) when significant depletion of lactobacilli with predominance of Gram-variable coccobacilli (*Gardnerella vaginalis*) and curved Gram-negative rods (*Mobiluncus* species), characteristic of bacterial vaginosis (Figure 1).

This microscopic assessment serves as the gold standard for BV diagnosis according to Nugent criteria, which quantifies bacterial morphotypes to determine vaginal flora status. The shift from lactobacilli dominance to mixed anaerobic flora represents the hallmark dysbiosis of bacterial vaginosis. The results indicated that we categorized 18 samples (72%) as BV-positive, 5 samples (20%) as intermediate, and 2 samples (8%) as health (Table 2).

Table 1. Characteristics of the participants.

Characteristic	Total (N)	Percentage (%)
Age (year)		
≤50	5	23.07
>50	20	76.92
Menopause age (year)		
≤50	13	53.85
>50	12	46.15
Menopause duration (year)		
≤2	5	23.07
>2	20	76.92
Estradiol level (pg/mL)		
≤20	23	84.61
>20	4	15.38
Occupation		
Housewife	10	38.46
Employee	15	61.53
Education		
Middle	10	38.46
Hight	15	61.53
Parity		
<4	15	61.53
≥4	10	38.46

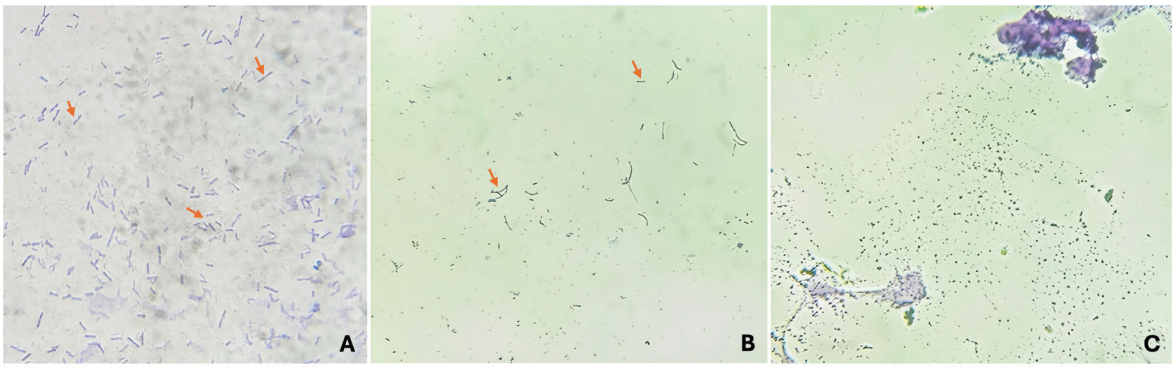


Figure 1. *Bacterial vaginosis assessment based on Nugent scoring. Arrows indicate of Lactobacillus sp (100x magnification), A: non BV (Lactobacillus-dominant), B: intermediate (Lactobacillus with others), C: BV (Lactobacillus-absent).*

Table 2. BV State and vaginal charge examination.

BV state	N (%)	pH		Glucose		Protein	
		<5	>5	Negative	Positive	≤2	>2
BV	18 (72%)	12 (66.7%)	6 (33.3%)	17 (94.4%)	1 (5.6%)	9 (50%)	9 (50%)
Intermediate	5 (20%)	4 (80%)	1 (20%)	5 (100%)	0 (0%)	4 (80%)	1 (20%)
Health	2 (8%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)	1 (50%)	1 (50%)
	p value	0.551		0.817		0.482	

Table 3. Variable analysis test with Anova.

Variable	BV (N=18)	Intermediate (N=5)	Health (N=2)	*p* value
Leucocyte	7.67±1.60	7.14±2.15	6.48±1.05	0.576
Lymphocyte	34.89±6.94	38.90±3.88	37.55±6.29	0.453
Eosinophil	3.50±1.94	3.66±1.24	3.30±1.13	0.970
Neutrophil	56.07±6.37	52.32±4.62	53.75±7.85	0.472
Monocyte	5.53±0.94	5.06±0.63	5.35±0.35	0.565
Basophil*	0.07±0.12	0.06±0.08	0.05±0.05	0.999
Platelet	281.28±61.39	231.20±14.62	231.50±14.85	0.143
PLR	8.66±3.60	5.96±0.26	6.22±0.46	0.156

**Kruskal wallis*

Discussion

The elevation of vaginal pH associated with the depletion of Lactobacillus in postmenopausal women creates a favourable environment for the proliferation of anaerobic pathogens. In postmenopausal women, decreased estrogen levels lead to a significant reduction in glycogen content within the vaginal epithelium. As a result, the growth of Lactobacillus species-key producers of lactic acid that help maintain an acidic vaginal environment-is impaired. The subsequent decline in Lactobacillus populations causes an elevation in vaginal pH, creating a more alkaline environment. This shift favours the proliferation of anaerobic and facultative anaerobic pathogens that are normally suppressed by the acidic milieu, thereby increasing the risk of bacterial vaginosis and other vaginal infections. The combined assessment of vaginal pH, glucose, and protein profiles offers valuable diagnostic insight and contributes to a deeper understanding of the complex pathophysiological mechanisms underlying bacterial vaginosis.[10-12]

In our study, we observed a higher prevalence of bacterial vaginosis (BV) among postmenopausal women compared to several previous reports.[4] Notably, our analysis did not reveal significant differences in vaginal pH, glucose, or protein levels among the BV, intermediate, and healthy groups. This finding may be attributed to the physiological changes associated with menopause, particularly the decline in estrogen levels, which leads to a naturally elevated vaginal pH and a reduction in Lactobacillus

species. These alterations can obscure the typical pH distinctions observed between BV and non-BV cases in premenopausal women.

The differential white blood cell (WBC) count reflects various components of the immune response. Neutrophils and lymphocytes, in particular, are key indicators of systemic inflammation and immune status. An increased neutrophil count often reflects acute inflammatory responses, while lymphocyte levels are associated with adaptive immunity. Changes in the ratio between these cell types-such as the neutrophil-to-lymphocyte ratio (NLR) or PLR-have been proposed as accessible biomarkers of systemic inflammation, including in gynaecologic and infectious conditions.[13]

A notable study published in *Scientific Reports* found that women with BV exhibited higher systemic inflammation markers, characterized by increased total WBC and lymphocyte counts suggest a link between BV and systemic inflammatory responses.[14] The analysis of hematological parameters in this study reveals no statistically significant differences among the BV, intermediate, and healthy groups (Table 3). Leukocyte counts were slightly higher in the BV group (7.67±1.60) compared to the intermediate (7.14±2.15) and healthy groups (6.48±1.05), but the difference was not significant (p=0.576). This finding suggests that systemic inflammatory responses in BV may not always be reflected in routine leukocyte counts. However, some studies have reported elevated leukocytes in BV cases, particularly in younger, premenopausal populations, possibly due to a more active immune response in the presence of bacterial dysbiosis.[13] Similarly, lymphocyte, eosinophil, neutrophil, monocyte, and basophil counts showed no significant variations among the groups, suggesting that BV-related inflammation may not elicit a marked systemic hematological response in postmenopausal women.

Notably, platelet levels were slightly higher in the BV group (281.28±61.39) compared to the intermediate (231.20±14.62) and healthy groups (231.50±14.85), though the difference was not statistically significant (p=0.143). The PLR, an emerging inflammatory marker, was also higher in the BV group (8.66±3.60) but did not reach statistical significance (p=0.156). PLR may reflect systemic inflammatory responses in various infectious and chronic conditions.[14,15] Its cut-off values vary depending on the condition, typically ranging from 121 to 232 cells/μL.[16] In our study, PLR values did not exceed the inflammatory thresholds reported in previous studies. Nevertheless, the noticeable differences observed between groups highlight the need for further investigation into its relevance in bacterial vaginosis, particularly in postmenopausal women.

Estrogen acts as an important immunomodulatory hormone, promoting anti-inflammatory responses and maintaining immune homeostasis. Its reduction leads to an increase in proinflammatory cytokines such as TNF-α, IL-1β, and IL-6, while decreasing anti-inflammatory mediators like IL-10. Additionally, menopause is characterized by immunosenescence, marked by a decline in naïve T cells, accumulation of memory T cells, and dysregulation of T-cell subsets, including increased effector T cells and decreased regulatory T cells. These changes contribute to a heightened risk of inflammatory and autoimmune conditions and may alter systemic and mucosal immune responses to infections, including bacterial vaginosis. Furthermore, menopause affects B-cell function, antibody production, and natural killer (NK) cell activity, collectively impacting the overall immune competence of postmenopausal women.[17,18]

The absence of significant hematological changes across the groups suggests that BV in postmenopausal women may not trigger the same systemic inflammatory response observed in younger populations. This could be due to the hormonal and immunological changes associated with menopause, which may alter the host response to bacterial imbalances.

Limitations

This study is limited by its small sample size and the use of a non-probability sampling technique, which may restrict the generalizability of the findings. However, considering the challenges in recruiting postmenopausal women for invasive procedures, the current sample represents the maximum achievable cohort under ethical and practical constraints.

Conclusion

The prevalence of bacterial vaginosis (BV) in this study was higher than that reported in previous studies; however, it was not associated with significant hematological alterations. Further research is needed to identify reliable systemic and vaginal biomarkers for BV diagnosis in this population, considering the potential influence of hormonal and immunological factors.

Ethical Approval

This study was approved by the health ethics commission of Hasanuddin University's faculty (B3/UN4.6.4.5.31-1/PP36/2024) adhered to the principles of the Declaration of Helsinki. All participants consented to engage in the research and publishing, we provided that their anonymity was guaranteed.

Funding

This study received financial support from the Centre for Higher Education Funding and Assessment, Ministry of Higher Education, Science, and Technology of the Republic of Indonesia.

Conflict of Interest

The authors declare no conflict of interest

CRediT authorship contribution statement

Yenti Purnamasari: conceptualization, methodology, data curation, writing: original draft; **Firdaus Hamid**:

supervision, project administration, validation, writing; review and editing; **Juminten Saimin:** investigation, formal analysis, resources, writing: review and editing; **Agussalim Bukhari:** visualization, supervision, writing; review and editing.

Acknowledgements

The authors extend their sincere gratitude to the postmenopausal women who generously contributed as research subjects, as well as the medical staff of the Obstetrics and Gynecology Clinic for their invaluable support and collaboration throughout the course of this study. The authors also thank the laboratory technicians for their technical assistance in sample processing and analysis, and the institutional ethics committee for their valuable approval and guidance.

References

[1] Verstraelen H, Swidsinski A. The biofilm in bacterial vaginosis: Implications for epidemiology, diagnosis and treatment: 2018 update. Curr Opin Infect Dis. 2019; 32(1): 38-42, doi: 10.1097/QCO.0000000000000516.

[2] Muzny CA, Schwebke JR. Pathogenesis of bacterial vaginosis: Discussion of current hypotheses. J Infect Dis. 2016; 214(Suppl 1): S1-S5, doi: 10.1093/infdis/jiw121.

[3] L. L. Stewart, L. A. Vodstrcil, J. Coombe, C. S. Bradshaw, J. S. Hocking. Prevalence of bacterial vaginosis in postmenopausal women: A systematic review and meta-analysis. Sex Health. 2022; 19(1): 17-26, doi: 10.1071/SH21083.

[4] O. T. Van Gerwen, S. E. Smith, C. A. Muzny. Bacterial vaginosis in postmenopausal women. Curr Infect Dis Rep. 2023; 25(1): 7-15, doi: 10.1007/s11908-022-00794-1.

[5] A. Challa, S. Sood, G. Kachhawa, A. D. Upadhyay, S. N. Dwivedi, S. Gupta. Diagnostic concordance between Amsel's criteria and the Nugent scoring method in the assessment of bacterial vaginosis. Sex Health. 2022; 18(6): 512-4, doi: 10.1071/SH21149.

[6] B. A. D. Genazzani, M. Milani. Postmenopausal Vaginal Atrophy : Non-hormonal therapy with a moisturizing bioadhesive vaginal gel. 08-0132_tr00E_M2[Internet]. 2003; 18:14-16. Available from: https://www.replens.eu/fr/wp-content/uploads/sites/2/2018/02/Atrophie-vaginale-post-m%C3%A9nopause.pdf.

[7] F. Palma, A. Volpe, P. Villa, A. Cagnacci. Vaginal atrophy of women in postmenopause. Results from a multicentric observational study: The AGATA study. Maturitas. 2016; 83: 40-4, doi: 10.1016/j.maturitas.2015.09.001.

[8] S. Cauci et al. Interrelationships of interleukin-8 with interleukin-1β and neutrophils in vaginal fluid of healthy and bacterial vaginosis positive women. Mol Hum Reprod. 2023; 9(1): 53-8, doi: 10.1093/molehr/gag003.

[9] R. P. Nugent, M. A. Krohn, S. L. Hillier. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J Clin Microbiol. 1991; 29(2): 297-301, doi: 10.1128/jcm.29.2.297-301.1991.

[10] S. Panda, A. Das, A. Singh, S. Pala. Vaginal pH: A marker for menopause. J Midlife Health. 2014; 5(1): 34, doi: 10.4103/0976-7800.127789.

[11] J. Saimin, Y. Purnamasari, S. A. Mulyawati, N. I. Purnamasari, Arismawati, F. Sandra. Vaginal acidity affects vaginal microbiota in postmenopausal women. Indones Biomed J. 2022; 14(2): 206-10, doi: 10.18585/inabj.v14i2.1836.

[12] A. Swidsinski et al. Antimicrobial susceptibility of microbiota in bacterial vaginosis using fluorescence In Situ hybridization. Pathogens. 2022; 11(4): 1-15, doi: 10.3390/pathogens11040456.

[13] B. J. Bain. Structure and function of red and white blood cells. Med. (United Kingdom). 2017; 45(4): 187-93, doi: 10.1016/j.mpmed.2017.01.011.

[14] C. Escalda, J. Botelho, J. J. Mendes, V. Machado. Association of bacterial vaginosis with periodontitis in a cross-sectional American nationwide survey. Sci Rep. 2021; 11(1): 1-9, doi: 10.1038/s41598-020-79496-4.

[15] T. Chen et al. Dysbiosis of the rat vagina is efficiently rescued by vaginal microbiota transplantation or probiotic combination. Int J Antimicro. Agents. 2021; 57(3): doi: 10.1016/j.ijantimicag.2021.106277.

[16] M. et all Vakhshoori. The impact of platelet-to-lymphocyte ratio on clinical outcomes in heart failure: a systematic review and meta-analysis. Ther Adv Vaccines. 2024; 18: 1-22, doi: 10.1177.

[17] Motta F, Di Simone N, Selmi C. The impact of menopause on autoimmune and rheumatic diseases. Clin Rev Allergy Immunol. 2025; 68(1): 1-15. Available from: https://doi.org/10.1007/s12016-025-09031-8.

[18] Averyanova M, Yureneva S, Kiseleva V, Yakushevskaya O, Iskusnykh M, Pavlova A, et al. Effect of menopausal hormone therapy on cellular immunity parameters and cytokine profile. Biomedicines. 2024; 12(8), doi: 10.3390/biomedicines12081892.

Journal of Associated Medical Sciences 2026; 59(1): 24-32

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

Monoclonal antibodies against hemoglobins for detecting thalassemia

Thanusak Tatu* **and Watcharapong Jugnam-ang**

Department of Clinical Immunology, Faculty of Medical Technology, Nation University, Maung, Lampang Province, Thailand.

ARTICLE INFO

Article history:
Received 11 June 2025
Accepted as revised 31 August 2025
Available online 11 September 2025

Keywords:
Hemoglobin, monoclonal antibody, polyclonal antibody, thalassemia carrier, thalassemia screening test.

ABSTRACT

Background: Hemoglobin is composed of globin polypeptide chains, which serve as immunogens to induce the production of antibodies.

Objectives: This review article aims to describe the use of antibodies against human hemoglobins for the identification of thalassemia and hemoglobinopathies.

Materials and methods: Literature review to discusses the nature of normal human hemoglobin, hemoglobin switching, thalassemia and hemoglobinopathies, laboratory diagnosis, general properties of antibodies, production of antibodies against human hemoglobins, and clinical applications of these antibodies in identifying thalassemia and hemoglobinopathies.Antibody-based detection of hemoglobin is highly useful in diagnosing thalassemia and hemoglobinopathies.

Results: Polyclonal antibodies against HbF have been applied in sandwich ELISA to accurately detect HbF levels. Monoclonal antibodies against HbH and Hb Bart's have been produced and utilized in sandwich ELISA for detecting α-thalassemia. In addition, monoclonal antibodies against hemoglobin containing α-globin chains were developed and applied in sandwich ELISA to identify infants with Hb Bart's hydrops fetalis, a condition in which no α-globin chains are produced. For detecting β-thalassemia carriers, monoclonal antibodies against HbA_2 were produced, and sandwich ELISA was employed to measure HbA_2 levels, which are elevated in these individuals.

Conclusion: Antibody-based diagnosis of thalassemia and hemoglobinopathies enhances the quality of screening platforms and makes diagnosis of these disorders more reliable.

Introduction

In human, hemoglobin is the major protein in erythrocyte playing a major role in transporting oxygen from lung to tissue and carbon dioxide from tissue to lung. Hemoglobin naturally is composed to globin polypeptide chain, imbedded inside with heme molecule.[1] Two types of globin chains are produced along human developmental stage, including α-like globin chain (ζ and α-globin chains) and β-like globin chains (ε, γ, δ and β-globin chains). Functionally active hemoglobin molecule of human is of tetrameric structure composing of 2 α-like globin chains and 2 β-globin chains. Normal synthetic ratio of α-like globin chain and β-like globin chain is approximately equal, namely balanced globin chain synthesis. Normal hemoglobin composition in humans varies across developmental stages. During the embryonic stage (1–3 months of gestation), when hematopoiesis occurs in the yolk sac, the predominant hemoglobin is

* **Corresponding contributor.**
Author's Address: *Department of Clinical Immunology, Faculty of Medical Technology, Nation University, Maung, Lampang Province, Thailand.*
E-mail address: tthanu@hotmail.com
doi: 10.12982/JAMS.2026.004
E-ISSN: 2539-6056

Hb Gower I ($\zeta_2\varepsilon_2$), followed by smaller amounts of Hb Gower II ($\alpha_2\varepsilon_2$), trace amounts of Hb Portland I ($\zeta_2\gamma_2$), and very limited amounts of Hb Portland II ($\zeta_2\beta_2$). In the fetal stage (4-6 months of gestation), with the liver and spleen serving as primary sites of hematopoiesis, approximately 90% of hemoglobin is HbF ($\alpha_2\gamma_2$), and about 10% is HbA ($\alpha_2\beta_2$). During the third trimester, HbA and HbF are produced in roughly equal proportions. At birth, a reciprocal shift occurs in the levels of HbF and HbA, and a small amount of HbA_2 ($\alpha_2\delta_2$) begins to appear. By around two years of age, the typical hemoglobin composition is approximately 95% HbA, 2.5% HbA_2, and less than 1% HbF.[1-3]

Thalassemia and hemoglobinopathies

Thalassemia is a group of genetic disorder caused by reduction or absence of globin chain production, leading to a condition termed imbalanced globin chain synthesis in which the normally synthesized globin chains become excessive. Hemoglobinopathy is also genetic disorder, but caused by production of abnormal globin chains, leading to production of abnormal hemoglobins. The abnormal hemoglobins may become either neutral or abnormal in physical and electrochemical properties.[3,5,6]

Several types of thalassemia have been identified, depending on globin chains that are reduced or absent. However, two main types of thalassemia have been found worldwide, including α-thalassemia and β-thalassemia. α-thalassemia is characterized by reduced or absent α-globin chain, while decreased or absent β-globin chain is unique in β-thalassemia. In α-thalassemia, α-globin chain containing hemoglobins, including Hb Gower II ($\alpha_2\varepsilon_2$), HbF ($\alpha_2\gamma_2$), HbA ($\alpha_2\beta_2$), and HbA_2 ($\alpha_2\delta_2$) are decreased or absent. In β-thalassemia, only HbA ($\alpha_2\beta_2$), the β-globin chain containing hemoglobin, is decreased. Two types of α-thalassemia are found; α^+ thalassemia or α-thalassemia 2 and α^0-thalassemia or α-thalassemia 1.[7] There are also 2 types of β-thalassemia; β^+-thalassemia and β^0-thalassemia.[8] Some α-globin chains are still produced in α^+-thalassemia, while no α-globin chains are produced in α^0-thalassemia. In the same way, some β-globin chains are still produced in β^+-thalassemia, while no β-globin chains are produced in β^0-thalassemia

Hemoglobinopathies, sometimes referred to as structural variants, are categorized into two main types: α-structural variants and β-structural variants.[9] α-hemoglobinopathies, or α-structural variants, involve abnormal hemoglobins formed by the assembly of abnormal α-globin chains with normal β-globin chains, such as Hb Constant Spring ($\alpha_2^{CS}\beta_2$) and Hb Pakse' ($\alpha_2^{PS}\beta_2$). As of now, approximately 841 structural variants have been identified worldwide.[10] Although some α-structural variants have a normal rate of synthesis, others exhibit reduced production, resulting in a phenotype resembling α^+-thalassemia. For example, Hb Constant Spring (CS) is an α-structural variant caused by a point mutation at codon 142 of the α2-globin gene (HBA2:c.427T>C). The mutated CS mRNA is unstable, leading to reduced synthesis of the CS α-globin chain. This makes Hb Constant Spring clinically like α^+-thalassemia. Clinical and hematological studies have shown that Hb Constant Spring presents a more severe clinical and hematological phenotype than deletional α^+-thalassemia. Patients with HbH disease due to Hb Constant Spring (HbH-CS disease) often experience more severe symptoms than those with deletional HbH disease and are particularly prone to severe hemolytic crises during infections. For this reason, Hb Constant Spring is classified as a severe α^+-thalassemia, and careful screening for thalassemia and hemoglobinopathy carriers is essential.[11-15] β-hemoglobinopathies, or β-structural variants, are abnormal hemoglobins formed by the combination of normal α-globin chains with abnormal β-globin chains, such as HbS (β^6Glu→Val) and HbE (β^26Glu→Lys). To date, approximately 958 β-structural variants have been identified worldwide.[10]

Abnormal globin genes that cause thalassemia and hemoglobinopathies are inherited in an autosomal recessive fashion. Those having these genes in heterozygous and doubly heterozygous forms are carriers or so-called traits of the disease. These carriers are clinically asymptomatic, requiring no medical care. Those homozygous or compound heterozygous for genes of severe types of thalassemia and hemoglobinopathies (- -/ - - for α-thalassemia and β^0/β^0 for β-thalassemia) are affected and need serious medical attention. Chance of having affected offspring of the heterozygous couples is definitely 25%. Therefore, screening for the carriers of thalassemia and hemoglobinopathies in the population is essential in controlling and preventing birth of the patients of this disease, especially in region rich of thalassemia and hemoglobinopathies.

Screening for carriers for thalassemia and hemoglobinopathies

Screening for carriers of α-thalassemia, β-thalassemia and abnormal hemoglobin such as HbE and Hb Constant Spring are crucial for preventing birth of babies with thalassemia diseases such as Hb Bart's hydrops fetalis (Homozygous α^0-thalassemia; - -/- -), transfusion dependent homozygous β^0/β^0-thalassemia, transfusion dependent compound heterozygous β^0/β^+ thalassemia and transfusion dependent compound heterozygous β^0/β^E thalassemia. Red blood cell indices including mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), one-tube osmotic fragility test (OFT), dichlorophenol indophenol precipitation (DCIP) test are among the tests performed to seek for carriers of α-thalassemia, β-thlassemia and HbE.[16-25] These screening techniques have quite high sensitivity, but all of them lack specificity with low positive predictive value (PPV).

Antibody

Antibodies (Ab), also known as immunoglobulins (Ig), are specialized serum proteins belonging to the γ-globulin fraction that identify and neutralize foreign invaders such as bacteria, viruses, and toxins. They are produced and secreted by plasma cells derived from B lymphocytes, which are part of the humoral immune system. Immunoglobulins are naturally Y-shaped proteins composed of two main regions: the Fc region at the C-terminal domain and the Fab region at the N-terminal domain. The Fc region binds to receptors on the surface of target cells, while the Fab region, or antigen-binding site, binds to epitopes on target antigens. Antibodies are formed by two types of polypeptide chains: heavy chains and light chains. The Fab region contains both heavy and light chains, whereas the Fc region is composed only of heavy chains. The tips of the "Y" structure contain variable regions that specifically bind to antigens, which represent the epitope molecules of invading pathogens (Figure 1).There are five major classes of immunoglobulins, distinguished by their heavy chains: IgG (γ-heavy chain), IgA (α-heavy chain), IgM (μ-heavy chain), IgE (ε-heavy chain), and IgD (δ-heavy chain). Both humans and mice possess all five classes, while rabbits lack IgD in their serum.[26-28] In general, antibodies can be classified into two main types: polyclonal antibodies (pAb) and monoclonal antibodies (mAb). The two differ in their production and specificity. Monoclonal antibodies are derived from a single B-cell clone, producing highly specific antibodies that bind to only one epitope on an antigen. In contrast, polyclonal antibodies are generated from multiple B-cell clones, resulting in a heterogeneous mixture of antibodies capable of binding to multiple epitopes on the same antigen. The production of polyclonal antibodies involves simpler procedures, whereas monoclonal antibody production requires the hybridoma technique, followed by limiting dilution to isolate a single hybridoma clone capable of producing the monoclonal antibody of interest. Because of their high specificity, monoclonal antibodies are particularly well suited for diagnostic applications and targeted therapies.[29]

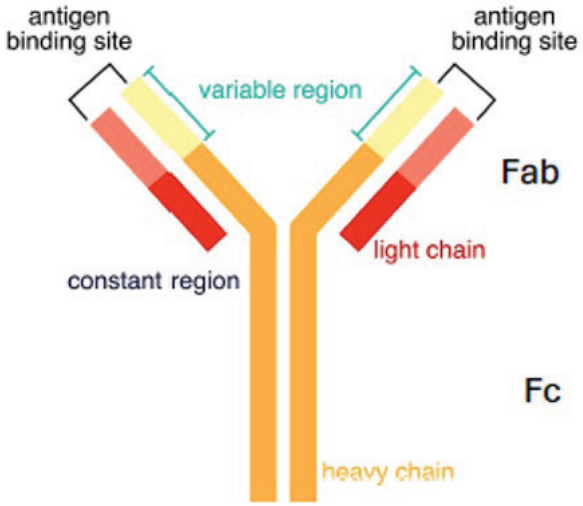


Figure 1. *Structure of antibody (Ab) or immunoglobulin (Ig).*
(modified from https://www.news-medical.net/life-sciences/Types-of-Antibodies.aspx)

Antibody-based detection of thalassemia and hemoglobinopathies

Antibody-based detection of hemoglobins can specifically identify hemoglobins that accurately indicate presence of thalassemia and hemoglobinopathies in blood samples. This capacity is superior to the conventional thalassemia screening tests which can only indicate chance of existing thalassemia and hemoglobinopathies in blood samples. Therefore, including the antibody-based detection of hemoglobins in screening protocol for thalassemia and hemoglobinopathies would increase effectiveness of the screening protocol of thalassemia and hemoglobinopathies.

As hemoglobins have globin polypeptide chains in the structure, they certainly can serve as immunogen to activate production of antibodies against them. The produced antibodies are then subsequently used to invent hemoglobin detection tools such as sandwich ELISA, flow cytometry, immunochromatographic strip test. Two types of antibodies against human hemoglobins have been produced, *i.e.* polyclonal antibody (pAb) and monoclonal antibody (mAb).

Production of pAb involves several laboratory techniques. As described by Kerdpoo and colleagues, the production of pAb against HbF ($\alpha_2\gamma_2$) involved the technique of HbF separation and purification by medium pressure liquid chromatography (MPLC), followed by immunizing the 2-month-old New Zealand White rabbits with purified HbF, mixed with complete Freund's adjuvant at the first injection and incomplete Freund's adjuvant in the second and third injections. Lastly, blood sample was collected from the marginal ear vein via incision technique and antiserum was obtained for further purification with protein G affinity chromatography (Figure 2).[30]

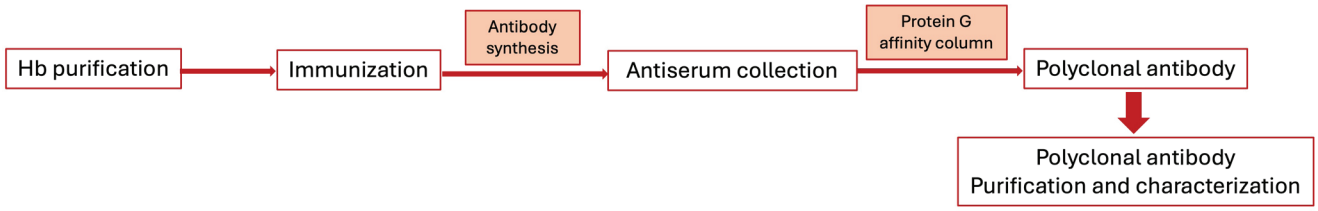


Figure 2. *Schematic presentation of procedure for production of polyclonal antibody for hemoglobin.*

Production of mAb against human hemoglobins also involves several steps as described by Pakdeepak and co-workers.[31] The first step is also separation and purification of hemoglobins by the MPLC, followed by injecting the purified hemoglobin mixed with complete Freund's adjuvant at the first injection and incomplete Freund's adjuvant in the second and third injection, into 2 six-week-old BALB/c mice, followed by spleen cell collection and fusion of spleen cells with P3-x63Ag8.653 myeloma cells to establish hybridoma clones capable of producing mAb specific to the immunized hemoglobin (Figure 3).

Figure 3. *Schematic presentation of procedure for production of hemoglobin-specific monoclonal antibody.*

Application of polyclonal antibody (pAb) in thalassemia and hemoglobinopathies

Polyclonal antibody by itself reacts with most types of hemoglobin. As shown in the study of Kerdpoo and colleagues that the pAb obtained after immunized the rabbits with purified HbF reacted with HbA ($\alpha_2\beta_2$), HbA_2($\alpha_2\delta_2$), HbE ($\alpha_2\beta_2^E$), HbF ($\alpha_2\gamma_2$), Hb Bart's (γ_4) and Hb Portand ($\zeta_2\gamma_2$).[30] With this reason, this pAb could not be applied in inventing the hemoglobin detection kit. However, Kerdpoo and co-workers used this pAb as a fluorescently labelled second antibody in the modified sandwich ELISA they invented and found that the new modified sandwich ELISA was effective in measuring HbF levels in the presence of minute amount of Hb Bart's in carriers of α-thalassemia (0.09-0.18 mg/mL).[32] Therefore, this pAb could be utilized to accurately measure HbF levels to identify a condition of hereditary persistence of fetal hemoglobin (HPFH) in area where α-thalassemia carriers are commonly encountered.

Application of monoclonal antibody (mAb) in thalassemia and hemoglobinopathies

In contrast to polyclonal antibodies (pAb), monoclonal antibodies (mAb) are highly suitable for detecting thalassemia and hemoglobinopathies due to their high specificity. Several studies have

attempted to produce monoclonal antibodies against hemoglobins, the levels or presence of which may indicate thalassemia or hemoglobinopathies.

For α-thalassemia, presence of HbH (β_4) or Hb Bart's (γ_4) indicates α-thalassemia. Shyamala and co-workers in 1992 produced mAb against HbH used for diagnosis of HbH disease by the technique of enzyme immunoassay.[33] The mAb they produced was highly specific to HbH and was able to quantify HbH level in HbH disease. However, there was no study utilizing mAb to HbH for detecting α-thalassemia carriers. It might be because HbH which is unstable has low level in the carriers of α-thalassemia. In contrast, significant amount of Hb Bart's was demonstrated in neonates with α^0-thalassemia carriers (--/αα).[7] Therefore, mAb to Hb Bart's was produced, followed by inventing the diagnostic tests utilizing this mAb. Monospecific antibody to Hb Bart's was produced by Garver and colleagues in 1984.[34] The established ELISA test utilizing this mAb to Hb Bart's showed that the carriers of α-thalassemia had average level of Hb Bart's was 6.10%, being significantly higher than that in normal individual whose Hb Bart's level was 0.25%. Monoclonal antibody against Hb Bart's was also produced by Tayapiwatana and co-workers in 2009 and immunochromatographic (IC) strip test was developed.[35] The IC strip test developed by Tayapiwatana and colleagues showed positive results in α^0-thalassaemia carriers (- -/αα), HbH disease (- -/- α), HbH-Constant Spring (H-CS) disease (- -/α^{CS}α), Hb Constant Spring EABart's disease (- -/α^{CS}α + $\beta^E\beta^N$), and homozygous α^+-thalassaemia (-α/-α). This IC strip test was not able to detect heterozygous α^+ thalassemia (- α/αα) and some α-structural variants such as Hb Westmead (HBA2: c.369C>G), Hb Jax (HBA2: c.44G>C), and Hb J-Buda [α61(E10)Lys → Asn, AAG > AAT].[36-38] The IC strip test did not show a negative result with all β-thalassemia and β-hemoglobinopathies. It could show positive results in individuals with elevated HbF levels or other coexisting hemoglobinopathies such as δβ-thalassemia and β-thalassemia (heterozygotes, homozygotes and HbE/β-thalassemia).[35,39,40] This IC strip test for Hb Bart's has been evaluated by several centers across the world. Wanapirak and colleagues evaluated this IC strip test in diagnosis of α^0-thalassemia in 499 pregnant women visiting antenatal care clinic at Maharaj Nakorn Chiang Mai hospital, Chiang Mai, Thailand and demonstrated 100% sensitivity and 89% specificity of this test. The falsely positive results observed in this study should be due to α^+-thalassemia (heterozygote or homozygote) which is also common in the northern Thailand.[41] In addition, Prayalaw and co-workers assessed this IC strip in 300 blood samples having positive thalassemia screening results and found that this IC strip had 100% sensitivity and 73.1% specificity in detecting α^0-thalassemia. This group postulated that all forms of α^+-thalassemin caused falsely positive results as this genotype also had Hb Bart's in blood.[42] Also, Sudjaroen and co-workers tested this IC strip in 414 pregnant women for α^0-thalassemia at Kudjab Hospital located in Udornthani Province, Thailand and found that this test had 92.6% sensitivity, 95.1% specificity and 94.9% efficiency.[43] This IC strip was also used by 4 laboratories in Thailand and Australia in screening for α-thalassemia by Winichagoon and colleagues who demonstrated 97% sensitivity and suggested this test to replace HbH inclusion body test in screening for α-thalassemia.[44] Recently, Bunkall and colleagues compared capability of the IC strip test and HbH inclusion body test in detecting α thalassemia in 67 blood samples. They showed that the IC strip test was extremely more sensitive than the HbH inclusion body test (76% vs 24% sensitivity, respectively) in screening for the α-thalassemia and blood samples stored in 4 °C was still good for the IC strip test.[45]

Beside mAb against Hb Bart's, mAbs against ζ-globin chain was also produced. Presence of the ζ-globin chain in hemolysate indicates the α^0-thalassemia of Southeast Asian (SEA) type.[46,47] Chui and co-workers produced murine hybridoma cells secreting mAb to the ζ-globin chain which was subsequently used in slot blot immunobinding assay in to screen for the SEA-α^0 thalassemia.[47] In 1993, Ireland and colleagues successfully used anti-ζ-immunobinding tetrazolium dye test to identify the α^0-thalassemia of SEA type in 225 blood samples they tested. [48] In 2008, Lafferty and colleagues employed commercial ζ-globin enzyme-linked immunosorbent assay (ELISA) to screen for SEA-α^0-thalassemia and found its sensitivity and specificity to be 1 and 0.94, respectively. [49] Pata and colleagues, in 2014, were successful in producing mAb to ζ-globin chain. This mAb was used to set up special platform of poly-l-lysine ELISA test which was proven to specifically detect the α^0-thalassemia of SEA type.[50] Immunostick coated with mAb to ζ-globin chain to detect the ζ-globin chain in hemolysate was subsequently developed by Pata and co-workers.[51] They demonstrated 100% sensitivity and 82% specificity of this novel immunostick test for the ζ-globin chain in identifying the SEA-α^0-thalassemia. Falsely positive results of this immunostick test was assumed to be due to cross reactivity of anti-ζ globin chain mAb with other globin chain.[51] In the same year, the IC strip test for detecting the ζ-globin chain was invented by Pata and colleagues and tested in combination with IC strip test for Hb Bart's in blood samples having MCV <80 fL and MCH <27 pg.[52] The analysis showed that the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of ζ-globin chain IC strip test in identifying the SEA-α^0 thalassemia in these blood samples were 100%, 65.2%, 90.7%, 100%, respectively. This result indicated that the IC strip test for the ζ-globin chain can only be used as the screening test, not diagnostic test.

This study then proposed that blood samples to be screened for the SEA-α^0 thalassemia must be first evaluated for mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH) and Hb typing results.

Only samples having low MCV, low MCH and Hb typing of A_2A were assayed by the IC strip test for the ζ-globin chain. Only samples having positive result for the IC strip test for the ζ-globin chain are further assayed by Gap-PCR.[52] Cross reaction of anti-ζ globin chain mAb to other globin chains was expected as a cause of falsely positive results.

In addition to identifying presence of HbH and Hb Bart's in α-thalassemia, however, Pakdeepak and colleagues showed by using mAb against α-globin chain containing hemoglobins (HbA; $\alpha_2\beta_2$, HbA_2; $\alpha_2\delta_2$, HbF;$\alpha_2\gamma_2$) that the levels of these hemoglobins in HbH disease were less than normal significantly.[31] The authors anticipated that this platform should be applied in detecting Hb Bart's hydrops fetalis babies in which no α-globin chain is produced at all.

As state previously, Hb Constant Spring (Hb Cs) is the abnormal hemoglobin that can cause severe transfusion-dependent Hb H disease. Therefore, identifying Hb Cs by antibody-based assay would improve efficiency in screening this abnormal hemoglobin. So-far, no study was found to describe production of antibody against Hb Constant Spring. However, with advancement of hemoglobin purification and hybridoma technologies at present, production of antibody against Hb Constant Spring is challenging.

For detection of β-thalassemia/hemoglobinopathies, mAb to HbA_2 was produced by several laboratories. Shyamala and co-workers successfully produced murine mAb against δ-globin chain of HbA_2 and the rapid ELISA test was invented.[53] They showed by this rapid ELISA test that the mean HbA_2 levels in β-thassemia carriers was 5.4%, being higher than normal individuals whose mean HbA_2 levels by this platform was 2.5%. Recently, Kuntaruk and colleagues produced murine mAb against HbA_2 which was utilized subsequently to set up the sandwich ELISA for measuring Hb A_2 in hemolysate.[54] They showed that under their developed sandwich ELISA, the mean HbA_2 levels in β-thalassemia carriers was 4.4%, being higher than the non β-thalassemia carriers (HbE carriers, HbE homozygote and α thalassemia carriers) whose mean HbA_2 levels by this platform ranged from 1.5% to 1.9%. They established 2.5% of HbA_2 level as cut-off values for identifying β-thalassemia carriers whose HbA_2 levels are higher than 2.5%. This 2.5% HbA_2 cut-off value was shown to be effective in detecting the β-thalassemia carriers by 100% sensitivity, 95% specificity, 82.3% positive predictive value and 100% negative predictive value. Therefore, when sandwich ELISA is performed, 2.5% cut-off HbA_2 values are used. However, for the routine laboratory diagnosis of the β-thalassemia carriers, the conventional HbA_2 cut-off value of 3.5% is still reliable.[55,56] Table 1 summarizes the studies that applied mAb to hemoglobin for detecting α-thalassemia and β-thalassemia.

Table 1. Summary of application of antibody-based detection of thalassemia/hemoglobinopathies.

Thalassemia	mAbs against	Invented diagnostic tools	Limitation	References
α	HbH	Sandwich ELISA	A	33
	Hb Bart's	Sandwich ELISA	A	34
	Hb Bart's	Immunochromatographic (IC) strip	B	35, 41, 42, 43, 44, 45
α	Hemoglobins that contain α-globin chain	Sandwich ELISA	B	31
	ζ-globin chain	Poly-l-Lysine ELISA	C	50
	ζ-globin chain	Immunostick test	C	51
	ζ-globin chain and Hb Bart's simultaneously	Immunochromatograpgic (IC) strip test	C	52
β	HbA_2	Sandwich ELISA	A	53, 54

Note: *A; many steps of operation, B: cannot detect specific type of α-thalassemia, C: not completely specific to SEA-α^0-thalassemia.*

The techniques for α-thalassemia listed in Table 1 had very high sensitivity for detecting the α^0-thalassemia. Therefore, in the national thalassemia screening Guideline of Thailand, they should be used in replacement of the conventional HbH inclusion body test which is the screening test for α-thalassemia. Modifying the α-thalassemia screening protocol using the mAb-based approach will increase efficiency and accuracy of α^0-thalassemia screening protocols. In contrast, the sandwich ELISA techniques were developed to quantify HbA_2 levels in blood lysates to help diagnosis of β-thalassemia carriers. Although this approach should replace cation-exchange HPLC or CZE in determining HbA_2 levels, the sandwich ELISA was suitable for only population-based screening for β-thalassemia carrier and testing in an individual is less possible. Therefore, at present, HbA_2-based diagnosis of β-thalassemia solely relies on HPLC-based or CZE-based enumeration of HbA_2 levels.

Limitations

Antibody-based identification of hemoglobins for the diagnosis of thalassemia and hemoglobinopathies currently focuses only on the qualitative analysis of

hemoglobins. Quantitative determination using this approach remains challenging, and further studies on this platform are encouraged.

Conclusion

Monoclonal antibodies (mAbs) against human hemoglobins and the ζ-globin chain become increasingly attractive for identifying α-thalassemia and β-thalassemia in both carrier and disease forms. With the invention of lateral flow IC strip or cassette-based immunological test kit utilizing these mAbs, the screening for carriers of thalassemia/ hemoglobinopathies will become reliable and greatly beneficial to the patients.

Funding

None

Conflict of interest

None

CRediT authorship contribution statement

Thanusak Tatu: conceptualization (lead), writing: original draft (lead), review and editing (equal); **Watcharapong Jugnamang**: writing: original draft (supporting), review and editing (equal).

References

[1] Perutz M. Hemoglobin structure. In: Bunn FF, Forget BG, eds. Hemoglobin: Molecular genetics and clinical aspects. Philadelphia: WB Sauners; 1986:13-35.

[2] Bunn HF, Forget BG, Ranney HM. Human Hemoglobins. Toronto: W.B. Saunders Company; 1977.

[3] Weatherall DJ, Clegg JB. The Thalassaemia Syndromes (4th Ed). Oxford: Blackwell Scientific; 2001.

[4] Wilber A, Nienhuis AW, Persons DA. Transcriptional regulation of fetal to adult hemoglobin switching: New therapeutic opportunities. Blood. 2011; 117(15): 3945-53. doi: 10.1182/blood-2010-11-316893.

[5] Fucharoen S, Winichagoon P. Thalassemia in SouthEast Asia: problems and strategy for prevention and control. Southeast Asian J Trop Med Public Health. 1992; 23(4): 647-55. PMID: 1298071.

[6] Fucharoen S, Winichagoon P. Hemoglobinopathies in Southeast Asia: molecular biology and clinical medicine. Hemoglobin. 1997; 21(4): 299-319. doi: 10.3109/03630269709000664.

[7] Harteveld CL, Higgs DR. Alpha-thalassaemia. Orphanet J Rare Dis. 2010; 5: 13. Doi: 10.1186/1750-1172-5-13

[8] Thein SL. The molecular basis of beta-thalassemia. Cold Spring Harb Perspect Med. 2013; 3(5): a011700. doi: 10.1101/cshperspect.a011700.

[9] Thom CS, Dickson CF, Gell DA, Weiss MJ. Hemoglobin variants: biochemical properties and clinical correlates. Cold Spring Harb Perspect Med. 2013; 3(3): a011858. doi: 10.1101/cshperspect.a011858.

[10] Laboratories of Computer Science & Engineering and Biochemistry & Molecular Biology at the Pennsylvania State University. Globin Gene Server. 2022; https://globin.bx.psu.edu/.

[11] Charoenkwan P, Taweephon R, Sae-Tung R, Thanarattanakorn P, Sanguansermsri T. Molecular and clinical features of Hb H disease in northern Thailand. Hemoglobin. 2005; 29(2): 133-140. PMID: 15921165.

[12] Fucharoen S, Viprakasit V. Hb H disease: clinical course and disease modifiers. Hematology Am Soc Hematol Educ Program. 2009: 26-34. doi: 10.1182/asheducation-2009.1.26.

[13] Hunt DM, Higgs DR, Winichagoon P, Clegg JB, Weatherall DJ. Haemoglobin Constant Spring has an unstable alpha chain messenger RNA. Br J Haematol. 1982; 51(3): 405-13. PMID: 7104225.

[14] Laig M, Pape M, Hundrieser J, et al. The distribution of the Hb constant spring gene in Southeast Asian populations. Hum Genet. 1990; 84(2): 188-90. doi: 10.1007/BF00208939.

[15] Laosombat V, Viprakasit V, Chotsampancharoen T, et al. Clinical features and molecular analysis in Thai patients with HbH disease. Ann Hematol. 2009; 88(12): 1185-92. doi: 10.1007/s00277-009-0743-5.

[16] Old JM. Screening and genetic diagnosis of haemoglobin disorders. Blood Rev. 2003; 17(1): 43-53. doi: 10.1016/s0268-960x(02)00061-9.

[17] Fucharoen S, Winichagoon P, Piankijagum A. Standardization on laboratory diagnosis of thalassemia and abnormal hemoglobin. South-east Asian J Trop Med Public Health. 1999; 30) Suppl 3): 90-8. PMID: 10926267.

[18] Leckngam P, Limweeraprajak E, Kiewkarnkha T, Tatu T. The Hb E (HBB: c.79G>A), Mean corpuscular volume, mean corpuscular hemoglobin cutoff points in double heterozygous Hb E/- -(SEA) alpha-thalassemia-1 carriers are dependent on hemoglobin levels. Hemoglobin. 2017; 41(1): 38-43. doi: 10.1080/03630269.2017.1295984.

[19] Manglani M, Lokeshwar MR, Vani VG, Bhatia N, Mhaskar V. 'NESTROFT'--an effective screening test for beta thalassemia trait. Indian pediatrics. 1997; 34(8): 702-7. PMID: 9492398.

[20] Munkongdee T, Chen P, Winichagoon P, Fucharoen S, Paiboonsukwong K. Update in Laboratory Diagnosis of Thalassemia. Front Mol Biosci. 2020; 7: 74. doi: 10.3389/fmolb.2020.00074.

[21] Sanchaisuriya K, Fucharoen S, Fucharoen G, et al. A reliable screening protocol for thalassemia and hemoglobinopathies in pregnancy: an alternative approach to electronic blood cell counting. Am J Clin Pathol. 2005; 123(1): 113-8. doi: 10.1309/fuf9evgq24v1pktp.

[22] Sanguansermsri T, Phumyu N, Chomchuen S, Steger HF. Screening for alpha-thalassemia-1 heterozygotes in expecting couples by the

combination of a simple erythrocyte osmotic fragility test and a PCR-based method. Community Genet. 1999; 2(1): 26-9. doi: 10.1159/000016180.

[23] Sirichotiyakul S, Maneerat J, Sa-nguansermsri T, Dhananjayanonda P, Tongsong T. Sensitivity and specificity of mean corpuscular volume testing for screening for alpha-thalassemia-1 and beta-thalassemia traits. J Obstet Gynaecol Res. 2005; 31(3): 198-201. doi: 10.1111/j.1447-0756.2005.00280.x.

[24] Sirichotiyakul S, Tantipalakorn C, Sanguansermsri T, Wanapirak C, Tongsong T. Erythrocyte osmotic fragility test for screening of alpha-thalassemia-1 and beta-thalassemia trait in pregnancy. Int J Gynaecol Obstet. 2004; 86(3): 347-350. doi: 10.1016/j.ijgo.2004.04.037.

[25] Sirichotiyakul S, Tongprasert F, Tongsong T. Screening for hemoglobin E trait in pregnant women. Int J Gynaecol Obstet. 2004; 86(3): 390-1. doi: 10.1016/j.ijgo.2004.04.041.

[26] Janeway CA, Travers P, Walport M. The structure of typical antibody molecule. In: Immunology: The Immune System in Heath and Disease 5th ed. New York: Garland Science; 2001.

[27] Feng L, Wang X, Jin H. Rabbit monoclonal antibody: potential application in cancer therapy. Am J Transl Res. 2011; 3(3): 269-274. PMID: 21633632.

[28] Rossi S, Laurino L, Furlanetto A, et al. Rabbit monoclonal antibodies: a comparative study between a novel category of immunoreagents and the corresponding mouse monoclonal antibodies. Am J Clin Pathol. 2005; 124(2): 295-302. doi: 10.1309/NR8H-N08G-DPVE-MU08.

[29] Actor JK. Assessment of immune parameters and immunodiagnostics. In: Actor JK, ed. Introductory Immunology; Basic Concepts for Interdisciplinary Applications (3rd Ed). Academic Press; 2023:169-188. doi: 10.1016/B978-0-12-420030-2.00019-6

[30] Kerdpoo S, Laopajon W, Kasinrerk W, Pata S, Tatu T. A modified sandwich ELISA for accurate measurement of HbF in alpha-thalassemia carriers containing Hb Bart's and Hb Portland 1. J Immunoassay Immunochem. 2018; 39(3): 323-36. doi: 10.1080/15321819.2018.1488726.

[31] Pakdeepak K, Pata S, Chiampanichayakul S, Kasinrerk W, Tatu T. Production and characterization of monoclonal antibodies against alpha-globin chain-containing human hemoglobins for detecting alpha-thalassemia disease. J Immunoassay Immunochem. 2016; 37(6): 564-71. doi: 10.1080/15321819.2016.1174135.

[32] Tatu T, Kiewkarnkha T, Khuntarak S, Khamrin S, Suwannasin S, Kasinrerk W. Screening for co-existence of alpha-thalassemia in beta-thalassemia and in HbE heterozygotes via an enzyme-linked immunosorbent assay for Hb Bart's and embryonic zeta-globin chain. Int J Hematol. 2012; 95(4): 386-93. doi: 10.1007/s12185-012-1039-4.

[33] Shyamala M, Kiefer CR, Moscoso H, Garver FA. A monoclonal antibody-linked immunoassay for hemoglobin H disease. Ann Hematol. 1992; 65(1): 37-40. doi: 10.1007/BF01715124.

[34] Garver FA, Singh H, Moscoso H, Kestler DP, McGuire BS, Jr. Identification and quantification of hemoglobins A_2 and Barts with an enzyme-labeled immunosorbent assay. Clin Chem. 1984; 30(7): 1205-8. PMID: 6203666.

[35] Tayapiwatana C, Kuntaruk S, Tatu T, et al. Simple method for screening of alpha-thalassaemia 1 carriers. Int J Hematol. 2009; 89(5): 559-67. doi: 10.1007/s12185-009-0331-4.

[36] Panyasai S, Permsripong N, Jaiping K, Khantarag P, Pornprasert S. Hematological and Molecular Characterization of Hb J-Buda [alpha61(E10) Lys --> Asn, AAG > AAT]. Indian J Hematol Blood Transfus. 2016; 32(Suppl 1): 311-4. doi: 10.1007/s12288-016-0641-7.

[37] Jiang F, Ju AP, Li J, et al. Hb Westmead (HBA2: c.369C>G): Hematological characteristics in heterozygotes with and without alpha(0)-thalassemia. Hemoglobin. 2020; 44(3): 153-5. doi: 10.1080/03630269.2020.1768109.

[38] Panyasai S, Jaiping K, Khantarag P, Nochod P, Satthakarn S. Unraveling thalassemia intermedia: Novel insights of a hemoglobin Jax [HBA2:c.44G>C] and deletional alpha(0)-thalassemia interaction phenotype. Hematol Transfus Cell Ther. 2025; 47(1): 103739. doi: 10.1016/j.htct.2025.103739.

[39] Chan G, Wong J, Tan T, Ghallyan N, Browett P. False positive Hb Bart's immunochromatographic strip test for alpha thalassaemia in delta-beta thalassaemia carriers. Int J Lab Hematol. 2023; 45(6): 984-6. doi: 10.1111/ijlh.14122.

[40] Theingi M, Khamphikham P, Pornprasert S. Consideration of immunochromatographic strip positive results for screening of alpha-thalassemia in elevated hemoglobin F samples. Scand J Clin Lab Invest. 2025; 85(5): 346-52. doi: 10.1080/00365513.2025.2524849.

[41] Wanapirak C, Piyamongkol W, Sirichotiyakul S, Tayapiwatana C, Kasinrerk W, Tongsong T. Accuracy of immunochromatographic strip test in diagnosis of alpha-thalassemia-1 carrier. J Med Assoc Thai. 2011; 94(7): 761-5. PMID: 21774279.

[42] Prayalaw P, Fucharoen G, Fucharoen S. Routine screening for alpha-thalassaemia using an immunochromatographic strip assay for haemoglobin Bart's. J Med Screen. 2014; 21(3): 120-5. doi: 10.1177/0969141314538611.

[43] Sudjaroen Y. Efficiency assessment of immunochromatographic strip test for the diagnosis of alpha-thalassemia-1 carriers. J Lab Physicians. 2015; 7(1): 4-10. doi: 10.4103/0974-2727.154779.

[44] Winichagoon P, Kumpan P, Holmes P, et al. Validation of the immunochromatographic strip for alpha-thalassemia screening: a multicenter study. Transl Res. 2015; 165(6): 689-95. doi: 10.1016/j.trsl.2014.10.013.

[45] Bunkall C, Ghallyan N, Elliott C, Van de Water N,

Chan G. Evaluation of an immunochromatographic strip test for alpha-thalassaemia screening. Int J Lab Hematol. 2018; 40(6): 691-6. doi: 10.1111/ijlh.12905.

[46] Luo HY, Clarke BJ, Gauldie J, Patterson M, Liao SK, Chui DH. A novel monoclonal antibody based diagnostic test for alpha-thalassemia-1 carriers due to the (-SEA/) deletion. Blood. 1988; 72(5): 1589-94. PMID: 3179441.

[47] Chui DH, Luo HY, Clarke BJ. Potential application of a new screening test for alpha-thalassemia-1 carriers. Hemoglobin. 1988; 12(5-6): 459-63. doi: 10.3109/03630268808991635.

[48] Ireland JH, Luo HY, Chui DH, Chu B, Yuen J, Hsia YE. Detection of the (--SEA) double alpha-globin gene deletion by a simple immunologic assay for embryonic zeta-globin chains. Am J Hematol. 1993; 44(1): 22-8. doi: 10.1002/ajh.2830440106.

[49] Lafferty JD, Crowther MA, Waye JS, Chui DH. A reliable screening test to identify adult carriers of the (--SEA) alpha zero-thalassemia deletion. Detection of embryonic zeta-globin chains by enzyme-linked immunosorbent assay. Am J Clin Pathol. 2000;114(6):927-931. doi: 10.1309/26G7-BQH4-93BV-UR0Q.

[50] Pata S, Khummuang S, Pornprasert S, Tatu T, Kasinrerk W. A simple and highly sensitive ELISA for screening of the alpha-thalassemia-1 Southeast Asian-type deletion. J Immunoassay Immunochem. 2014; 35(2): 194-206. doi: 10.1002/ajh.2830380311.

[51] Pata S, Pongpaiboon M, Laopajon W, et al. Immunostick test for detecting zeta-globin chains and screening of the Southeast Asian alpha-thalassemia 1 Deletion. Biol Proceed Online. 2019; 21: 15. doi: 10.1186/s12575-019-0104-2.

[52] Pata S, Laopajon W, Pongpaiboon M, et al. Impact of the detection of zeta-globin chains and hemoglobin Bart's using immunochromatographic strip tests for alpha0-thalassemia (--SEA) differential diagnosis. PLoS One. 2019; 14(10): e0223996. doi: 10.1371/journal.pone.0223996

[53] Shyamala M, Kiefer CR, Moscoso H, Garver FA. Application of a monoclonal antibody specific for the delta chain of hemoglobin A2 in the diagnosis of beta thalassemia. Am J Hematol. 1991; 38(3): 214-9. doi: 10.1002/ajh.2830380311.

[54] Kuntaruk S, Tatu T, Keowkarnkah T, Kasinrerk W. Sandwich ELISA for hemoglobin A2 quantification and identification of beta-thalassemia carriers. Int J Hematol. 2010; 91(2): 219-28. doi: 10.1007/s12185-009-0490-3.

[55] Committee for Developing Direction of Laboratory Diagnosis of Thalassemia and Hemoglobinopathies, Clinical Research Center, Department of Medical Sciences, Ministry of Public Health. Direction of Laboratory Diagnosis of Thalassemia and Hemoglobinopathies. Nonthaburi, Mudded Ltd. Co., 2009 (in Thai).

[56] Colaco S, Colah R, Nadkarni A. Significance of borderline HbA(2) levels in beta thalassemia carrier screening. Sci Rep. 2022; 12(1): 5414. doi: 10.1038/s41598-022-09250-5

Journal of Associated Medical Sciences 2026; 59(1): 33-39

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

Analytical errors in the laboratory of a general hospital of the Thai Red Cross Society: A 3-year experience

Srivilai Trakulkaseamsiri* **and Kittipong Chumchujan**

Department of Laboratory, Queen Savang Vadhana Memorial Hospital, Thai Red Cross Society, Chonburi Province, Thailand.

ARTICLE INFO

Article history:
Received 21 May 2025
Accepted as revised 4 September 2025
Available online 11 September 2025

Keywords:
Laboratory error, laboratory testing process, pre-analytical phase.

ABSTRACT

Background: Laboratory test results are critical to clinical decision-making, including patient screening, diagnosis, treatment planning, and monitoring therapeutic responses. Therefore, laboratory errors can significantly impact patient care and outcomes.

Objectives: To identify and quantify the types and frequencies of errors in the total testing process of the medical laboratory department, with the goal of reducing preventable errors and improving overall quality.

Materials and methods: This retrospective descriptive study analyzed error records, incident reports and nonconformities from January 2021, to December, 2023. Errors were categorized by testing phase: pre-analytical, analytical and post-analytical. Specimen sources were classified as outpatient, inpatient, or external (outside hospital). Data were summarized using descriptive statistics, including frequencies, percentages and Six Sigma performance scale values.

Results: A total of 2,261,729 specimens were received during the study period. The overall error rate was 0.197% (4,460 errors). Error rates by phase were 0.116% for pre-analytical, 0.077% for analytical, and 0.004% for post-analytical. Six Sigma performance scores were 4.60 (pre-analytical), 4.70 (analytical), and 5.50 (post-analytical), all within acceptable quality thresholds (>4.15, >3.85 and >4.80, respectively). The most common pre-analytical error was clotted samples (1,233 cases). Analytical errors were dominated by unacceptable internal quality control results (1,088 cases). The most frequent post-analytical error was excessive turnaround time (54 cases). Pre-analytical errors occurred most frequently in inpatient specimens (76.90%), with clotted samples accounting for 41.49% of those cases. Six Sigma values by specimen source were 4.20 for inpatients, 4.90 for outpatients and 4.90 for external sources.

Conclusion: Most laboratory errors occurred in the pre-analytical phase, primarily due to specimen collection issues related to quality and volume, particularly in the inpatient setting. Targeted preventative measures-especially in pre-analytical processes-are essential to minimize errors and improve patient safety. Analytical errors were primarily due to unacceptable IQC, EQA. For quality control improvement, training program laboratory staff on quality control, sigma metrics, risk assessment and quality goal index (QGI), including selection and application of Westgard's rules, is important. The integration of sigma metrics with QGI, risk management and systematic quality control can enhance laboratory performance and reliability of analytical results.

* **Corresponding contributor.**
Author's Address: *Department of Laboratory, Queen Savang Vadhana Memorial Hospital, Thai Red Cross Society, Chonburi Province, Thailand.*
E-mail address: srivilaisiri@gmail.com
doi: 10.12982/JAMS.2026.005
E-ISSN: 2539-6056

Introduction

The Medical Laboratory Department at Queen Savang Vadhana Memorial Hospital, under the Thai Red Cross Society, performs approximately 3,000,000 tests annually. It offers diagnostic services in hematology, clinical microscopy, clinical biochemistry, microbiology, immunology, molecular biology and blood banking. Most of these analyses are conducted using fully automated analyzers, with manual testing comprising less than 5% of the total. The laboratory is accredited under ISO 15189:2022 and ISO 15190:2020 standards, underscoring its commitment to quality and safety in laboratory medicine.

Accurate, precise and high-quality laboratory testing is essential to the healthcare system and plays a critical role in clinical decision-making. Physicians depend on laboratory results for diagnosis, monitoring and prognostication. However, several factors, both within and beyond the laboratory's control, can compromise test quality. Such errors can result in misleading results and inappropriate clinical decisions, leading to repeated testing, patient discomfort, wasted resources and increased healthcare costs. In severe cases, diagnostic errors may even contribute to patient complications or fatalities.[1]

Errors in the laboratory testing process may occur at any stage, namely the pre-analytical, analytical and post-analytical phases. Studies have shown that the majority of these errors occurred during the pre-analytical phase (46-68%), followed by the post-analytical (18-47%) and analytical phases (7-13%).[2,3] The high incidence of pre-analytical errors is largely attributed to human factors and process steps occurring outside the laboratory's direct oversight.[2-5] These errors often necessitated specimen recollection, resulting in patient discomfort and diagnostic delays.[6-12] Currently, laboratory error rates are commonly assessed using the Six Sigma scale, calculated with Westgard's calculator, to quantify process quality and identify areas for improvement.[13] Previous studies in Thailand and internationally have documented varying types and frequencies of laboratory errors. The present study aims to investigate the nature and distribution of laboratory errors at each testing phase, Early detection and mitigation of such errors are crucial to ensuring timely and accurate patient care.

Materials and methods

This descriptive study was conducted at the Medical Laboratory of Queen Savang Vadhana Memorial Hospital over a three-year period (January 2021, to December 2023). The study focused on errors recorded across the total testing process in various laboratory units, including clinical microscopy, hematology, clinical chemistry, immunology, molecular biology, microbiology and phlebotomy.

The study population consisted of documented error records, including risk incident reports, patient complaints, nonconformities with regulatory requirements, quality control failures, equipment maintenance issues, and other relevant deviations in laboratory operations.

Inclusion criteria

All errors that occurred during the pre-analytical, analytical and post-analytical phases of the medical laboratory testing process were included.

Exclusion criteria

Errors unrelated to the medical laboratory testing process, such as administrative or clerical errors not affecting test quality, were excluded.

Research instruments

Laboratory errors were classified into the following three phases:

1) Pre-analytical phase: Includes clotted specimens, hemolyzed samples, insufficient sample volume, incorrect test requests, misidentification, unlabeled specimens, incorrect or wrong sample collection and errors in specimen registration. The source of the sample-whether from inpatients, outpatients, or external facilities-was also recorded.

2) Analytical phase: Includes discrepancies between current and previous test results, quality control failures and instrument malfunctions.

3) Post-analytical phase: Includes excessive turnaround times, incorrect test reporting, incomplete results, failure to communicate critical values. All data were recorded using Microsoft Excel for further statistical analysis.

Statistical analysis

Data were analyzed separately for the pre-analytical, analytical and post-analytical phases. Sample sources were categorized as originating from inpatients, outpatients or external healthcare facilities. Descriptive statistics were used to summarize the frequency and percentage of errors. Additionally, Six Sigma metrics were calculated using Westgard's calculator[13] to evaluate the quality performance of each testing phase.

Results

Over the three-year study period (2021-2023), a total of 2,261,729 specimens were received by the laboratory. Of these, 4,460 errors were recorded, representing an overall error rate of 0.197%. The distribution of errors across the three laboratory phases was as follows: 2,615 pre-analytical phase errors (0.116%); 1,738 analytical phase errors (0.077%); 107 post-analytical phase errors (0.004%). Performance based on the Six Sigma scale showed values of 4.60 for the pre-analytical phase, 4.70 for the analytical phase, and 5.50 for the post-analytical phase, all exceeding the minimum acceptable thresholds (>4.15, >3.85 and

>4.80, respectively, (Table 1)[13,18,19]

As detailed in Table 2, the most common pre-analytical errors were clotted samples (47.15%) and hemolyzed samples (22.64%). In the analytical phase, internal quality control (IQC) results outside acceptable ranges accounted for 62.60% of errors, followed by external quality assessment (EQA) failures (18.64%). For post-analytical errors, excessive turnaround time was the most frequent (50.47%), followed by incorrect test results (28.97%) and incomplete reports (17.76%).

Among the 2,615 pre-analytical errors, 2,011 (76.90%) were associated with inpatient specimens, 575 (21.99%) with outpatient specimens, and 29 (1.11%) with samples from external sources. The error rate for inpatient specimens was 0.37%, with a corresponding Six Sigma value of 4.20. Outpatient and external samples had error rates of 0.04% and Six Sigma values of 4.90, respectively (Table 3).

Detailed breakdowns of pre-analytical errors by source are presented in Table 4. In inpatient samples, clotted specimens (41.49%) and hemolyzed specimens (14.53%) were the most common errors. In outpatient samples, hemolysis was the leading issue (7.84%), followed by clotted specimens (5.51%). For external samples, incorrect patient identification (0.38%) and hemolysis (0.27%) were the primary issues.

Table 1. Percentage of error in pre-analytical, analytical and post-analytical phase.

Year (N)	Pre-analytical phase		Analytical phase		Post-analytical phase		Total	
	N (%)	Sigma	N (%)	Sigma	N (%)	Sigma	N (%)	Sigma
2021 (675,019)	1,021 (0.151)	4.50	484 (0.072)	4.70	34 (0.005)	5.40	1,539 (0.228)	4.40
2022 (763,927)	702 (0.092)	4.70	527 (0.069)	4.70	37 (0.005)	5.40	1,266 (0.166)	4.50
2023 (822,783)	892 (0.108)	4.60	727 (0.088)	4.70	36 (0.004)	5.50	1,655 (0.201)	4.40
Total(2,261,729)	2,615 (0.116)	4.60	1,738 (0.077)	4.70	107 (0.004)	5.50	4,460 (0.197)	4.40

Table 2. Type of error in pre-analytical, analytical and post-analytical phase.

Type of error	2021		2022		2023		Total	
	N	%	N	%	N	%	N	%
Pre-analytical phase								
Clotted sample	487	31.64	334	26.38	412	24.89	1,233	47.15
Hemolyzed sample	251	16.31	140	11.06	201	12.15	592	22.64
Inadequate sample	103	6.69	52	4.11	130	7.85	285	10.90
Incorrect test request	38	2.47	87	6.87	60	3.63	183	7.00
Incorrect identification	74	4.81	39	3.08	60	3.63	173	6.62
Unlabeled sample	22	1.43	17	1.34	14	0.85	53	2.03
Incorrect sample	22	1.43	21	1.66	9	0.54	52	1.99
Wrong sample received	10	0.66	4	0.33	2	0.12	16	0.61
Collected sample with wrong patient	7	0.45	5	0.39	1	0.06	13	0.49
Error in patient registration	7	0.45	3	0.23	3	0.18	13	0.49
Subtotal	1,021	66.34	702	55.45	892	53.90	2,615	100
Analytical phase								
IQC unacceptable range	312	20.27	406	32.07	370	22.36	1,088	62.60
EQA unacceptable range	98	6.37	114	9.01	112	6.77	324	18.64
Test results not correlated with previous results	53	3.45	3	0.23	231	13.95	287	16.52
Equipment malfunction	21	1.36	4	0.32	14	0.85	39	2.24
Subtotal	484	31.45	527	41.63	727	43.93	1,738	100

Table 2. Type of error in pre-analytical, analytical and post-analytical phase. (Continue)

Type of error	2021		2022		2023		Total	
	N	%	N	%	N	%	N	%
Post-analytical phase								
Excessive turnaround time	17	1.11	19	1.50	18	1.09	54	50.47
Incorrect test result	10	0.65	14	1.11	8	0.48	32	29.91
Incomplete test report	6	0.39	4	0.31	9	0.54	19	17.76
Failure to notify critical value	1	0.06	0	0	1	0.06	2	1.86
Subtotal	34	2.21	37	2.92	36	2.17	107	100
Total	**1,539**		**1,266**		**1,655**			

Note: IQC: internal quality control, EQA: external quality assessment.

Table 3. The source of error in pre-analytical phase.

Source	Error (N)			Sample (N)			Total			Sigma
	2021	2022	2023	2021	2022	2023	Error	Sample	%	
Inpatients	755	568	688	174,581	184,060	191,360	2,011	550,001	0.37	4.20
Outpatients	262	124	189	477,070	548,501	604,697	575	1,630,268	0.04	4.90
Outside-hospital	4	10	15	23,368	31,366	26,726	29	81,460	0.04	4.90
Total	**1,021**	**702**	**892**	**675,019**	**763,927**	**822,783**	**2,615**	**2,261,729**	**0.12**	**4.60**

Table 4. Percentage of pre-analytical error inpatient, outpatient and outside-hospital samples.

Type of error	Inpatients		Outpatients		Outside hospital	
	N	%	N	%	N	%
Clotted samples	1,085	41.49	144	5.51	4	0.15
Hemolyzed samples	380	14.53	205	7.84	7	0.27
Inadequate samples	206	7.88	79	3.02	0	0.00
Incorrect identification	116	4.44	47	1.80	10	0.38
Incorrect test request	155	5.93	28	1.07	2	0.08
Incorrect samples	24	0.92	26	0.99	2	0.08
Collected sample from wrong patient	8	0.31	5	0.19	0	0.00
Wrong sample received	7	0.27	9	0.34	0	0.00
Error in patient registration	1	0.03	12	0.47	0	0.00
Unlabeled sample	29	1.10	20	0.76	4	0.15
Total	**2,011**	**76.90**	**575**	**21.99**	**29**	**1.11**

Discussion

This study found that pre-analytical errors accounted for 53.90% to 66.34% of total annual errors over the three-year period-consistent with previous research that reported rates ranging from 42.80% to 71.80%.[1,14-16] One study even reported pre-analytical errors as high as 80.95%.[17] In our study, the overall pre-analytical error rate was 0.12%, slightly lower than previously reported rates of 0.16-0.17%.[18,19] The majority of pre-analytical errors originated from inpatient specimens (76.90%), with clotted samples being the most common (41.49%). These findings align with Arechep *et al.*, who reported that 82.00% of pre-analytical errors were from inpatients and predominantly due to clotted blood.[19] Clotting can result from improper blood-to-anticoagulant ratios or

insufficient mixing post-collection. Hemolysis, another frequent error, often results from poor collection technique, such as not removing the needle before transferring blood into the tube or using excessive suction. The Six Sigma score for overall pre-analytical errors was 4.60, which is like values reported in other studies (Six Sigma score 4.50).[18,19]

Analytical errors constituted 31.45% to 43.93% of total annual errors. Contrary to several reports where post-analytical errors were more common than analytical errors,[1-4, 16-18] our findings align with others who also observed a higher proportion of analytical errors.[14,15] The overall analytical error rate in this study was 0.077%, lower than the 2.39% reported by Choosongsang *et al.*, who analyzed data over five years.[18] The leading analytical error in our study was IQC failure. These errors often require repeat testing and can delay result reporting. The Six Sigma value for analytical errors was 4.70, lower than the 5.40 reported in some prior studies.[18] Result from the root cause analysis showed that the root cause of internal quality control (IQC) and external quality assessment (EQA) being out of the acceptable range were due to laboratory staff not following work instructions and standard procedures, insufficient knowledge and skills of control material preparation for IQC and unknown material preparation for EQA. For improvement, training programs to enhance laboratory staff's competency on quality control and sigma metrics are required, along with the selection of suitable quality control rules for testing programs and interpretation of quality control results. Then, implementation with sigma metrics to evaluated performance on test items and choosing suitable quality control rules for testing programs must follow. Testing with sigma metric <4 was assessed using quality goal index (QGI) combined with risk assessment by risk matrix. Testing with moderate or high-risk impact levels must be examined by root cause analysis for conducting quality control plans. QGI analysis reveals issues with precision, accuracy or both. This must be followed by the implementation of specific quality control improvement and reassessment.

Post-analytical errors accounted for only 2.17% to 2.92% of total annual errors, significantly lower than rates of 11.68% to 24.30% reported elsewhere.[14-17] The overall post-analytical error rate was 0.004%, again lower than the 0.06% found by Choosongsang *et al.*[18] The higher Six Sigma score of 5.50 in this study may be attributed to the hospital's computerized system that ensures automated transfer of results from analyzers to the Laboratory Information System (LIS) and subsequently to the Hospital Information System (HIS), thereby minimizing transcription errors, and LIS show warning message when current analytical results not correlated previous analytical results in same patients to prevent mistake results reporting.

Limitations

As a retrospective study, data collection was limited by the completeness and consistency of existing records. Manual documentation practices may have led to underreporting or misclassification of errors. Future research should utilize a centralized electronic Healthcare Risk Management System on Cloud (HRMS on Cloud) for real-time error tracking. Additionally, data should be stratified by inpatient ward types (e.g., general vs. special wards) to better account for the differences in collection technique and staff training.

Conclusion

The overall laboratory error rate in this study was 0.197%, with the highest incidence in the pre-analytical phase (0.116%). This relatively low rate compared with previous reports may reflect the effectiveness of personnel responsible for specimen screening prior to analysis. Analytical and post-analytical error rates were 0.077% and 0.004%, respectively. Six Sigma scores for all three phases were within acceptable quality standards: 4.60 (pre-analytical), 4.70 (analytical) and 5.50 (post-analytical).

Most pre-analytical errors stemmed from improper specimen collection-both in quality and volume-particularly among inpatient samples, which were typically collected by nursing staff. These errors often necessitated recollection, causing patient dissatisfaction, delays and increased workload and cost. Preventative strategies should include structured training for nursing and phlebotomy staff, focusing on proper tube selection, collection techniques and adherence to the correct order of draw.

Analytical errors were primarily due to unacceptable IQC and EQA results. Although these were resolved before patient testing, they contributed to time loss and resource consumption. For quality control improvement, training program for laboratory staff on quality control, sigma metrics, risk assessment and quality goal index (QGI), including selection and application of Westgard's rules, is critical. The integration of sigma metrics with QGI, risk management and systematic quality control can enhance laboratory performance and reliability of analytical results. Quality control improvement provides cost reduction and resource optimization due to decreased unnecessary repeated tests, control material and reagent wastage, reducing turnaround time for analysis results and reducing the clinical testing workload.

Post-analytical errors were rare, with turnaround delays being most frequent issue. These errors were

caused by manual handling and transcription risks. Implementation of LIS-HIS system integration, an automated warning system that cross-checks previous patient results, and dashboards for turnaround monitoring, resulted in an enhancement in the accuracy

and timeliness and overall reliability of reporting. Overall, these findings underscore that continuous monitoring, targeted staff training, and system-level interventions are essential to minimize laboratory errors, improve efficiency, and ensure high-quality patient care.

Ethical approval

This study was approved by the Institutional Review Board of Queen Savang Vadhana Memorial Hospital (COE No. 035/2567 / IRB No. 042/2567 / Study code: 042/2567).

Funding

This study was funded by Queen Savang Vadhana Memorial Hospital, Thai Red Cross Society.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Srivilai Trakulkaseamsiri: conceptualization, methodology, performed research protocol, formal analysis, writing: original draft, review and editing, supervision, project administration, funding acquisition, and served as the corresponding author; **Kittipong Chumchujan**: investigation, data curation, performed graphic abstract. Both authors approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Acknowledgements

The authors gratefully acknowledge the financial support provided by Queen Savang Vadhana Memorial Hospital, Thai Red Cross Society.

References

[1] Vongsakulyanon A. Prevention of clinical laboratory test error: Pre-analytical error. Rama Med J. 2019; 42(2): 49-62. doi: 10.33165/rmj.2019.42.2.142535. [in Thai].

[2] Plebani M. The detection and prevention of errors in laboratory medicine. Ann Clin Biochem. 2010; 47(2): 101-10. doi: 10.1258/acb.2009.009222.

[3] Hammerling JA. A review of medical errors in laboratory diagnostics and where we are today. Lab Med. 2012; 43(2): 41-44.

[4] Pratoomtone P. Types and frequency of laboratory errors in Sawanpracharak Hospital, Nakornsawan province: One year after laboratory accreditation. Bull Chiang Mai Assoc Med Sci. 2010; 13: 70-76. [in Thai].

[5] Atay A, Demir L, Cuhadar S, Saglam G, Unal H, Aksum S, et al. Clinical biochemistry laboratory rejection rates due to various types of preanalytical errors. Biochem Med (Zagreb). 2014; 24: 376-82.

[6] Giavarina D, Lippi G. Blood venous sample collection: Recommendations overview and a checklist to improve quality. Clin Biochem. 2017; 50(10-11): 568-73. doi: 10.1016/j.clinbiochem.2017.02.021.

[7] Sanford KW, McPherson RA. Preanalysis. In: McPherson RA, Pincus MR, Editors. Henry's clinical diagnosis and management by laboratory methods. Philadelphia: Elsevier; 2011: 24-36.

[8] Adcock DM, Favaloro EJ, Lippi G. Critical pre-examination variables in the hemostasis laboratory and their quality indicators. Clin Biochem. 2016; 49(18): 1315-20. doi: 10.1016/j.clinbiochem.2016.08.022.

[9] Comes M, van Dongen-Lases E, Grankvist K, Ibarz M, Kristensen G, Lippi G, et al. Order of blood draw: Opinion paper by the European Federation for Clinical Chemistry and Laboratory Medicine (EFLM) Working Group for the Preanalytical Phase (WG-PRE). Clin Chem Lab Med. 2017; 55(1): 27-31. doi: 10.1515/cclm-2016-0426.

[10] Cornes M. Case report of unexplained hypocalcaemia in a slightly haemolysed sample. Biochem Med (Zagreb). 2017; 27: 426-9. doi.org/10.11613/BM.2017.046.

[11] Mrazek C, Lippi G, Keppel MH, Felder TK, Oberkofler H, Haschke-Becher E, et al. Errors within the total laboratory testing process, from test selection to medical decision-making: a review of causes, consequences, surveillance and solutions. Biochem Med (Zagreb). 2020; 30(2): 020502: 1-19. doi: 10.11613/BM.2020.020502.

[12] World Health Organization. WHO guidelines on drawing blood: best practices in phlebotomy [Internet]. Geneva: World Health Organization; 2010 [cited2024Oct30].Availablefrom:https://iris.who.int/bitstream/handle/10665/44294/9789241599221_eng.pdf?sequence=1.

[13] Westgard JO. Six sigma Calculators. [Internet]. 2009 [cited 2024 Oct 24] Available from:https://www.westgard.com/six-sigma-calculators.htm.

[14] Kimengech KK, Waithaka SK, Onyuka J, Kigondu CS. Determination of errors that compromise the quality of laboratory service in a tertiary hospital. Asian J Med Sci [Internet]. 2017 [cited 2024 Oct 27]; Available from: https://www.pdfs.semanticscholar.org/76C7/1b40af4649bf31c11222f439fca02ee429a5.pdf.

[15] Abdollahi A, Saffar Hi, Saffar Ha. Types and frequency of errors during different phase of testing at a clinical medical laboratory of a teaching hospital in Tehran, Iran. North Am J Med Sci [Internet]. 2014; 6: 224-8 [cited : 2024 OCT 28] Available from https://pmc.ncbi.nlm.nih.gov/articles/PMC4049056.

[16] Tola EK, Dabi YT, Dano GT. Assessment of types and frequency of errors in diagnostic laboratories among selected hospitals in East Wollega Zone, Oromia, Ethiopia. Pathol Lab Med Int [Internet]. 2022; 14: 1-6 [cited 2024 Oct 28]. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC4049056.

[17] Zaidi SSH, Sana M, Ghafoor MT. Assessment of total laboratory errors in clinical chemistry

laboratory: experience at a tertiary care hospital. Liaquat Natl J Prim Care [Internet]. 2022; 4(1): 6-10 [cited 2024 Oct 28]. Available from: https://journals.lnh.deu.pk/lnjpc/pdfl9da637ba-75bd-45bd-8029-6db074353181.pdf.

[18] Choosongsang P, Wannapong N, Choosongsang P, Wilairat B, Treerut P, Musigavon P, et al. Evaluation of analytical errors in the clinical chemistry laboratory of Songklanagarind hospital: A 5-year experience. J Med Tech Phy Ther. 2017; 29(1): 25-33. [in Thai].

[19] Arechep N, Wongkrajang P, Sutasanasuang J, Laiwejpithaya S. Error rate of pre-analytical process at laboratory in Thailand medical school. J Med Tech Assoc Thai. 2024; 52(1): 8933-43. [in Thai].

Journal of Associated Medical Sciences 2026; 59(1): 40-46

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 relationship between tinnitus characteristics, stress and depression in patients with subjective tinnitus

Kittiphorn Luengrungrus[1*] and Arnat Wannasri[2]

[1]Otolaryngology Unit, Nopparat Rajathanee Hospital, Department of Medical Services, Ministry of Public Health, Bangkok, Thailand.
[2]Nopparat Rajathanee Hospital, Department of Medical Services, Ministry of Public Health, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 12 May 2025
Accepted as revised 30 July 2025
Available online 11 September 2025

Keywords:
Tinnitus characteristics, stress, depression.

ABSTRACT

Background: Tinnitus is a perception of a noise in the ears. Individuals with tinnitus symptoms may experience disrupt daily life activities and mental health issues such as stress and depression.

Objectives: This study aimed to examine the relationship between the tinnitus symptoms and mental health conditions – stress and depression – in individuals with subjective tinnitus.

Materials and methods: This was an analytical cross-sectional study, involved 139 participants with subjective tinnitus aged between 18 and 70 years. All participants underwent clinical hearing tests. The tinnitus severity level was assessed by the Tinnitus Handicap Inventory in Thai version (THI), while the Stress Test 5 (ST-5) and the nine-question depression screening test (9Q) were used to assess stress and depression, respectively.

Results: After adjusting for tinnitus-related variables, the overall THI scores showed significantly positively associated with both the ST-5 and 9Q scores. Additionally, gender was also significantly positively associated with the 9Q scores. In contrast, other tinnitus-related factors including age, affected ear, tinnitus pattern and duration and average hearing thresholds in the affected ear, showed no statistically significant correlations with the ST-5 and 9Q scores.

Conclusion: This study highlights a significant association between the THI scores and both ST-5 and 9Q, indicating that tinnitus severity is closely linked to the development of stress and depression. Therefore, it is essential to monitor and manage psychological distress in patients with tinnitus.

Introduction

Tinnitus is recognized as a condition in which individuals perceive sound that is not generated by an external source.[1] The global incidence of tinnitus ranges from 5.10 to 42.70%, with its occurrence increasing in association with noise exposure and advancing age.[2] The pathophysiological mechanisms of tinnitus may result from dysfunction in the peripheral or central auditory systems or from somatosensory dysfunction.[3,4] Tinnitus can occur in individuals regardless of their hearing condition, with prevalence rates ranging from 85 to 96%.[4] Furthermore, tinnitus can significantly impact the daily lives of patients, with approximately 20% experiencing severe effects that lead to mental health issues.[2] Studies show that 31.29 to 65.00% of individuals with tinnitus experience stress,[5,6] while 48.80 to 74.00% suffer from depression.[7–9]

*** Corresponding contributor.**
Author's Address: *Otolaryngology Unit, Nopparat Rajathanee Hospital, Department of Medical Services, Ministry of Public Health, Bangkok, Thailand.*
E-mail address: kittiphorn.audiology.nrh@gmail.com
doi: 10.12982/JAMS.2026.006
E-ISSN: 2539-6056

The clinical diagnostic approach includes taking a case history, conducting audiology assessments, evaluating the severity of tinnitus, and applying management strategies that vary according to the severity of symptoms in each patient. According to the internal statistical data of Nopparat Rajathanee Hospital, approximately 80% of patients visiting the outpatient clinic for hearing tests report tinnitus. Furthermore, 25-30% of these patients return with recurring tinnitus symptoms.[10] Some patients might experience or develop anxiety related to the tinnitus condition, which negatively affects their daily lives. Recognizing the importance of mental health for tinnitus patients, alongside audiological assessments, individuals should be screened and monitored for their mental health status. This study utilized: the Tinnitus Handicap Inventory (THI) – a 25 item self-administered questionnaire to assess the severity of tinnitus in Thai version,[11] Stress Test 5 (ST-5) – a five question to assess current stress level[12] and the nine-question depression screening test (9Q) – a widely used self-administered questionnaire to assess depression.[13] Therefore, this study aimed to examine the relationship between tinnitus symptoms and mental health status – stress and depression – in patients with subjective tinnitus.

Materials and methods

Participants

This cross-sectional analytical study included 139 patients aged 18-70 years with subjective tinnitus who received care at the Department of Otorhinolaryngology, Nopparat Rajathanee Hospital. Inclusion criteria were subjective tinnitus in at least one ear, age 18-70 years, and adequate communication skills, including the ability to read and write in Thai. Exclusion criteria were: conductive or mixed hearing loss, retrocochlear pathology, diagnosed psychiatric disorders, or inability to cooperate with hearing tests or complete all three questionnaires. Written informed consent was obtained from all participants prior to enrollment.

Ethics committee

This study received approval from the Research and Ethics Committee of Nopparat Rajathanee Hospital on July 3, 2024 (Certificate number: 41/2567).

Examination and evaluation

Audiological assessments included otoscopy to examine the ear canal and tympanic membrane, followed by pure tone audiometry to measure air conduction (250-8,000 Hz) and bone conduction (500-4,000 Hz) thresholds. Participants completed the Thai versions of the THI, ST-5, and 9Q questionnaires, which were self-administered with clarification available if needed.

The THI evaluates the severity of tinnitus impact on daily life, with total scores ranging from 0 to 100, classified into five categories: slight (0-16), mild (18-36), moderate (38-56), severe (58-76), and catastrophic (78-100).[11] The ST-5 assesses the level of stress, where higher total scores indicate greater stress severity, categorized as low (0-4), moderate (5-9), high (10-14), and severe stress (15-21).[12] The 9Q screens for depressive symptoms, with total scores interpreted as minimal (0-6), mild (7-12), moderate (13-18), and severe depression (19 or above).[13] The entire process was completed within the same day, with no follow-up required.

Statistical analysis

Descriptive statistics, including frequency counts, percentages, means, and standard deviations, were used to summarize characteristics data, hearing thresholds, THI scores, ST-5 scores, and 9Q scores. To examine the relationships between tinnitus symptoms, stress, and depression three main models were performed separately using simple and multiple linear regression applied to investigate the crude odd ratios for a single independent variable, while multiple linear regression was used to adjust for potential confounders, including sex, age, tinnitus ear, pattern of tinnitus, duration of tinnitus, average hearing threshold and THI scores.

Results

Among 139 tinnitus patients (62 males, 77 females, mean age 50.37 years, and SD±13.14 years) found that the left ear exhibited tinnitus more frequently than the right ear. Most participants experienced tinnitus for a duration of 0-3 months, accounting for 64.03% (89 cases). The average hearing thresholds are 30.99 and 33.78 dB HL for the frequency ranges of 500-2,000 Hz and 500-4,000 Hz, respectively, indicating mild hearing loss. The overall mean THI-score was 44.12 (SD=26.38), classified as moderate tinnitus severity level. Meanwhile, the mean ST-5 score was 4.57 (SD=3.71), indicating a moderate stress level, and the mean 9Q score was 5.39 (SD=4.90), suggesting no to very mild depression (Table 1).

Table 1. Characteristics of tinnitus patients (N=139).

Characteristics (N=139)	N (%)
Sex	
Male	62 (44.60%)
Female	77 (55.40%)
Age (years), mean±SD	50.37±13.14
Tinnitus ear	
Right ear	68 (48.92%)
Left ear	71 (51.08%)
Pattern of tinnitus	
Intermittent tinnitus	61 (43.88%)
Continuous tinnitus	78 (56.12%)
Duration of tinnitus (days), mean±SD	273.31±493.54
0-3 months	89 (64.03%)
>3-6 months	5 (3.60%)
>6-12 months	20 (14.39%)
>12 months	25 (17.99%)
Average hearing threshold (dB HL), mean±SD	
Frequency 500-2,000 Hz	30.99±22.28
Frequency 500-4,000 Hz	33.78±22.20
THI score, mean±SD	
Function subscale	19.01±12.79
Emotion subscale	13.64±10.88
Catastrophic subscale	11.47±5.01
Overall THI score	44.12±26.38
ST-5 score, mean±SD	4.57±3.71
9Q score, mean±SD	5.39±4.90

Simple linear regression revealed that individuals with continuous tinnitus reported significantly higher stress scores than those with intermittent tinnitus (ß=1.334, 95% CI: 0.096-2.572). THI scores were also positively associated with stress, with each 1-point increase corresponding to a 0.095-point rise in stress scores (ß=0.095, 95% CI: 0.078-0.113). After adjusting for tinnitus-related variables, the association between continuous tinnitus and stress was attenuated and no longer significant (ß=0.308, 95% CI: -0.685-1.301), while the THI–stress relationship remained significant and stable (ß=0.094, 95% CI: 0.075-0.113). Other factors—including age, sex, affected ear, tinnitus pattern and duration, and average hearing thresholds—were not significantly associated with stress scores (Table 2).

The unadjusted analysis showed a significant positive association between tinnitus severity and depression, with each 1-point increase in THI score linked to a 0.128-point rise in depression score (ß=0.128, 95% CI: 0.106-0.151). This association remained unchanged after adjusting for tinnitus-related variables (ß=0.128, 95% CI: 0.104-0.152). Female gender was also significantly associated with higher depression scores in both models—2.070 points higher than males in the unadjusted model (ß=2.070, 95% CI: 0.448-3.692), decreasing to 1.312 points after adjustment (ß=1.312, 95% CI: 0.063-2.574). Other factors, including age, affected ear, tinnitus pattern and duration, and average hearing thresholds, showed no significant association with depression scores (Table 3).

Table 2. Association between tinnitus characteristics and stress scores (ST-5).

Variable	Unadjusted results		Adjusted results	
	ß	95% CI	ß	95% CI
Age	-0.024	-0.071-0.024	-0.025	-0.062-0.012
Sex				
Male	Reference		Reference	
Female	0.910	-0.337-2.156	0.245	-0.745-1.236
Tinnitus ear				
Right ear	Reference		Reference	
Left ear	0.278	-0.970-1.526	-0.092	-1.062-0.877
Pattern of tinnitus				
Intermittent tinnitus	Reference		Reference	
Continuous tinnitus	1.334*	0.096-2.572	0.308	-0.685-1.301
Duration of tinnitus	-0.001	-0.002-0.001	0.000	-0.001-0.001
Average hearing threshold at frequency 500-2,000 Hz	0.012	-0.016-0.040	0.017	-0.100-0.135
Average hearing threshold at frequency 500-4,000 Hz	0.007	-0.021-0.035	-0.022	-0.141-0.096
THI scores	0.095*	0.078-0.113	0.094*	0.075-0.113

***Note**: *p<0.05.*

Table 3. Association between tinnitus characteristics and depression scores (9Q).

Variable	Unadjusted results		Adjusted results	
	ß	95% CI	ß	95% CI
Age	-0.039	-0.101-0.024	-0.046	-0.093-0.001
Sex				
Male	Reference		Reference	
Female	2.070*	0.448-3.692	1.312*	0.063-2.574
Tinnitus ear				
Right ear	Reference		Reference	
Left ear	1.048	-0.592-2.689	0.437	-0.792-1.666
Pattern of tinnitus				
Intermittent tinnitus	Reference		Reference	
Continuous tinnitus	1.423	-0.222-3.067	-0.106	-1.365-1.153
Duration of tinnitus	-0.001	-0.002-0.001	0.001	-0.001-0.002
Average hearing threshold at frequency 500-2,000 Hz	0.022	-0.015-0.059	-0.056	-0.205-0.093
Average hearing threshold at frequency 500-4,000 Hz	0.018	-0.019-0.055	0.061	-0.089-0.211
THI scores	0.128*	0.106-0.151	0.128*	0.104-0.152

***Note**: *p<0.05.*

Discussion

Relationship between tinnitus characteristics and stress

Simple linear regression analysis revealed that individuals with continuous tinnitus reported stress scores that were 1.334 points higher than those with intermittent tinnitus. However, this study did not observe a relationship between the characteristics of tinnitus and stress scores after adjusting for potential confounders. This is consistent with previous study investigating the relationship between the characteristics of tinnitus and problems, which indicated that stress was presented in both intermittent and continuous tinnitus meaning that tinnitus status was not significantly associated with stress.[14] Another study measuring cortisol to assess HPA axis reactivity found that individuals with tinnitus exhibited a blunted cortisol response, similar to other stress-related conditions, suggesting a link between tinnitus and stress.[15] Unfortunately, this study did not classify tinnitus status. As there is limited study on the relationship between tinnitus, and with limited research on its relationship with stress, further studies are needed for stronger evidence. Additionally, THI scores showed a significant positive correlation with stress, indicating that greater tinnitus severity is associated with higher stress levels. These findings support earlier studies linking tinnitus severity to stress. Assessing both tinnitus severity and mental health is crucial to mitigate its impact on quality of life. This association persists across different stress measurement tools.[16,17]

Other variables, including sex, age, tinnitus ear, duration of tinnitus, and average hearing thresholds in the affected ear, showed no statistically significant correlation with ST-5 scores. Although some previous studies have reported results consistent with these findings, others have demonstrated contradictory outcomes, highlighting the complexity and variability of factors associated with stress in tinnitus patients. Jae Hee Lee *et al*.[16] reported that neither tinnitus loudness nor frequency was significantly associated with tinnitus-related distress. However, younger patients showed significantly higher levels of stress and anxiety, indicating that reduced tolerance to tinnitus may exacerbate psychological distress in this group. Furthermore, Murtaza *et al*.[17] found no significant link between stress and tinnitus duration, type, or perception, but did observe a correlation with the affected ear, suggesting lateralization may influence stress responses. Discrepancies between their findings and those of the present study may stem from differences in psychological assessment tools and statistical methods employed. Moreover, several studies have demonstrated that hearing loss contributes to increased stress, with greater impairment linked to higher stress levels in individuals with tinnitus.[18-21]

Relationship between tinnitus characteristics and depression

In the present study, THI scores were found to be significantly associated with depression scores (9Q), indicating that greater tinnitus severity is correlated with higher prevalence and intensity of depressive symptoms. Several previous studies have demonstrated a consistent positive association between tinnitus severity, measured by THI scores, and depressive symptoms, despite using different depression assessment tools.[22-27] These consistent findings across various tools reinforce the robust link between tinnitus severity and depression.

Previous studies have linked tinnitus to mental health disorders, particularly through the limbic system, which regulates emotion and behavior. Functional imaging shows limbic activation in both tinnitus and anxiety, suggesting neurobiological overlap with auditory processing.[28] Additionally, elevated cortisol levels—common in anxiety and depression—are often found in tinnitus patients, possibly due to inner ear glucocorticoid receptors. This supports a feedback loop where cortisol exacerbates tinnitus symptoms.[29]

Besides gender was another demographic factor that significantly associated with depression. Females were more likely to experience depression compared to males, which may reflect gender-specific differences in neural processing. Recent investigations have highlighted the roles of estrogen and progesterone in cochlear function. Estrogen and progesterone modulate cochlear function and neural signaling, influencing excitatory–inhibitory balance in the auditory pathway. These hormonal effects may contribute to tinnitus onset and severity.[30] However, it is worth nothing that the relationship between gender and depression is controversial. Some studies found that tinnitus severity is closely associated with depressive symptoms, with a stronger link observed in males. These findings underscore the need to explore how gender influences the link between tinnitus severity and depression.[31,32]

Other variables—including age, tinnitus ear, duration of tinnitus, and average hearing thresholds in the affected ear—were not significantly associated with 9Q scores. This finding is consistent with previous researches [7,26] reported no statistically significant association between depressive symptoms and tinnitus-related characteristics. One previous study found that while the affected ear of tinnitus was significantly associated with depression, there was no significant correlation between depression and the duration and type of tinnitus.[17] However, several other studies have demonstrated significant associations between tinnitus characteristics and depression. For example, Zhang *et al*. reported that higher pure-tone average thresholds were significantly correlated with the depression scale, linking hearing loss to increased

depression severity.[8] In addition, Prolonged tinnitus has been consistently associated with an elevated risk of depression, indicating that its chronic nature may substantially contribute to psychological distress.[33,34]

Limitations

This study lacks analysis of other factors that may affect stress and depression, such as medical conditions, medication use, and socioeconomic status. Future studies should include these variables and use longitudinal designs to clarify the temporal and causal relationships between tinnitus and psychological distress.

Conclusion

This study investigated tinnitus-related factors influencing stress and depression. After adjusting for relevant covariates, the overall THI score was positively correlated with both ST-5 and 9Q scores, indicating that greater tinnitus severity is associated with higher stress and depression levels. These findings highlight the importance of assessing symptom severity and mental health in individuals with tinnitus to mitigate potential impacts on quality of life.

Ethical approval

This study was approved by the Research and Ethics Committee of Nopparat Rajathanee Hospital on July 3, 2024 (Certificate No. 41/2567). All participants provided written informed consent.

Funding

None

Conflict of interest

The authors declared no potential conflicts of interest.

CRedit authorship contribution statement

Kittiphorn Luengrungrus: conceptualization, methodology, investigation, data curation, data dollection, formal analysis, writing – original draft; **Arnat Wannasri**: supervision, validation, writing – review and editing.

Acknowledgements

We would like to express our heartfelt gratitude to all participants for their involvement in this study. Our appreciation also extends to the research collaborators for their invaluable support in data collection and analysis.

References

[1] Tyler RS, Noble W, Coelho C, Roncancio ER, Jun HJ. Tinnitus and hyperacusis. In: Katz J, editor. Handbook of clinical audiology. 7th Ed. Philadelphia: Wolters Kluwer Health; 2015. p. 647-72.

[2] McCormack A, Edmondson-Jones M, Somerset S, Hall D. A systematic review of the reporting of tinnitus prevalence and severity. Hear Res. 2016; 337: 70-9. doi: 10.1016/j.heares.2016.05.009.

[3] Hackenberg B, Döge J, O'Brien K, Bohnert A, Lackner KJ, Beutel ME, et al. Tinnitus and its relation to depression, anxiety, and stress—a population-based cohort study. J Clin Med. 2023; 12(3): 1169. doi: 10.3390/jcm12031169.

[4] Scheffer AR, Mondelli MFCG. Tinnitus and hearing survey: cultural adaptation to Brazilian Portuguese. Braz J Otorhinolaryngol. 2021; 87(1): 28-34. doi: 10.1016/j.bjorl.2019.06.009.

[5] Ciminelli P, Machado S, Palmeira M, Carta MG, Beirith SC, Nigri ML, et al. Tinnitus: The sound of stress? Clin Pract Epidemiol Ment Health. 2018; 14(1):264-9.doi:10.2174/1745017901814010264.

[6] Musleh A, Alharthy AKH, Alzahrani MYM, Bin Maadhah SA, Al Zehefa IA, AlQahtani RY, et al. Psychological impact and quality of life in adults with tinnitus: A cross-sectional study. Cureus. 2024; 16(1): e51976. doi: 10.7759/cureus.51976.

[7] Zöger S, Svedlund J, Holgers KM. Relationship between tinnitus severity and psychiatric disorders. Psychosomatics. 2006; 47(4): 282-8. doi: 10.1176/appi.psy.47.4.282.

[8] Zhang L, Du H, You H. Correlation between the degree of hearing loss and the levels of anxiety and depression in patients with tinnitus. Noise Health. 2023; 25(118): 195-201. doi: 10.4103/nah.nah_46_23.

[9] Tunkel DE, Bauer CA, Sun GH, Rosenfeld RM, Chandrasekhar SS, Cunningham ER, et al. Clinical practice guideline: tinnitus. Otolaryngol Head Neck Surg. 2014; 151(2 Suppl): S1-40. doi: 10.1177/019 4599814545325.

[10] Department of Otolaryngology, Nopparat Rajathanee Hospital. Outpatient statistics of patients receiving hearing examinations, Fiscal year 2024. Bangkok: Nopparat Rajathanee Hospital; 2024. (in Thai).

[11] Limviriyakul S, Supavanich W. The validity and reliability of tinnitus handicap inventory Thai version. J Med Assoc Thai [Internet]. 2012; 95(11): 1433-40. Available from: https://pubmed.ncbi.nlm.nih.gov/ 23252210/.

[12] Department of Mental Health, Thailand. Stress Test-5 (ST-5) [Internet]. 2019 [cited 2024 May 12]. Available from: https://dmh.go.th/test/Download/view.asp?id=18.

[13] Department of Mental Health, Thailand. 9-Question Depression Screening Test (9Q) [Internet]. 2019 [cited 2024 May 12]. Available from: https://dmh.go.th/test/Download/view.asp?id=22

[14] Kim SY, Jeon YJ, Lee JY, Kim YH. Characteristics of tinnitus in adolescents and association with psychoemotional factors. Laryngoscope. 2017; 127(9): 2113-9. doi: 10.1002/lary.26334.

[15] Hébert S, Lupien SJ. The sound of stress: Blunted cortisol reactivity to psychosocial stress in tinnitus sufferers. Neurosci Lett. 2007; 411(2): 138-42. doi: 10.1016/j.neulet.2006.10.028.

[16] Lee JH, Ra JJ, Kim YH. Adequacy of the simplified version of the Tinnitus Handicap Inventory (THI-S) to measure tinnitus handicap and relevant distress. Korean J Audiol. 2014; 18(1): 19-27. doi: 10.7874/kja.2014.18.1.19.

[17] Murtaza M, Saqulain G, Rao RA. Characteristics of tinnitus and its association with clinico-demographic factors and mental health. J Riphah Coll Rehabil Sci. 2024; 12(1): 3-9. doi: 10.53389/JRCRS.2024120102.

[18] Ratnayake SAB, Jayarajan V, Bartlett J. Could an underlying hearing loss be a significant factor in the handicap caused by tinnitus? Noise Health. 2009; 11(44): 156-60. doi: 10.4103/1463-1741.53362.

[19] Wang Y, Zhang JN, Hu W, Li JJ, Zhou JX, Zhang JP, et al. The characteristics of cognitive impairment in subjective chronic tinnitus. Brain Behav. 2018; 8(3): e00918. doi: 10.1002/brb3.918.

[20] Waechter S. Association between hearing status and tinnitus distress. Acta Otolaryngol. 2021; 141(4): 381-5. doi: 10.1080/00016489.2021.

[21] Aazh H, Salvi R. The Relationship between severity of hearing loss and subjective tinnitus loudness among patients seen in a specialist tinnitus and hyperacusis therapy clinic in UK. J Am Acad Audiol. 2019; 30(8): 712-9. doi: 10.3766/jaaa.17144.

[22] Khilnani AK, Thaddanee R, Chodvadiya M, Khilnani G. Correlation of severity of tinnitus with severity of anxiety and depression in adults in a tertiary care hospital at Bhuj, Gujarat, India. Int J Otorhinolaryngol Head Neck Surg. 2018; 4(4): 1044. doi: 10.18203/issn.2454-5929.ijohns 20182710.

[23] Meijers SM, Lieftink AF, Stegeman I, Smit AL. Coping in chronic tinnitus patients. Front Neurol. 2020; 11: 570989. doi: 10.3389/fneur.2020.570989.

[24] Meijers SM, Rademaker M, Meijers RL, Stegeman I, Smit AL. Correlation between chronic tinnitus distress and symptoms of depression: A systematic review. Front Neurol. 2022; 13: 870433. doi: 10.3389/fneur.2022.870433.

[25] Zeman F, Koller M, Langguth B, Landgrebe M, Figueiredo R, Aazevedo A, et al. Which tinnitus-related aspects are relevant for quality of life and depression: results from a large international multicentre sample. Health Qual Life Outcomes. 2014; 12(1): 7. doi: 10.1186/1477-7525-12-7.

[26] Xu Y, Yao J, Zhang Z, Wang W. Association between sleep quality and psychiatric disorders in patients with subjective tinnitus in China. Eur Arch Oto-Rhino-Laryngology. 2016; 273(10): 3063-72. doi: 10.1007/s00405-016-3906-8.

[27] Weidt S, Delsignore A, Meyer M, Rufer M, Peter N, Drabe N, et al. Which tinnitus-related characteristics affect current health-related quality of life and depression? A cross-sectional cohort study. Psychiatry Res. 2016; 237: 114-21. doi: 10.1016/j.psychres.2016. 01.065.

[28] Lockwood AH, Salvi RJ, Coad ML, Towsley ML, Wack DS, Murphy BW. The functional neuroanatomy of tinnitus: Evidence for limbic system links and neural plasticity. Neurology. 1998; 50(1): 114–20. doi: 10.1212/ wnl.50.1.114.

[29] Yehuda R, Boisoneau D, Mason JW, Giller EL. Glucocorticoid receptor number and cortisol excretion in mood, anxiety, and psychotic disorders. Biol Psychiatry. 1993; 34(1-2): 18-25. doi: 10.1016/ 0006-3223(93)90252-9.

[30] Lee SS, Han K Do, Joo YH. Association of perceived tinnitus with duration of hormone replacement therapy in Korean postmenopausal women: a cross-sectional study. BMJ Open. 2017; 7(7): e013736. doi: 10.1136/bmjopen-2016-013736.

[31] Han TS, Jeong JE, Park SN, Kim JJ. Gender differences affecting psychiatric distress and tinnitus severity. Clin Psychopharmacol Neurosci. 2019; 17(1): 113-20. doi: 10.9758/cpn.2019.17.1.113.

[32] Gomaa MAM, Elmagd MHA, Elbadry MM, Kader RMA. Depression, Anxiety and Stress Scale in patients with tinnitus and hearing loss. Eur Arch Otorhinolaryngol. 2014; 271(8): 2177-84. doi: 10.1007/s00405-013-2715-6.

[33] Malakouti SK, Nojomi M, Mahmoudian S, Alifattahi N, Salehi M. Comorbidity of chronic tinnitus and mental disorders. Int Tinnitus J [Internet]. 2011; 16(2): 118-22. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22249870

[34] Niemann U, Brueggemann P, Boecking B, Mazurek B, Spiliopoulou M. Development and internal validation of a depression severity prediction model for tinnitus patients based on questionnaire responses and socio-demographics. Sci Rep. 2020; 10(1): 4664. doi: 10.1038/s41598-020-61593-z.

Journal of Associated Medical Sciences 2026; 59(1): 47-53

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

Performance comparisons of three rapid screening methods for the G6PD deficiency test in newborns

Jiranan Neamyanon[1], Aungkana Saejeng[2], Phaithoon Wongwian[3], Suparporn Kiti[3], Satitpong Nanjai[3], Thaworn Jaiyasen[3], Sakorn Pornprasert[1*]

[1]Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]The Office of Disease Prevention and Control Region 1 Chiang Mai, Chiang Mai Province, Thailand.
[3]Department of Medical Technology, Nakornping Hospital, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 24 June 2025
Accepted as revised 8 September 2025
Available online 12 September 2025

Keywords:
Glucose-6-phosphate dehydrogenase deficiency, rapid test, performance, point-of-care, newborns.

ABSTRACT

Background: Glucose-6-phosphate dehydrogenase (G6PD) deficiency is an X-linked recessive disorder that affects over 400 million people worldwide. The deficit causes individuals susceptible to hemolysis during oxidative stress. In newborns, G6PD deficiency can lead to hyperbilirubinemia, bilirubin-induced neurologic dysfunction, and kernicterus, making early detection and screening crucial.

Objectives: This study aimed to compare the diagnostic performance of three rapid screening tests for G6PD deficiency in newborns: the fluorescent spot test (FST), G6PD rapid test kit, and SD Biosensor, using spectrophotometry as the gold standard.

Materials and methods: Blood samples from 70 newborns were tested using these three methods. The diagnostic performances, including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and efficiency of each method were analyzed.

Results: Both the FST and G6PD rapid test kit exhibited higher specificity, PPV, and efficiency compared to the SD Biosensor. Nonetheless, the SD Biosensor exhibited superior sensitivity and NPV, but it was unable to identify G6PD activity in 16.4% of instances due to elevated hemoglobin concentrations.

Conclusion: The FST and G6PD rapid test kit are reliable and suitable for G6PD deficiency screening in newborns, especially in settings with limited resources, due to their high efficiency, specificity, and rapid results. The SD Biosensor remains a valuable tool in clinical contexts requiring high sensitivity. For newborns with high hemoglobin levels, the FST or G6PD rapid test is recommended for accurate screening. Further studies with larger sample sizes are necessary to confirm the reliability of these tests in diverse populations.

Introduction

Glucose-6-phosphate dehydrogenase (G6PD) deficiency is the most prevalent human enzyme deficit, exhibiting an X-linked recessive inheritance pattern, affecting over 400 million individuals globally, particularly among those of African, Asian, and Mediterranean ancestry.[1] The global prevalence of G6PD-deficient variants mirrors the geographical distribution of malaria, supporting the hypothesis that G6PD deficiency provides some level of protection against malaria.[2,3] In Thailand, several studies indicated

* **Corresponding contributor.**
Author's Address: *Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: sakornmi001@gmail.com
doi: 10.12982/JAMS.2026.007
E-ISSN: 2539-6056

a significant prevalence of G6PD deficiency, impacting roughly 16.9% of the population.[4-7] The previous reports revealed that the prevalence of G6PD deficiency among neonates at Rajavithi Hospital Bangkok, Thailand from May 1995 to July 1998 was 5.1%[7] and at King Memorial Chulalongkorn Hospital, Bangkok, Thailand from February to August 2021 was 6.5%.[8]

G6PD plays a crucial role in the first step of the hexose monophosphate pathway, also known as the pentose phosphate pathway (PPP). In this process, G6PD converts glucose-6-phosphate (G-6-P) into 6-phosphogluconolactone (6-PG), while also reducing nicotinamide adenine dinucleotide phosphate (NADP) to NADPH. In red blood cells, the PPP is the only source of NADPH, which is vital for protecting the cells from oxidative damage. This protection is mainly carried out by glutathione in its reduced form (GSH). The NADPH is essential for the regeneration of GSH by the enzyme glutathione reductase. Therefore, the primary function of NADPH in red blood cells is to maintain the reduced state of glutathione, which helps prevent oxidative damage to the cells. Individuals with deficiency are particularly susceptible to hemolytic events during infections or after exposure to fava beans and certain medications.[9] The G6PD deficiency in newborns presents an increased risk of hyperbilirubinemia, which may rapidly escalate and lead to bilirubin-induced neurologic dysfunction (BIND), potentially resulting in kernicterus, a permanent and devastating neurological damage. Therefore, screening for G6PD deficiency and close monitoring of affected infants are important.[10]

Bilirubin is primarily produced from the breakdown of hemoglobin in red blood cells. While small amounts of bilirubin serve as antioxidants, excessive accumulation can be cytotoxic, necessitating timely excretion from the body.[11] In neonates, bilirubin production is significantly higher than in adults due to the higher turnover and shorter lifespan of their erythrocytes.[12] However, their ability to eliminate bilirubin is less efficient than that of adults.[13] The combination of high bilirubin production and limited elimination capacity puts newborns at risk of developing neonatal hyperbilirubinemia, which if severe and untreated, can lead to kernicterus.[14] Fortunately, kernicterus is almost always preventable, but it requires timely and accurate detection of hyperbilirubinemia in newborns.[15] One of the most common causes of neonatal hyperbilirubinemia is G6PD deficiency.[16]

In hospital laboratories, the measurement of G6PD enzyme levels in blood for diagnosing G6PD deficiency can be performed using several methods. The gold standard for quantifying G6PD activity in red blood cells is the spectrophotometric assay, which measures G6PD activity by detecting the formation of NADPH, based on the difference in absorbance of the sample at 340 nm over time.[17,18] However, due to various limitations, such as budget constraints, many hospitals are unable to perform enzyme activity tests for G6PD deficiency using standard methods. In several hospitals, the fluorescent spot test (FST), which is endorsed by the International Committee of Standardization in Hematology (ICSH)[19], is still commonly used as a reliable screening method for G6PD deficiency. Additionally, many hospitals are increasingly adopting rapid test and biosensor devices for G6PD screening. In particular, the biosensor, which allows for quantitative measurement of G6PD enzyme levels, has gained prominence in screening practices. The SD Biosensor (STANDARD G6PD test, Suwon, Korea) is a new point-of-care device that provides a quantitative G6PD measurement, adjusted for hemoglobin levels, making it accessible for use in smaller clinics and laboratories.[20,21] Nevertheless, there is still limited research on the effectiveness of biosensor and rapid test methods. Further studies are needed to evaluate the accuracy and efficiency of these methods for G6PD deficiency screening, to establish their reliability as standard diagnostic tools. The aim of this study is to evaluate the efficacy of three rapid test kits for screening G6PD deficiency in neonates: FST, G6PD rapid test kit, and SD Biosensor, utilizing data from the spectrophotometry method as the reference.

Materials and methods

Blood samples

At Nakornping Hospital, 0.5 mL whole blood samples of newborns were collected in ethylenediaminetetraacetic acid (EDTA) anticoagulated tubes. The G6PD activity was measured using the G6PD kit (GPD0204, Mindray, Shenzhen, China) and an automated UV enzymatic analyzer (BS-360E, Mindray). The principle of the test kit is mainly based on light absorption. The NADP is reduced to NADPH by G6PD with the presence of its specific substrate, G-6-P. The enzyme activity can be determined by measuring the changes in absorbance rate at 340 nm due to the reduction of NADP. Based on the manufacturer's instructions, the blood samples which had G6PD enzymatic activity <3.8 U/gm Hb were diagnosed as G6PD deficiency. This technique was used as a reference method. The over-left blood samples were aliquoted into microcentrifuge tubes and sent to the Associated Medical Sciences-Clinical Service Center (AMS-CSC), Chiang Mai University, Chiang Mai, Thailand. The whole blood samples were stored in a refrigerator at 4 °C until used. The samples size was calculated using the Yamane's formula[22] as follow:

$$n = N / (1 + Ne^2)$$

n: represents the required sample size.

N: represents the total population size of neonates requiring G6PD activity assessment at Nakornping Hospital each month (83 samples).

e: represents the desired margin of error (0.05).

The required sample size (n) was 69 samples. As a result, whole blood samples from 70 babies were taken for this study and the flowchart illustrating the experimental process is shown in Figure 1.

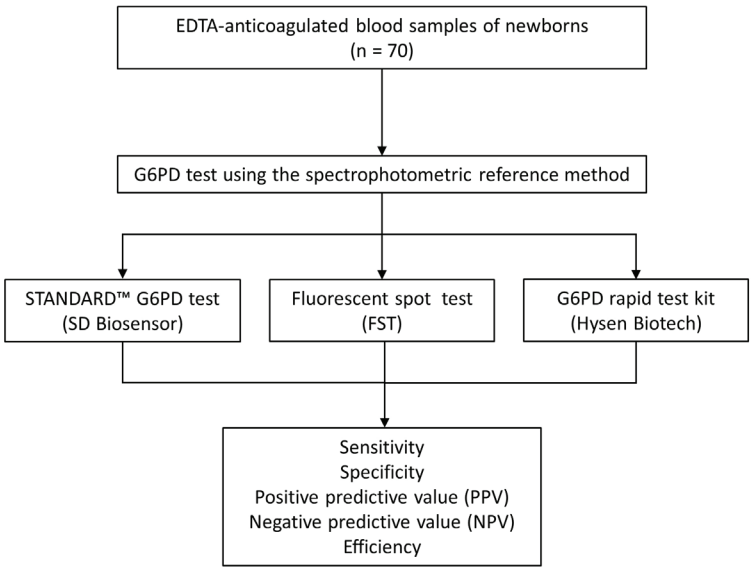


Figure 1. *The flowchart illustrating the experimental process.*

G6PD activity measurement using the SD Biosensor

The principle of SD Biosensor (STANDARD G6PD test, Suwon, Korea) is based on a colorimetric detection system for the automatic calculation of G6PD activity on the code chip for each test device. G6PD catalyzes the first step in the PPP, oxidizing G-6-P to 6-PG and reducing NADP to NADPH. When NADPH is generated by G6PD, the 5-bromo-4-3-indolyl-phosphate (BCIP) and nitro blue tetrazolium (NBT) are reduced by the diaphorase reaction to yield a violet color. The rate of the color production is directly proportional to the concentration of G6PD present in the specimen. The color intensity can be measured through reflectance photometry of the reduced BCIP and NBT. For the test, begin by removing the chip and inserting it into the device. Pipette 10 µL of whole blood using a STANDARD Ezi tube+ into the buffer tube that contains the extraction buffer. Mix with the STANDARD Ezi tube+ 8-10 times. Then, apply 10 µL of the mixed specimen using the STANDARD Ezi tube+ to the application hole of the test device. The result will appear on the screen after 2 minutes. The measuring ranges of the tests were as follows: total Hb 4-25 gm/dL and G6PD 0-20 U/g Hb. Interpretation of the test results was done according to the manufacturers' reference ranges: G6PD activity ≤4.0 U/g Hb was interpreted as G6PD deficiency for both males and females, G6PD activity ≥4.1 U/gm Hb was interpreted as normal for males while G6PD activity between 4.1 and 6.0, and ≥6.1 U/gm Hb was interpreted as G6PD intermediate and normal, respectively for females.

Detection of G6PD deficiency using the fluorescent spot test (FST)

The principle of FST is based on the G6PD enzyme's ability to oxidize G-6-P into 6-PG and reduce NADP to NADPH. The amount of NADPH generated is directly proportional to G6PD activity, observable through its fluorescence when excited by UV light at 340 nm.[23] For the test, 5 µL of whole blood was pipetted into test tube that contained 100 µL of G6PD screening reagent (Sigma-Aldrich, MO, USA) and positive control tube that contained 100 µL of NADP free reagent. The reagents were incubated at room temperature, and 10 µL of the mixture was spotted onto Whatman filter paper after incubation for 5 and 15 minutes. The spots were air-dried and then examined under ultraviolet light with a wavelength range of 340 nm. The blood samples with control reagent showed no fluorescence. Positive results, indicating G6PD deficiency, occur when the spot shows no fluorescence at both 5 and 15 minutes. Negative results, indicating normal G6PD activity, are determined when the spot fluoresces under UV light at both 5 and 15 minutes. In addition, the intermediate G6PD activity can be interpreted when the spot shows no fluorescence at 5 minutes but fluorescence at 15 minutes

Detection of G6PD deficiency using G6PD rapid test kit

The principle of G6PD rapid test kit (Hysen Biotech, Seoul, South Korea) is based on the G-6-P substrate colorimetric method. In the presence of NADP, G-6-P is oxidized by G6PD enzyme in the sample, resulting in the production of 6-PG and NADPH. The light yellow NBT is reduced to insoluble blue-purple crystalline formazan in the reaction between NADPH and phenazine methosulfate (PMS). The experiment was conducted in accordance with the manufacturer's instructions. Specifically, 300 µL of buffer was introduced into the tube containing the reagent pad, which was subsequently shaken to dissolve the colored substance from the pad into the buffer. Following this, 10 µL of whole blood samples were added to the tube. The sample-reagent mixture was incubated for 10 minutes. The card was then taken out and the reacted buffer was deposited into the designated hole of the card. After approximately 1 minute, the liquid was fully absorbed, allowing for

result interpretation. Positive results, indicative of G6PD deficiency, manifest as pink, white, or yellow coloration in the center of the card. Conversely, negative results, signifying normal G6PD activity, present as blue, blue-purple, or black coloration in the center of the detection card.

Statistical analysis

Statistical analyses were performed using Microsoft Excel version 2021. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and efficiency of each approach were estimated based on the definitions of positive and negative results, which correspond to G6PD deficient and non-deficient outcomes, respectively. The formulas were as follows:

Sensitivity = true positive/(true positive + false negative) × 100%

Specificity = true negative/(true negative + false positive) × 100%

Positive predictive value (PPV) = true positive/(true positives + false positive) × 100%

Negative predictive value (NPV) = true negative/(true negative + false negative) × 100%

Efficiency = (true positive + true negative)/(true positive + true negative + false positive + false negative) × 100%

Results

Venous blood samples in EDTA anticoagulated tubes were collected from 70 newborns for this study. Their characteristics and hematological data, including total Hb and reticulocyte counts and G6PD enzymatic activity assessed by gold standard method are shown in Table 1. Based on the ultraviolet (UV) spectrophotometric method, a gold standard method, the G6PD deficiency was found in 22 (31.4%) newborns, while 48 (68.6%) newborns had normal G6PD enzyme activity.

The screening results for G6PD deficiency using the FST and G6PD rapid test kit of 70 newborns showed identical outcomes, detecting G6PD deficiency in 22 newborns, while 48 newborns had normal G6PD enzyme activity. Among those, there was one false positive and one false negative case. Thus, the sensitivity, specificity, PPV, NPV, and efficiency were calculated, with the results shown in Table 2.

Among the 70 samples, the SD Biosensor was conducted on 67 neonates due to insufficient blood samples from three individuals for testing. Moreover,

Table 1. Characteristics and hematological data of samples.

Characteristics and hematological data (N=70)		Reference ranges
Age (days)	3.5±2.9 (1-17)	1-28
Gender: Male/Female	45 / 25	
Total Hb (g/dL)	17.3±2.3 (11.5-21.9)	17.0-20.0
Reticulocyte counts (%)	5.0±1.9 (1.7-8.9)	2.0-6.0
G6PD activity measured by gold standard method (U/gm Hb)	18.6±12.9 (1.3-45.5)	≥3.8

Note: *The values are presented as mean±SD (ranges).*

Table 2. Performances of three methods.

	Reference Method		Total	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Efficiency
	Deficiency	Normal						
FST (N=70)								
Deficiency	21	1	22					
Normal	1	47	48	95.4	97.9	95.4	97.9	97.1
Total	22	48	70					
G6PD rapid test kit (N=70)								
Deficiency	21	1	22					
Normal	1	47	48	95.4	97.9	95.4	97.9	97.1
Total	22	48	70					
SD Biosensor (N=56)								
Deficiency	20	4	24					
Normal	0	32	32	100.0	88.9	83.3	100.0	92.8
Total	20	36	56					

results were indeterminate for 11 of 67 (16.4%) individuals because their total Hb levels exceeded the upper detectable limits (>25 gm/dL) of the SD Biosensor. Based on the gold standard method, 36 of 56 remaining samples had normal G6PD enzyme activity and 20 were G6PD deficiency.

The G6PD deficiency analysis conducted by SD Biosensor revealed that 32 samples had normal G6PD enzyme activity, while 20 samples demonstrated G6PD deficiency. Furthermore, one female identified by SD Biosensor demonstrated intermediate enzyme activity (4.7 u/gm Hb). However, the reference method lacks a defined threshold for intermediate activity, and the G6PD enzyme activity of this sample, analyzed by the gold standard method at 30.3 U/g Hb, was within normal values. Consequently, this case was categorized into the G6PD false positive group for performance evaluations. Additionally, three other false positive cases were identified by SD Biosensor. Thus, the sensitivity, specificity, PPV, NPV, and efficiency were calculated, with the results presented in Table 2.

The performances of each test were analyzed by using the diagnostic results obtained from the gold standard method as references. The results showed that the FST and G6PD rapid test kit had a higher specificity, PPV, and efficiency than the SD Biosensor. However, the former methods had a lower sensitivity and NPV than the later method (Table 2).

Discussion

In this study, we evaluated the performance of FST (Sigma-Aldrich), G6PD rapid test kit (Hysen Biotech), and SD Biosensor (STANDARD™ G6PD Test) for screening G6PD deficiency in newborns by comparing their diagnostic results with those of enzymatic assay, which is considered the gold standard method. The results indicated that both FST and the G6PD rapid test kit exhibited higher specificity, PPV and efficiency compared to the SD Biosensor (Table 2). In contrast, the SD Biosensor demonstrated the highest sensitivity and NPV (Table 2), consistent with previous studies on biosensor-based methods. Pal et al found that the SD Biosensor performed similarly to a reference assay, with a sensitivity of 95.5-100% and specificity of 97%, making it a reliable option for diagnosing G6PD deficiency in both males and females across diverse clinical settings, including resource-limited areas.[20] Similarly, the study by Adu-Gyasi *et al*. reported that the CareStart G6PD deficiency RDT, a biosensor-based method, demonstrated a sensitivity of 100% and a specificity of 72.1%.[24] In the present study, the overall detection rate of the SD Biosensor was 83.6% for all neonates, with 11 tests (16.4%) failing to detect G6PD activity due to high Hb levels. This failure is particularly relevant in newborns, who tend to have higher Hb levels that exceed the detectable range of the SD Biosensor (Hb 4-25 gm/dL). An elevated Hb level (>25 gm/dL) not only affects the absorbance of the colorimetric reaction but can also retard the rate of G6PD enzymatic activity, potentially leading to false positive results as found in four cases (1 intermediate and 3 deficiencies) (Table 2). The G6PD activity in these four cases, as evaluated by the reference method, exceeded the cutoff value (3.8 U/g Hb) for G6PD deficiency, ranging from 19.1 to 30.3 U/g Hb. Consequently, to prevent uncorrected outcomes, the device displayed unanalyzed results in samples with elevated Hb levels, as observed in 11 samples. This makes it less optimal for use as a screening tool in newborns, as accurate screening in this population is crucial. Despite these limitations, the SD Biosensor remains a viable option for detecting G6PD deficiency in other contexts where high sensitivity is essential. However, additional confirmation may be required to avoid false positives.

Both the FST and G6PD rapid test kit are highly suitable for clinical use due to their excellent specificity, PPV, and efficiency; ease of use; and quick turnaround times, approximately 20 minutes and 15 minutes, respectively, for each sample. These methods are particularly advantageous in settings where rapid screening and diagnosis are required, such as newborn care or regions with limited healthcare access.

Several studies have evaluated the performance of the FST in screening for G6PD deficiency. Jiang et al confirmed that the FST is a reliable and convenient screening method, demonstrating high sensitivity (92-100%) and specificity (98%).[25] Similarly, Keihanian *et al*. reported that the FST performed well in clinical settings, with a sensitivity of 91.4%, specificity of 99.9%, NPV of 99.4%, and PPV of 97.7%.[26] Overall, these studies indicate that the FST provides acceptable sensitivity and specificity for detecting G6PD activity in newborns. However, there are certain challenges with this approach. For example, the process requires scientific expertise and UV light.

A study by Tinley *et al*. evaluated the BinaxNOW® G6PD test, a rapid qualitative enzyme chromatographic test (ECT) that detects G6PD activity by reducing nitro blue tetrazolium dye to a blue formazan product, like the G6PD rapid test kit. The test demonstrated high sensitivity (98%) and specificity (97-98%).[27]

In clinical settings, the enzymatic assay remains the gold standard for confirming G6PD deficiency.[26] However, for hospitals with limited equipment or resources, both FST and the G6PD rapid test kit are reliable alternatives. They provide high sensitivity and specificity, and their quick results make them an excellent choice for rapid screening in newborn care settings. On the other hand, the G6PD rapid test kit offers a faster turnaround time compared to FST, making it an attractive option for hospitals with budgetary constraints. Furthermore, even though FST is a qualitative method, it can potentially identify samples with G6PD intermediate activity. Consequently, in clinical practice, the G6PD test should be initially conducted using the FST or G6PD rapid test kit, followed by

quantification of G6PD enzymatic activity in deficient samples using the SD Biosensor or an automated UV enzymatic analyzer at the central laboratory.

Limitations

The limitation of this study is a small sample size (70 newborns), which may limit the generalizability of the results to larger or more diverse populations. Therefore, to verify the validity of these tests across a range of demographics, more research with larger sample sizes and/or a multi-center study is required.

Conclusion

The FST and G6PD rapid test kit are reliable and suitable for G6PD deficiency screening in newborns, especially in settings with limited resources, due to their high efficiency, specificity, and rapid results. The SD Biosensor remains a valuable tool in clinical contexts requiring high sensitivity. For newborns with high hemoglobin levels, FST and G6PD rapid test kit are recommended for accurate screening. Further studies with larger sample sizes are necessary to confirm the reliability of these tests in diverse populations.

Ethical approval

This study was approved by the Ethics Committee of the Faculty of Associated Medical Sciences at Chiang Mai University (approval No. AMSEC67EM-035). In addition, it was also subsequently submitted to the Nakornping Hospital Ethics Committee for approval regarding research involving human subjects (approval No. NKP165/67).

Funding

This this work was supported by a part of the research grant from the Faculty of Associated Medical Sciences, Chiang Mai University, Thailand.

Conflict of interest

The authors state no conflict of interests.

CRediT authorship contribution statement

Jiranan Neamyanon: conceptualization, investigation, methodology, writing original draft; **Aungkana Saejeng:** conceptualization, resources, review and editing; **Phaithoon Wongwian:** blood and data collections, investigation, review and editing; **Suparporn Kiti:** blood and data collections, investigation, review and editing; **Satitpong Nanjai:** blood and data collections, investigation, review and editing; **Thaworn Jaiyasen:** blood and data collections, investigation, review and editing; **Sakorn Pornprasert:** conceptualization, project administration, validation, writing, reviewing and editing.

Acknowledgements

The authors thank the technicians at the Associated Medical Sciences Clinical Service Center, Faculty of Associated Medical Sciences, Chiang Mai University, and the Department of Laboratory, Nakornping Hospital, Chiang Mai, Thailand for their assistance. This study was supported by grants from the Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand.

References

[1] Cappellini MD, Fiorelli G. Glucose-6-phosphate dehydrogenase deficiency. Lancet. 2008; 371 (9606): 64-74. doi: 10.1016/S0140-6736(08)60073-2.

[2] Vulliamy T, Othman A, Town M, Nathwani A, Falusi A, Mason P, et al. Polymorphic sites in the African population detected by sequence analysis of the glucose-6-phosphate dehydrogenase gene outline the evolution of the variants A and A-. Proc Natl Acad Sci U S A. 1991; 88(19): 8568-71. doi: 10.1073/pnas.88.19.8568.

[3] Howes RE, Battle KE, Satyagraha AW, Baird JK, Hay SI. G6PD deficiency: global distribution, genetic variants and primaquine therapy. Adv Parasitol. 2013; 81: 133-201. doi: 10.1016/B978-0-12-407826-0.00004-7.

[4] Nuchprayoon I, Sanpavat S, Nuchprayoon S. Glucose-6-phosphate dehydrogenase (G6PD) mutations in Thailand: G6PD Viangchan (871G>A) is the most common deficiency variant in the Thai population. Hum Mutat. 2002; 19(2): 185. doi: 10.1002/humu.9010.

[5] Nuinoon M, Krithong R, Pramtong S, Sasuk P, Ngeaiad C, Chaimusik S, et al. Prevalence of G6PD deficiency and G6PD variants amongst the southern Thai population. Peer J. 2022; 10:e14208. doi: 10.7717/peerj.14208.

[6] Charoenkwan P, Tantiprabha W, Sirichotiyakul S, Phusua A, Sanguansermsri T. Prevalence and molecular characterization of glucose-6-phosphate dehydrogenase deficiency in northern Thailand. Southeast Asian J Trop Med Public Health. 2014; 45(1): 187-93.

[7] Ratrisawadi V, Horpaopan S, Chotigeat U, Sangtawesin V, Kanjanapattanakul W, Ningsanond V, et al. Neonatal screening program in Rajavithi Hospital, Thailand. Southeast Asian J Trop Med Public Health. 1999; 30 (Suppl 2):28-32.

[8] Pimpakan T, Mungkalasut P, Tansakul P, Chanda M, Jugnam-Ang W, Charucharana S, et al. Effect of neonatal reticulocytosis on glucose 6-phosphate dehydrogenase (G6PD) activity and G6PD deficiency detection: a cross-sectional study. BMC Pediatr. 2022; 22(1): 678. doi: 10.1186/s12887-022-03740-1.

[9] WHO. Glucose-6-phosphate dehydrogenase deficiency. WHO Working Group. Bull World Health Organ. 1989; 67(6): 601-11.

[10] Lee HY, Ithnin A, Azma RZ, Othman A, Salvador A, Cheah FC. Glucose-6-phosphate dehydrogenase deficiency and neonatal hyperbilirubinemia: insights on pathophysiology, diagnosis, and gene variants in disease heterogeneity. Front Pediatr. 2022; 10:

875877. doi: 10.3389/fped.2022.875877.
[11] Stocker R, Yamamoto Y, McDonagh AF, Glazer AN, Ames BN. Bilirubin is an antioxidant of possible physiological importance. Science. 1987; 235(4792): 1043-6. doi: 10.1126/science.3029864.
[12] Brouillard RP. Measurement of Red Blood Cell Life-Span. JAMA. 1974; 230(9): 1304-5.
[13] Dennery PA, Seidman DS, Stevenson DK. Neonatal hyperbilirubinemia. N Engl J Med. 2001; 344(8): 581-90. doi: 10.1056/NEJM200102223440807.
[14] Ip S, Chung M, Kulig J, O'Brien R, Sege R, Glicken S, et al. An evidence-based review of important issues concerning neonatal hyperbilirubinemia. Pediatrics. 2004; 114(1): e130-53. doi: 10.1542/peds.114.1.e130.
[15] Liu H, Liu W, Tang X, Wang T. Association between G6PD deficiency and hyperbilirubinemia in neonates: a meta-analysis. Pediatr Hematol Oncol. 2015; 32(2): 92-8. doi: 10.3109/08880018.2014.887803.
[16] Huang CS, Chang PF, Huang MJ, Chen ES, Chen WC. Glucose-6-phosphate dehydrogenase deficiency, the UDP-glucuronosyl transferase 1A1 gene, and neonatal hyperbilirubinemia. Gatroenterology. 2002; 123(1): 127-33. doi: 10.1053/gast.2002.34173.
[17] Ley B, Bancone G, von Seidlein L, Thriemer K, Richards JS, Domingo GJ, et al. Methods for the field evaluation of quantitative G6PD diagnostics: a review. Malar J. 2017; 16(1): 361. doi: 10.1186/s12936-017-2017-3.
[18] Pfeffer DA, Ley B, Howes RE, Adu P, Alam MS, Bansil P, et al. Quantification of glucose-6-phosphate dehydrogenase activity by spectrophotometry: A systematic review and meta-analysis. PLoS Med. 2020; 17(5): e1003084. doi: 10.1371/journal.pmed.1003084.
[19] Saejeng A, Khantikul N, Boonin P, Anantasuk P, Srithep S, Pornprasert S. Modified Fluorescent Spot Test for screening G6PD deficiency among malaria patients in remote health care services, Northern Thailand. LPHJ. 2018; 14(2): 12-21.
[20] Pal S, Bansil P, Bancone G, Hrutkay S, Kahn M, Gornsawun G, et al. Evaluation of a Novel 1 Quantitative Test for Glucose-6-Phosphate Dehydrogenase Deficiency: Bringing Quantitative Testing for Glucose-6-Phosphate Dehydrogenase Deficiency Closer to the Patient. Am J Trop Med Hyg. 2019; 100(1): 213-21. doi: 10.4269/ajtmh.18-0612.
[21] Sadhewa A, Satyagraha AW, Alam MS, Adissu W, Anvikar A, Bancone G, et al. Performance of quantitative point-of-care tests to measure G6PD activity: An individual participant data meta-analysis. PLoS Negl Trop Dis. 2025; 19(3): e0012864. doi: 10.1371/journal.pntd.0012864.
[22] Yamane T. Statistics: An Introductory Analysis. 3rd ed. New York: Harper and Row; 1973.
[23] Beutler E, Blume KG, Kaplan JC, Löhr GW, Ramot B, Valentine WN. International Committee for Standardization in Haematology: recommended screening test for glucose-6-phosphatedehydrogenase (G-6-PD) deficiency. Br J Haematol. 1979; 43(3): 465-7. doi: 10.1111/j.1365-2141.1979.tb03774.x.
[24] Adu-Gyasi D, Asante KP, Newton S, Dosoo D, Amoako S, Adjei G, et al. Evaluation of the diagnostic accuracy of CareStart G6PD deficiency Rapid Diagnostic Test (RDT) in a malaria endemic area in Ghana, Africa. PLoS One. 2015; 10(4): e0125796. doi: 10.1371/journal.pone.0125796.
[25] Jiang J, Ma X, Song C, Lin B, Cao W, Wu S, et al. Using the fluorescence spot test for neonatal screening of G6PD deficiency. Southeast Asian J Trop Med Public Health. 2003; 34 (Suppl 3): 140-2.
[26] Thielemans L, Gornsawun G, Hanboonkunupakarn B, Paw MK, Porn P, Moo PK, et al. Diagnostic performances of the fluorescent spot test for G6PD deficiency in newborns along the Thailand-Myanmar border: A cohort study. Wellcome Open Res. 2018; 3(1). doi: 10.12688/wellcomeopenres.13373.1.
[27] Tinley KE, Loughlin AM, Jepson A, Barnett ED. Evaluation of a rapid qualitative enzyme chromatographic test for glucose-6-phosphate dehydrogenase deficiency. Am J Trop Med Hyg. 2010; 82(2): 210-4. doi: 10.4269/ajtmh.2010.09-0416.

Journal of Associated Medical Sciences 2026; 59(1): 54-61

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 exergaming on balance among children with intellectual disability

Sivapriya S[1*], Punitha P[1], Ahamed Ashwaq HA[2]

[1]*Department of Occupational Therapy, Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Saveetha Nagar Thandalam, Chennai, Tamil Nadu, India.*
[2]*Occupational Therapist, Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Saveetha Nagar Thandalam, Chennai, Tamil Nadu, India.*

ARTICLE INFO

Article history:
Received 9 February 2025
Accepted as revised 29 July 2025
Available online 13 September 2025

Keywords:
Intellectual disability, balance, exergaming, occupational therapy, pediatric balance scale.

ABSTRACT

Background: Intellectual disability (ID) is associated with balance and motor coordination challenges, leading to increased fall risk and limitations in daily activities. Traditional occupational therapy interventions have shown efficacy in addressing these issues, but there is a growing interest in integrating modern technologies, such as exergaming, into therapeutic programs. Exergaming combines physical activity with interactive video games, potentially enhancing motivation and engagement among children with ID. This study focuses on assessing the impact of exergaming on improving balance in children with ID, an area that remains underexplored in the Indian context.

Objectives: This study aimed to determine the effect of an exergaming intervention on improving balance in clients with intellectual disabilities.

Materials and methods: The study involved 30 children aged 6-12 years diagnosed with mild to moderate ID. Participants were randomly divided into two groups: the experimental group (N=15) received exergaming-based balance training, and the control group (N=15) underwent conventional occupational therapy. Both groups participated in 36 sessions over three months, each lasting 45 minutes. Balance was assessed using the Pediatric Balance Scale (PBS) before and after the intervention. The PBS includes 14 tasks that evaluate balance through various daily activities, with scores ranging from 0 (unable to perform) to 4 (independent).

Results: The results demonstrated a significant improvement in balance for both groups. However, the experimental group, which engaged in exergaming, showed a more substantial improvement in PBS scores, increasing from a mean of 29.07 to 41.93. In contrast, the control group's mean PBS score increased from 28.13 to 33.73. Statistical analysis confirmed that the improvement in the experimental group was significantly more significant (p<0.001), indicating the added benefit of exergaming in enhancing balance among children with ID.

Conclusion: The results indicate that exergaming is an effective intervention for improving balance in children with intellectual disabilities, as it offers engaging, challenging, and play-based therapy. Enhancing balance and postural control may also lead to broader physical and cognitive benefits, such as improved cardiovascular health, muscular strength, and cognitive processing. These improvements can further support functional independence, promoting inclusivity and better community engagement. Ultimately, this study highlights the potential of exergaming to empower individuals with intellectual disabilities, fostering their overall well-being and quality of life.

***Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Saveetha Nagar Thandalam, Chennai, Tamil Nadu, India.*
E-mail address: sivapriyas.scot@saveetha.com
doi: 10.12982/JAMS.2026.008
E-ISSN: 2539-6056

Introduction

Occupational therapists provide support to individuals with intellectual disabilities (ID) throughout

their lives and in a variety of practice settings. However, more needs to be done by the profession to address the specific needs of this marginalized population. ID is one kind of developmental disability that frequently co-occurs with other developmental disorders like autism spectrum disorder (ASD) or cerebral palsy.[1]

The World Health Organization reports that approximately 2% of Indians of all ages have an intellectual handicap based on surveys conducted among the country's general population. However, if the issue is limited to children (those under the age of 18), then roughly 3% of all the children in the same community who are under the age of 18 will have an intellectual handicap. Based on a comprehensive analysis of individual prevalence studies conducted over the past six decades, the established prevalence of ID in India is 2%.[2]

A strategic element of coordination capacities called balance ability depends on several variables. Individuals with ID frequently experience falls that result in injuries, which can be attributed to their poor balancing abilities, which could be a sign of incomplete development. The most noticeable impairments in balance and motor skills are seen in inactive adults with ID, who may be more vulnerable to a loss of basic functioning and have less autonomy in daily life activities.[3]

Reduced postural stability has been observed in individuals with ID.[4-8] This has been linked to increased body sway and/or higher sway velocity during quiet stance,[9] an increased fall risk,[10,11] abnormal gait patterns,[12,6,13] delayed responses to balance perturbations,[14] and using fewer effective strategies for maintaining balance. Postural instability has also been linked to a disrupted central integration of multisensory information in this population.[15]

Exergaming is the practice of mixing video game technology with physical activity to promote an active and healthy lifestyle.[16] Exergames, often known as active video games, have user interfaces that call for active participation and physical effort from players. Players can have fun while exercising with these fitness games, which are made to track body motion. The effectiveness of these games for rehabilitation or exercise has been the subject of numerous studies.[17]

A waiting-list control group was contrasted with the exergaming intervention. Microsoft's Xbox Kinect implemented the exergaming intervention (Microsoft, Redmond, WA). This game console has an input device that detects motion. Users move their bodies to control and interact with the console.[23] Active video games that double as exercise are referred to as exergaming. Exergaming has the potential to support preschool children's PA, even if it is screen-based. Exergaming has been employed increasingly recently in school-based settings as a creative and enjoyable way to encourage an active lifestyle with encouraging and hopeful outcomes.[18-20]

Materials and methods

The study was ethically approved by the Institutional Review Board of Saveetha College of Occupational Therapy (SCOT/ISRB/061/2023) and was conducted at Aadhuraa Special School, Kanchipuram, India. A quasi-experimental quantitative study design was employed, involving 30 children with mild to moderate intellectual disabilities aged 6 to 12 years. The participants were divided equally into experimental and control groups, with 15 children in each group. The experimental group received balance training using exergaming intervention and conventional occupational therapy consisting of 45-minute sessions held three times per week over approximately 3 months, totaling around 36 sessions. The control group received only conventional occupational therapy for the same duration. The Pediatric Balance Scale was used to measure balance levels in both groups through pre-test and post-test assessments.

Participants

Thirty children aged 6 to 12 years, 10 females and 20 males diagnosed with mild to moderate intellectual disability, participated in this study to enhance their balance skills. Initially, all participants underwent a pre-assessment of their balance using the Pediatric Balance Scale. Following the pre-assessment conducted during the first two sessions of the study, the children were randomly allocated into an experimental group (N=15) and a control group (N=15). The inclusion criteria for this study were 1) children diagnosed with mild (Intelligence Quotient 55-70) and moderate ID (Intelligence Quotient 40-55), 2) children aged from 6 to 12 years, both males and females included, and 3) children with poor balance scores above 20. The exclusion criteria for this study were 1) children with poor comprehension skills, 2) children with poor visual foundation skills, and 3) children using wheelchairs and walking aids.

Instruments

The Pediatric Balance Scale (PBS) is a refined version of the Leaflet Scale designed to assess functional balance in school children with mild to moderate disabilities.21 This assessment consists of 14 tasks that repeat daily activities, and each task is evaluated on a five-point scale (0 to 4). A score of 0 indicates no ability to act independently, while a score of 4 indicates high independence. The qualification criteria assess factors such as the duration of position maintenance, the distance to the upper extremities, and the time required to complete the task with a maximum score of 56.[22] A reliability analysis conducted with a group of twenty 5- to 15-year-old children with varying degrees of movement disorders showed high interrater reliability (ICC=0.997) and reliability of test-retest (ICC=0.998).[21] A valid test including 30 children with cerebral palsy, 4 to 10 years old and classified

GMFCS levels I-III, there is a strong relationship between PBS and movement (r=0.82, p<0.001) and self-care (r=0.73) showed, p<0.001) dimensions of the Disability Evaluation Questionnaire (PDEI).[22] In addition, a descriptive study with 23 children aged 6 to 15 years with cerebral palsy (hemiplegia or diplegia) showed a significant correlation between PBS and the Selective Lower Extremity Control Evaluation.

Procedure

The experimental group participated in a balance training program that combined exergaming with conventional occupational therapy. In contrast, the control group received standard occupational therapy sessions. Both groups underwent 36 sessions, each lasting 45 minutes, conducted three times per week over 12 weeks. To assess the efficacy of the exergaming intervention in enhancing balance among children with intellectual disabilities, the Pediatric Balance Scale was administered post-intervention to both groups.

The exergaming intervention utilized an interactive system designed to improve balance using dynamic light and sound cues. The system consisted of three towers, each approximately 6 feet tall and arranged in a triangular formation, with 10 lights per tower. Lights were illuminated in random sequences, accompanied by corresponding tones, prompting participants to touch the lit targets. The program advanced through three graded levels:

Grade 1: Participants touched the lights while walking.
Grade 2: Participants touched the lights while hopping.
Grade 3: Participants performed the task while standing on one leg.

Participants earned one point for each correct light touched. Upon achieving 8 or more points, they progressed to the next grade.

The experimental group's exergaming intervention not only targeted motor control but also engaged cognitive domains such as executive function (e.g., planning, attention, and working memory), visual-motor coordination, and decision-making. For example, during the exergaming tasks, children had to remember target sequences, shift attention between lights, and respond to visual and auditory cues stimulating their cognitive flexibility and processing speed. In contrast, the control group's task-oriented interventions emphasized structured physical activities such as reaching tasks and balance exercises without the same level of cognitive engagement.

The control group received standard occupational therapy in a special school setting, which focused on balance training through conventional methods. This included recreational activities aimed at improving standing and sitting balance and performing activities of daily living (ADLs). Interventions included task-oriented training, perceptual exercises, task-specific reaching, visual feedback, balance biofeedback, and multisensory training. Table 1 listed below the intervention procedure for experimental group and control group.

Table 1. Intervention procedure for experimental group and control group.

Step	Component	Experimental group	Control group
1	Type of intervention	Exergaming + conventional occupational therapy	Conventional occupational therapy only
2	Session duration	45 minutes	45 minutes
3	Frequency	3 sessions per week	3 sessions per week
4	Total sessions	36 sessions over 12 weeks	36 sessions over 12 weeks
5	Exergaming setup	Interactive system with 3 towers (6 feet tall) arranged in a triangle, each with 10 lights; lights flash randomly with sound cues	Not applicable
6	Exergaming tasks	Grade 1: Walking while touching lit targets Grade 2: Hopping to targets Grade 3: Touching targets while standing on one leg	Not applicable
7	Progression criteria	Earn ≥8 points to progress to next grade level	Not applicable
8	Balance training activities	Integrated balance training with exergaming tasks like dynamic and static balance training by using hip and ankle strategies and leaning and reaching forward, lateral and backward.	Task-oriented interventions, perceptual exercises, visual feedback training, balance biofeedback, and multisensory training

Table 1. Intervention procedure for experimental group and control group (continued).

Step	Component	Experimental group	Control group
9	Exergaming with traditional occupational therapy	Exergaming was blended with traditional occupational therapy by incorporating balance boards, foam surfaces, and interactive light cues to enhance gross motor, balance, sensory, and cognitive skills in a game-like setup that boosts motivation and participation.	Children in the control group received conventional occupational therapy focusing on balance and coordination using activities like standing on foam pads, stepping over obstacles, and catching balls, guided directly by the therapist without the use of digital or exergaming tools.
10	Functional focus	Balance improvement via interactive gaming and traditional methods	Balance training integrated with activities of daily living (ADLs)
11	Outcome measure	Pediatric Balance Scale (PBS), pre- and post-intervention	Pediatric Balance Scale (PBS), pre- and post-intervention

Figure 1 illustrates the customized exergaming tower designed for children with intellectual disability. The setup includes vertically aligned light units with motion sensors. These lights are programmed to illuminate in specific sequences to encourage targeted motor responses such as reaching, stepping, or shifting weight. The tower's height and light placement were adjusted based on each child's functional ability as shown in Figure 1A. The child actively engaging with the exergaming tower is demonstrated in Figure 1B. The task requires the child to maintain balance on a foam surface or balance board while responding to randomly activated lights by touching or stepping toward them. This dynamic task integrates both balance control and cognitive processing such as attention, visual tracking, and response inhibition.

Figure 1. *Exergaming intervention setup. A: exergaming tower setup, B: child performing balance task using exergaming.*

Statistical analysis

This study used inferential statistical methods to evaluate the effect of a physical activity intervention on the balance of children with intellectual disabilities. Descriptive statistics, including mean and standard deviation, were used to summarize data and evaluate the distribution of participants. Due to the small sample size (30 people) and the nature of the data, non-parametric statistical methods were used. Wilcoxon signed rank scores were used to evaluate differences between pre-and post-test test scores in the experimental and control groups. The analysis showed a significant improvement in balance scores for the control group (Z=-3.425, p=0.001). Accordingly, the experimental group showed substantially increased balance scores (Z=3.425, p=0.001). The Mann-Whitney U test was used to compare the effectiveness of the interventions between the control and experimental groups. The results after the test showed a significant difference between the two groups (Z=-4.646, p<0.001), which shows that the intervention of physical activity has a significant effect in improving the remainder were treated with conventional occupational therapy. A significance level of p<0.05 was used for all statistical analyses.

Results

The study demonstrated that the exergaming intervention led to a significant improvement in balance among children with intellectual disabilities when compared to conventional occupational therapy. The experimental group, which engaged in exergaming,

showed a substantial increase in their Pediatric Balance Scale (PBS) scores from a pre-test mean of 29.07 to a post-test mean of 41.93. Conversely, the control group, which received standard occupational therapy, showed a more modest improvement, with scores increasing from 28.13 to 33.73. Statistical analysis confirmed that the improvement in the experimental group was significantly more significant than that in the control group.

Table 2 shows that the *p* value of 0.001 is less than 0.05, so the alternate hypothesis is accepted. Hence, there is a statistically significant difference between the pre-test and post-test scores in the PBS control group. This suggests that the intervention received by the control group had a significant improvement.

Table 3 shows the *p* value of the experimental group is 0.001 is less than 0.05, alternate hypothesis is accepted. Hence, there is a highly statistically significant difference in the experimental group between pre-test and post-test scores of PBS. This suggests that the intervention received by the experimental group had significant improvement.

Table 4 shows a *p* value of 0.00 is less than 0.05; an alternate hypothesis is accepted. Hence, there is a highly statistically significant difference in post-test scores between the experimental and control groups of the PBS. This suggests that the intervention received by the experimental group showed more improvement than the control group.

Table 2. Statistical analysis of pre-test and post-test in control group.

Test	Mean	SD	N	Z value	*p* value
Pre-test	28.1333	2.89992	15	-3.425	0.001*
Post-test	33.7333	2.49189	15		

***Note**: *significant at 5% alpha level.*

Table 3. Statistical analysis of pre-test and post-test in the experimental group.

Test	Mean	SD	N	Z value	*p* value
Pre-test	29.0667	2.25093	15	-3.425	0.001*
Post-test	41.9333	1.43759	15		

***Note**: *significant at 5% alpha level.*

Table 4. Statistical analysis between the post-test scores of the control and experimental group.

Group	Mean	SD	N	Z value	*p* value
Control group	33.7333	2.49189	15	-4.645	0.00*
Experimental group	41.9333	1.43759	15		

***Note:** * Significant at 5% alpha level*

Discussion

This study aimed to evaluate the effectiveness of exergaming interventions in enhancing balance among children with intellectual disabilities. Specifically, the study aimed to compare the effects of exergaming with conventional occupational therapy on balance improvement in a sample of 30 children, who were evenly divided into control and experimental groups. This investigation supports prior research by Hilton *et al*. which emphasized the necessity of a control group in determining the efficacy of balance interventions.[17]

The current study analyzed pre-test and post-test scores using the Pediatric Balance Scale (PBS) to assess improvements in balance and postural control abilities among the participants. The analysis revealed a statistically significant improvement in the control group's balance, as indicated by the pre-test and post-test scores (p=0.001). This improvement can be attributed to the conventional occupational therapy provided to the control group, consistent with findings from Lourenco *et al*. which highlighted the positive impact of organized and systematic interventions on motor proficiency, including balance, in children with autism spectrum disorder (ASD).[24]

The results in the experimental group demonstrated a statistically significant difference between pre-test and post-test scores (p=0.001), suggesting that the exergaming intervention was particularly effective in enhancing balance and motor skills. This result has been supported by the previous articles by Nikolaos *et al*. which showed that exergaming interventions effectively stimulate a wide range of fundamental motor skills.[25] Hanley *et al*. conducted a study on autism and exergaming: impacts on cognitive processes and repetitive actions in that study, two pilot studies examined the potential behavioral and cognitive advantages of exergaming.[26] A dozen kids with Pilot I involved teens with autism spectrum disorders and involved a control task and an acute bout of Dance Dance Revolution (DDR);

Pilot II involved ten more teenagers and an acute bout of cybercycling. Executive function and repeated behaviors were evaluated both before and after each activity. After the exercise, compared to the control condition, repeated behaviors decreased, and Digits Backwards performance improved. Although Hanley et al. studied children with autism spectrum disorders, the findings on executive function and reduced repetitive behaviors provide indirect support for the cognitive benefits of interactive exergaming.[26] While our primary outcome was balance improvement, these cognitive domains are often impaired in children with intellectual disabilities as well and stimulating them through exergaming may contribute to improved postural control. Schmidt *et al*. determined a study on effect of exergaming on executive function in attention deficit hyperactivity disorder.[27] The study was a parallel group randomized trial, 51 children between 8-12 years diagnosed with ADHD were assigned either to an 8-week exergaming intervention group or a waiting-list control group. The core executive functions such as inhibition, switching, updating, parent ratings of symptoms, and motor abilities were assessed before and after the intervention. Finding of the study was exergaming intervention group improved in specific executive functions, general psychopathology as well as motor abilities compared to control group. Claudia list Hilton *et al*. conducted a pilot study investigated the impact of Makoto arena training on children with autism spectrum disorder (ASD).[17] Results showed improved response speed, Executive Function (EF), and motor skills. Strong correlations were found between EF and motor scores. Participants demonstrated increased reaction speed (effect size = 1.18). Significant enhancements were observed in working memory, metacognition, strength, and agility. Exergaming, particularly the Makoto arena, could complement standard interventions for ASD children with motor and EF deficits.

Moreover, the comparison between the post-test scores of the control and experimental groups (p=0.00) revealed a significantly greater improvement in the experimental group. This underscores the superior efficacy of exergaming interventions over conventional therapy in improving balance among children with intellectual disabilities. These findings are consistent with Hilton *et al*.,[17] which demonstrated the positive effects of exergaming on various outcomes, including motor skills and executive function, in children with developmental disorders.

Limitations

The study is subject to limitations due to its small sample size and the limited number of sessions, which may constrain the generalizability of the findings. Additionally, the reliance on a convenience sampling method introduces potential selection bias and affects the sample's representativeness. These limitations should be carefully considered when assessing the applicability of the results to broader populations.

The study did not statistically compare pre-test PBS scores between the two groups to confirm baseline equivalence. While mean values were similar, a formal test (e. g., Mann-Whitney U test) could strengthen the assumption of comparability.

Although random assignment was used, the distribution of participants based on the severity of intellectual disability (mild vs moderate), age, and gender was not stratified. This could have influenced the intervention outcomes.

Pre- and post-assessments were conducted by the research team, and blinding was not implemented. This could introduce observer bias and affect the objectivity of the outcome evaluation.

Conclusion

The findings of this study underscore the efficacy of exergaming as an innovative therapeutic intervention for improving balance in children with intellectual disabilities. Exergaming was more effective than conventional occupational therapy alone, suggesting that it could be a valuable addition to therapeutic programs aimed at enhancing motor skills and functional independence in this population. This approach improves physical outcomes and provides a more engaging and motivating form of therapy. The results advocate for the broader application of exergaming in occupational therapy and suggest that further research should explore its long-term benefits and potential in various pediatric populations.

Ethical approval

This study was approved by the Institutional Review Board of Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences (Approval No: SCOT/ISRB/061/2023).

Funding statement

None

Conflict of interest

The authors have no potential conflicts of interest to disclose.

CRedit authorship contribution statement

Sivapriya S: conceptualization, methodology, writing; original draft, statistical analysis; **Punitha P:** supervision, validation, review and editing; **Ahamed Ashwaq HA:** data collection, study execution at the school, manuscript editing.

Acknowledgements

The authors would like to thank Aadhuraa Special School for providing the study setting, and all participants and their families for their valuable participation.

References

[1] Johnson KR, Blaskowitz M, Mahoney WJ. Occupational therapy practice with adults with intellectual disability: What more can we do? Open J Occup Ther. 2019; 7: 1-15. doi: 10.15453/2168-6408.1573.

[2] Russell PSS, Nagaraj S, Vengadavaradan A, Russell S, Mammen PM, Shankar SR, et al. Prevalence of intellectual disability in India: A meta-analysis. World J Clin Pediatr. 2022; 11: 206-14. doi: 10.5409/wjcp.v11.i2.206.

[3] Dehghani M, Gunay M. The effect of balance training on static and dynamic balance in children with intellectual disability. J Appl Environ Biol Sci. 2015; 5: 127-31. doi: 10.5958/2249-7315.2015.00108.2

[4] Blomqvist S, Olsson J, Wallin L, Wester A, Rehn B. Adolescents with intellectual disability have reduced postural balance and muscle performance in trunk and lower limbs compared to peers without intellectual disability. Res Dev Disabil. 2013; 34: 198-206. doi: 10.1016/j.ridd.2012.07.008.

[5] Rine RM, Dannenbaum E, Szabo J. 2015 section on pediatrics knowledge translation lecture: Pediatric vestibular-related impairments. Pediatr Phys Ther. 2016; 28: 2-6. doi: 10.1097/PEP.0000000000000226.

[6] Boonyong S, Siu K-C, van Donkelaar P, Chou L-S, Woollacott MH. Development of postural control during gait in typically developing children: the effects of dual-task conditions. Gait Posture. 2012; 35: 428-34. doi: 10.1016/j.gaitpost.2011.11.002.

[7] Karachi C, Grabli D, Bernard FA, Tandé D, Wattiez N, Belaid H, et al. Cholinergic mesencephalic neurons are involved in gait and postural disorders in Parkinson disease. J Clin Invest. 2010; 120: 2745-54. doi: 10. 1172/JCI42642.

[8] Zur O, Ronen A, Melzer I, Carmeli E. Vestibulo-ocular response and balance control in children and young adults with mild-to-moderate intellectual and developmental disability: a pilot study. Res Dev Disabil. 2013; 34: 1951-7. doi: 10.1016/j.ridd.2013.03.007.

[9] Dellavia C, Pallavera A, Orlando F, Sforza C. Postural stability of athletes in special olympics. Percept Mot Skills. 2009; 108: 608-22. doi: 10.2466/PMS.108.2.608-622.

[10] Chiba Y, Shimada A, Yoshida F, Keino H, Hasegawa M, Ikari H, et al. Risk of fall for individuals with intellectual disability. Am J Intellect Dev Disabil. 2009; 114: 225-36. doi: 10.1352/1944-7558-114.4:225-236.

[11] Wajda DA, Motl RW, Sosnoff JJ. Dual task cost of walking is related to fall risk in persons with multiple sclerosis. J Neurol Sci. 2013; 335: 160-3. doi: 10.1016/ j.jns.2013.09.021.

[12] Almuhtaseb S, Oppewal A, Hilgenkamp TIM. Gait characteristics in individuals with intellectual disabilities: a literature review. Res Dev Disabil. 2014; 35: 2858-83. doi: 10.1016/j.ridd.2014.07.017.

[13] Joseph C, Leavy B, Franzén E. Predictors of improved balance performance in persons with Parkinson's disease following a training intervention: analysis of data from an effectiveness-implementation trial. Clin Rehabil. 2020; 34: 837-44. doi: 10.1177/0269215520917199.

[14] Oppewal A, Hilgenkamp TIM. The association between gait and physical fitness in adults with intellectual disabilities: Gait and physical fitness in adults with intellectual disabilities. J Intellect Disabil Res. 2018; 62: 454-66. doi: 10.1111/jir.12484.

[15] Blomqvist S, Wester A, Rehn B. Postural muscle responses and adaptations to backward platform perturbations in young people with and without intellectual disability. Gait Posture. 2014; 39: 904-8. doi: 10.1016/j.gaitpost. 2013.11.018.

[16] Bogost I. The rhetoric of exergaming. In: DAC '05: Proceedings of the Digital Arts and Cultures Conference; 2005 Dec 1–3; Copenhagen, Denmark. p.51. https://share.google/PggimpB9BECJOQTlL

[17] Hilton CL, Cumpata K, Klohr C, Gaetke S, Artner A, Johnson H, et al. Effects of exergaming on executive function and motor skills in children with autism spectrum disorder: a pilot study. Am J Occup Ther. 2014; 68: 57-65. doi: 10.5014/ajot.2014.008664.

[18] Gao Z, Pope Z, Lee JE, Stodden D, Roncesvalles N, Pasco D, et al. Impact of exergaming on young children's school day energy expenditure and moderate-to-vigorous physical activity levels. J Sport Health Sci. 2017; 6: 11-6. doi: 10.1016/j.jshs.2016.11.008.

[19] Gao Z, Hannon JC, Newton M, Huang C. Effects of curricular activity on students' situational motivation and physical activity levels. Res Q Exerc Sport. 2011; 82: 536-44. doi: 10.1080/02701367.2011.10599786.

[20] Gao Z, Chen S, Pasco D, Pope Z. Effects of active video games on physiological and psychological outcomes among children and adolescents: A meta-analysis. Obes Rev. 2015; 16: 783-94.doi: 10.1111/obr.12287

[21] Franjoine MR, Gunther JS, Taylor MJ. Pediatric balance scale: a modified version of the berg balance scale for the school-age child with mild to moderate motor impairment. Pediatr Phys Ther. 2003; 15: 114-28. doi: 10.1097/01.PEP.0000068117.48023.18.

[22] Ac Duarte N, Grecco LA, Franco RC, Zanon N, Oliveira CS. Correlation between Pediatric Balance Scale and functional test in children with cerebral palsy. J Phys Ther Sci. 2014; 26: 849-53. doi: 10.1589/jpts.26.849

[23] Benzing V, Schmidt M. Exergaming for children and adolescents: Strengths, weaknesses, opportunities and threats. J Clin Med. 2018; 7: 422. doi: 10.3390/jcm7110422.

[24] Lourenço C, Esteves D, Corredeira R, Seabra A. The effect of a trampoline-based training program on the muscle strength of the inferior limbs and motor proficiency in children with autism spectrum

disorders. J Phys Educ Sport. 2015; 15(3): 479-86 doi: 10.7752/jpes.2015.03073.

[25] Vernadakis N, Papastergiou M, Zetou E, Antoniou P. The impact of an exergame-based intervention on children's fundamental motor skills. Comput Educ. 2015; 83: 90-102. doi: 10.1016/j.compedu. 2015.01. 001.

[26] Ferguson MB, Anderson-Hanley PC, Mazurek MO, Parsons S, Warren Z. Game interventions for autism spectrum disorder. Games Health. 2012; 1: 248-53. doi: 10.1089/g4h.2012.0717.

[27] Benzing V, Schmidt M. The effect of exergaming on executive functions in children with ADHD: A randomized clinical trial. Scand J Med Sci Sports. 2019; 29: 1243-53. doi: 10.1111/sms.13446.

Journal of Associated Medical Sciences 2026; 59(1): 62-72

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

Gamma ray interaction of germanoborate glasses for radiation shielding applications

Natthakridta Chanthima[1,2], Supakit Yonphan[1,2], Chalermpon Mutuwong[1,2], Wuttichai Chaiphaksa[1,2*], Jakrapong Kaewkhao[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 June 2025
Accepted as revised 29 August 2025
Available online 16 September 2025

Keywords:
Glass, gadolinium, radiation shielding, WinXCom.

ABSTRACT

Background: Ionizing radiation is essential in medical imaging and therapy, but it also poses health risks to medical staff and patients. Traditional shielding materials like lead are effective but toxic, while concrete lacks transparency. Therefore, glass systems incorporating high-Z oxides offer a promising alternative by combining optical clarity with enhanced radiation shielding performance.

Objectives: This study aims to develop and investigate germanoborate glasses doped with Gd_2O_3, with a particular focus on their photon attenuation properties. The findings are intended to development of transparent, lead-free radiation shielding materials suitable for medical and industrial applications.

Materials and methods: The Gd_2O_3 added germanoborate glasses in the system (60-x)B_2O_3-GeO_2-x Gd_2O_3 (with x=10, 20, 30 and 40 mol%) were synthesized via melting at 1400 °C followed by melt quenching method. The densities were measured by Archimedes' method. The mass attenuation coefficients (μ_m), the effective atomic number (Z_{eff}), the effective electron density (N_{eff}), and half-value layers (HVL) were computed using the WinXCom program (NIST XCOM Database) to assess shielding properties. PHITS Monte Carlo simulations were employed to calculate the effective dose rate.

Results: The addition of Gd_2O_3 increased the glass density from 3.7546 gm/cm^3 to 5.4604 gm/cm^3. Replacing B_2O_3 with Gd_2O_3 enhanced the mass attenuation coefficients (μ_m), effective atomic number (Zeff) and effective electron density (Neff). HVL values decreased, Pb equivalent values increased, and the effective dose was reduced.

Conclusion: The incorporation of Gd_2O_3 into germanoborate glasses significantly enhances their gamma-ray attenuation capabilities, confirming their effectiveness as lead-free shielding materials. The developed glasses exhibit optimized radiation protection properties, highlighting their potential for transparent shielding applications in medical, industrial, and nuclear fields. and infliximab exert protective effects on liver function in acetic acid induced UC model. These findings suggest potential benefits of these therapies in mitigating liver damage associated with UC, highlighting the importance of considering liver health in UC management.

*** 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: 10.12982/JAMS.2026.009
E-ISSN: 2539-6056

Introduction

Ionizing radiation is essential in modern diagnostic imaging and radiation therapy. However, its use involves potential risks of radiation dose received by both

healthcare professionals and patients. Accordingly, the development and implementation of effective radiation protection methods are essential to minimize exposure and ensure safety in environments where radiation is routinely used.[1] The effectiveness of a shielding material is largely determined by its density and atomic composition. Materials with high atomic number (Z) and high density present more electrons per unit volume, leading to a higher probability of photon interactions. In the energy range used for medical imaging and therapy, the dominant photon interactions are the photoelectric effect and Compton scattering High-Z, high-density materials strongly enhance photoelectric absorption, while Compton scattering depends mainly on total electron density.[2] Lead (Pb) is a commonly used shielding material due to its high density, low cost, and ease of fabrication. Nevertheless, the toxic nature of lead poses serious environmental and health concerns, prompting the need for alternative, ecofriendly shielding materials.[3] As a result, lead free shielding materials have gained considerable attention. Concrete is another traditional bulk shielding material. It is inexpensive, easily cast into walls or blocks, and has good mechanical strength. Ordinary concrete is effective for gamma shielding partly because it can incorporate heavy aggregates to raise its density. However, its opacity limits its use in certain medical applications that require visibility.

Glass materials have emerged as a promising alternative due to their transparency, chemical versatility, and ability to be tailored for specific radiation shielding applications. By incorporating heavy metal oxides into the glass matrix, their density and photon attenuation capabilities can be significantly enhanced. Thus, by incorporating heavy metal oxides into transparent glasses, it is possible to create shields that approach the performance of lead while still allowing light to pass through.[4] Borate glasses are valued as host materials because they form stable, dense networks with desirable thermal and mechanical performance. They exhibit high thermal stability and excellent glass forming ability. In practice, borate glasses can incorporate modifiers to further increase density and hardness without crystallization. These features make borate matrices resistant to cracking and thermal shock, and well suited for optical transparency.[5]

Adding GeO_2 is known to improve network connectivity and optical transmission in borate glasses. GeO_2 is a relatively dense glass former (ρ = 4.25 g/cm^3) characterized by a high refractive index and a broad infrared transmission window. GeO_2 doped glasses generally exhibit lower phonon energies (800-975 cm^1) compared to borate-based glasses. As a result, the incorporation of GeO_2 into a glass matrix effectively reduces the overall host phonon energy. As a result, germanoborate glasses can exhibit higher refractive index and broader optical transparency, as well as increased glass transition temperature, all of which strengthen the network.[4] For radiation shielding, Gadolinium oxide is an especially attractive dopant for gamma shielding. Gd has a high atomic number (Z=64) and density (7.895 g/cm^3), which contribute to enhanced photoelectric absorption and Compton scattering interactions with gamma photons. Importantly, Gd_2O_3 does not induce radiation related color centers in glass. In fact, several studies have reported that Gd_2O_3 acts as a heavy metal modifier, increasing the glass density and enhancing gamma-ray attenuation while maintaining optical transparency.[6] Moreover, PHITS Monte Carlo simulation was employed to simulate radiation transport and assess absorbed dose distributions in the context of modern diagnostic imaging and radiation therapy.

In this work, the Gd_2O_3 added germanoborate glasses were fabricated using a melt quenching technique. The obtained samples were analyzed for their density, structure, and radiation properties. The photon attenuation parameters were determined using WinXCom software, which extracts data from the NIST database. WinXCom is an accepted reference for X-ray and gamma ray interaction coefficients and has been widely used in glass shielding studies.[7]

Materials and methods

Theoretical background

The gamma ray attenuation in materials follows the exponential Beer-Lambert law:

$$I = I_0 e^{-\mu x}, \quad (1)$$

where I and I_0 are the transmitted and incident intensities, μ is the linear attenuation coefficient (cm^{-1}), x and is the thickness (cm).The mass attenuation coefficient (μ_m) is calculated by normalizing μ to the material's density, and provides a density independent measure of attenuation:

$$\mu_m = \sum_i w_i \left(\frac{\mu}{\rho}\right)_i, \quad (2)$$

When w_i is the fractional weight, and μ_m is the mass attenuation coefficient of each element. Theoretical calculations can be performed using the WinXCom software, which provides tabulated photoncross sections based on composition and energy.[8]

The total atomic cross section, (σ_a), can be evaluated from the values of the total mass attenuation coefficients by the following

$$\sigma_a = \frac{\mu}{N_A \sum_i \frac{w_i}{A_i}}, \quad (4)$$

Similarly, the total electronic cross section, (σ_e), is given by the following formula

$$\sigma_e = \frac{1}{N_A}\sum_i \left(\sum_i \frac{f_i A_i}{Z_i} \right) w_i \quad ,(5)$$

where f_i and Z_i are respectively the fractional abundance and the atomic number of each element. For compound materials, the effective atomic number (Z_{eff}), is defined as the ratio between the total atomic effective cross section and the total electronic effective cross section.

$$Z_{eff} = \frac{\sigma_a}{\sigma_{el}} \quad ,(6)$$

The effective electron density (N_{eff}) defines the electrons number per unit mass, which is related to the Z_{eff}, is formulated as follows.[9]

$$N_{eff} = N_A \frac{Z_{eff}}{\sum_i \frac{w_i}{A_i}} \quad ,(7)$$

The half value layer (HVL), defined as the thickness of material required to reduce the intensity by half, it is given by:

$$HVL = \frac{0.693}{\mu} \quad ,(8)$$

where μ is the linear attenuation coefficient. HVL is a practical parameter for assessing and comparing the shielding effectiveness of different materials.[10]

Lead equivalent thickness refers to the thickness of lead that would provide an identical shielding performance against radiation as the material under investigation, under the same specified conditions. The degree of shielding can be evaluated through the material's transmission factor. For photons of a particular energy passing through a certain thickness of material or lead, the relationship between the linear attenuation coefficient (LAC) and thickness allows for the estimation of the equivalent lead thickness.[11] This can be expressed as:

$$d_{pb} = \frac{\mu_{material}}{\mu_{pb}} \times d_{material} \quad ,(9)$$

where μ_{pb} denotes the LAC of lead, d_{pb} represents the calculated lead equivalent thickness, $\mu_{material}$ is the LAC of the material, and $d_{material}$ is the thickness of that material.

Sample preparations

The Gd_2O_3 added germanoborate glasses to radiation shielding were prepared by melt quenching technique in compositions of (60-x)B_2O_3-GeO_2-xGd_2O_3 (with x=10, 20, 30 and 40 mol%). Each batch was melted in alumina crucibles by an electrical furnace for about 3 hrs, at 1400 °C and poured into the graphite mold at room temperature and annealed at 500 °C for 3 hrs. All glass samples were cut and polished into a size about 1 cm × 1.5 cm x 0.3 cm for more investigations.

WinXCom program

WinXCom is a software implementation of the NIST XCOM photon cross-section database. It provides theoretical mass attenuation coefficients and partial interaction cross sections (photoelectric absorption, Compton scattering, pair production) for any element, compound or mixture over a wide photon energy range (approximately 1 keV to 100 GeV). Users simply specify the material's composition (by chemical formula or elemental mix) and the energy range of interest, and WinXCom computes the attenuation data. The output includes not only the mass attenuation coefficients but also allows derivation of composite parameters such as the effective atomic number and effective electron density, which are standard metrics for quantifying gamma-ray shielding effectiveness. The main advantages of WinXCom are its speed and versatility: it can rapidly generate attenuation data for a vast range of materials and energies, greatly aiding the design and optimization of novel radiation shields. For these reasons it has become an essential tool in shielding research.[12]

PHITS Monte Carlo simulation

To investigate the absorbed radiation dose, the adult male mesh-type reference computational phantom (MRCP-AM) was implemented in PHITS. The MRCP-AM dataset comprises MRCP-AM.node (nodal data), MRCP-AM.ele (element connectivity), MRCP-AM.cell (geometric configuration, spatial location, and density information), and MRCP-AM.material (material composition), as illustrated in Figure 1. The primary objective was to evaluate the absorbed and effective dose rates within the MRCP-AM phantom, both prior to and following the inclusion of shielding glass, using Monte Carlo simulation techniques. The phantom model was imported into PHITS via the [Surface], [Cell], and [Materials] sections to accurately define the anatomical structure of the entire human body.[13] The irradiation source and anterior-posterior (AP) geometry was defined within the [Source] section, with the source placed at coordinates (0, 0, -60). The shielding glass was defined using the [Surface], [Cell], and [Materials] sections, with dimensions of 60.0 cm × 1.0 cm × 200.0 cm and located at (0, 0, -40). Organ-specific absorbed dose distributions were calculated using the [T-Deposit] section with a "reg" mesh, while effective dose rates, based on ICRP Publication 103 recommendations for AP irradiation geometry,[14] were determined using the [T-Track] section with an "xyz" mesh and the "-202" weighting factor in the [Multiplier] section.[15] A total of 100 million primary photons were simulated to ensure sufficient statistical accuracy.

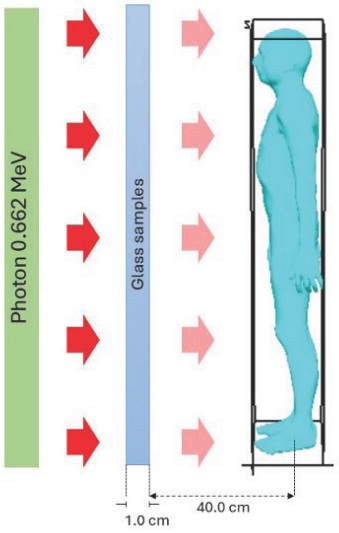


Figure 1. *The MRCP-AM models from PHITS.*[16]

Result and discussions

Physical properties

The physical properties of the Gd_2O_3 added germanoborate glass system with the composition (60-x) B_2O_3-GeO_2-xGd_2O_3 (where x=10, 20, 30, and 40 mol%) were systematically investigated were shown in Figure 2. It was observed that the glass density increased from 3.7546 gm/cm^3 (Gd10) to 5.4604 g/cm^3 (Gd40) as the Gd_2O_3 content increased. This trend is primarily due to the replacement of B_2O_3, a low-density component, with Gd_2O_3 which contains Gd atoms of significantly higher atomic mass (157.25 gm/mol) compared to B (10.81 gm/mol), thereby increasing the overall density of the glass network. Meanwhile, the molar volume increased from 30.0711 cm^3/mol (Gd10) to 37.3287 cm^3/mol (Gd30), reflecting the network expansion due to the formation of non-bridging oxygen (NBO) sites induced by Gd_2O_3. However, a slight decrease in molar volume was observed at 36.7679 cm^3/mol for Gd40, possibly indicating network compaction. This behavior suggests that at high concentrations, Gd^{3+} ions acting as network modifiers with a relatively large ionic radius may disrupt the open structure typically promoted by NBO formation and instead lead to a denser packing.[17,18] These findings demonstrate the structural impact of Gd_2O_3 substitution on the glass network connectivity and packing in Gd_2O_3 added germanoborate glasses.

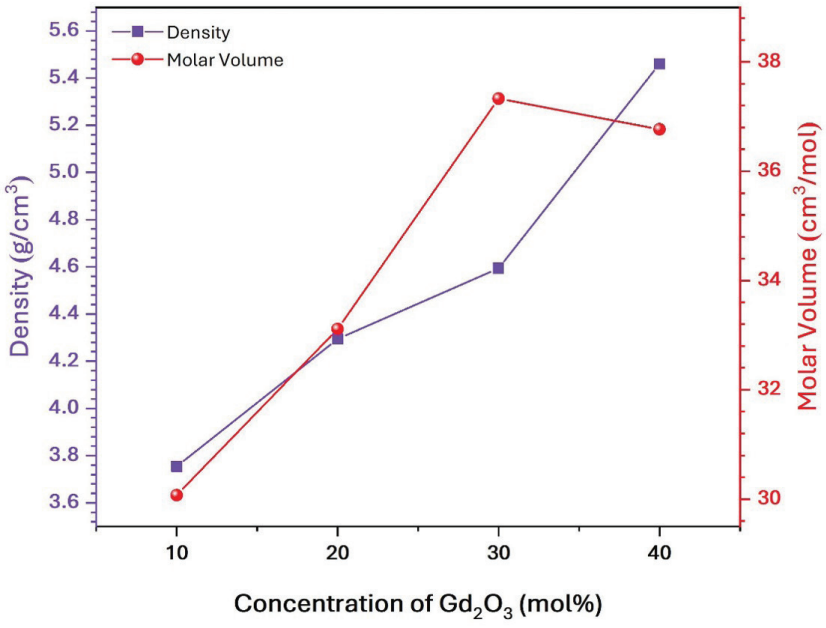


Figure 2. *The density and molar volume of Gd_2O_3 added germanoborate glasses.*

Radiations shielding properties

Figure 3. illustrates the dominant photon interaction mechanisms in germanoborate glass added with 40 mol% Gd_2O_3 over the photon energy range of 0.015-15 MeV. The interactions are categorized into three primary sub processes the photoelectric effect, Compton scattering, and pair production. In the low energy region (below 0.1 MeV), the photoelectric effect is the predominant attenuation mechanism. This is consistent with the increased probability of photon absorption in high atomic number (Z) materials such as gadolinium (Z=64). Notably, a distinct absorption edge, or K-edge, appears near the binding energy of the K-shell electrons of Gd, reflecting a sharp increase in photon attenuation efficiency in this energy region. As photon energy increases into the intermediate range (0.1-1 MeV), Compton scattering becomes the principal interaction, characterized by the elastic scattering of photons. This process is closely related to the materials electron density and plays a critical role in applications requiring intermediate energy radiation shielding, particularly in medical diagnostics and industrial radiography. At higher photon energies (>1.02 MeV), pair production becomes increasingly significant. In this process, a high energy photon is converted into an electron-positron pair when interacting near the nucleus. The rising attenuation coefficient in this region indicates the material's capability to manage high energy photons, which is essential for shielding applications involving linear accelerators and other high energy radiation sources. Overall, the total attenuation curve represents the cumulative contribution of all interaction mechanisms, exhibiting distinct variations as a function of photon energy. The addition of Gd_2O_3 significantly improves photon attenuation efficiency across all energy regions, particularly at lower energies due to the strong influence of the photoelectric effect and the prominent K-edge of Gd.

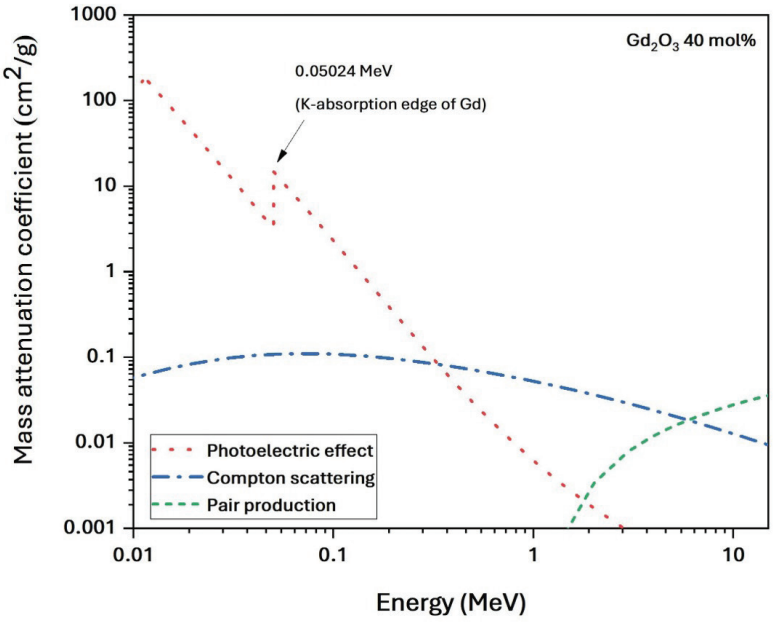


Figure 3. *The partial interaction of Gd_2O_3 added germanoborate glasses.*

The mass attenuation coefficient (μ_m) of germanoborate glasses with the composition (60-x) B_2O_3-GeO_2-xGd_2O_3 (where x=10, 20, 30, 40 mol%) was evaluated across a photon energy range of 0.015 to 15 MeV using the WinXCom software were present in Figure 4. The results revealed that μ_m decreases with increasing photon energy, consistent with the dominant interaction mechanisms governing photon attenuation namely, the photoelectric effect, Compton scattering, and pair production. At low photon energies (<0.2 MeV), the photoelectric effect is the dominant interaction, resulting in a high μ_m value that rapidly decreases as energy increases. A pronounced peak in μ_m was observed around 0.05024 MeV, corresponding to the K-absorption edge of gadolinium (Gd), where photon energy is sufficient to eject K-shell electrons from Gd atoms, significantly enhancing attenuation. Following this sharp rise, μ_m declines steadily in the intermediate energy range (0.2-1 MeV), where Compton scattering becomes the predominant interaction.

In the high energy region (>1.02 MeV), pair production becomes the leading mechanism, characterized by the conversion of high energy photons into electron positron pairs, altering the attenuation behavior based on interaction probability. Comparative analysis of the glasses indicated that increasing the Gd_2O_3 content leads to higher μ_m values, especially in the low energy region. This enhancement is attributed to the high atomic number (Z=64) and density of Gd, which significantly increases the probability of the photoelectric effect. However, at higher photon energies, the influence of Gd content on μ_m becomes less significant, as the

dominant interactions. Compton scattering and pair production are less sensitive to atomic number.[19] These findings demonstrate that germanoborate glasses added with higher concentrations of Gd_2O_3 are particularly effective in attenuating gamma radiation, making them promising candidates for radiation shielding applications.

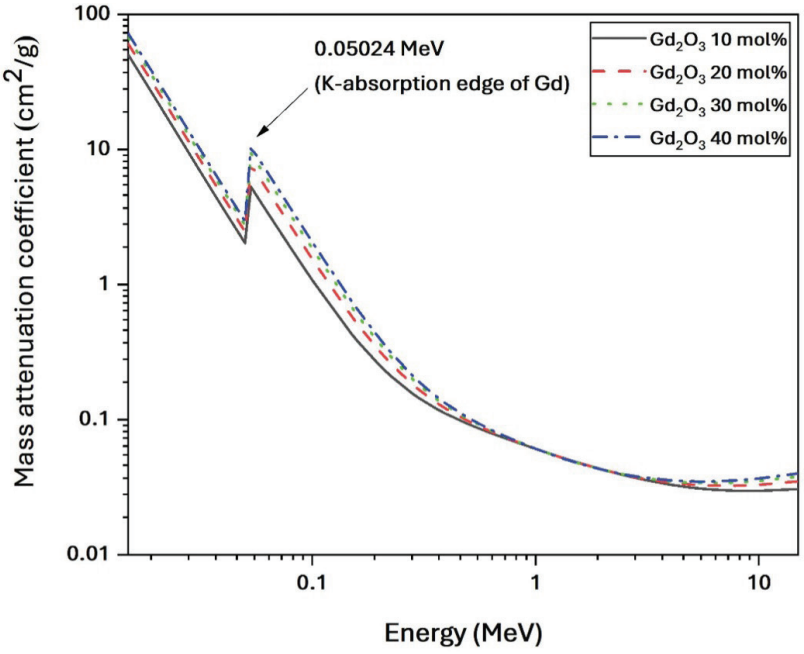


Figure 4. *The mass attenuation coefficient of Gd_2O_3 added germanoborate glasses.*

Figure 5. illustrates the variation of the effective atomic number (Zeff) of germanoborate glasses with different Gd_2O_3 contents (10-40 mol%) over a photon energy range of 0.015 to 15 MeV. The Zeff values show a strong dependence on photon energy, governed by the underlying photon interaction mechanisms. At lower energies (below 0.1 MeV), Zeff exhibits relatively high values, particularly for glasses with higher Gd_2O_3 content. This enhancement is due to the photoelectric effect, which strongly favors high atomic number constituents such as gadolinium (Z = 64). As a result, the substitution of B_2O_3 with Gd_2O_3 significantly elevates the effective atomic number in this region. A distinctive peak in Zeff is observed around 0.05024 MeV, which corresponds to the K-absorption edge of gadolinium. At this energy, photons possess sufficient energy to eject K-shell electrons, leading to a sharp rise in the interaction cross section and a corresponding sudden rise in Zeff. This localized increase reinforces the role of Gd in enhancing photoelectric absorption near its characteristic absorption edge. Beyond the absorption edge, Zeff values gradually decline with increasing photon energy, reaching a minimum in the intermediate range (approximately 0.2-2 MeV). This behavior corresponds to the dominance of Compton scattering, which is less sensitive to the atomic number and more dependent on the electron density of the material. In this range, the Zeff values for different Gd_2O_3 contents tend to converge, reflecting the diminished influence of high Z elements. At higher photon energies (>1.02 MeV), Zeff shows a slight upward trend, attributed to the onset of pair production, a process that becomes increasingly significant at high energies and is again influenced by the atomic number of the absorbing material. Overall, the variation of Zeff across the energy spectrum closely aligns with the trends observed in the μ_m, particularly in the photoelectric region. The incorporation of Gd_2O_3 not only elevates Zeff at low energies but also enhances the glass's photon interaction potential across a broad energy range, confirming its suitability for applications requiring efficient gamma ray shielding.

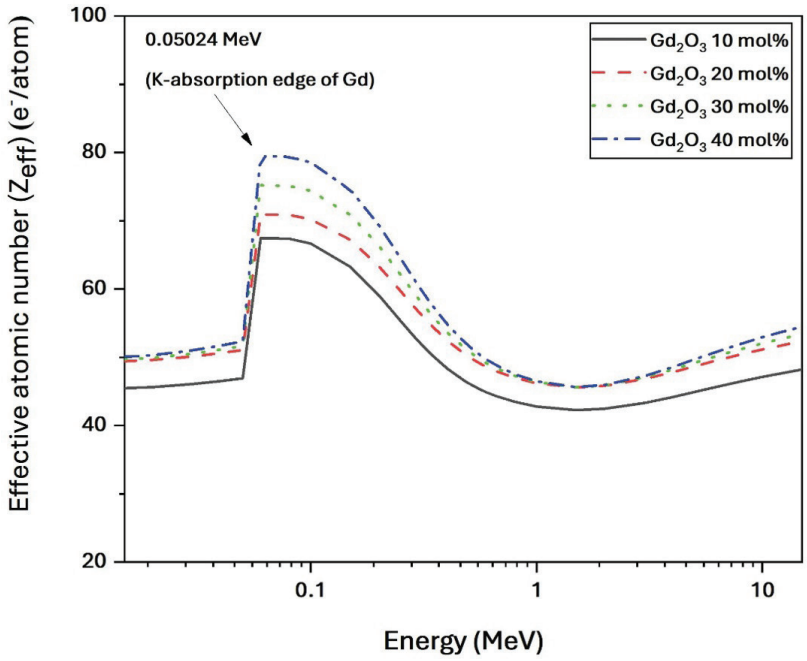


Figure 5. *The effective atomic number of Gd_2O_3 added germanoborate glasses.*

The variation of effective electron density (Neff) in Gd_2O_3 added germanoborate glasses, as shown in Figure 6., generally follows a trend similar to that of Um and Zeff, reflecting the dominant photon interaction mechanisms across energy ranges. However, it is noteworthy that (<0.05 MeV) the Neff values show inversion from the 10 mol% and 40 mol% Gd_2O_3 samples, with the 10 mol% composition exhibiting the highest Neff values. This behavior may result from the complex interplay between the increase in atomic number and the corresponding mass increase, which affects the electron density per unit mass. Furthermore, a pronounced peak is evident near the K-absorption edge of Gd, attributed to the photoelectric effect. Overall, Neff values decrease gradually with increasing photon energy, ranging from higher values at low energies to lower values at high energies, consistent with the energy dependent nature of radiation-matter interactions.[19]

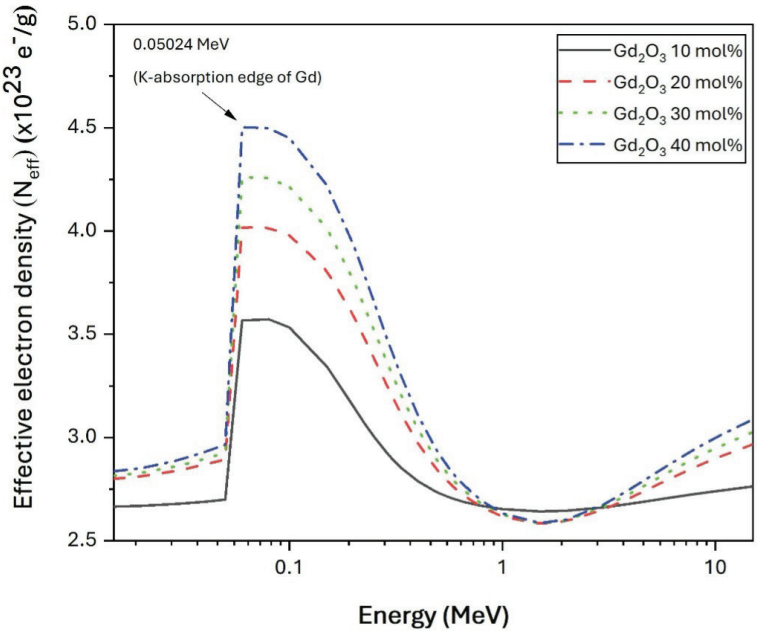


Figure 6. *The effective electron density of Gd_2O_3 added germanoborate glasses.*

Figure 7. illustrates the Half Value Layer (HVL) values of Gd_2O_3 added germanoborate glasses at a photon energy of 662 keV, as a function of Gd_2O_3 concentration (10-40 mol%). It is evident that HVL values decrease progressively with increasing Gd_2O_3 content, indicating enhanced gamma ray attenuation capability. This trend can be attributed to the high atomic number (Z = 64) of gadolinium, which significantly increases the probability of photon interactions particularly the photoelectric effect and Compton scattering within the glass matrix. For comparative, the HVL values of the Gd_2O_3 added glasses are plotted alongside standard shielding materials, including commercial window glass,[20] ordinary concrete,[21] and PbO Based Glass (50% PbO).[22] The results clearly demonstrate that all Gd_2O_3 added compositions possess superior shielding performance compared to both window glass and ordinary concrete. However, their HVL values remain higher than that of the PbO based glass, which is known for its excellent attenuation properties due to the high atomic number of lead (Z=82). Notably, the germanoborate glass sample containing 40 mol% Gd_2O_3 exhibits the lowest HVL among the studied compositions, highlighting its potential as a promising lead-free gamma ray shielding material.

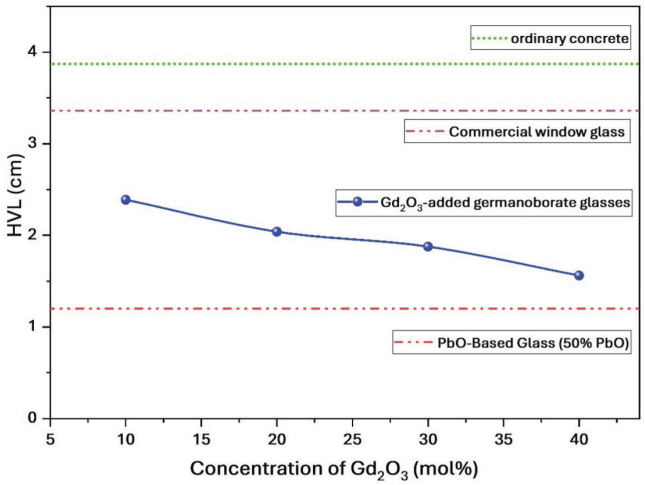


Figure 7. *The Half Value Layer (HVL) values of Gd_2O_3 added germanoborate glasses.*

The Pb equivalent of Gd_2O_3 added germanoborate glasses was determined for a thickness of 0.3 cm, as shown in Figure 8. From the results, the Pb equivalent of Gd_2O_3 added germanoborate glasses increases with the addition of Gd_2O_3. This trend indicates an enhancement in the materials radiation shielding capability, as higher Pb equivalent values correspond to greater attenuation effectiveness against ionizing radiation. When compared to standard reference shielding materials. The Pb equivalent values of the glasses are notably higher than those of commercial window glass,[21] but remain slightly lower than those of PbO-based glass compositions.[22]

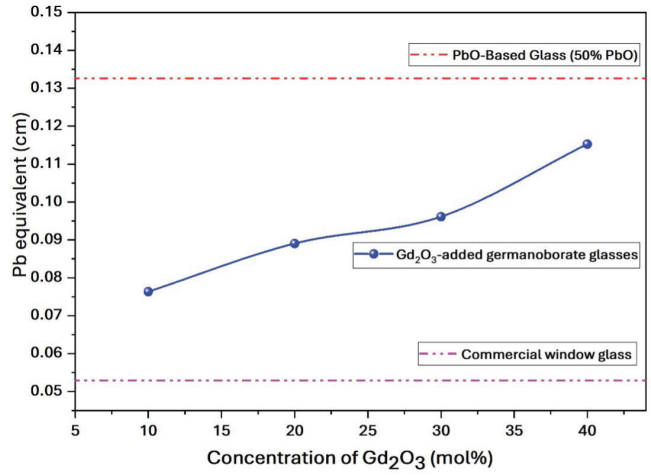


Figure 8. *The Pb equivalent of Gd_2O_3 added germanoborate glasses.*

The effective dose rate simulation

Figure 9 illustrates the effective dose rate (μSv/h) from gamma ray interactions with a computational phantom at an energy of 0.662 MeV. Simulations were performed using the PHITS software under anterior-posterior (AP) irradiation geometry to evaluate the shielding performance of Gd_2O_3 added glass at 40 mol%. The simulation results demonstrate a reduction in effective dose rate intensity when the Gd_2O_3 glass is used as a shielding material. Specifically, regions depicted in white, representing higher dose rates, are significantly diminished in the presence of 40 mol% Gd_2O_3 glass, while darker regions, indicating lower dose rates, become more prominent. This visual transition from white to deep blue or black suggests a decrease in gamma ray penetration, confirming the enhanced attenuation capability of the Gd_2O_3 glass. Figure 8(A) shows the effective dose rate distribution without any glass shielding material, with a maximum dose rate of 31.57 μSv/hr and a minimum of 0.1040 μSv/h. While Figure 8(B) which represents Gd_2O_3 glass at 40 mol%, shows a reduction in high dose regions, with dose rates ranging from 0.1315 μSv/hr (minimum) to 29.59 μSv/hr (maximum). These results highlight the effectiveness of Gd_2O_3 added glass in reducing radiation exposure in medical or shielding applications.

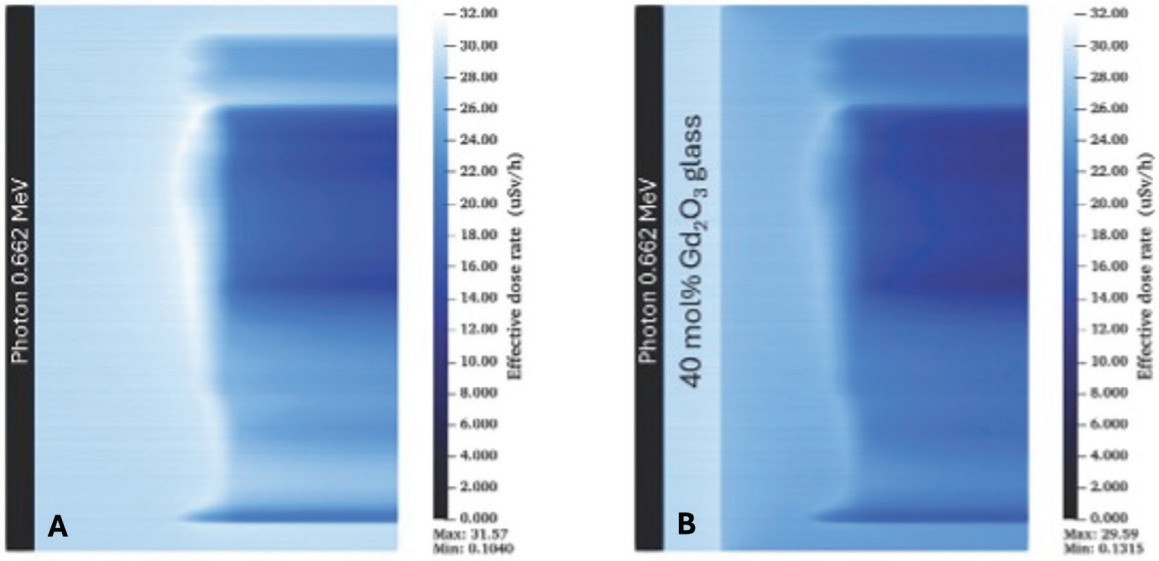


Figure 9. *The effective dose rate (μSv/h) simulation of glass in the energy range of 0.662 MeV. A: no glass, B: with Gd_2O_3 glasses at 40 mol%.*

Conclusion

In this work, the radiation shielding performance of Gd_2O_3 added germanoborate glasses with the composition (60-x)B_2O_3-GeO_2-Gd_2O_3 (x=10, 20, 30 and 40 mol%) was systematically investigated. The glasses were synthesized using the conventional melt quenching technique. Physical characterization revealed that increasing Gd_2O_3 content led to a rise in density from 3.7546 to 5.4604 gm/cm^3, while the molar volume initially increased. Radiation shielding parameters, including the mass attenuation coefficient (μ_m), effective atomic number (Zeff), and effective electron density (Neff), were evaluated over the energy range of 0.015-15 MeV using the WinXCom program. The results showed that all parameters followed the expected photon interaction mechanisms, which are dominated by the photoelectric effect at low energies, Compton scattering at intermediate energies, and pair production at high energies. Increasing the concentration of Gd_2O_3 resulted in higher values of these parameters, indicating improved radiation attenuation capability. From the HVL results, all Gd_2O_3 added germanoborate glasses exhibit superior shielding performance compared to commercial window glass and ordinary concrete, although their performance remains lower than that of PbO basedglass. Pb equivalent of glasses increases with Gd_2O_3 content, is significantly higher than that of commercial window glass, but remains slightly lower than PbO-based glass compositions. For the effective dose rate, the incorporation of 40 mol% Gd_2O_3 into the glass led to a noticeable reduction in high dose regions. These findings suggest that the incorporation of Gd_2O_3 significantly enhances the glass radiation shielding performance. Overall, the developed germanoborate glasses exhibit favorable physical and shielding properties, demonstrating strong potential for use as lead free transparent materials in radiation protection applications across medical, industrial, and nuclear fields.

Funding

National Research Council of Thailand for supporting this research (project number TSRI_68_3.3)

Conflict of interest

The authors declare no conflict of interest

CRediT authorship contribution statement

Natthakridta Chanthima: writing: original draft, validation, resources, investigation, funding acquisition, conceptualization; **Supakit Yonphan:** writing: review and editing, software, investigation; **Chalermpon Mutuwong**: methodology, investigation, software, conceptualization; **Wuttichai Chaiphaksa:** writing: review and editing, validation, conceptualization; **Jakrapong Kaewkhao:** writing: review and editing, supervision, conceptualization.

Acknowledgements

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_68_3.3)

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. JBPE. 2024; 14: 229-44. doi: 10.31661/jbpe.v0i0.2404-1742.

[2] Geist WH, Santi PA, Swinhoe Editors MT. Nondestructive assay of nuclear materials for safeguards and security.

[3] Bektasoglu M, Ali Mohammad M. Investigation of radiation shielding properties of TeO_2–ZnO–Nb_2O_5–Gd_2O_3 glasses at medical diagnostic energies. Ceram Int. 2020; 46(10): 16217-23. doi: 10.1016/j.ceramint.2020.03.177.

[4] 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. Vol. 10, Heliyon. Elsevier Ltd. 2024; doi: 10.1016/j.heliyon.2024.e40249.

[5] Hordieiev Y, Zaichuk A. Structural, thermal, and radiation shielding properties of antimony-doped zinc borate glasses. Sci Rep. 2025; 15(1): 16158. doi: 10.1038/s41598-025-96015-5.

[6] 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. Vol. 5, Materials Today: Proceedings. 2018;https://doi.org/10.1016/j.matpr.2018.02.062

[7] Gerward L, Guilbert N, Jensen KB, Levring H. WinXCom - A program for calculating X-ray attenuation coefficients. In: Radiation Physics and Chemistry. 2004; pp 653-4. doi: 10.1016/j.radphyschem.2004.04.040.

[8] Özdoğan H, Üncü YA, Akman F, Polat H, Kaçal MR. Investigation of gamma ray shielding characteristics of binary composites containing polyester resin and lead oxide. Polymers (Basel). 2024; 16(23): 3324. doi: 10.3390/polym16233324.

[9] 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: 10.1016/j.anucene.2014.08.015.

[10] Shamshad L, Rooh G, Limkitjaroenporn P, Srisittipokakun N, Chaiphaksa W, Kim HJ, et al. A comparative study of gadolinium-based oxide and oxyfluoride glasses as low energy radiation shielding materials. Prog Nucl Energy. 2017; 97: 53-9. doi: 10.1016/j.pnucene.2016.12.014.

[11] Limkitjaroenporn P, Cheewasukhanont W, Kothan S, Kaewkhao J. Development of new high transparency Pb-free radiation shielding glass. Integr Ferroelectr. 2021; 214(1): 181-204. doi:10.1080/10584587.2020.1857196.

[12] Gerward L, Guilbert N, Bj(rn Jensen K, Levring H. X-ray absorption in matter. Reengineering XCOM [Internet]. Vol. 60, Radiat Phys Chem. 2001; doi:10.1016/S0969-806X(00)00324-8.

[13] Kim CH, Yeom YS, Petoussi-Henss N, Zankl M, Bolch WE, Lee C, et al. ICRP Publication 145: Adult mesh-type reference computational phantoms. Ann ICRP. 2020; 49(3): 13-201. doi: 10.1177/0146645319893605. PMID: 33231095.

[14] Petoussi-Henss N, Bolch WE, Eckerman KF, Endo A, Hertel N, Hunt J, et al. Conversion coefficients for radiological protection quantities for external radiation exposures. Ann ICRP. 2010; 40(2-5): 1-257. doi: 10.1016/j.icrp.2011.10.001.

[15] Sato T, Iwamoto Y, Hashimoto S, Ogawa T, Furuta T, Abe SI, et al. Recent improvements of the particle and heavy ion transport code system–PHITS version 3.33. J Nucl Sci Technol. 2024; 61(1): 127-35. doi: 10.1080/00223131.2023.2275736.

[16] Chaiphaksa W, Yonphan S, Mutuwong C, Kothan S, Intachai N, Kaewkhao J. Experimental, theoretical, and Monte Carlo simulation study of radiation shielding properties of La2O3-Added lithium borate glasses. Radiat Phys Chem. 2024; 224: 111994. doi: 10.1016/j.radphyschem.2024.111994.

[17] Madshal MA, El-Damrawi G, Abdelghany AM, Abdelghany MI. Structural studies and physical properties of Gd2O3-doped borate glass. J Mater Sc: Mater Electron. 2021; 32(11): 14642-53. doi: 10.1007/s10854-021-06022-1.

[18] Chandra SRM, Prabhakar RA, Krishnamurthy GK, Appa RB. Physical and optical properties of PbO-Sb2O3-B2O3 glasses doped with Gd2O3. In: Materials Today: Proceedings. Elsevier Ltd. 2016; pp 3970-5. doi: 10.1016/j.matpr.2016.11.058.

[19] 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. 20205; 544: 120171. doi: 10.1016/j.jnoncrysol.2020.120171.

[20] 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 Mater Sci Process. 2021; 127(10): doi: 10.1007/s00339-021-04889-9.

[21] Singh K, Singh H, Sharma V, Nathuram R, Khanna A, Kumar R, et al. Gamma-ray attenuation coefficients in bismuth borate glasses. NIM-B. 2002; 194(10): 1-6. doi: 10.1016/S0168-583X(02)00498-6.

[22] 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: 10.1007/s41779-023-00951-2.

Journal of Associated Medical Sciences 2026; 59(1): 73-79

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 ciprofloxacin, prednisolone, and infliximab effects on liver functions in acetic acid-induced ulcerative colitis rat model

Olusoji Adebusoye Oyesola, Adebowale Raymond Adeojo, Oluwaseye Emmanuel Olayemi*, Eunice Oluwabunmi Ojo-Adebayo, Emmanuel Taiwo George

Department of Physiology, Faculty of Basic Medical Sciences, Olabisi Onabanjo University, Ago-Iwoye, Ogun State, Nigeria.

ARTICLE INFO

Article history:
Received 25 April 2025
Accepted as revised 21 August 2025
Available online 18 September 2025

Keywords:
Ulcerative colitis, ciprofloxacin, prednisolone, infliximab.

ABSTRACT

Background: Ulcerative colitis (UC) is a chronic inflammatory bowel disease characterized by recurring episodes of gastrointestinal inflammation and damage, affecting millions worldwide. Emerging evidence suggests a complex relationship between UC and liver dysfunction, with increased risk of hepatotoxicity and liver-related disorders.

Objectives: To investigate the effects of ciprofloxacin, prednisolone, and infliximab on liver function in an acetic acid-induced UC rat model.

Materials and methods: Fifty male Sprague-Dawley rats were divided into five groups of ten rats each: control, UC, ciprofloxacin, prednisolone, and infliximab group. UC was induced with 2 ml of 4% acetic acid solution transrectal. Liver function tests (AST, ALT, ALP), oxidative stress markers (MDA), protein synthesis (total protein, albumin), and bilirubin processing were evaluated, and histological examination of liver tissues was performed.

Results: The UC group showed significant increase in liver function enzymes (ALP, AST, ALT), oxidative stress markers (MDA) and significant decrease in total protein, albumin, total bilirubin and conjugated bilirubin level when compared with the control group. Respective treatment groups demonstrated improved liver enzyme levels, reduced oxidative stress, enhanced protein synthesis and bilirubin processing. Histological examination revealed improved liver architecture and reduced inflammation in treatment groups when compared with the UC group.

Conclusion: This study provides evidence that ciprofloxacin, prednisolone, and infliximab exert protective effects on liver function in acetic acid induced UC model. These findings suggest potential benefits of these therapies in mitigating liver damage associated with UC, highlighting the importance of considering liver health in UC management.

Introduction

Ulcerative colitis (UC) is a chronic and debilitating inflammatory bowel disease characterized by recurring episodes of gastrointestinal inflammation and damage.[1,2] This condition affects millions worldwide, significantly impacting quality of life and morbidity. Emerging evidence suggests a complex relationship between UC and liver dysfunction, with studies indicating an increased risk of hepatotoxicity and liver-related disorders in UC patients.[3,4]

The liver plays a critical role in metabolizing

* **Corresponding contributor.**
Author's Address: *Department of Physiology, Faculty of Basic Medical Sciences, Olabisi Onabanjo University, Ago-Iwoye, Ogun State, Nigeria.*
E-mail address: olayemi.oluwaseye@oouagoiwoye.edu.ng
doi: 10.12982/JAMS.2026.010
E-ISSN: 2539-6056

drugs and toxins, regulating immune responses, and maintaining overall health. However, its dysfunction can exacerbate the severity of UC, leading to adverse reactions and compromised therapeutic outcomes.[5] The liver's role in UC is multifaceted, involving the metabolism of pro-inflammatory cytokines, detoxification of bacterial products, and regulation of immune cell function.[6]

Current treatments for UC, including ciprofloxacin, prednisolone, and infliximab, have shown promise in mitigating intestinal inflammation. Ciprofloxacin target pathogenic bacteria, while prednisolone reduce inflammation and immune activity.[2] Infliximab, particularly those targeting tumor necrosis factor-alpha (TNF-α), have revolutionized UC management by modulating immune responses.[7]

However, the impact of these treatments on liver function in the context of UC remains poorly understood. This knowledge gap is concerning, as ciprofloxacin and prednisolone are known to have hepatotoxic potential.[8] Infliximab, while generally safe, may modulate liver immune responses in unpredictable ways, potentially affecting liver function. To address this critical knowledge gap, this study aims to investigate the effects of ciprofloxacin, prednisolone, and infliximab on liver function in a rat model of acetic acid-induced ulcerative colitis.

Materials and methods

Fifty adult male Sprague-Dawley rats seven weeks old (weight = 180-200g) obtained from Animal House of the Olabisi Onabanjo University, Sagamu Campus, Ogun State, Nigeria were housed in a standard environment conditions with 12 hours light/dark cycle and air filtration. Animals have free access to water and feed. Experimental protocols met the Guideline of Animal Experimentation approved by Olabisi Onabanjo University Teaching Hospital Human Research Ethics Committee (approval number: OOUTH/HREC/746/2023 AP). The animals were divided into five groups of ten rats each:

Group A: Normal control group
Group B: Ulcerative colitis group
Group C: Ciproflaxin-treated group
Group D: Prednisolone-treated group
Group E: Infliximab-treated group

Induction of ulcerative colitis

The induction of ulcerative colitis was done according to the method described by Al-Rejaie *et al.*,[9]. After 24 hours fast, the rats were anesthetized using 1.75 mL ketamine (1 gm/mL) and 0.25 mL Xylazine (1 gm/mL) intra-peritoneal prior to ulcerative colitis induction. 2 mL of 4% acetic acid solution was administered transrectally using a (2.7 mm) soft pediatric catheter. Rats were held horizontally for 2 min to prevent acetic acid leakage. Control group rats undergo the same procedure using equal volume of normal saline instead of acetic acid solution.

Administration of ciprofloxacin, prednisolone and infliximab

100 mg of ciprofloxacin was dissolved in 2.5 mL of distilled water to give a solution of 40 mg/ml, 40 mg/kg of ciprofloxacin was administered every 72 hours for forty-two days orally,[10]. 10 mg of prednisolone was disoled in 5 mL of distilled water to give a solution of 2 mg/mL, 2 mg/kg of prednisolone was administered every 72 hours for forty-two days orally,[11]. Five mg/kg body weight of infliximab was administered bi-weekly for forty-two days intra-peritoneal.[12]

Determination of liver function test

Serum alanine aminotransferase (ALT), aspartate aminotransferase (AST) and alkaline phosphatase (ALP) activities were carried out according to the method of Reitman and Frankel.[13] For Serum Alanine amino-transferase (ALT) determination, diluted sample (0.01 mL) was mixed with phosphate buffer (100mM, pH 7.4), L-alanine (200mM) and the mixture was incubated for exactly 30 minutes at 37 °C. An aliquot (0.5 mL) of 2, 4 dinitrophenylhydrazine (2mM) was added to the reaction mixture and allowed to stand for exactly 20 minutes at 25 °C. NaOH (0.4 M, 5 mL) was then added, and the absorbance was read against reagent blank after 5 minutes at 546 nm. Reagent blank was prepared as described above by replacing sample with 0.1 mL of distilled water.

For determination of Aspartate Aminotransferase (AST) activity, 0.1 mL of diluted sample was mixed with phosphate buffer (100 mM, pH 7.4), L-asparatate (100 mM), and α-Oxoglutarate (2 mM) and the mixture was incubated for exactly 30 minutes at 37 °C. An aliquot of 2, 4 dinitrophenylhydrazine (2 mM, 0.5 mL)) was then added to the reaction mixture and allowed to stand for exactly 20 minutes at 25 °C. Then, 0.5 mL of NaOH (0.4 M) was added and the absorbance was read against reagent blank after 5 minutes at 546 nm. Reagent blank was prepared as described above replacing sample with 0.1 mL of distilled water.

Alkaline phosphatase (ALP) activities were determined using 2.2 mL of 0.1 M carbonate buffer (pH 10.1), 0.1 mL of 0.1 M $MgSO_4.7H_20$, and 0.2 mL of the sample were mixed and incubated at 37 °C for 10 minutes. Thereafter, 0.5 mL of 19 mM of paranitrophenyl phosphate was added and again incubated at 37 °C for 10 minutes. NaOH (2.0 mL) was added and mixed, and read against blank at 400 nm.

Serum protein concentration was determined using the Biuret method.[14] Serum albumin was determined by the method of Doumas *et al.*,[15] based on the ability of bromocresol green (BCG), pH 4.2, to bind albumin which results in the formation of a complex that exhibits optical properties different from dye.

Following the method described by Tietz,[16] serum total bilirubin and conjugated bilirubin was determined

based on the principle of bilirubin in the serum that is coupled with diazotized sulphanilic acid to form azobilirubin.

Determination of malondialehyde activity (lipid peroxidation)

The MDA activity was measured by the double heating method.[17] One ml of tissue homogenate was combined with 2 mL of TCA-TBA-HCL and mixed thoroughly. The solution was heated for 15 minutes in a boiling water bath. After cooling, the fluorescent precipitate was removed by centrifugation at 1000g for 10 mins. The absorbance of the sample was determined at 535 nm against a blank that contains all the reagents minus the sample. The malondialdehyde concentration of the sample was calculated using an extinction coefficeient of 1.56×10^{-5} M^{-1} Cm^{-1}. Calculation of lipid peroxidation:

$$\text{MDA(nmol/ml)} = \boldsymbol{OD}\sum\times\frac{\boldsymbol{V}}{\boldsymbol{v}}$$

OD = Absorbance (optical density) of sample
∑ = Molar extinction coefficient
V = Total volume of the reacting sample
V = Volume of the sample

Histological analysis

The organs were carefully removed, and then the fat was removed. After weighing, they were fixed in 10% formal saline right away. The tissues were fixed, placed in progressively higher alcohol grades, and finally cleansed in xylene. After embedding them in paraffin, 5µm serial slices were produced. Hematoxylin and eosin were used to stain the sections. The slides were observed under a light microscope for the histological changes and morphometrical examination.

Statistical analysis

All the values are expressed as mean±standard error of mean (SEM). Analysis of data was done using Graph Pad Prism version 5 for Windows. Differences between groups were analyzed by one-way ANOVA followed by Bonferroni *post-hoc* test. Differences were considered significant at $p<0.05$.

Results

Results of liver function enzymes presented in figure 1 showed a significant increase in AST, ALT, and ALP in the ulcerative colitis group B (114.00±10.39, 75.66±10.26, and 79.00±9.50, respectively) when compared with the control group A (79.33±5.20, 39.00±1.52, and 38.66±14.43). Elevated AST, ALT and ALP levels indicate liver damage or dysfunction, hepatocellular damage or necrosis, and cholestasis or bone disorders respectively. Treatment group C, D, E showed a significant decrease in AST (88.00±2.64, 84.00±3.05, and 72.66±5.33), ALT (32.00±1.00, 29.33±9.66, and 37.00±1.00), ALP (40.66±1.20, 33.33±2.02, 39.33±6.88) when compared with group B respectively. No significant difference was observed within the respective treatment group.

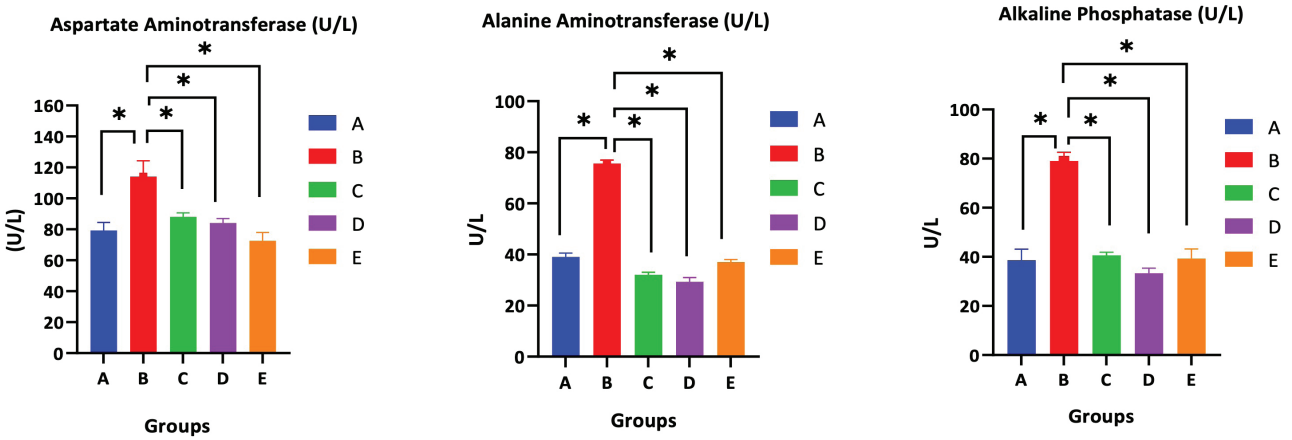
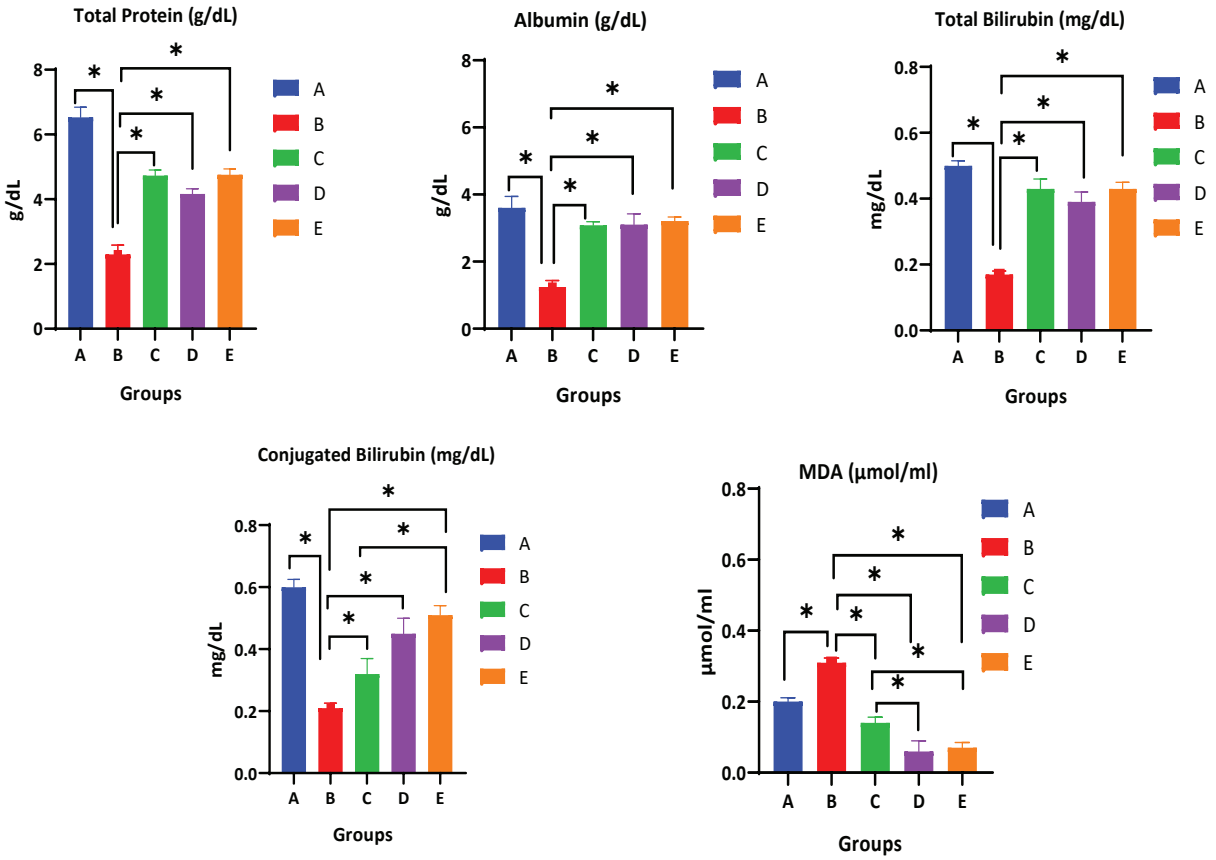


Figure 1. *The effect of ciprofloxacin, prednisolone and infliximab on liver function enzymes in acetic acid induced ulcerative colitis rat model. A-E: experimental groups, *p≤0.05.*

Results of some liver function markers and marker of oxidative stress presented in figure 2 showed a significant decrease in Total Protein, Albumin, Total Bilirubin, Conjugated Bilirubin in group B (2.30±1.28, 1.24±0.99, 0.17±0.10, 0.21±0.16) when compared with group A (6.53±0.31, 3.60±0.34, 0.50±0.15, 0.60±0.25, respectively). The decrease in total protein and albumin may be attributed to inflammation-induced suppression of protein synthesis, malabsorption of amino acids due to intestinal inflammation, and increased catabolism of proteins due to systemic inflammation while total bilirubin and conjugated bilirubin decrease may be attributed to reduced hepatic uptake and conjugation of bilirubin due to inflammation and impaired bile duct function and bile flow. Also, there was a significant increase in liver MDA in group B (0.31±0.13) when

compared with group A (0.20±0.11) which indicates lipid peroxidation and cellular damage.

Treated group C, D, E showed a significant increase in TP (4.73±0.75, 4.16±1.16, 4.76±0.78), ALB (3.08±0.11, 3.10±0.32, 3.21±0.12), TB (0.43±0.06, 0.39±0.03, 0.43±0.02), CB (0.32±0.05, 0.45±0.05, 0.51±0.03) and significant decrease in MDA (0.14±0.16, 0.06±0.03, 0.07±0.05) when compared with group B respectively. However, there was a significant decrease in conjugated bilirubin in group C compared to group E. More so, a significant increase was observed in liver MDA in group C compared to group D and E respectively.



shows the effect of antibiotics, anti-inflammatory agent and monoclonal agent on some liver function markersand marker of oxidative stress in acetic acid induced ulcerative colitis rat model.

*LEGEND: A= Group A, B= Group B, C= Group C, D= Group D, E= Group E *= statistically significant at p≤0.05*

Figure 2. *The effect of ciprofloxacin, prednisolone and infliximab on some liver function markers and marker of oxidative stress in acetic acid induced ulcerative colitis rat model. A-E: experimental groups, *p≤0.05.*

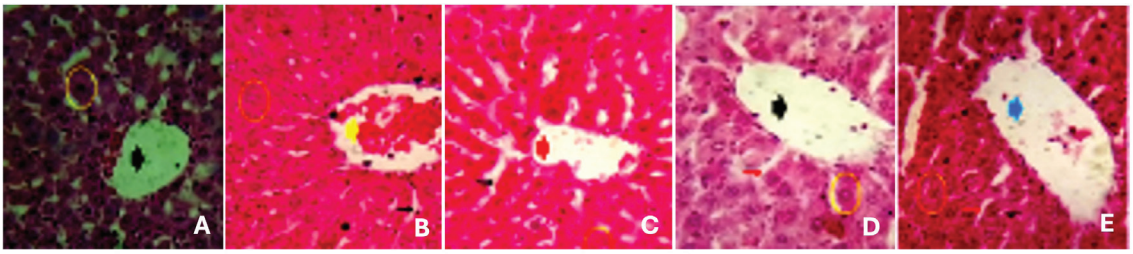


Figure 3. *Photomicrograph of hepatic tissue showing the effect of ciprofloxacin, prednisolone and infliximab in acetic acid induced ulcerative colitis rat model. (H&E Staining, x400 magnification).*

A: well differentiated and organized hepatocytes (yellow circle), sinusoids with Kupfer cells (black thin arrow) and clear central vein (black thick arrow),

B: congested central vein (yellow thick arrow), loss of hepatocytes (yellow circle) and constricted sinusoids (black thin arrow),

C: Normal sinusoids (black thin arrow), clear central vein (red thick arrow) and slight loss of hepatocytes (yellow circle),

D: clear central vein (black thick arrow), constricted sinusoids (red thin arrow) and hepatocytes (yellow circle),

E: clear irregular central vein (blue thick arrow), hepatocytes (yellow circle) and constricted sinusoids (red thin arrow).

Discussion

Aspartate transaminase (AST), a liver enzyme that catalyzes the conversion of aspartate and α-ketoglutarate to oxaloacetate and glutamate.[18,19] Elevated AST levels indicate liver damage or dysfunction.[20] Alanine Aminotransferase (ALT), a cytosolic enzyme primarily found in hepatocytes, catalyzing the conversion of alanine and α-ketoglutarate to pyruvate and glutamate.[18,19] Elevated ALT levels indicate hepatocellular damage or necrosis.[20] Alkaline Phosphatase (ALP), a membrane-bound enzyme, plays a crucial role in bile duct function and bone mineralization.[21] Elevated ALP levels indicate cholestasis or bone disorders.[22]

In this study, AST, ALT, and ALP level (Figure 1) significantly increased in the ulcerative colitis group compared to the control group, suggesting liver impairment, bile duct dysfunction or inflammation. Previous study showed that this elevation may be attributed to the systemic inflammation and oxidative stress associated with ulcerative colitis,[23] release of pro-inflammatory cytokines, such as TNF-α and IL-6, which induce hepatocellular damage.[24] Additionally, the acetic acid-induced colitis model triggers oxidative stress, leading to liver damage.[25]

Treatment with ciprofloxacin, prednisolone, and infliximab significantly decreased AST, ALT, and ALP levels compared to the ulcerative colitis group. Based on previous findings, ciprofloxacin which is an antibiotic may likely reduce bacterial translocation and subsequent liver damage,[26] prednisolone may decrease inflammation and oxidative stress[27] and infliximab target TNF-α to reduce inflammation and improved liver function.[28]

Malondialdehyde (MDA) levels (Figure 2), a marker of oxidative stress, significantly increased in the ulcerative colitis group compared to the control group. Previous study has showed that increase MDA indicates lipid peroxidation and cellular damage.[29] prednisolone and infliximab treatment groups showed significant decreased MDA levels compared with UC group and ciprofloxacin treated group, suggesting reduced oxidative stress. The reduction in MDA levels may be attributed to the antioxidant properties of prednisolone, infliximab targeting TNF-α also mitigate oxidative stress by reducing inflammation according to Rutgeerts *et al*.[28]

Total protein and albumin are essential proteins synthesized by the liver, playing critical roles in maintaining blood volume, transporting hormones, vitamins, and minerals, and regulating fluid balance.[16] Decreased levels of total protein and albumin (Figure 2) in the ulcerative colitis group compared to the control group. These decrease according to previous studies may be attributed to inflammation-induced suppression of protein synthesis, malabsorption of amino acids due to intestinal inflammation, and increased catabolism of proteins due to systemic inflammation.[30,31] Treatment with ciprofloxacin, prednisolone, and infliximab significantly increased total protein and albumin levels compared to the ulcerative colitis group. This may be due to reduced inflammation and oxidative stress,[27] improved intestinal absorption of amino acids,[26] enhanced protein synthesis and reduced catabolism.[28]

Total bilirubin and conjugated bilirubin are breakdown products of heme, primarily processed by the liver.[32] Decreased levels in the ulcerative colitis group (Figure 2). The decrease may be attributed to reduced hepatic uptake and conjugation of bilirubin due to inflammation and impaired bile duct function and bile flow.[22] Treatment groups showed increased total bilirubin and conjugated bilirubin levels, suggesting improved liver function and bilirubin processing. This may be due to reduced inflammation and oxidative stress, enhanced hepatic uptake and conjugation of bilirubin,[28] and improved bile duct function and bile flow.

Histological examination of liver tissues (Figure 3) revealed significant alterations in the ulcerative colitis group compared to the control group. The control group exhibited well-differentiated and organized hepatocytes, patent sinusoids with intact Kupffer cells, and a clear central vein, indicative of normal liver architecture. In contrast, the ulcerative colitis group showed a congested central vein indicating increased blood pressure and inflammation, loss of hepatocytes suggesting hepatocellular damage or apoptosis, constricted sinusoid indicating reduced blood flow and increased oxidative stress. These changes are consistent with liver dysfunction in inflammatory bowel disease (IBD).[24,33] The systemic inflammation and oxidative stress associated with ulcerative colitis likely contribute to liver damage.

Ciprofloxacin treatment groups demonstrated improved histological features such as a normal sinusoid indicating restored blood flow and reduced inflammation, clear central vein suggesting improved blood pressure and reduced congestion, and a slight loss of hepatocytes which indicates partial recovery of hepatocellular damage. Ciprofloxacin likely reduced bacterial translocation and subsequent liver damage, leading to improved histological features.[34]

Prednisolone treated group showed a clear central vein indicating reduced inflammation and improved blood pressure, constricted sinusoid which may indicate persistent oxidative stress, hepatocytes indicating preserved liver cell architecture. Prednisolone reduced inflammation and oxidative stress, preserving hepatocellular structure and function.[27]

Infliximab treatment group showed a clear irregular central vein which suggests improved blood pressure and reduced congestion, hepatocytes indicating preserved liver cell architecture, constricted sinusoid which may indicate persistent oxidative stress.[25] Infliximab targeting TNF-α reduced systemic inflammation and oxidative stress, leading to improved liver histology.[28]

Conclusion

The results of this study demonstrate that treatment with ciprofloxacin, prednisolone, and infliximab regulate liver enzyme levels, reduced oxidative stress, and improve protein synthesis and bilirubin processing. Histological examination revealed improved liver architecture, reduced inflammation, and oxidative stress in treatment groups. These findings suggest that these treatments particularly infliximab may be beneficial in reducing bacterial translocation, decreasing inflammation and oxidative stress, enhancing protein synthesis, and improving bilirubin processing.

Ethical approval

The National Research Council's guidelines for the care and use of laboratory animals were followed in all animal studies and methodology. Ethical approval was obtained from Olabisi Onabanjo University Teaching Hospital Human Research Ethics Committee (OOUTH-HREC), with the number OOUTH/HREC/746/2023AP.

Funding

This study received no external funding.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Oyesola OA and **Adeojo AR**: conceptualization and method; **Adeojo AR**, **Olayemi OE**, **Ojo-Adebayo EO**, and **George ET**: data collection and analysis; Adeojo **AR**, **Olayemi OE**, **Ojo-Adebayo EO**, and **George ET**: interpretation of results; **Oyesola OA**: supervision, writing: original draft. All authors contributed to writing and revising the manuscript.

Acknowledgements

We would like to acknowledge no one for their contribution to this research.

References

[1] Gajendran, M., Loganathan, P., Jimenez, G., Catinella, A. P., Ng, N., Umapathy, C., et al. A comprehensive review and update on ulcerative colitis. Dis Mon. 2019; 65(12): 100851. doi:10.1016/j.disamonth.2019.02.004.

[2] Kucharzik T, Koletzko S, Kannengiesser K, Dignass A. Ulcerative colitis-diagnostic and therapeutic algorithms. DtschArztebl Int. 2020; 117(33-34): 564-74. doi:10.3238/arztebl.2020.0564.

[3] Wieser V, Gerner R, Moschen AR, Tilg H. Liver complications in inflammatory bowel diseases. DigDis.2013;31(2):233-8.doi:10.1159/000353377.

[4] Schröder, T., Schmidt, K. J., Olsen, V., Möller, S., Mackenroth, T., Sina, C., et al. Liver steatosis is a risk factor for hepatotoxicity in patients with inflammatory bowel disease under immunosuppressive treatment. Eur J Gastroenterol Hepatol. 2015; 27(6): 698-704. doi:10.1097/MEG.0000000000000350.

[5] Chen M, Suzuki A, Borlak J, Andrade RJ, Lucena MI. Drug-induced liver injury: Interactions between drug properties and host factors. J Hepatol. 2015; 63(2): 503-14. doi:10.1016/j.jhep.2015.04.016.

[6] Robinson MW, Harmon C, O'Farrelly C. Liver immunology and its role in inflammation and homeostasis. Cell Mol Immunol. 2016; 13(3): 267-76. doi:10.1038/cmi.2016.3.

[7] Melero I, Grimaldi AM, Perez-Gracia JL, Ascierto PA. Clinical development of immunostimulatory infliximab and opportunities for combination. Clin Cancer Res. 2013; 19(5): 997-1008. doi:10.1158/1078-0432.CCR-12-2214.

[8] Devarbhavi H, Andrade RJ. Drug-induced liver injury due to antimicrobials, central nervous system agents, and nonsteroidal anti-inflammatory drugs. Semin Liver Dis. 2014; 34(2): 145-61. doi:10.1055/s-0034-1375956.

[9] Al-Rejaie SS, Abuohashish HM, Al-Enazi MM, Al-Assaf AH, Parmar MY, Ahmed MM. Protective effect of naringenin on acetic acid-induced ulcerative colitis in rats. World J Gastroenterol. 2013; 19(34): 5633-44. doi:10.3748/wjg.v19.i34.5633.

[10] Collins HH, Cross AS, Dobek A, Opal SM, McClain JB, Sadoff JC. Oral ciprofloxacin and a infliximab to lipopolysaccharide protects leukopenic rats from lethal infection with Pseudomonas aeruginosa. J Infect Dis. 1989; 159(6): 1073-82. doi:10.1093/infdis/159.6.1073.

[11] Witaicenis A, Luchini AC, Hiruma-Lima CA, Felisbino SL, Garrido-Mesa N, Utrilla P, et al. Suppression of TNBS-induced colitis in rats by 4-methylesculetin, a natural coumarin: comparison with prednisolone and sulphasalazine. Chem Biol Interact. 2012; 195(1): 76-85. doi:10.1016/j.cbi.2011.11.004.

[12] Sandborn, W. J., & Hanauer, S. B. Infliximab in the treatment of Crohn's disease: a user's guide for clinicians. The American journal of gastroenterology. 2002; 97(12): 2962-72. doi.org/10.1111/j.1572-0241.2002.07093.x.

[13] REITMAN S, FRANKEL S. A colorimetric method for the determination of serum glutamic oxalacetic and glutamic pyruvic transaminases. Am J Clin Pathol. 1957; 28(1): 56-63. doi:10.1093/ajcp/28.1.56.

[14] Plummer D.T. An Introduction to Practical Biochemistry. 2nd Edition, London; McGraw-Hill: 1978: pp 144-145.

[15] Doumas BT, Watson WA, Biggs HG. Albumin standards and the measurement of serum albumin with bromcresol green. Clin Chim Acta. 1971; 31(1): 87-96. doi:10.1016/0009-8981(71)90365-2

[16] Tietz, N.W. Clinical Guide to Laboratory Tests (ELISA). 3rd Edition, Philadelphia; W.B. Saunders, Co: 1995: pp 22-23.

[17] Stocks J, Dormandy TL. The autoxidation of human

red cell lipids induced by hydrogen peroxide. BJ Haem. 1971; 20(1): 95-111.

[18] Bergmeyer HU, Editor. Methods of enzymatic analysis. Elsevier; 2012.

[19] Schiff ER, Maddrey WC, Reddy KR, editors. Schiff's Diseases of the Liver. John Wiley & Sons; 2017.

[20] Kalas MA, Chavez L, Leon M, Taweesedt PT, Surani S. Abnormal liver enzymes: A review for clinicians. World J Hepatol. 2021; 13(11): 1688.

[21] Vimalraj S. Alkaline phosphatase: Structure, expression and its function in bone mineralization. Gene. 2020; 754: 144855. doi:10.1016/j.gene.2020.144855.

[22] Lowe D, Sanvictores T, Zubair M, John S. Alkaline phosphatase. StatPearls. 2023.

[23] Losurdo, G., Brescia, I. V., Lillo, C., Mezzapesa, M., Barone, M., Principi, M., et al. Liver involvement in inflammatory bowel disease: What should the clinician know?. World J Hepatol. 2021; 13(11): 1534-51. doi:10.4254/wjh.v13.i11.1534.

[24] Björnsson HK, Björnsson ES. Hepatotoxicity in inflammatory bowel disease: Immunomodulators, biologics, and beyond. Clin Liver Dis (Hoboken). 2024; 23(1): e0199. Published 2024 Jun 14. doi:10.1097/CLD.0000000000000199.

[25] Owusu, G., Obiri, D. D., Ainooson, G. K., Osafo, N., Antwi, A. O., Duduyemi, B. M., et al. Acetic Acid-Induced Ulcerative Colitis in Sprague Dawley Rats Is Suppressed by Hydroethanolic Extract of Cordia vignei Leaves through Reduced Serum Levels of TNF-α and IL-6. Int J Chronic Dis. 2020; 2020: 8785497. Published 2020 Feb 6. doi:10.1155/2020/8785497.

[26] Abraham B, Quigley EMM. Ciprofloxacin and probiotics in inflammatory bowel disease: when to use them? [published correction appears in Frontline Gastroenterol. 2020;11(3): e1. doi: 10.1136/flgastro-2018-101057corr1]. Frontline Gastroenterol. 2020; 11(1): 62-9. doi:10.1136/flgastro-2018-101057.

[27] Jeuring, S. F. G., Biemans, V. B. C., van den Heuvel, T. R. A., Zeegers, M. P., Hameeteman, W. H., Romberg-Camps, M. J. L., et al. Corticosteroid sparing in Inflammatory bowel disease is more often achieved in the immunomodulator and biological Era-results from the Dutch population-based IBDSL cohort. Am J Gastroenterol. 2018; 113(3): 384-95. https://doi.org/10.1038/ajg.2017.482.

[28] Sands BE, Hanauer SB, Colombel JF, Sandborn W, Schreiber S, Danese S, Klopocka M, Kulynych R, Kierkus J, Soltysiak A, Smolinski P. P492 Subcutaneous infliximab (CT-P13 SC) as maintenance therapy for ulcerative colitis: a phase 3, randomized, placebo-controlled study: results of the LIBERTY-UC study. J Crohn's Colitis. 2023; 17(Supplement_1): i623-4.

[29] Jarmakiewicz-Czaja S, Ferenc K, Filip R. Antioxidants as protection against reactive oxidative stress in inflammatory bowel disease. Metabolites. 2023; 13(4): 573.

[30] Balestrieri P, Ribolsi M, Guarino MP, Emerenziani S, Altomare A, Cicala M. Nutritional aspects in inflammatory bowel diseases. Nutrients. 2020; 12(2): 372.

[31] Vidal-Lletjós S, Beaumont M, Tomé D, Benamouzig R, Blachier F, Lan A. Dietary protein and amino acid supplementation in inflammatory bowel disease course: what impact on the colonic mucosa?. Nutrients. 2017; 9(3): 310.

[32] Kalakonda A, Jenkins BA, John S. Physiology, Bilirubin. In: StatPearls. Treasure Island (FL): StatPearls Publishing; 2022.

[33] Gaspar R, Branco CC, Macedo G. Liver manifestations and complications in inflammatory bowel disease: A review. World J Hepatol. 2021; 13(12): 1956-67. doi:10.4254/wjh.v13.i12.1956.

[34] Fu L, Qian Y, Shang Z, Sun X, Kong X, Gao Y. Ciprofloxacin enhancing drug-induced liver injury assessed for causality using Roussel Uclaf causality assessment method: Emerging role of gut microbiota dysbiosis. Front Med (Lausanne). 2022; 9: 972518. doi:10.3389/fmed.2022.972518.

Journal of Associated Medical Sciences 2026; 59(1): 80-86

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Case Report: Two cases of Hb Malay (HBB: c.59A>G) found in Northern Thailand

Moe Theingi[1], Chedtapak Ruengdit[2], Manoo Punyamung[2], Sakorn Pornprasert[1*]

[1]*Division of Clinical Microscopy, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*

[2]*Associated Medical Sciences Clinical Service Center, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*

ARTICLE INFO

Article history:
Received 5 July 2025
Accepted as revised 16 September 2025
Available online 19 September 2025

Keywords:
Hb Malay, b+-thalassemia, iron deficiency, hemoglobinopathy, diagnosis.

ABSTRACT

Background: Hemoglobin (Hb) Malay is a common β-hemoglobinopathy in Malaysia resulting from an AAC to AGC mutation at codon 19, which produces an abnormal β-globin chain and manifests as a β^{+}-thalassemia phenotype characterized by mild anemia and increased HbA_2 levels. Although Hb Malay is commonly prevalent in Southern Thailand, there have been few reported cases in Northern Thailand.

Objectives: This study aims to report two cases of Hb Malay detected in Chiang Mai, Northern Thailand, highlighting the diagnostic complexities.

Materials and methods: Two female patients, aged 50 and 67, presenting with anemia were investigated. Initial hematological profiles, Hb analysis by HPLC (Case 1) and CE (Case 2), and iron studies were performed. Due to hypochromic-microcytic anemia and elevated HbA_2 levels in both cases, genomic DNA was extracted. Multiplex real-time PCR with HRM analysis was performed to detect common α^{0}-thalassemia deletions. Further genetic analysis was conducted using next-generation sequencing (NGS) targeting *HBA1, HBA2, and HBB* genes.

Results: Both cases were identified to carry the Hb Malay variant. Case 1, a 50-year-old female with mild anemia, was diagnosed with double heterozygosity for Hb Malay (β^{Malay}/β^{A}) and α^{+}-thalassemia ($-\alpha^{3.7}/\alpha\alpha$). Case 2, a 67-year-old female with severe anemia and iron deficiency, was diagnosed with heterozygosity for Hb Malay (β^{Malay}/β^{A}).

Conclusion: The diagnosis of Hb Malay can be complicated especially when it coexists with other thalassemia traits or iron deficiency. Therefore, better understanding of the hematological and clinical characteristics, as well as the laboratory detection of this hemoglobin variant, would be beneficial for genetic counseling, particularly in areas with a high prevalence of thalassemia, hemoglobinopathy, and iron deficiency such as Northern Thailand.

Introduction

β-thalassemia is a diverse group of inherited disorders prevalent in East Asian populations, characterized by point mutations or deletions affecting β-globin chain production. The two types are β^{0}-thalassemia, which involves the complete absence of β-globin chain production, and β^{+}-thalassemia, resulting in reduced synthesis. Numerous β-globin gene mutations have been identified across countries such as India, Thailand, and China.[1,2] In general, these mutations are population-specific, meaning each

* **Corresponding contributor.**
Author's Address: *Division of Clinical Microscopy, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: sakornmi001@gmail.com
doi: 10.12982/JAMS.2026.011
E-ISSN: 2539-6056

ethnic group exhibits its set of common mutations.

In Thailand, the prevalence of β-thalassemia ranges from 3% to 9%. Common β-globin gene mutations include codons 41/42(–CTTT) (HBB: c.126_129delCTTT) and codon17(A>T) (HBB: c.52A>T), which are observed nationwide. Additionally, HbE (HBB: c.79 G>A) is frequently found in Southern and Northeastern Thailand, while Hb Malay (HBB: c.59A>G) is more common in Southern Thailand.[3] Hb Malay is caused by an AAC to AGC mutation in codon 19, leading to Asn-Ser replacement and the production of an abnormal β-globin chain which first case was confirmed by DNA probe hybridization.[1,4] The mutation activates a cryptic RNA splice site in exon 1 of the β-globin gene, leading to abnormal RNA processing and β^+-thalassemia phenotype.[5]

Routine diagnostic methods such as dichlorophenol indophenol precipitation (DCIP) test, high-performance liquid chromatography (HPLC), and capillary electrophoresis (CE) are widely used to detect Hb variants. Screening for β-thalassemia typically relies on HbA_2 level. However, borderline HbA_2 samples should be performed DNA analysis to confirm carriers, as missed cases may increase the risk for at-risk couples having affected offspring.[6] Hb Malay is one of the commonest β-globin gene mutations among individuals with borderline HbA_2 (3.0-3.9%).[6] Moreover, it can co-elute or co-migrate with HbA on HPLC chromatogram and CE electrophoregram, complicating diagnosis.[1,4] Reverse-phase HPLC, peptide analysis, and DNA studies, and advanced molecular techniques such as Multiplex Amplification Refractory Mutation System (M-ARMS) PCR, allele-specific PCR, and reverse dot-blot allele-specific oligonucleotide (ASO) hybridization are used to detect Hb Malay.

Hb Malay has been reported as highly prevalent (2.0%) in Southern Thailand.[7] However, a previous study in Northern Thailand documented four cases of HbE-β^+-thalassemia (Hb Malay), of which only one case was clearly identified as co-inheritance with rare α^+-thalassemia mutation (cap +14 C>G in the α_2-globin promoter).[8] One of the present cases involves the co-inheritance of Hb Malay with the common α^+-thalassemia deletion ($-\alpha^{3.7}/\alpha\alpha$), a combination not previously reported in Northern Thailand. Therefore, this study reports two cases of Hb Malay detected in our thalassemia laboratory at the Associated Medical Sciences-Clinical Service Center (AMS-CSC), Chiang Mai University, Chiang Mai, Thailand. The findings expand the documented spectrum of Hb Malay-associated thalassemia in the Northern region, highlighting the necessity for comprehensive molecular testing to ensure accurate diagnosis and effective genetic counseling.

Case report

Case 1

A 50-year-old woman presented with mild anemia detected during a health screening. Her hematological profile showed Hb 11.4 gm/dL, mean corpuscular volume (MCV) 78.0 fL, and mean corpuscular hemoglobin (MCH) 25.9 pg. Hemoglobin analysis using HPLC (VARIANT β-thalassemia Short Program; Bio-Rad Laboratories, Hercules, CA, USA) revealed HbA 84.3%, HbA_2 5.4%, and HbF 0.8% (Figure 1A), which suggests that she has a β-thalassemia trait. Iron studies were within normal range: Ferritin 100 ng/mL, Serum iron 89 µg/dL, TIBC 283 µg/dL, and Transferrin saturation 31% (Table 1). These findings indicated mild microcytic-hypochromic anemia in the absence of iron deficiency.

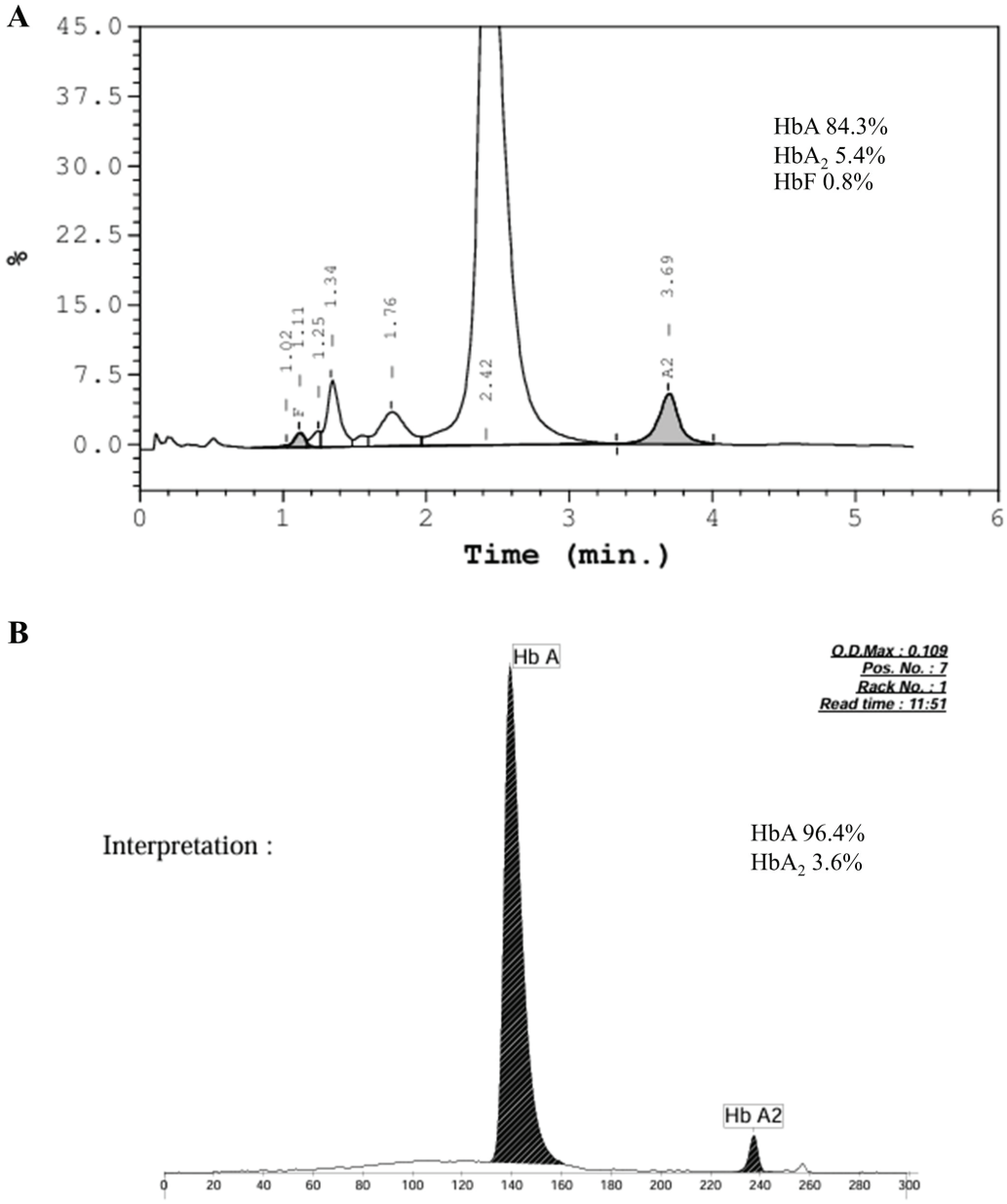


Figure 1. *The HPLC chromatogram (A) for Case 1, and CE electrophoregram (B) for Case 2, without seeing the separate peak of Hb Malay.*

Table 1. The characteristics and hematological parameters of cases.

Characteristics &hematological parameters	Case No. 1	Case No. 2
Age (years)	50	67
Gender	Female	Female
α-globin genotype	$-\alpha^{3.7}/\alpha\alpha$	αα/αα
β-globin genotype	β^{Malay}/β^{A}	β^{Malay}/β^{A}
RBCs ($x10^{12}$/L)	4.4	4.1
Total Hb (gm/dL)	11.4	6.4
PCV (L/L)	0.35	0.22
MCV (fL)	78.0	54.0
MCH (pg)	25.9	16.0
MCHC (gm/L)	325	290
RDW (%)	13.7	33.9
Ferritin (ng/mL)	100.0	12.1
Serum iron (mg/dL)	89.0	8.0
TIBC (mg/dL)	283	401
Transferrin saturation (%)	31.0	2.0
HbA (%)	84.3*	$96.4^{\#}$
HbA_2 (%)	5.4*	$3.6^{\#}$
HbF (%)	0.8*	$0.0^{\#}$

***Note**: Normal range of adults: red blood cell counts (RBCs) 4.2-6.1x10^{12}/L, total Hb 12.0-18.0 gm/dL, packed cell volume (PCV) 0.37-0.52 L/L, mean corpuscular volume (MCV) 80-100 fL, mean corpuscular Hb (MCH) 27.0-31.0 pg, mean corpuscular Hb concentration (MCHC) 320-360 gm/L, red cell distribution width (RDW) 11.0-16.0%, HbA 95.0-98.0%, HbA_2 1.5-3.5%, HbF 0.0-1.0%, ferritin 4.6-204.0 ng/mL, serum iron 50.0-170.0 mg/dL, total iron-binding capacity (TIBC) 259-388 mg/dL, transferrin saturation 20.0-50.0%. *Hb analysis by HPLC, #Hb analysis by CE.*

Case 2

A 67-year-old woman presented with severe anemia discovered during a routine laboratory screening. Her Hb was 6.4 gm/dL, with marked microcytosis (MCV 54.0 fL), low MCH (16.0 pg), and a red cell distribution width (RDW) of 33.9%. Iron studies showed severe deficiency results (Table 1). However, Hb analysis by CE (Capillarys 2 Flex piercing, Sebia, Evry, France) showed HbA 96.4% and elevated HbA_2 level (3.6%) (Figure 1B).

Due to the presence of microcytic-hypochromic anemia and elevated HbA_2 levels in both cases, genomic DNA was extracted from peripheral blood samples using the NucleoSpin® kit (Macherey-Nagel, KG., Duren, Germany) following the manufacturer's instructions. Subsequently, a single-tube multiplex real-time PCR with EvaGreen and high-resolution melting (HRM) analysis was performed to detect common α^0-thalassemia $--^{SEA}$, $--^{THAI}$, and $--^{CR}$ deletions, as previously described.[9] Neither patient carried these deletions. Further genetic analysis was performed using next-generation sequencing (NGS, BGI Group, Shenzhen, China), targeting the coding regions of the *HBA1*, *HBA2*, and *HBB* genes as previously described.[10] Both cases were found to carry the AAC to AGC mutation at codon 19 of the β-globin gene, leading to the Hb Malay variant (Figure 2). Furthermore, Case 1 had a single α-globin gene deletion ($-\alpha^{3.7}/\alpha\alpha$), whereas Case 2 did not. Thus, Case 1 was diagnosed as double heterozygosity for Hb Malay (β^{Malay}/β^{A}) and α^+-thalassemia ($-\alpha^{3.7}/\alpha\alpha$), while Case 2 was diagnosed with heterozygosity for Hb Malay (β^{Malay}/β^{A}) (Table 1).

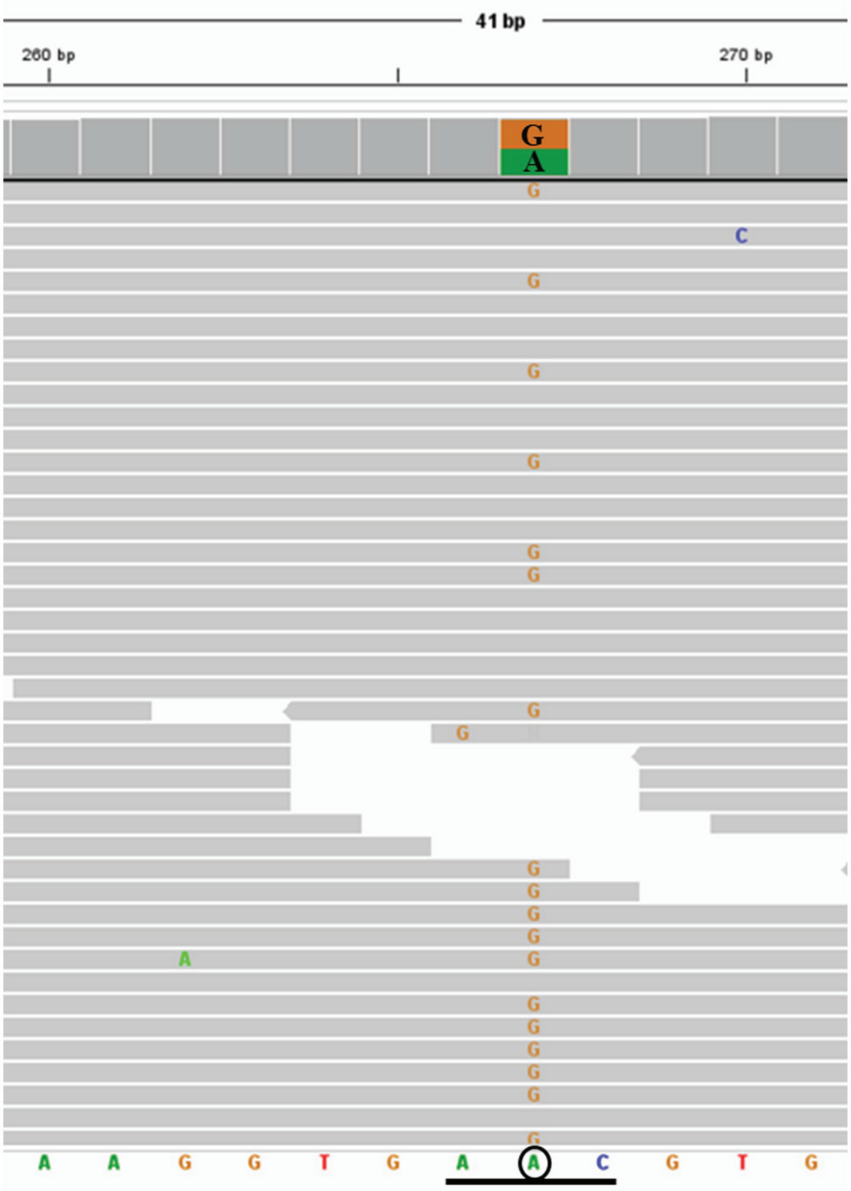


Figure 2. *Representative of NGS results exported from the integrative genomics viewer (IGV) of Hb Malay (HBB: c.59A>G).*

Discussion

Hemoglobin Malay was first reported in 1989 in a 22-year-old Malay patient.[4] The clinical features, hematological and biochemical findings in this patient and a sibling with homozygous Hb Malay showed mild to moderate hemolytic features consistent with thalassemia traits. Due to its HbA-like electrophoretic behavior with conventional electrophoresis, it was initially identified as β-thalassemia trait. However, unstable hemoglobin was detected in the patient and other family members, and a Hb variant eluting slightly later than the normal β-chains was seen by reverse-phase HPLC. The new variant was subsequently confirmed by hybridization of amplified DNA and later designated as Hb Malay.[1,4] In Thailand, hemoglobin variants are common and genetically heterogeneous among its population. Hb Malay is more prevalent in Southern Thailand and its detection in Northern cases may be related to historical migration or intermarriage with Southern populations. Even though an individual whose Hb typing appears normal or suggests a β-thalassemia trait with elevated HbA_2 (>3.5%), the presence of β-globin variants should be considered including Hb Malay.[6]

In Case 1, a 50-year-old female, her slightly low Hb, MCV and MCH values raised suspicion of thalassemia. Besides, her Hb analysis showed an elevated HbA_2 level (5.4%), suggestive of a β-thalassemia trait. The previous study reported that samples with double heterozygous for β- and α-thalassemia are usually diagnosed as typical β-thalassemia carriers and α-thalassemia is usually ignored.[11] However, there is no detectable common α^0-thalassemia $--^{SEA}$, $--^{THAI}$, and $--^{CR}$ deletions in Case 1. Nonetheless, subsequent molecular analysis,

conducted to identify other Hb variants, revealed double heterozygosity for Hb Malay and α⁺-thalassemia (-α$^{3.7}$). This case highlighted the possibility of misdiagnosis without further molecular study. In the previous study, the heterozygosity of Hb Malay without α-thalassemia co-inheritance showed MCV and MCH with 72 fL and 23 pg respectively. However, these values increase up to 75 fL and 25 pg in double heterozygosity of Hb Malay and α⁺-thalassemia,[7] which was consistent with our Case 1. This phenomenon could be due to the reducing the imbalance of α- and β-globin ratio.[12]

In contrast, Case 2 showed severe anemia with marked microcytic-hypochromic erythrocytes. Although her Hb analysis did not indicate any risk of thalassemia, her HbA_2 level was slightly increased, raising the possibility of β-thalassemia trait or a Hb variant. On the other hand, she was 67 years old, and her iron profile showed significantly deficient (Table 1) indicating iron deficiency anemia. Previous study has shown that iron deficiency anemia could reduce the level of HbA_2, potentially masking β-thalassemia trait.[13] Although initial hematological findings did not strongly indicate thalassemia, molecular analysis confirmed heterozygosity for Hb Malay. This case highlights the diagnostic complexity when iron deficiency coexists with hemoglobinopathies. Previous studies have reported that patients with heterozygosity for Hb Malay are typically characterized by mild anemia or normal Hb level in some cases.[1,14] However, Case 2 exhibited severe anemia (Hb 6.4 gm/dL). Thus, severe anemia in Case 2 was primarily due to iron deficiency anemia. Despite severe iron deficiency anemia, her red blood cell (RBC) count was nearly normal. This phenomenon may be attributed to the coexistence of β-thalassemia trait and iron deficiency anemia as previously reported.[15]

As Hb Malay exhibits HbA-like electrophoretic behavior, its diagnosis using conventional electrophoresis methods is challenging. Negligently interpreting such samples as normal could lead to at-risk couples giving birth to affected offspring. For accurate diagnosis of Hb Malay, techniques such as reverse phase HPLC, M-ARMS PCR, allele-specific PCR, and reverse dot-blot ASO hybridization can be applied. Furthermore, the NGS offers a comprehensive and robust diagnostic tool.

Conclusion

The diagnosis of Hb Malay can be complex, especially when accompanied by concurrent thalassemia characteristics or iron deficiency, resulting in misleading hematological profiles without molecular confirmation. Our findings highlight the importance of integrating hematological and molecular analyses for accurate diagnosis of this Hb variant, especially in regions where thalassemia, hemoglobinopathy and iron deficiency are prevalent such as Northern Thailand. Timely and accurate identification is essential for effective clinical management, prevention strategies, and genetic counseling.

Ethical approval

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

Funding

Moe Theingi was supported by the CMU Presidential Scholarship from Chiang Mai University, Chiang Mai, Thailand, since academic year 2024. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

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

CRediT authorship contribution statement

Moe Theingi: manuscript preparation; **Chedtapak Ruengdit:** conducting the experiments; **Manoo Punyamung:** conducting the experiments; **Sakorn Pornpraset**: designed and conducted all of the experiments, manuscript preparation. All authors have read and approved of the final manuscript.

Acknowledgements

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

References

[1] Yang KG, Kutlar F, George E, Wilson JB, Kutlar A, Stoming TA, et al. Molecular characterization of β-globin gene mutations in Malay patients with Hb E-β-thalassemia and thalassemia major. Br J Haematol. 1989; 72(1): 73-80. doi: 10.1111/j.1365-2141.1989.tb07655.x.

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

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

[4] George E, Huisman TH, Yang KG, Kutlari F, Wilson JB, Kutlar A, et al. First observation of haemoglobin Malay alpha 2B2 26 (B1) Asn----Ser--a case report. Med J Malaysia. 1989; 44(3): 259-62.

[5] Yamsri S, Singha K, Prajantasen T, Taweenan W, Fucharoen G, Sanchaisuriya K, et al. A large cohort of β⁺-thalassemia in Thailand: Molecular, hematological and diagnostic considerations. Blood Cells Mol. Diseases. 2015; 54(2):164-9. doi: 10.1016/j.bcmd.2014.11.008.

[6] Bahar R, Shahida S N, Hassan mn, Marini R, mohd noor nh, Shafini M Y, et al. The diagnosis of beta thalassemia with borderline HbA2 level among

Kelantan population. J Blood Disord Transfus. 2017; 8: 5). doi: 10.4172/2155-9864.1000396.

[7] Tepakhan W, Kanjanaopas S, Sreworadechpisal K, Penglong T, Sripornsawan P, Wangchauy C, et al. Molecular epidemiology and hematological profiles of hemoglobin variants in southern Thailand. Sci Rep. 2024;14(1): 9255. doi: 10.1038/s41598-024-59987-4.

[8] Jomoui W, Satthakarn S, Panyasai S. Molecular understanding of unusual HbE-β(+)-thalassemia with Hb phenotype similar to HbE heterozygote: simple and rapid differentiation using HbE levels. Ann Med. 2023; 55(2): 2267054. doi: 10.1080/07853890.2023.2267054.

[9] Ruengdit C, Punyamung M, Intasai N, Pornprasert S. Single-tube multiplex real-time PCR with EvaGreen and high-resolution melting analysis for diagnosis of alpha0-thalassemia--^{SEA},--^{THAI}, and--^{CR} type deletions. PLoS One. 2023;18(11): e0293838. doi: 10.1371/journal.pone.0293838.

[10] Li R, Shen X, Chen H, Peng D, Wu R, Sun HY. Developmental validation of the MGIEasy Signature Identification Library Prep Kit, an all-in-one multiplex system for forensic applications. Int J Legal Med. 2021;135(3): 739-53. doi: 10.1007/s00414-021-02507-0.

[11] Al Hasan M, Nourin T. Co-inheritance of Alpha & Beta-Thalassemia: Case Report of a Bangladeshi Individual. Haematology Journal of Bangladesh. 2024; 8(1): 32-8. doi: 10.37545/haematoljbd2024117

[12] Winichagoon P, Fucharoen S, Chen P, Wasi P. Genetic Factors Affecting Clinical Severity in β-Thalassemia Syndromes. J Pediatr Hematol Oncol. 2000; 22(6): 573-80. doi: 10.1097/00043426-200011000-00026.

[13] Passarello C, Giambona A, Cannata M, Vinciguerra M, Renda D, Maggio A. Iron deficiency does not compromise the diagnosis of high HbA(2) β thalassemia trait. Haematologica. 2012; 97(3): 472-3. doi: 10.3324/haematol.2011.052357.

[14] Asri AS, Samsuddin MH, Jalil N, Mohamad Tahir N, Hashim H, Azma RZ, et al. Characterization of hemoglobin Malay phenotypes in tertiary hospitals. Hemoglobin. 2024; 48(3): 153-60. doi: 10.1080/03630269.2024.2380 873.

[15] Verma S, Gupta R, Kudesia M, Mathur A, Krishan G, Singh S. Coexisting iron deficiency anemia and beta thalassemia trait: effect of iron therapy on red cell parameters and hemoglobin subtypes. ISRN Hematol. 2014; 2014: 293216. doi: 10.1155/2014/293216.

Journal of Associated Medical Sciences 2026; 59(1): 87-97

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 reliability and validity of the Thai version of the Visual Cognitive Assessment Test (VCAT: Th)

Sutinun Juntorn[1*†], Thitiya Wangkawan[1†], Parlinzhaadaa Phatcharapisitgul[2], Peeradech Thichanpiang[1], Tharadon Rodkaeo[1], Paweena Sangsawee[1]

[1]Division of Occupational Therapy, Faculty of Physical Therapy, Mahidol University, Nakhon Pathom, Thailand.
[2]Department of Psychiatry, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 24 July 2025
Accepted as revised 17 September 2025
Available online 23 September 2025

Keywords:
Cognitive screening test, visual-based cognitive screening tool, mild cognitive impairment, mild Alzheimer disease, dementia.

ABSTRACT

Background: The global prevalence of dementia is projected to nearly double every 20 years, significantly impacting the quality of life for both individuals with dementia and their caregivers. Early detection is therefore critical, particularly in the context of Thailand's rapidly aging population. The Visual Cognitive Assessment Test (VCAT) is a visual-based cognitive screening tool designed to overcome educational limitations commonly associated with traditional assessments. Widely adopted across Southeast Asia, the VCAT demonstrates high sensitivity in detecting cognitive impairments from the earliest stages.

Objectives: This study aims to evaluate the validity and reliability of the Thai version of the Visual Cognitive Assessment (VCAT: Th)

Materials and methods: This study employed a cross-sectional, psychometric design and included 98 participants, consisting of 31 healthy controls (HC), 31 individuals with mild cognitive impairment (MCI), and 36 individuals with mild Alzheimer's disease (AD). The HC group was recruited from community-dwelling individuals residing in Bangkok and its surrounding metropolitan areas, while the MCI and mild AD groups were recruited from the Faculty of Medicine Ramathibodi Hospital, Mahidol University. All participants completed the Thai version of the Visual Cognitive Assessment Test (VCAT: Th), which was culturally adapted from the original version to suit the Thai context. The collected data was then used to evaluate internal consistency, test–retest reliability, concurrent validity, and known-group validity.

Results: The VCAT: Th demonstrated strong psychometric properties. Internal consistency was highly acceptable (α=0.78), and test-retest reliability was excellent (ICC=0.95). Concurrent validity showed a strong and significant correlation with MoCA scores (r=0.79,p<0.001). Known-group validity revealed significant group differences with a large effect size (η^2=0.53).

Conclusion: The VCAT: Th is a culturally adapted cognitive assessment tool that effectively identifies impairments typically present in the early stages of cognitive decline. It demonstrates acceptable psychometric properties and is appropriate for use within the Thai population.

† **These authors contributed equally to the work.**
* **Corresponding contributor.**
Author's Address: Division of Occupational Therapy, Faculty of Physical Therapy, Mahidol University, Nakhon Pathom, Thailand.
E-mail address: sutinun.jun@mahidol.ac.th
doi: 10.12982/JAMS.2026.012
E-ISSN: 2539-6056

Introduction

Dementia is a major contributor to dependency among the elderly population. According to the World Health Organization, an estimated 55 million individuals were living with dementia worldwide in 2024,[1] with

more than 60% residing in low- and middle-income countries. Each year, nearly 10 million new cases are identified globally, underscoring the increasing public health burden posed by dementia. According to the Diagnostic and Statistical Manual of Mental Disorders, 5th Ed. (DSM-5), dementia is characterized by a decline in one or more cognitive domains, including executive function, complex attention, language, learning and memory, perceptual-motor function, or social cognition. This cognitive impairment interferes significantly with an individual's ability to perform daily activities and occupational tasks, thereby diminishing overall quality of life. Individuals with dementia often face limitations in independent living due to these deficits. Previous studies have shown that individuals with dementia experience difficulties with scheduling appointments, locating personal items, remembering recent events, and following television programs. Additionally, they may struggle with managing finances, shopping, completing routine tasks, taking medications, maintaining household cleanliness, as well as driving and using public transportation.[2-4] Beyond the individuals affected, dementia imposes a significant burden on caregivers and family members, leading to broader social and economic consequences for both families and national healthcare systems. Managing cognitive decline and maintaining quality of life (QoL) are critical in dementia care.[5] Occupational therapists play a key role in early detection, which helps delay dementia onset and enables timely intervention. Occupational therapists contribute to the early identification of dementia by monitoring changes in individuals' functional abilities and occupational performance during everyday activities. Through careful observation and the use of standardized assessments, they can differentiate cognitive changes from pathological decline. These evaluations provide essential insights that support timely diagnosis and enable the initiation of appropriate interventions.[6]

In Thailand, the screening test for cognitive impairment such as the Montreal Cognitive Assessment (MoCA) and the Mini-Mental State Examination (MMSE) are commonly used. However, these screening test have notable limitations, particularly regarding educational dependence, as they require reading, spelling, and arithmetic skills. Furthermore, they are not suitable for individuals with hearing impairments, a common issue in the aging population. The Visual Cognitive Assessment Test (VCAT) addresses these limitations by using visually based tasks, making it less dependent on education and auditory skills. The VCAT originally was developed in Singapore by Kandiah,[7] and has been implemented in several Asian countries, including Malaysia, Indonesia, and the Philippines.[8] Its use has since extended beyond Southeast Asia, with ongoing validation studies being conducted in a range of international settings, such as Brazil, Canada, China, Korea, and India.[9] However, the VCAT remains a novel assessment tool and has not yet been widely recognized in Thailand. The original version of the VCAT also provides verbal instructions in English. Additionally, some of the pictures used in the assessment are not culturally relevant to the Thai context. Therefore, this study aims to examine the psychometric properties of the Thai version of the VCAT to ensure its applicability within the Thai context.

The Visual Cognitive Assessment Test (VCAT) is a visual-based cognitive screening tool designed for early detection of cognitive impairment. It evaluates five core cognitive domains: episodic memory, attention and working memory, executive function, visuospatial ability, and language. All components of the assessment are delivered through an image-based format, with visual stimuli specifically selected and validated for clarity and recognize ability in older adult populations.[7] The VCAT has demonstrated high accuracy in differentiating individuals with cognitive impairment from cognitively healthy controls, as well as in distinguishing between mild cognitive impairment (MCI) and mild Alzheimer's disease (AD). Compared with the Mini-Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA), the VCAT demonstrates superior discriminative capacity and performs in discriminating between healthy people, MCI, particularly in identifying early-stage dementia. Its diagnostic utility is especially evident in contexts requiring precise detection during the early phases of cognitive decline.[10]

At present, the Visual Cognitive Assessment Test (VCAT) has not been implemented in Thailand, primarily due to certain test items being culturally inappropriate. In response to this limitation, the present study aims to evaluate the validity and reliability of the Thai version of the Visual Cognitive Assessment Test (VCAT: Th), which has been culturally adapted to align with the sociocultural characteristics of the Thai population.

Materials and methods

Design and participation

This study employed a cross-sectional, psychometric design and included a total of 98 participants, comprising 31 healthy controls (HC), 36 individuals with mild cognitive impairment (MCI), and 31 individuals with mild Alzheimer's disease (AD). The sample size was calculated using the G*Power program. The effect size (Cohen's d) was derived from the mean and standard deviation reported in a comparative study of Visual Cognitive Assessment Test (VCAT) scores between normal participants and those with cognitive impairment.[11] The resulting effect size was d=0.8. By setting the significance level (α error probability) at 0.05 and statistical power (1-β error probability) at 0.80, the required sample size was determined to be 26 participants per group. To account for an anticipated dropout rate of 20%, the adjusted sample size was increased to approximately 31 participants per group. In addition, five

participants with MCI were recruited to participate in the pilot testing phase of the assessment. Inclusion criteria for all groups required participants to be aged 50 years or older, no significant visual or motor impairments, Thai nationality, and proficient in Thai language skills in listening, reading, and speaking. Participants in the HC group were recruited from community-dwelling individuals residing in Bangkok and its surrounding metropolitan areas. Eligibility for this group required a MoCA score of 25 or above. While individuals in the MCI and mild AD groups were recruited from the Faculty of Medicine Ramathibodi Hospital, Mahidol University. For the MCI group, inclusion was based on a clinical diagnosis of mild cognitive impairment made by cognitive neurologists in accordance with the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5).[12] For the mild AD group, inclusion was based on a diagnosis of mild Alzheimer's disease following the criteria established by the National Institute on Aging and the Alzheimer's Association (NIA-AA).[13,14] All participants were diagnosed by psychiatrists in a clinical setting. The demographics characteristics of participants are summarized in Table 3.

Procedures

Cultural adaptation and psychometric properties studies

The adaptation of the original Visual Cognitive Assessment Test (VCAT) into the Thai version involved a four-step process. The first step focused on modifying test content that was culturally inappropriate with the Thai context. Although the original version of the Visual Cognitive Assessment Test (VCAT) has been implemented in several countries,[7,9,15] where cultural and lifestyle contexts are generally similar, certain contextual differences exist in Thailand particularly regarding food culture. As such, specific modifications to food-related test items were necessary to ensure cultural appropriateness and relevance for Thai people. To address this, the research team obtained formal permission from the original VCAT developers to make the required content adaptations for the Thai version, ensuring alignment with the sociocultural context of Thailand. One key modification was made to the Delayed Memory: Objects item. Specifically, the image of a hamburger was replaced with grilled chicken, and the accompanying cue images of French fries and cola in the test booklet were substituted with somtam (papaya salad) and sticky rice -foods that are widely recognized among the Thai population. In this method, the research team adapted the pictures to align with the sociocultural context of Thailand prior to expert evaluation. These modifications are detailed in Table 1.

The second step involved evaluating content validity. In this phase, three experts independently assessed each item using the Index of Item-Objective Congruence (IOC) method.[16] All experts had more than ten years of professional experience, comprising two occupational therapists specializing in cognition and one professional psychiatric nurse. The results demonstrated a high level of agreement among the experts, with all items receiving an IOC score of 1.00. (Table 2)

Table 1. The content modifications of the Visual Cognitive Assessment Test (VCAT) to align with the Thai cultural context.

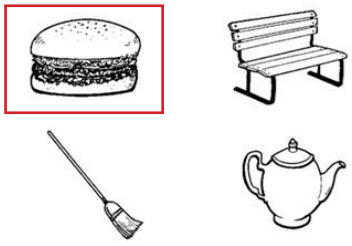
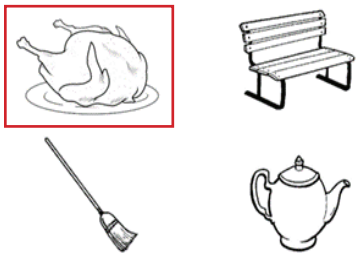
Original version	Culturally adapted version for the Thai context
Delayed memory: objects Visual stimuli for assessment	
	
Cue images in test booklet	
	

Table 2. Index of Item–Objective Congruence (IOC) for the items and scoring levels of the VCAT: Th.

VCAT-Th items	IOC
Memory-Shapes	1
Memory-Objects	1
Memory-Scene	1*
Attention-Shape cancellation	1
Executive function-Gears	1
Executive function-Pattern	1
Executive function-Category	1
Visuospatial-Cube	1
Visuospatial-Grid	1
Language-Picture naming	1
Language-Semantic fluency	1

**Note: The IOC value was obtained after revision according to the experts' recommendations.*

The third step consisted of a pilot trial of the VCAT: Th with a sample of five participants clinically diagnosed with mild cognitive impairment (MCI). Feedback from this trial was used to further refine and adjust the assessment. This process resulted in the finalized version of the VCAT: Th.

The final step was to examine the psychometric properties of the VCAT: Th, including (1) internal consistency, (2) test–retest reliability, (3) concurrent validity, and (4) known-group validity.

Psychometric evaluation

The psychometric properties of the Thai version of the Visual Cognitive Assessment Test (VCAT: Th) were examined using data from 93 participants, comprising 31 healthy controls (HC), 31 individuals with mild cognitive impairment (MCI), and 31 individuals with mild Alzheimer's disease (mild AD). Internal consistency was evaluated using the all-item scores of VCAT: Th derived from all participants.

Test-retest reliability was assessed exclusively in the MCI group. This decision was based on the primary purpose of the VCAT: Th, which was developed as a screening tool for individuals with early-stage cognitive impairment. Administering the test-retest procedure in the healthy control group would likely result in a ceiling effect, as most individuals in this group are expected to achieve maximum or near-maximum scores. Such score clustering at the upper limit would eliminate meaningful variance, thereby precluding an accurate estimation of reliability. Conversely, administering the test–retest in the mild AD group may yield artificially low reliability estimates, given the inherent clinical instability and day-to-day fluctuations in cognitive performance associated with the disease. Thus, the MCI group was deemed the most appropriate population in which to evaluate test-retest reliability. Participants in this group completed the VCAT: Th twice, with a one-week interval between assessments. Intraclass correlation coefficients (ICCs) were calculated using total, subdomain and item scores from both time points to determine stability over time.

To evaluate concurrent validity, the MoCA was administered to participants in the MCI and mild AD groups. The total scores from the MoCA and VCAT: Th were then compared using correlation analysis to determine the degree of association between the two assessments.

To assess known-groups validity, VCAT: Th scores were statistically compared across the three diagnostic groups-healthy controls (HC), mild cognitive impairment (MCI), and mild Alzheimer's disease (AD)- to evaluate the instrument's ability to distinguish between differing levels of cognitive impairment.

Measurements

The Visual Cognitive Assessment Test (VCAT) is a visually mediated cognitive screening instrument developed to facilitate the early identification of cognitive decline. Made up of 11 items, the VCAT is designed to be completed in approximately 15.7±7.3 minutes. It yields a total score ranging from 0 to 30, with higher scores indicating better cognitive performance. The development of the VCAT was informed by widely accepted diagnostic criteria for mild cognitive impairment (MCI) and Alzheimer's disease (AD), with the test structured to assess five core cognitive domains: episodic memory, attention and working memory, executive function, visuospatial abilities, and language.[7] Notably, the assessment places particular emphasis on episodic memory, given its strong association with progression to AD, thereby enhancing the tool's sensitivity in early detection of the disease. In addition, greater weight is allocated to executive

functioning to improve the identification of cognitive profiles associated with vascular dementia.[10]

The VCAT demonstrated strong discriminative ability in distinguishing individuals with cognitive impairment from those without, yielding an area under the curve (AUC) of 85.5%. In comparison, the MoCA and MMSE produced AUCs of 85.3% and 79.5%, respectively. Furthermore, the VCAT effectively differentiated between individuals with MCI and those with mild AD, with an AUC of 84.2%, outperforming the MoCA (AUC=81.8%) and the MMSE (AUC=76.5%). Based on these findings, VCAT cut-off scores were established as follows: 23-30 for normal cognition, 18-22 for MCI, and 0-17 for dementia.[10]

Data analysis

Data analysis for this study was conducted using SPSS for Windows, version 23. Continuous variables were first assessed for normality using the Shapiro-Wilk test. For variables that followed a normal distribution, such as age, between-group comparisons were conducted using one-way analysis of variance (ANOVA). For non-normally distributed continuous variables, including MoCA scores, the Kruskal-Wallis test was applied. Categorical variables such as gender and education level were analyzed using chi-square (χ^2) tests of independence.

Internal consistency of the VCAT: Th was evaluated using Cronbach's alpha coefficient. Test-retest reliability was assessed using the Intraclass Correlation Coefficient (ICC), based on the ICC (3,1) model (two-way mixed-effects, single measurement).[17] Given that the Shapiro-Wilk test indicated non-normal distribution for the VCAT scores (p<0.05), Spearman's rank correlation coefficient was used to assess concurrent validity with the MoCA. While known-groups validity was evaluated through the Kruskal-Wallis test.

Ethical considerations

This study received ethical approval from the Human Research Ethics Committee of Mahidol University, under the reference number COA No. MU-MOU 2023/176.0612, granted in December 2023. Participant rights were rigorously safeguarded throughout the research process, and all collected data were treated with strict confidentiality. Personal identifiers were excluded from all analyses; participant responses were fully anonymized and utilized solely in aggregated form for the purpose of this study.

Results

The sample included 93 participants. No significant gender differences were found between groups, though females were more prevalent in the HC group (67.70%), the MCI group (77.40%), and the mild AD group (67.70%). In terms of age, participants in the mild AD group were significantly older (mean=77.90, SD=7.12) than those in the HC (mean=58.36, SD=6.33) and MCI groups (mean=73.81, SD=6.31). Most participants in the HC and MCI groups held a bachelor's degree, whereas the majority in the mild AD group had completed a high vocational certificate. The three diagnostic groups showed significant differences in total MoCA scores, with the highest scores observed in the HC group (mean=26.29, SD=1.13), followed by the MCI group (mean=23.55, SD=3.95) and then the mild AD group (mean=15.87, SD=5.62). (Table 3)

Table 3. Demographic characteristics by group (N=93).

Characteristic or Test	HC (N=31)	MCI (N=31)	Mild AD (N=31)	p value
	N (%)	N (%)	N (%)	
Gender[1]				0.625
Male	10 (32.30)	7 (22.60)	10 (32.30)	
Female	21 (67.70)	24 (77.40)	21 (67.70)	
Education[1]				<0.001
Elementary school	5 (16.10)	9 (29.00)	3 (9.70)	
Junior high school	2 (6.50)	2 (6.50)	5 (16.10)	
Senior high school	3 (9.70)	5 (16.10)	1 (3.20)	
Vocational certificate	0 (0.00)	0 (0.00)	1 (3.20)	
High vocational certificate	2 (6.50)	0 (0.00)	10 (32.30)	
Bachelor Degree	14 (45.20)	11 (35.50)	3 (9.70)	
Master Degree	5 (16.10)	4 (12.90)	1 (3.20)	

Table 3. Demographic characteristics by group (N=93). (Continue)

Characteristic or Test	HC (N=31)	MCI (N=31)	Mild AD (N=31)	p value
	N (%)	N (%)	N (%)	
	M (SD)	M (SD)	M (SD)	
Age[2] (years)	58.36 (6.33)	73.81(6.31)	77.90 (7.12)	<0.001
MoCA[3]	26.29 (1.13)	23.55 (3.95)	15.87 (5.62)	<0.001

Note: [1]: χ2 test, [2]: ANOVA, [3]: Kruskal–Wallis test.

The internal consistency of the VCAT: Th was found to be highly acceptable (α=0.78)[18]. Test-retest reliability of the VCAT: Th was evaluated using intraclass correlation coefficients (ICCs) calculated for the total score, subdomain scores, and individual item scores. The total score demonstrated excellent reliability[19], with excellent ICC values for the total VCAT-Th score (ICC= 0.95). Subdomain scores also showed good reliability, with ICCs ranging from moderate to excellent values (ICC = 0.64-0.89): 0.89 for episodic memory, 0.76 for attention/working memory, 0.88 for executive function, 0.88 for visuospatial function, and 0.64 for language. At the item level, ICCs similarly indicated good reliability, reflecting moderate to excellent agreement across repeated administrations (ICC=0.56-0.95). (Table 4)

Table 4. Test-retest reliability of VCAT:Th as demonstrated by intraclass correlation coefficients (ICC) and 95% confidence interval (CI) (N=31).

	ICC (3,1) (95% CI)	*p* value
VCAT-Th total score	0.95	0.000
VCAT-Th subdomains		
Episodic Memory	0.89	0.000
Attention/Working Memory	0.76	0.000
Executive Function	0.88	0.000
Visuospatial	0.88	0.000
Language	0.64	0.004
VCAT-Th items		
Memory-Shapes	0.81	0.000
Memory-Objects	0.88	0.000
Memory-Scene	0.56	0.013
Attention-Shape cancellation	0.76	0.000
Executive function-Gears	0.81	0.000
Executive function-Pattern	0.76	0.000
Executive function-Category	0.86	0.000
Visuospatial-Cube	0.63	0.004
Visuospatial-Grid	0.85	0.000
Language-Picture naming	0.80	0.000
Language-Semantic Fluency	0.95	0.000

Note: VCAT:Th: Visual Cognition Assessment Test: Thai version.

Concurrent validity

Spearman correlation coefficients between the total VCAT: Th scores and the total MoCA scores reported significantly strong correlations (r=0.79, p <0.001)[20]. These results provide for the hypothesis regarding the concurrent validity of VCAT: Th.

Known-group validity

The results of the Kruskal-Wallis test assessing known-group validity and effect sizes are presented in Table 5. The total VCAT-Th score demonstrated a significant difference among groups, with a large effect size (η^2=0.53).[21] Subdomain scores also showed significant differences, with effect sizes ranging from small to large (η^2 = 0.16-0.54). At the item level, most VCAT-Th items showed significant differences with small to large effect sizes (η^2=0.10-0.56), except for the Executive Function-Gears item, which did not reach statistical significance.

Table 5. Kruskal-Wallis test and effect size for known-group validity (N=93).

		Range	M (SD)			Chi square (χ^2)	Effect size (η^2)
			HC (N=31)	MCI (N=31)	Mild AD (N=31)		
VCAT:Th	total score	0-30	26.06 (2.66)	22.39 (4.49)	13.36 (6.22)	48.82***	0.53
VCAT:Th	subdomains						
	Episodic Memory	0-13	11.61 (1.61)	10.13 (2.32)	4.90 (3.10)	49.80***	0.54
	Attention/WM	0-3	2.58 (0.99)	1.90 (1.35)	0.77 (1.26)	26.22***	0.29
	Executive Function	0-6	4.52 (1.26)	3.81 (1.74)	2.45 (1.59)	21.64***	0.24
	Visuospatial	0-3	2.48 (0.63)	2.13 (0.89)	1.55 (1.06)	14.53***	0.16
	Language	0-5	4.84 (0.37)	4.42 (0.89)	3.36 (1.02)	35.79***	0.39
VCAT:Th items							
	Memory-Shapes	0-2	1.13 (0.99)	0.77 (0.81)	0.29 (0.59)	13.15***	0.14
	Memory-Objects	0-8	7.48 (1.06)	6.39 (1.98)	2.55 (2.14)	51.66***	0.56
	Memory-Scene	0-3	3.00 (0.00)	2.97 (0.18)	2.07 (0.96)	40.26***	0.44
	Attention-Shape cancellation	0-3	1.13 (0.99)	0.77 (0.81)	0.29 (0.59)	26.22***	0.29
	Executive function-Gears	0-3	1.87 (1.23)	1.68 (1.19)	1.65 (1.08)	0.557	0.01
	Executive function-Pattern	0-2	1.65 (0.49)	1.42 (0.67)	0.48 (0.72)	33.76***	0.37
	Executive function-Category	0-1	1.00 (0.00)	0.71 (0.46)	0.65 (0.49)	12.98**	0.14
	Visuospatial-Cube	0-1	0.61 (0.50)	0.39 (0.50)	0.19 (0.40)	11.28**	0.12
	Visuospatial-Grid	0-2	1.87 (0.50)	1.74 (0.68)	1.36 (0.92)	9.09**	0.10
	Language-Picture naming	0-3	2.97 (0.18)	2.94 (0.25)	2.68 (0.54)	10.91**	0.12
	Language-Semantic Fluency	0-2	1.87 (0.34)	1.48 (0.77)	0.68 (0.87)	30.03***	0.33

*Note: **p<0.01, ***p<0.001*

Discussion

This study aimed to examine the validity and reliability of the Thai version of the Visual Cognitive Assessment Test (VCAT: Th), which was culturally modified to suit the Thai population. The findings demonstrated excellent content validity, with an IOC of 1.00, indicating that the assessment effectively measures cognitive abilities as intended and is appropriate for use in the Thai context. This strong alignment may be attributed to the fact that individuals with dementia or MCI commonly exhibit pronounced deficits in domains such as visuospatial skills, attention, delayed recall, language, and executive function-all of which are core cognitive components of the VCAT.

In the process of cultural adaptation and psychometric evaluation, modifications were applied to the *Delayed Memory: Objects* item. Specifically, the image of a hamburger was replaced with grilled chicken, while the accompanying cue images of French fries and cola in the test booklet were substituted with somtam (papaya salad) and sticky rice, which are more culturally familiar to the Thai population. Testing instruments developed predominantly within Western paradigms frequently overlook cultural nuances, variations in educational background, and differences in language use, all of which may substantially influence test outcomes and compromise the accuracy of diagnostic conclusions. This challenge carries serious consequences, as reliance on instruments that neglect cultural factors may result in the erroneous diagnosis of intact patients with cognitive disorders, while simultaneously failing to identify individuals with actual cognitive impairment.[22]

The internal consistency of the VCAT: Th, indicated by a Cronbach's alpha coefficient of 0.78, reflects a good level of reliability. This finding suggests that the test items are sufficiently correlated and consistently measure the underlying cognitive constructs. These results align with those of Low *et al.*,[10] who reported a comparable internal consistency (α=0.74) for the original version of the VCAT in a study involving healthy controls, individuals with MCI, and patients with mild AD.

The cognitive domains assessed by the VCAT-visuospatial ability, attention, delayed recall, language, and executive function-are known to be among the first areas affected in individuals with MCI. Prior research has consistently shown that individuals with MCI demonstrate significantly lower performance in visual attention and visuospatial tasks compared to cognitively healthy individuals.[3,23] Moreover, previous studies examining the risk of Alzheimer's disease have demonstrated that individuals with MCI perform significantly worse on visual memory tasks compared to cognitively normal individuals. These findings suggest not only an increased risk of progression to Alzheimer's disease in the MCI group but also provides clear evidence of visual memory deficits associated with early cognitive decline.[24,25] Language difficulties have also been documented in previous studies involving individuals with dementia or MCI. These studies indicated that such individuals often exhibit early impairments in delayed recall and language-related functions, including word retrieval, accurate word usage, syntactic construction, and the ability to convey clear and coherent information.[26] In addition, executive function deficits have been identified as a prominent feature in individuals with MCI. Research has shown that impairments in executive functioning are frequently observed in this population.[27,28] These deficits can substantially affect an individual's ability to manage daily activities and maintain independence.[29] Accordingly, each item within the VCAT: Th has been carefully designed to target specific cognitive domains commonly affected in individuals with MCI, ensuring the tool's sensitivity in detecting early-stage cognitive impairments.

The findings from the test–retest reliability analysis demonstrated good reliability, indicating that the VCAT: Th produces stable and consistent results over time. The moderate to excellent levels of agreement across repeated administrations suggest that the instrument can capture cognitive performance reliably, even when reassessed after a short interval. These results confirm that the VCAT: Th is a standardized tool for assessing cognitive abilities.

The results of the concurrent validity analysis demonstrated a strong and statistically significant correlation between the total scores of the VCAT: Th and the MoCA, indicating that both instruments assess similar cognitive constructs. This finding supports the convergent validity of the VCAT: Th and affirms its capability to evaluate global cognitive function in a manner consistent with internationally standardized assessment tools. These results are in line with the study by Low *et al.*,[10] which also reported a strong correlation between the original VCAT and MoCA scores (r=0.81), further affirming the validity of the VCAT across diverse cultural settings.

The findings from the known-groups validity analysis revealed that both the total score and subdomain scores of the VCAT: Th was significantly different between diagnostic groups. At the individual item level, most items showed statistically significant differences, except for the Executive Function-Gears item, which did not reach significance. This lack of difference may be attributable to the dichotomous response format of the item, consisting of only two options (i.e., an arrow rotating to the left or to the right), which could have increased the likelihood of chance-level responding among participants with MCI and mild AD. Consequently, no statistically significant differences in scores were observed across the three groups. These results align with previous research by Low and colleagues,[10] who examined the construct validity of the original VCAT in healthy individuals,

MCI, and mild AD. The results indicated that the VCAT demonstrated superior discriminative ability compared to the MMSE and performed comparably to the MoCA in differentiating between individuals with mild AD and MCI, as well as distinguishing cognitively healthy individuals from those with cognitive impairment.

The VCAT was designed to assess multiple domains of cognitive functioning, with particular emphasis on memory and executive function. This targeted focus provides a key advantage over other cognitive screening tools, particularly in the detection of Alzheimer's disease (AD) and vascular dementia (VaD), the two most prevalent types of dementia.[7] Prior research has indicated that older adults, irrespective of amyloid-beta accumulation, consistently demonstrate significantly lower performance on assessments of visuospatial memory and executive functioning compared to younger adults. These findings suggest that aging is associated with a general decline in brain function, especially in regions such as the prefrontal cortex and medial temporal lobe, which play critical roles in memory and learning.[30]

Limitations

This study has several limitations that should be acknowledged. First, there was a gender imbalance in the sample, with a greater proportion of female participants compared to males. Second, participants with MCI and mild AD were recruited exclusively from a single institution located at the Faculty of Medicine Ramathibodi Hospital, Mahidol University, which may limit the generalizability of the findings to broader clinical populations. Third, there were differences in the distribution of educational levels across the diagnostic groups, which may have influenced cognitive performance and affected the interpretation of group comparisons.

Conclusion

The findings of this study support that the VCAT: Th is a cognitive assessment tool capable of detecting cognitive impairments commonly observed from the early stages of cognitive impairment. It is culturally appropriate for use in the Thai context and demonstrates acceptable psychometric properties. Furthermore, the VCAT: Th functions as a standardized screening instrument, exhibiting diagnostic accuracy comparable to that of internationally recognized cognitive assessment tools.

Ethical approval

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki and complied with all applicable institutional and national regulations governing research involving human participants. Ethical approval was granted by the Human Research Ethics Committee of Mahidol University (COA No. MU-MOU 2023/176.0612) prior to data collection. All participants received detailed information regarding the study's objectives and procedures, and written informed consent was obtained before enrollment. Participation was entirely voluntary, and participants retained the right to withdraw from the study at any point without consequence.

Funding

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

Conflict of interest

The authors declare no conflicts of interest related to the research, authorship, or publication of this article.

CRediT authorship contribution statement

Sutinun Juntorn: conceptualization, methodology, data curation, formal analysis, writing: original draft, supervision, project administration, correspondence; **Thitiya Wangkawan**: methodology, data curation, investigation, writing: review and editing; **Parlinzhaadaa Phatcharapisitgul**: data curation, validation; **Peeradech Thichanpiang**: supervision, writing: review and editing; **Tharadon Rodkaeo**: data curation; **Paweena Sangsawee**: data curation.

Future direction and clinical implementation

Future research on the VCAT: Th should extend its application to populations with hearing impairments. Given that traditional cognitive screening tools often rely heavily on verbal instructions, individuals who are deaf or hard of hearing may be systematically disadvantaged, resulting in inaccurate diagnostic outcomes. The VCAT, as a visually based assessment, offers unique potential for reducing such barriers and ensuring more equitable cognitive evaluation. Furthermore, the VCAT: Th provides occupational therapists and clinicians with a culturally adapted, visually based tool for the screening and assessment of cognitive function. Finally, use of the VCAT: Th in clinical settings may facilitate interdisciplinary collaboration, where results can be incorporated into comprehensive evaluations of cognitive function that inform medical, psychological, and occupational therapy interventions.

Acknowledgements

The authors would like to express their sincere gratitude to all participants for their valuable cooperation in this study. Appreciation is also extended to the Faculty of Medicine Ramathibodi Hospital, Mahidol University, for the support and facilitation provided during the data collection process.

Reference

[1] World Health Organization. Dementia in refugees and migrants: epidemiology, public health implications

and global responses. Geneva: World Health Organization; 2024.

[2] Aretouli E, Brandt J. Everyday functioning in mild cognitive impairment and its relationship with executive cognition. Int J Geriatr Psychiatry. 2010; 25(3): 224-33.doi: 10.1002/gps.2325.

[3] Chang CW, Su KC, Lu FC, Cheng HM, Cheng CY. Visual function and visual perception among senior citizens with mild cognitive impairment in Taiwan. Healthcare (Basel). 2021; 10(1): 20. doi: 10.3390/healthcare10010020.

[4] Feng YR, Meuleners L, Stevenson M, Heyworth J, Murray K, Fraser M*, et al*. The impact of cognition and gender on speeding behaviour in older drivers with and without suspected mild cognitive impairment. NPJ Aging. 2021; 16: 1473-83.doi: 10.2147/cia.S319129.

[5] Bennett S, Laver K, Voigt-Radloff S, Letts L, Clemson L, Graff M*, et al*. Occupational therapy for people with dementia and their family carers provided at home: a systematic review and meta-analysis. BMJ open. 2019; 9(11): e026308. doi: 10.1136/bmjopen-2018-026308.

[6] Graff MJL, Vernooij-Dassen MJM, Thijssen M, et al. Community based occupational therapy for patients with dementia and their care givers: randomised controlled trial. BMJ. 2006; 333(7580): 1196. doi: 10.1136/bmj.39001.688843.BE.

[7] Kandiah N, Zhang A, Bautista DC, Silva E, Ting SKS, Ng A*, et al*. Early detection of dementia in multilingual populations: Visual Cognitive Assessment Test (VCAT). J Neurol Neurosurg Psychiatry. 2016; 87(2): 156-60. doi: 10.1136/jnnp-2014-309647.

[8] Lim L, Ng TP, Ong AP, Tan MP, Cenina AR, Gao Q*, et al*. A novel language-neutral Visual Cognitive Assessment Test (VCAT): validation in four Southeast Asian countries. Alzheimers Res Ther. 2018;10(1): 6. doi: 10.1186/s13195-017-0333-z.

[9] Ng KP, Ong GCF, Saffari SE, Bernardes MM, de Carvalho AB, Schilling LP*, et al*. Visual Cognitive Assessment Test (VCAT): A language-neutral test to detect mild cognitive impairment and dementia in multinational cohorts. Alzheimers Dement. 2024; 20(S3): e090328. doi: 10.1002/alz.090328.

[10] Low A, Lim L, Lim L, Wong B, Silva E, Ng KP*, et al*. Construct validity of the Visual Cognitive Assessment Test (VCAT)-a cross-cultural language-neutral cognitive screening tool. Int Psychogeriatr. 2020; 32(1): 141-9. doi: 10.1017/s1041610219000504.

[11] Ng Y, Chin CJ, Danial M, Albart S, Suppiah PD, Ganasegeran K,*et al*. Validation of the Visual Cognitive Assessment Test (VCAT) for the early diagnosis of cognitive impairment in multilingual population in Malaysia. Psych. 2022; 4(1): 38-48. doi: 10.3390/psych4010003.

[12] American Psychiatric Association. Diagnostic and statistical manual of mental disorder [cited 2024 December 11]. Available from: https://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425596.

[13] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, *et al*. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011; 7(3): 270-9. doi: 10.1016/j.jalz.2011.03.008.

[14] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Jr., Kawas CH*, et al*. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011; 7(3): 263-9. doi: 10.1016/j.jalz.2011.03.005.

[15] Lim L, Ng TP, Ong AP, Tan MP, Cenina AR, Gao Q*, et al*. A novel language-neutral Visual Cognitive Assessment Test (VCAT): validation in four Southeast Asian countries. Alzheimers Res Ther. 2018;10(1):6. doi: 10.1186/s13195-017-0333-z.

[16] Waltz CF, Strickland OL, Lenz ER. Measurement in nursing and health research. (5th Ed.). New York: Springer Publishing Company; 2016.

[17] Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull. 1979; 86(2): 420-8. doi: 10.1037/0033-2909.86.2.420.

[18] Taber KS. The use of Cronbach's Alpha when developing and reporting research instruments in science education. Res Sci Educ. 2018; 48(6): 1273-96.doi: 10.1007/s11165-016-9602-2.

[19] Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016; 15(2): 155-63.doi: https://doi.org/10.1016/j.jcm.2016.02.012.

[20] Prion S, Haerling KA. Making sense of methods and measurement: Spearman-Rho ranked-order correlation coefficient. Clin Simul Nurs. 2014; 10(10): 535-6.doi: 10.1016/j.ecns.2014.07.005.

[21] Heesch KC, Mâsse LC. Lack of time for physical activity: perception or reality for African American and Hispanic women? Women Health. 2004; 39(3): 45-62. doi: 10.1300/J013v39n03_04.

[22] Daga M, Mohanty P, Krishna R, Priya S. Statistical validation in cultural adaptations of cognitive tests: a multi-regional systematic review. arXiv [Preprint]. 2025. doi:https://doi.org/10.48550/arXiv.2504.13495.

[23] Frtusova J, Rivest J, Doody A, Leach L. Selective and divided visual attention in mild cognitive impairment as measured with the Baycrest Attention Test and Traditional Neuropsychological Tests. Alzheimers Dement. 2010; 6: e30-e30. doi: 10.1016/j.jalz.2010.08.095.

[24] Alescio-Lautier B, Michel BF, Herrera C, Elahmadi A, Chambon C, Touzet C*, et al*. Visual and visuospatial

short-term memory in mild cognitive impairment and Alzheimer disease: Role of attention. Neuropsychologia. 2007; 45(8): 1948-60.doi: https://doi.org/10.1016/j.neuropsychologia.2006.04.033.

[25] Oltra-Cucarella J, Sánchez-SanSegundo M, Lipnicki DM, Crawford JD, Lipton RB, Katz MJ*, et al*. Visual memory tests enhance the identification of amnestic MCI cases at greater risk of Alzheimer's disease. Int Psychogeriatr. 2019; 31(7): 997-1006. doi: 10.1017/s104161021800145x.

[26] Mueller KD, Hermann B, Mecollari J, Turkstra LS. Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. J Clin Exp Neuropsychol. 2018; 40(9): 917-39.doi: 10.1080/13803395.2018.1446513.

[27] Jung YH, Park S, Jang H, Cho SH, Kim SJ, Kim JP*, et al*. Frontal-executive dysfunction affects dementia conversion in patients with amnestic mild cognitive impairment. Sci Rep. 2020; 10(1): 772. doi: 10.1038/s41598-020-57525-6.

[28] Reinvang I, Grambaite R, Espeseth T. Executive dysfunction in MCI: Subtype or early symptom. Int J Alzheimers Dis. 2012; 2012 :936272. doi: 10.1155/2012/936272.

[29] Brandt J, Aretouli E, Neijstrom E, Samek J, Manning K, Albert MS*, et al*. Selectivity of executive function deficits in mild cognitive impairment. Neuropsychology. 2009; 23(5): 607-18. doi: 10.1037/a0015851.

[30] Oh H, Madison C, Haight TJ, Markley C, Jagust WJ. Effects of age and β-amyloid on cognitive changes in normal elderly people. Neurobiol Aging. 2012; 33(12): 2746-55. doi: 10.1016/j.neurobiolaging.2012.02.008.

Journal of Associated Medical Sciences 2026; 59(1): 98-108

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

Cross-cultural adaptation and psychometric evaluation of the Communication Function Classification System-Thai version for individuals with cerebral palsy in Thailand

Pim Chaisook[1], Phuanjai Rattakorn[1*], Supaporn Chinchai[1], Wannipa Bunrayong[2]

[1]Communication Sciences and Disorders Division, Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 13 March 2025
Accepted as revised 19 September 2025
Available online 30 September 2025

Keywords:
Cerebral palsy, classification, communication, psychometrics, speech-language pathologists.

ABSTRACT

Background: Cerebral palsy (CP) is a neurodevelopmental disorder that affects motor functions and is often accompanied by impairments such as speech and communication challenges. Functional classification systems are increasingly used to describe daily activities and participation, aiming to guide comprehensive and effective treatment planning that improves patients' quality of life. The Communication Function Classification System (CFCS) is a tool designed to describe communication functions in the daily lives of individuals with CP. Nonetheless, the absence of a Thai version of the CFCS highlighted the need for this study.

Objectives: This study aimed to translate and culturally adapt the CFCS into Thai (CFCS-TH) and to evaluate its psychometric properties, including validity and reliability, in individuals with CP in Thailand.

Materials and methods: The study was conducted in two phases: (1) cross-cultural adaptation and translation of the CFCS into Thai using a six-step forward and backward translation method, and (2) psychometric evaluation of the CFCS-TH, focusing on content validity, inter-rater reliability, and intra-rater reliability. Validity was assessed using the content validity index (CVI), while reliability was measured using weighted kappa statistics (κ). A total of 35 individuals with CP, aged 2-18 years, participated in the study. The raters included one speech-language pathologist (SLP), two physical therapists (PTs), and 35 parents, each of whom rated the level of communication function of the individuals with CP twice.

Results: The cross-cultural adaptation of the CFCS to Thai was carried out by ensuring that the language and terminology were appropriate for Thai users while maintaining the conceptual integrity and purpose of the tool. The CFCS-TH translation steps included vocabulary and syntax selection, equivalency review, and revisions from translators, an expert committee, and pretest users, followed by approval from the instrument developers. The CFCS-TH demonstrated excellent content validity, with an item-level CVI and scale-level CVI of 1.0, indicating that the CFCS-TH could measure its intended construct. Inter-rater reliability was good between the SLP and PTs (κ=0.71), good between the SLP and parents (κ=0.66), and fair between PTs and parents (κ=0.55). When separating the data by the first and second classification rounds, agreement between the SLP and PTs remained good with a slight decrease (κ=0.73 and 0.69). The SLP-parent agreement remained consistently good across both rounds (κ =0.61 and 0.66), while

*** 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: phuanjai.rattakorn@cmu.ac.th
doi: 10.12982/JAMS.2026.013
E-ISSN: 2539-6056

the PTs-parents agreement improved from fair to good (k=0.48 and 0.62). Intra-rater reliability was excellent for the SLP (k=0.92) and very good for PTs (k = 0.91) and parents (k=0.86).

Conclusion: The CFCS-TH is a valid and reliable tool for describing the communication functions of individuals with CP in Thailand. This tool provides a standardized framework for assessing communication performance to support clinical and research efforts aimed at intervention planning and improving the quality of life for individuals with CP.

Introduction

Cerebral palsy (CP) is a neurodevelopmental disorder caused by brain injuries or lesions during development, leading to abnormalities in motor function, movement, and muscle tone. Individuals with CP often experience co-occurring impairments, including challenges in speech, language, and communication. These challenges may stem from dysarthria and other comorbidities, such as epilepsy, intellectual disabilities, and sensory impairments, which can vary in severity, ranging from mild speech difficulties to an inability to speak.[1] However, individuals with CP can still communicate through alternative means of communication that supplement or replace spoken language.[2,3]

Modern medicine emphasizes a holistic approach to healthcare, prioritizing comprehensive care that enhances overall quality of life. The International Classification of Functioning, Disability, and Health (ICF), developed by the World Health Organization (WHO), provides a framework that categorizes individual health status or disease into three components: body structures and functions, activities, and participation. This framework has led to the widespread adoption of functional classification systems.[4, 5] For individuals with CP, classification systems often focus on communication, manual functions, and mobility. Commonly used tools include the Gross Motor Function Classification System (GMFCS)[6], the Manual Ability Classification System (MACS),[7] and the Communication Function Classification System (CFCS).[8]

Corresponding to the activity and participation components of the ICF, the Communication Function Classification System (CFCS) is a widely recognized tool developed to classify the communication performance of individuals with CP in their daily lives. The CFCS categorizes communication performance into five distinct levels, ranging from effective sender and receiver with both familiar and unfamiliar partners to seldom effective sender and receiver even with familiar partners.[9] Unlike traditional assessments that focus solely on speech or language impairments, the CFCS emphasizes functional communication, encompassing all methods such as vocalizations, facial expressions, gestures, sign language, symbols, writing, and alternative and augmentative communication (AAC). This system complements traditional speech and language assessments by enabling clinicians, educators, and parents to better understand and plan interventions tailored to the communication needs of individuals with CP.

The CFCS demonstrates high inter-rater reliability among professionals (weighted kappa score: k=0.66, 95% CI: 0.55–0.78)[10] and fair reliability between professionals and parents (k=0.49, 95% CI: 0.40-0.59). Intra-rater reliability among professionals is very good (k=0.82, 95% CI: 0.74-0.90).[10] Widely adopted worldwide, the CFCS has been translated and validated in multiple languages, including Dutch,[11] Farsi,[12] Chinese,[13] Korean,[14] and Turkish,[15] demonstrating high validity and reliability. The CFCS can integrate well with the GMFCS and MACS, helping to enhance interdisciplinary communication and therapeutic planning, making it widely used in hospitals, rehabilitation centers, and public health data collection.[16,17]

Despite the successful translations of the GMFCS and MACS into Thai, a standardized version of the Thai CFCS has not yet been established.[18,19] Therefore, the present study aimed to cross-culturally adapt the original English version of the Communication Function Classification System (CFCS) into Thai (CFCS-TH) for individuals with CP. Additionally, the study sought to evaluate the reliability and validity of the CFCS-TH in individuals aged 2-18 years with CP. The study was executed in two phases: a) the cross-cultural adaptation and translation of the CFCS into Thai (CFCS-TH); and b) the psychometric evaluation of the CFCS-TH, emphasizing its validity and reliability.

Materials and methods

Phase 1: Cross-cultural adaptation and translation of the CFCS-TH

The cross-cultural adaptation of the Communication Function Classification System into Thai (CFCS-TH) followed established guidelines proposed by Beaton *et al.*,[20] which involve six sequential s to ensure linguistic and conceptual equivalence.

1) Forward translation: Two forward translations were independently produced-one by an experienced bilingual speech-language pathologist (SLP) and the other by a professional language translator with no prior knowledge of the CFCS or medical background.

2) Synthesis of translations: An expert committee synthesized the two Thai versions into a single, harmonized draft. The committee consisted of the bilingual SLP translators from Step 1, the language

professional translators from Step 1, the principal investigator (i.e., SLP), and a research advisor, ensuring the translated version retained its original meaning and maintained linguistic appropriateness.

3) Backward translation: The synthesized Thai version was back-translated into English by two independent professional translators who had no prior exposure to the original CFCS and were not involved any of the previous steps.

4) Expert committee review: A five-member expert panel-including two forward translators from Step 1 (i.e., the bilingual SLP and a language professional), two backward translators from Step 3 (i.e., two language professionals), and an additional SLP with over 10 years of clinical experience in pediatric communication disorders who was not involved in earlier steps-evaluated the semantic, idiomatic, and conceptual equivalence of the translated content. Each item was rated using a three-level scale: -1 (not equivalent), 0 (uncertain), and +1 (clearly equivalent to the original CFCS). Suggestions for revision were also provided. The researcher then calculated the Index of Item-Objective Congruence (IOC) for each item. Items with an IOC score below 0.5 were revised according to feedback from the committee.

5) Pre-testing: Five testers-two SLPs who were not involved in any earlier steps, two occupational therapists (OTs), and one parent of an individual with CP-each of whom had at least three months of experience caring for individuals with CP, participated in the review and pilot testing of the CFCS Thai version from the previous step. Subsequently, cognitive interviews were conducted to assess the utility, clarity, and language appropriateness of the instrument, and participants also rated the overall ease of use on a five-point Likert scale, ranging from 'very easy' to 'very difficult'. All suggestions received were incorporated into the Thai pre-final version.

6) Developer review and approval: The revised pre-final version was back-translated again by an experienced translator not involved in previous steps. Both the Thai pre-final version and the new back-translated English pre-final version were submitted to the original CFCS developers for review. The finalized Thai version (CFCS-TH) was subsequently used in Phase 2 of this study to assess content validity, inter-rater reliability, and intra-rater reliability.

Phase 2: Psychometric evaluation of the CFCS-TH (validity and reliability)

Content Validity

Five expert SLPs, all of whom had not participated in any prior steps, evaluated the content of the instrument to determine its alignment with the intended measurement objectives, thereby determining its content validity. The item-level content validity index (I-CVI) and scale-level content validity index using the average method (S-CVI/Ave) were calculated based on expert ratings using a 4-point scale: 1 as not relevant, 2 as needs major revision, 3 as needs minor revision, and 4 as relevant to the measurement objective.[21,22]

Inter-rater and intra-rater reliability

As a result of the restrictions caused by the COVID-19 pandemic, access to individuals with CP and their parents was limited. Consequently, the sample size of this study was determined using the methodology for Cohen's kappa statistics that was proposed by Bujang et al. [23]. This method specifies that a five-level categorical scale with a minimum weighted kappa (k_2) of 0.4, a null hypothesis kappa (k_1) of 0, an alpha level of 0.05, and a power of 80% requires a minimum sample size of 15 participants. The target sample size was increased to 35 participants to accommodate potential classification imbalances and participant departure (estimated at 15%).

Participants, study design, and setting

Participants in Phase 1

Phase 1 included four participant groups, comprising speech-language pathologists (SLPs), translators, occupational therapists (OTs), and a parent of an individual with CP, each contributing to specific steps based on defined roles and qualifications.

Speech-Language Pathologists (SLPs): Four SLPs were directly involved in the translation process. One bilingual SLP conducted the forward translation in Step 1, participated in the synthesis in Step 2, and served as a member of the expert panel in Step 4. Another SLP from the Rajanukul Institute, who had over five years of pediatric experience, participated in Step 4 as an expert reviewer. Two others from Chiang Mai University, each with at least three months of experience caring for an individual with CP, were involved in pre-testing in Step 5. Separately, the principal investigator who oversaw the entire translation process, is also an SLP.

Translators: Four professional translators were involved in Phase 1. One translator, without a medical background, conducted the forward translation in Step 1, participated in the synthesis in Step 2, and later contributed to the expert review in Step 4. Two other translators, also without medical background, completed the backward translation in Step 3 and also participated in Step 4. The fourth independent translator carried out the pre-final back-translation in Step 6. None of the translators had prior exposure to the CFCS, and all were selected for their professional translation expertise.

Occupational Therapists (OTs): Two OTs from Chiang Mai University joined the pre-testing in Step 5, each with at least three months of experience caring for an individual with CP.

Parent of the Individual with CP: One parent participated in Step 5, meeting the criteria of Thai fluency and caregiving experience of at least three months.

Participants in Phase 2

Content-validity panel: The content validity of the CFCS-TH was evaluated by a panel of five SLPs from various institutions, including the Department of Communication Sciences and Disorders at Ramathibodi Hospital and the Sirindhorn National Medical Rehabilitation Institute, as well as private clinics and non-profit organizations specializing in augmentative and alternative communication (AAC). These experts had a minimum of 10 years of experience working with individuals with communication difficulties, including cerebral palsy, and utilizing AAC devices. Expert feedback was solicited primarily by postal mail and email. Despite follow-up reminders sent through both email and private social media messaging after the submission deadline, one expert did not respond. Consequently, the final analysis was based on the feedback provided by the four experts who completed the evaluation.

Reliability study participants: Participants in the CFCS-TH reliability study comprised two groups: individuals with CP and raters (parents, one SLP, and two PTs). Purposive sampling was employed to recruit the individuals with CP and their parents. The individuals with CP (N=35) were recruited from the Rajanagarindra Institute of Child

Development (N=15) and Srisangwan Chiangmai School (n = 20) with the following inclusion criteria: (a) individuals with CP aged between 2 and 18 years, (b) use of Thai as their primary language, and (c) provision of parental or guardian consent for participation in the study. The mean age of participants was 11 years and 1 month. The majority of subjects exhibited spastic diplegia or spastic hemiplegia. Table 1 reports general demographic data of the individuals with CP.

The parent raters consisted of 35 caregivers of individuals with CP, including 20 parents or family members and 15 homeroom teachers assigned to the boarding individuals, who met the following inclusion criteria: 1) ability to communicate in Thai, 2) fluency in reading Thai, and 3) at least three months of caregiving experience with an individual with CP. Table 2 presents the demographic characteristics of the parent raters.

Professional raters included one physical therapist (PT) from each setting and one speech-language pathologist (SLP), who also served as the principal investigator.

Table 1. General information of individuals with cerebral palsy.

Information	Number	Percentage (%)
Sex (N=35)		
Male	18	51.43
Female	17	48.57
Age (N=35)		
Under 5 years	3	8.57
5-7 years	3	8.57
7-12 years	13	37.14
12-18 years	16	45.72
Types of cerebral palsy (N=35)		
Spastic hemiplegia	7	20
Spastic diplegia	21	60
Spastic triplegia or quadriplegia	3	8.57
Other types of CP	4	11.43
Communication partners in daily life (multiple responses allowed)		
Family members (N=35)	35	100
Teachers and peers (N=35)	27	77.14
Other Individuals in daily life (unfamiliar) (N=35)	17	48.57

Table 2. General information of parents or caregivers.

Information	Number	Percentage (%)
Sex (N=35)		
Male	6	17.14
Female	29	82.86
Age (N=35)		
20-29 years	8	22.86
30-39 years	9	25.71
40-49 years	12	34.29
50 years and above	6	17.14
Education Level (N=35)		
Primary education	7	20
Lower or upper secondary education	4	11.43
Higher vocational certificate	1	2.86
Bachelor's degree	19	54.28
Postgraduate degree	4	11.43
Relationship to the individual with cerebral palsy (N=35)		
Father or mother	16	45.71
Relatives	4	11.43
Homeroom teacher	15	42.86

Research methods

The communication function level of each participating individual with CP was independently classified by one parent, one SLP, and one PT using the CFCS-TH on two occasions, with a 2-to-7-week interval between sessions. Before rating, parents received a 10- to 15-minute briefing on the tool's purpose and content, had the opportunity to ask questions, and then independently reviewed and completed the classification during the scheduled session. The professional raters independently classified the individual with CP's communication level following each therapy session.

After all data were collected, the classification results were analyzed to determine the level of agreement using weighted kappa statistics.[24] Interrater reliability analysis was conducted across three rater pairings: SLP vs PT, SLP vs. caregiver, and PT vs. caregiver. Further examination was performed by dividing interrater reliability data according to the first and second ratings. The level of agreement between each rater group was determined to increase the overall statistical power of the reliability calculations. Intra-rater reliability analysis was assessed to evaluate the consistency of each rater's classifications between the first and second ratings. The analysis was performed separately for each rater group (SLPs, PTs, and parents) to determine the stability of their classifications over time.

Ethical considerations

Ethical approval for this study was obtained from the Human Research Ethics Committees of the Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai (Approval ID: AMSEC-65FB-003), and from the Human Research Ethics Committees of Suan Prung Psychiatric Hospital, Chiang Mai (Approval ID: SPH. IRB005/2565SCs_Ful). Data collection commenced only after receiving formal approval.

Statistical analysis

Evaluation of equivalence: In step 4 of the translation procedure, the item-objective congruence index (IOC) was used to assess the equivalence of the CFCS-TH. Items with an IOC score below 0.50 were revised according to recommendations from the expert committee.[25]

Content validity: The content validity index (CVI) was calculated at both the item level (I-CVI) and scale level (S-CVI/Ave) based on expert assessments. Items with an I-CVI below 0.78 were revised, [21, 22] and the S-CVI/Ave was required to meet a minimum threshold of 0.80.[21]

Reliability: Inter-rater and intra-rater reliability were determined by calculating weighted kappa statistics. The expected inter-rater reliability was considered acceptable at a fair level of agreement (≥0.41),[10-15] while intra-rater reliability was considered good (≥0.61)[10-

[15] based on weighted kappa values.[24] Agreement levels were classified as follows: <0.20 indicated poor agreement, 0.21-0.40 slight agreement, 0.41-0.60 fair agreement, 0.61-0.80 good agreement, 0.81-0.91 very good agreement, and >0.92 excellent agreement.[24]

Results

Phase 1: Cross-cultural translation of the CFCS-TH

The six-step cross-cultural adaptation process began with forward translations conducted by a bilingual SLP and a lay translator, ensuring linguistic accuracy by incorporating perspectives from both an expert in the field and a non-expert translator while preserving the conceptual intent of the original CFCS. Step 2 yielded a synthesized version that integrated input from both translators, resulting in a culturally appropriate preliminary draft. In Step 3, back-translation confirmed preliminary semantic equivalence with the original version and was reviewed prior to further validation steps. The subsequent steps focused on evaluating the equivalence, usability, and final approval of the translated version.

In Step 4, the Item-Objective Congruence (IOC) scores from expert reviews confirmed that the translated items were well aligned with the original instrument in both meaning and structure. Semantic equivalence ranged from 0.60 to 1.00, and conceptual equivalence scores were between 0.80 and 1.00. However, idiomatic equivalence ranged from 0.40 to 1.00, with one item requiring revision for clarity by modifying the phrase “in a typical manner” to “generally” and restructuring the sentence to improve coherence.

In Step 5, feedback from five testers indicated that they understood the classification system as intended, with usability ratings ranging from “moderately easy” (one parent) to “easy” (all four professionals: two SLPs and two OTs). Based on the testers’ suggestions, the pre-final version was revised by improving phrasing, reorganizing content to reduce misinterpretation, and eliminating ambiguous terms.

In Step 6, the original developers reviewed the back-translated pre-final version of the CFCS-TH and raised questions about the wording of certain terms. After clarification was provided, they accepted the explanations and officially approved the translation on November 10, 2022. This approval confirmed both the linguistic accuracy and cultural relevance of the Thai version and marked the successful conclusion of the cross-cultural adaptation process.

Phase 2: Psychometric evaluation of the CFCS-TH

Content validity of the CFCS-TH

The item-level content validity index (I-CVI) was calculated as 1.00 for all items, and the scale-level content validity index (S-CVI/Ave) was also 1.00. These results indicated that the CFCS-TH demonstrated excellent content validity, confirming that the tool accurately measures its intended construct.

Reliability of the CFCS-TH

Inter-rater reliability

Inter-rater reliability was initially analyzed using the combined classification data from both rating rounds. The weighted kappa values indicated good agreement between the SLP and PT (ƙ=0.71, 95% CI: 0.60-0.81), good agreement between the SLP and parents (ƙ=0.66; 95% CI: 0.54-0.77), and fair agreement between the PT and parents (ƙ = 0.55, 95% CI: 0.44-0.67), as shown in Table 3. These combined results were consistent with previous validation studies of other CFCS language versions and were included in the original thesis report.

To further investigate potential differences across time points, inter-rater reliability was also analyzed by separating the data from the first and second classifications. In the first classification round, weighted kappa values indicated good agreement between the SLP and PT (ƙ=0.73, 95% CI: 0.60-0.87), fair agreement between the SLP and parents (ƙ=0.61, 95% CI: 0.44-0.77), and moderate agreement between the PT and parents (ƙ=0.48, 95% CI: 0.32-0.65), as shown in Table 3.

In the second classification round, the agreement between the SLP and PT decreased slightly (ƙ=0.69, 95% CI: 0.53-0.84), the SLP-parent agreement remained stable (ƙ=0.66, 95% CI: 0.54-0.77), and the PT-parent agreement improved (ƙ=0.62, 95% CI: 0.47-0.77), as shown in Table 3. These results provide a more nuanced view of inter-rater agreement over time and offer further insight into the consistency and responsiveness of the CFCS-TH when used across different rater pairings.

Intra-rater reliability

The intra-rater reliability of the SLP was excellent (ƙ=0.92; 95% CI: 0.84-0.99). The PT also demonstrated excellent intra-rater reliability (ƙ=0.91, 95% CI: 0.83-1.00). Caregiver intra-rater reliability was substantial (ƙ=0.86, 95% CI: 0.73-0.99), indicating consistent classifications across sessions. Communication function classification results by raters between the first and second ratings are presented in Table 4.

Table 3. Summary of inter-rater reliability for CFCS-TH classifications between rater groups with different classification rounds.

Rater pairing	Classification round	Number of rating pairs	Weighted kappa value (k; 95% CI)	Interpretation
SLP vs PTs	Combined	70	0.71 (0.60-0.81)	Good
SLP vs PTs	First	35	0.73 (0.60-0.87)	Good
SLP vs PTs	Second	35	0.69 (0.53-0.84)	Good
SLP vs Parents	Combined	70	0.66 (0.54-0.77)	Good
SLP vs Parents	First	35	0.61 (0.44-0.77)	Good
SLP vs Parents	Second	35	0.66 (0.54-0.77)	Good
PTs vs Parents	Combined	70	0.55 (0.44-0.67)	Fair
PTs vs Parents	First	35	0.48 (0.32-0.65)	Fair
PTs vs Parents	Second	35	0.62 (0.47-0.77)	Good

Note: SLP: speech-language pathologist, PTs: physical therapists, parents: caregivers of the individuals with CP. Weighted kappa (k) values were used to assess inter-rater reliability across rater pairs, based on CFCS-TH classification results between the combined, first, and second classification rounds. The interpretation of k values was based on criteria: <0.20=poor agreement, 0.21-0.40=slight, 0.41-0.60=fair, 0.61-0.80=good, 0.81-0.91=very good, and >0.92=excellent agreement.

Table 4. Summary of intra-rater reliability for CFCS-TH classifications across rater groups.

Rater	Number of Rating pairs	Weighted kappa value (k; 95% CI)	Interpretation
SLP	35	0.92 (0.84–0.99)	Very good
PTs	35	0.91 (0.83–1.00)	Very good
Parents	35	0.86 (0.73–0.99)	Very good

Note: SLP: speech-language pathologist, PTs: physical therapists, parents: caregivers of the individuals with CP. Weighted kappa (k) values were used to assess inter-rater reliability across rater pairs, based on CFCS-TH classification results between the combined, first, and second classification rounds. The interpretation of k values was based on criteria: <0.20=poor agreement, 0.21-0.40 slight, 0.41-0.60= fair, 0.61-0.80=good, 0.81-0.91=very good, and >0.92=excellent agreement.

Discussion

The purposes of this study were to culturally adapt the Communication Function Classification System (CFCS) into Thai and to evaluate its psychometric properties through a structured two-phase approach. Phase 1 involved the cross-cultural adaptation process, which resulted in the finalized Thai version of the CFCS (CFCS-TH). Phase 2 assessed the content validity, inter-rater reliability, and intra-rater reliability of the adapted tool. The findings from both phases are examined in the following sections, concerning relevant empirical studies.

Phase 1: Cross-cultural adaptation of the CFCS-TH

The cross-cultural adaptation of the CFCS into Thai (CFCS-TH) in this study placed particular emphasis on ensuring that sentence structures and terminology were appropriate for both professionals and laypersons while maintaining the intended meaning of the original classification system. The translation process was carefully structured to preserve the conceptual integrity of the tool, with the forward-backward translation approach playing a crucial role in ensuring accuracy and balancing the perspectives of an SLP and a layperson.

In Step 4, the Item-Objective Congruence (IOC) scores from expert reviews confirmed that the translated items were well aligned with the original instrument in both meaning and structure. Most terms achieved direct semantic equivalence. However, one phrase required minor revision to ensure idiomatic clarity. Expert feedback also helped ensure that the terminology was not only linguistically appropriate but also clinically relevant, thereby enhancing the overall validity of the CFCS-TH.

In Step 5, results from the cognitive interviews indicated that testers with varied professional and personal perspectives, including one parent, two SLPs, and two OTs, generally understood the classification system as intended and offered suggestions to improve its clarity and usability. Of the five testers, four professionals (80%) rated the CFCS-TH as “easy to apply”, while the caregiver (20%) rated it as “moderately easy to apply”. These usability ratings closely align with

those reported by Soleymani et al.[12], who evaluated the Farsi version of the CFCS and found that 70% of testers rated it as "very easy to apply", 20% as "easy to apply", and 10% as "fairly easy to apply". This similarity in rating patterns supports the cross-cultural applicability and face validity of the CFCS across different language contexts. Although some descriptors involved technical language, participants reported that layout features, such as the font emphasis and use of examples, facilitated understanding of the classification levels. Feedback suggested that users were better able to distinguish among levels when they clearly understood the key variables. Based on these findings, it is recommended that a brief orientation or instructional guide be provided prior to implementation.

In Step 6, the original developers reviewed the back-translated pre-final version of the CFCS-TH and approved the Thai version without requesting any modifications, although they raised a few minor questions regarding specific word choices. This approval not only confirmed the linguistic accuracy and cultural relevance of the translation but also marked the successful conclusion of the cross-cultural adaptation process, supporting the tool's readiness for psychometric evaluation in Phase 2.

Phase 2: Psychometric evaluation of the CFCS-TH

Content validity of the CFCS-TH

The content validity evaluation of the CFCS-TH faced a limitation due to the unavailability of one expert reviewer, resulting in an analysis based on the opinions of four experts. Despite this, the number of reviewers was considered adequate for a valid assessment of content validity.[21,22] The I-CVI was 1.00 across all components, exceeding the threshold for necessary revision, while the S-CVI/Ave was also 1.00, meeting established quality criteria. Despite being based on four experts rather than five as originally planned, a hypothetical scenario assuming the lowest possible rating from the fifth expert still resulted in an I-CVI above 0.78 and an S-CVI/Ave of 0.80, both of which satisfy established validity standards. Therefore, it can be concluded that the CFCS-TH is a valid tool that effectively measures communication functions in line with its intended purpose.

In other language versions of the CFCS, concurrent validity evaluation has typically been conducted by comparing CFCS classifications with the social function subscale of the Pediatric Evaluation of Disability Inventory (PEDI)[11, 13, 14]. However, this method could not be used in this study due to the lack of a Thai version of PEDI or any other standard Thai-language tool for comparison.

Reliability of the CFCS-TH

Inter-rater reliability

When analyzing inter-rater reliability, weighted kappa statistics revealed that the CFCS-TH exhibited moderate to good reliability across different rater groups. These findings indicate that the CFCS-TH has acceptable inter-rater reliability, aligning with the original CFCS, which also demonstrated moderate to good agreement. When compared to other language adaptations of the CFCS, inter-rater reliability findings ranged from moderate to excellent. The results of this study closely matched those reported in the CFCS-NL (Dutch version) by Zwart *et al.*,[11] where inter-rater reliability among SLPs was also classified as good (k=0.78, 95.5% CI: 0.66–0.89). The findings were further corroborated by studies of the CFCS-FS (Farsi version) by Soleymani *et al.*,[12] and the CFCS-KR (Korean version) by Choi *et al.*,[14] both of which demonstrated very good agreement between SLPs and therapists (k=0.81-0.87). These levels of agreement indicate that the CFCS-TH is a sufficiently robust classification tool for evaluating communication functions across diverse evaluator types.

Multiple factors may contribute to the observed variation in inter-rater reliability between the CFCS-TH and other language versions. One such factor is participant age, as previous studies have reported increased inter-rater reliability in individuals with CP aged five years and older. For instance, Hidecker *et al.*,[10] observed higher reliability among older individuals with CP, while Wang *et al.*,[13] found that inter-rater agreement was greater in individuals aged 4-16 years compared to those aged 2-4 years. Studies of the CFCS Korean version[14] and CFCS Turkish version[15] similarly noted higher reliability when participants were at least four years old. Given that the majority of participants in this study were older than seven years, it is likely that participant age contributed to the observed level of inter-rater reliability.

Parental education level was another influential factor. In this study, 80% of parents had at least a lower secondary education, which may have contributed to the good agreement between SLPs and parents. This aligns with findings by Mutlu *et al.*,[15] who reported higher inter-rater reliability when parents had higher education levels. More educated parents were better able to comprehend the definitions and criteria underlying the communication classifications, thereby producing more consistent assessments.[14,26] However, highly educated parents may also have elevated expectations of their children, which could influence their rating decisions and, consequently, impact inter-rater reliability.[26]

The lower agreement between PTs and parents in this study may have been influenced by limited prior exposure to the CFCS-TH among raters. [27,28] Neither the PTs nor parents had prior experience using the CFCS-TH, which may have contributed to variability in their interpretation of communication function levels. Previous studies, such as those by Choi *et al.*,[14] have found that experienced raters demonstrated higher inter-rater reliability. Similarly, Soleymani *et al.*,[12] reported improved reliability after providing additional

orientation and training with the CFCS-FS. These findings suggest that future applications of the CFCS-TH could benefit from structured training sessions or familiarity-building activities before use.[28]

To provide a more detailed perspective, the present study conducted an additional analysis by separating inter-rater reliability data based on the first and second assessment sessions. The resulting weighted kappa values offered further insight into the stability of agreement patterns over time. Agreement between the SLP and PT remained substantial across both sessions (κ=0.73 and κ=0.69, respectively), indicating stable interpretability of the CFCS-TH among professional raters.

The SLP-parent pairing demonstrated moderate agreement in both rounds (κ=0.61 and κ=0.66), suggesting that minimal exposure to the instrument may have enhanced parents' ability to apply the classification criteria in a manner more aligned with professional judgment. A similar pattern was observed between PTs and parents, whose agreement improved from fair (κ=0.48) in the first round to moderate (κ=0.62) in the second. This improvement may reflect a familiarity effect, wherein lay raters became more adept at applying CFCS-TH categories with repeated use.

Taken together, these results suggest that, although professional raters initially exhibit more consistent agreement, parents can achieve comparable reliability with increased exposure and appropriate orientation. These findings are consistent with prior evidence indicating that factors such as rater experience, participant age, and parental education level play a significant role in inter-rater agreement. As noted previously, the majority of individuals with CP participating in this study were older than seven years, and most parents had at least a lower-secondary level of education, both of which are factors known to contribute positively to reliability outcomes.

In conclusion, this study found that the CFCS-TH produced moderate to good inter-rater reliability across diverse rater groups, including both professionals and parents. These results are consistent with previous validations of the original and adapted CFCS versions, supporting the cross-cultural robustness of the instrument. Furthermore, the additional session-based analysis confirmed the temporal consistency of classification outcomes and revealed a potential familiarity effect among layperson raters. Collectively, these results affirm the CFCS-TH as a reliable and practical tool for assessing communication function in individuals with CP. With appropriate training or orientation, it holds promise for widespread use across clinical, educational, and community-based contexts.

Intra-rater reliability

The intra-rater reliability of the CFCS-TH was found to be very good and excellent, consistent with previous studies on the original CFCS and its various language adaptations (e.g., Dutch, Farsi, Korean, and Turkish), which also reported high levels of intra-rater reliability.[10-15] The methodologies employed in previous studies on translated versions of the CFCS exhibited certain variations, including differences in the types of raters involved, the number of raters per study, and the time intervals between assessments. The typical rater groups across these studies have included SLPs,[11,12,14] PTs,[14] Ots,[12] multidisciplinary professionals,[10] and parents.[12,14] The number of raters per group has varied significantly, with some studies involving a single rater per classification group,[14] while others have included between two and four raters,[11,12] and some studies have involved as many as 48 raters.[10] The time interval between the first and second classification assessments has also differed, ranging from retesting within two weeks,[14] to a minimum of two weeks,[10,12] and at least seven weeks apart in some cases.[11] Despite these methodological differences, the high consistency in results across studies confirms that both the CFCS-TH and other translated versions are robust tools with strong intra-rater reliability.

Additionally, a common finding in all studies,[10-12,14] is that intra-rater reliability among parents tends to be lower than that of professional raters, which is likely due to the same influencing factors previously identified in the inter-rater reliability analysis, including participant age, parental education level, and prior familiarity with the CFCS-TH.

Limitations

A key limitation of this study was the absence of a standardized Thai-language assessment tool for criterion validity testing, which is commonly used in the evaluation of CFCS adaptations in other languages. Additionally, the study was conducted during the COVID-19 pandemic, which hindered access to the target sample and led to a smaller sample size as a result of adjustments in statistical analysis. Some individuals with CP were students residing at Srisangwan Chiangmai School, meaning their communication functions were necessarily classified by their homeroom teachers rather than their parents or guardians. The sample was also imbalanced, particularly in the 2–4 years age group, which may have impacted inter-rater reliability and limited the generalizability of the findings to younger individuals. For individuals with CP in this age group, a collaborative evaluation involving clinicians and parents may be necessary, or additional factors should be considered to improve inter-rater reliability.

Conclusion

The Communication Function Classification System for Individuals with Cerebral Palsy-Thai Version (CFCS-TH) was demonstrated to be a valid and reliable medical classification tool for assessing the communication performance levels of individuals with CP in their daily lives in Thailand. The findings

of this study support its application in both clinical and research settings by providing a standardized framework for evaluating communication functions and facilitating effective intervention planning.

Ethical approval

This study received ethical approval from the Research Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University (AMSEC-65FB-003), as well as from the Suan Prung Psychiatric Hospital, Chiang Mai (SPH.IRB005/2565SCs_ful) for research involving human participants. Prior to enrollment in this study, all participants and caregivers were provided with all pertinent research information, and written informed consent was obtained.

Funding

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

Conflicts of interest

The authors declare no conflicts of interest.

CRediT authorship contribution statement

Pim Chaisook: conceptualization, methodology, project administration, formal analysis, investigation, data curation, writing: original draft, visualization; **Phuanjai Rattakorn**: supervision, project administration, data curation, writing: review and edit, funding acquisition; **Supaporn Chinchai**: supervision, project administration, writing: review and edit, funding acquisition; **Wannipa Bunrayong**: supervision, project administration, funding acquisition.

Acknowledgements

The authors would like to express their sincere gratitude to the Clinical Service Center, Faculty of Associated Medical Sciences, Chiang Mai University, the Rajanagarindra Institute of Child Development, and Srisangwan Chiangmai School for their valuable support and assistance in participant recruitment and data collection. We also extend our appreciation to all participants and their families for their kind cooperation and meaningful contributions to this study.

References

[1] Gulati S, Sondhi V. Cerebral palsy: an overview. Indian J Pediatr. 2018; 85: 1006-16. doi: 10.1007/s12098-017-2475-1.

[2] Pennington L. Cerebral palsy and communication. J Paediatr Child Health. 2008; 18: 405-9. doi: 10.1016/j.paed.2008.05.013.

[3] McFadd ED, Hustad KC. Communication modes and functions in children with cerebral palsy. J Speech Lang Hear Res. 2020; 63: 1776-92. doi: 10.1044/2020_jslhr-19-00228.

[4] Stucki G, Grimby G. Applying the ICF in medicine. J Rehabi Med. 2004; 44: 5-6. doi: 10.1080/16501960410023573.

[5] Ogoke CC. Cerebral palsy-clinical and therapeutic aspects [Internet]. London: IntechOpen; 2018 [cited 2021 Jan 20]. Available from: https://www.intechopen.com/books/7072

[6] Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. Dev Med Child Neurol. 1997; 39(4): 214-23. doi:10.1111/j.1469-8749.1997.tb07414.x.

[7] Eliasson AC, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Öhrvall AM, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. Dev Med Child Neurol. 2006; 48(7): 549-54. doi:10.1017/S0012162206001162.

[8] Rosenbaum P, Eliasson AC, Hidecker MJC, Palisano RJ. Classification in childhood disability: focusing on function in the 21st century. J Child Neurol. 2014; 29: 1036-45. doi: 10.1177/0883073814533008.

[9] Hidecker MJC, Paneth N, Rosenbaum PL, Kent RD, Lillie J, John B. Eulenberg, et al. Communication Function Classification System (CFCS) [Internet]. Wyoming: Communication function classification systems; 2011. [cited 2021 June 29]. Available from: http://cfcs.us/wp-content/uploads/2014/02/CFCS_universal_2012_06_06.pdf.

[10] Hidecker MJC, Paneth N, Rosenbaum PL, Kent RD, Lillie J, Eulenberg JB, et al. Developing and validating the communication function classification system for individuals with cerebral palsy. Dev Med Child Neurol. 2011; 53: 704-10. doi: 10.1111/j.1469-8749.2011.03996.x.

[11] Zwart KEV, Geytenbeek JJ, de Kleijn M, Oostrom KJ, Gorter JW, Hidecker MJC, et al. Reliability of the Dutch-language version of the communication function classification system and its association with language comprehension and method of communication. Dev Med Child Neurol. 2016; 58: 180-8. doi: 10.1111/dmcn.12839.

[12] Soleymani Z, Joveini G, Baghestani AR. The communication function classification system: cultural adaptation, validity, and reliability of the Farsi version for patients with cerebral palsy. Pediatr Neurol. 2015; 52: 333-7. doi: 10.1016/j.pediatrneurol.2014.10.026.

[13] Yan-na W, Wei S. Reliability and validity of Chinese communication function classification system (Fudan). Chin J Evid Pediatr 2017; 12: 321-7. doi: 10.3969/j.issn.1673-5501.2017.05.001.

[14] Choi JY, Hwang EH, Rha D, Park ES. Reliability and validity of the Korean-language version of the communication function classification system in children with cerebral palsy. Child Care Health. Dev. 2018; 44: 140-6. doi: 10.1111/cch.12507.

[15] Mutlu A, Kaya Kara Ö, Livanelioğlu A, Karahan S, Alkan H, Yardımcı BN, et al. Agreement between parents and clinicians on the communication function levels and relationship of classification systems of children with cerebral palsy. Disabil Health J. 2018;11(2):281-6. doi:10.1016/j.dhjo.2017.11.001.

[16] Alghamdi MS, Chiarello LA, Palisano RJ, McCoy SW. Understanding participation of children with cerebral palsy in family and recreational activities. Res Dev Disabil. 2017; 69: 96-104. doi: 10.1016/j.ridd.2017.07.006.

[17] Küçükakkaş O, Yılmaz Bairamov N, Yurdakul OV, İnce B. Evaluation of factors affecting communication functions in children with cerebral palsy. Turk J Pediatr Dis. 2021; 15: 110-16. doi: 10.12956/tchd.747013.

[18] Ramrit S, Emasithi A, Amatachaya S, Siritaratiwat W. Reliability of GMFCS-E&R and GMFCS-FR Thai version in children with cerebral palsy. J Med Tech Phy Ther. 2014; 26: 67-75 (in Thai). Available from: https://he01.tci-thaijo.org/index.php/ams/article/view/66350.

[19] Manual Ability Classification System. Manual ability classification system for children with cerebral palsy 1-18 years. [Internet]. [cited 2021 June 6]. Available from: https://macs.nu/files/MACS_Thai_2016.pdf.

[20] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine J 2000; 25: 3186-91. doi: 10.1016/j.spinee.2000.09.013.

[21] Phatthanasombatsuk M. Validation of nursing and social science research instruments. The Southern College Network Journal of Nursing and Public Health 2021; 8: 329-43 (in Thai). Available from: https://he01.tci-thaijo.org/index.php/scnet/article/view/240643/169173

[22] Polit DF, Beck CT, Owen SVJRin, health. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health 2007; 30: 459-67. doi: 10.1002/nur.20199.

[23] Bujang MA, Baharum N. Guidelines of the minimum sample size requirements for Kappa agreement test. Epidemiol Biostat Public Health. 2017; 14: e122671-7. doi: 10.2427/122671.

[24] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33(1): 159-74. doi:10.2307/2529310.

[25] Chaihanawirote U, Vantum C. Evaluation of content validity for research instrument. Journal of Nursing and Health Sciences. 2017; 11: 105-11 (in Thai).

[26] Ezeizabarrena MJ, Kovacevic M. Editorial: Parental questionnaires as a reliable instrument for the assessment of child language development. Front Psychol. 2024; 15: 1471747. doi: 10.3389/fpsyg.2024.1471747.

[27] Virella D, Pennington L, Andersen GL, Andrada MG, Greitane A, Himmelmann K, et al. Classification systems of communication for use in epidemiological surveillance of children with cerebral palsy. Dev Med Child Neurol. 2016; 58(3): 213-316. doi: 10.1111/dmcn.12802.

[28] Cunningham BJ, Rosenbaum P, Hidecker MJC. Promoting consistent use of the communication function classification system (CFCS). Disabil. 2016; 38: 195-204. doi: 10.3109/09638288.2015.1027009.

Journal of Associated Medical Sciences 2026; 59(1): 109-120

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 systematic review of the effect of the integration of speech therapy with music therapy on swallowing for patients with Parkinson's disease

Natwipa Wanicharoen[1], Vich Boonrod[2], Erin Songdechagraiwud[3], Palita Yaemsuan[1*]

[1]Communication Sciences and Disorders Division, Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Department of Music, Faculty of Humanities, Naresuan University, Phitsanulok Province, Thailand.
[3]Chonburi Hospital, Chonburi Province, Thailand.

ARTICLE INFO

Article history:
Received 12 July 2025
Accepted as revised 29 September 2025
Available online 14 October 2025

Keywords:
Parkinson's disease, swallowing, speech therapy, music therapy, systematic review.

ABSTRACT

Background: Dysphagia, or swallowing impairment, is a common complication in individuals with Parkinson's disease. Speech therapy is a non-pharmacological intervention that employs behavioral techniques to address this condition. A growing body of research also suggests that music therapy may have beneficial effects on swallowing function in this population. However, the effectiveness of integrating speech and music therapy for improving swallowing in individuals with Parkinson's disease has not yet been systematically evaluated.

Objectives: The objective of this systematic review was to examine the effectiveness of integrating speech therapy with music therapy for swallowing in patients with Parkinson's disease.

Materials and methods: This study was registered in PROSPERO (CRD420251118370) and performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). A total of two studies met the inclusion criteria. The authors systematically reviewed four electronic databases: PubMed, CINAHL Complete, the Cochrane Library, and Google Scholar. A hand search of the reference list of related articles was performed to identify any additional eligible studies. The assessment of bias risk was conducted using the Joanna Briggs Institute's (JBI) critical appraisal tools.

Results: A total of two studies investigated the effects of speech therapy in combination with music therapy. One study implemented an interdisciplinary choral singing therapy program that included vocal exercise and choral singing. The other study applied a conventional speech therapy protocol-specifically, Lee Silverman Voice Treatment (LSVT LOUD®) and vocal exercises-combined with music therapy incorporating motor, breathing, vocal, and singing exercises. The integration of speech and music therapy demonstrated promising results in improving swallowing function in individuals with Parkinson's disease.

Conclusion: The findings highlight the potential benefits of integrating speech and music therapy to improve swallowing function in individuals with Parkinson's disease. However, due to the limited number of studies and methodological heterogeneity, definitive conclusions regarding the most effective intervention components cannot yet be drawn. While encouraging, the current evidence underscores the need for further high-quality research particularly randomized controlled trials to establish standardized protocols. Advancing this interdisciplinary approach may enhance rehabilitation outcomes and improve quality of life for this population.

*** 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: palita.y@cmu.ac.th
doi: 10.12982/JAMS.2026.014
E-ISSN: 2539-6056

Introduction

The prevalence of Parkinson's disease is projected to exceed one million individuals in the United States by 2030, based on U.S. Census Bureau population estimates, and more than 10 million people worldwide.[1] Parkinson's disease is a neurological condition characterized by involuntary or uncontrolled movements such as tremors in the hands, arms, legs, jaw, or head, along with muscle stiffness, impaired balance, and coordination difficulties. Symptoms typically emerge gradually and worsen progressively over time. As the condition progresses, individuals may encounter difficulties with mobility and communication.[2] This includes dysphagia, or swallowing impairment, which is a common non-motor symptom observed in individuals with Parkinson's disease across all stages, including the prodromal phase when diagnostic criteria are not yet met[3] as well as the early stages of the disease.[4]

Individuals with Parkinson's disease may exhibit dysphagia in the oral, pharyngeal, and esophageal phases.[5-7] Contributing factors include muscular incoordination or weakness and sensory dysfunction, with repetitive tongue pumping being the most frequent manifestation. Additional difficulties may involve delayed initiation of the pharyngeal swallow, poor lip seal and oral containment, pharyngeal residue, and airway invasion.[6]

Studies reported differences in the prevalence of dysphagia in Parkinson's disease. For instance, a meta-analysis by Kalf *et al.*[8] examined the prevalence of oropharyngeal dysphagia. The findings showed that while using subjective outcomes yielded an estimated prevalence of oropharyngeal dysphagia of 35% (95% confidence interval; CI=28-41), using objective data yielded an estimate of 82% (95% CI=77-87). However, due to variations in the methods used to assess swallowing functions, the prevalence of dysphagia in individuals with Parkinson's disease varies. Recently, a meta-analysis of Gong *et al.*[9] studies the prevalence of dysphagia in Parkinson's disease across continents. The findings showed that Oceania has the highest frequency of dysphagia in Parkinson's disease (56.3%), followed by Africa (39.5%), Asia (38.6%), Europe (36.1%), and America (28.9%).

Additionally, studies have investigated quality of life (QoL) related to swallowing in patients with Parkinson's disease in various countries. For example, in Brazil, a study by Carneiro *et al.*[10] used the Swallowing Quality of Life Questionnaire (SWAL-QOL). The findings showed that patients with Parkinson's disease have considerably lower ratings on all SWAL-QOL domains than normal controls. The disease's progression worsens swallowing QOL, particularly in terms of eating duration, symptom frequency, and sleep. This occurs primarily in the later stages of the disease. In Hong Kong, a study by Chan *et al.*[11] used Chinese version of the Swallowing Quality-of-Life Questionnaire (C-SWAL-QOL). The findings showed that the most severely affected scales included sleep, fatigue, eating duration, and communication. Dysphagia, in particular, hampers medicine intake and causes malnutrition and aspiration pneumonia, which is a leading cause of death in individuals with Parkinson's disease.[12] In clinical practice, patients should receive early screening, diagnosis, and treatment to avoid complications from dysphagia.[9]

Numerous studies have demonstrated that speech therapy is a non-pharmacological intervention employing behavioral techniques to address dysphagia in patients with Parkinson's disease. Early evidence from Robertson and Thomson[13] showed that intensive speech therapy improved speech intelligibility and voice quality, with patients also reporting secondary improvements in swallowing, despite the therapy not being specifically designed for dysphagia. Subsequent studies have provided physiological confirmation of these cross-system effects. Evidence suggests that intensive speech therapy exerts beneficial influences on swallowing across multiple phases. In the oral phase, increased tongue and submental activity enhance bolus propulsion, consistent with reports of reduced oral transit time following intensive voice therapy.[14] In the pharyngeal phase, greater hyolaryngeal excursion and improved tongue base retraction facilitate bolus transit through the pharyngoesophageal segment, thereby reducing aspiration risk.[14,15] For airway protection, enhanced glottal closure and stronger expiratory drive contribute to safer swallowing, with recent findings also demonstrating improved reflexive cough capacity.[15,16] These overlapping mechanisms support the rationale for investigating speech therapy in relation to dysphagia, even though its primary target is communication.[17]

Currently, a growing body of research suggests that music therapy particularly therapeutic singing may support swallowing function in individuals with Parkinson's disease. In the early stages, Stegemöller *et al.*[18] reported that a structured singing program enhanced motor and neuromuscular outcomes, which were interpreted as reflecting prolonged laryngeal elevation and improved airway protection, although swallowing-related quality of life remained unchanged. In later stages, Yeo *et al.*[19] demonstrated that a therapist led singing intervention helped preserve swallowing function and improved self-reported swallowing outcomes compared to usual care. Consistent with these findings, a systematic review highlighted that music therapy programs especially those incorporating singing can yield benefits across motor, communication, and swallowing domains in Parkinson's disease, suggesting their potential as a complementary rehabilitation strategy.[20]

Mechanistically, music therapy engages vocal-respiratory tasks that overlap with the physiology of swallowing. Singing involves sustained phonation and

pitch modulation, which strengthen oral motor control, stimulate the suprahyoid and laryngeal muscles to facilitate hyolaryngeal excursion, and enhance breath-swallow coordination and glottal closure. These mechanisms mirror some effects of conventional speech or voice therapy; however, music therapy distinguishes itself by embedding exercises within rhythm, melody, and familiar songs, thereby fostering emotional engagement and long-term adherence.[18-20] Emerging studies have begun to explore the integration of speech and music therapy for dysphagia in patients with Parkinson's disease, reflecting a growing interest in multidisciplinary approaches.[21,22] For example, a study by Funderburke investigated an 8-week interdisciplinary choral singing therapy (ICST) program that combined music and voice therapy.[21] Participants' swallowing evaluations improved, showing increases in angle of elevation and magnitude of displacement. Similarly, a study by Mohseni *et al*.[22] investigated a 4-week telerehabilitation program that integrated speech and music therapy for dysphagia in patients with Parkinson's disease. The combined therapy group demonstrated greater improvements on the Swallowing Disturbance Questionnaire (SDQ) and Dysphagia Handicap Index (DHI) compared to groups receiving speech or music therapy alone.

While early reports suggest potential benefits of such combined interventions, the effectiveness of integrating speech therapy with music therapy for swallowing in patients with Parkinson's disease has not yet been systematically evaluated. To address this gap, the objective of this systematic review was to investigate the effectiveness of integrating speech and music therapy on swallowing outcomes in this population. The review question was: "What is the effectiveness of speech therapy integrated with music therapy on swallowing outcomes in patients with Parkinson's disease?" The findings aim to provide an up-to-date synthesis of the existing evidence, supporting clinical decision-making and encouraging interdisciplinary collaboration to optimize patient care.

Method

This study was registered in PROSPERO (CRD420251118370) and performed according to PRISMA.[23]

Data sources and searches

The literature search was performed by two independent authors (N.W and V.B). The four electronic databases included PubMed, CINAHL Complete, the Cochrane Library, and Google Scholar. A hand search of the reference list of related articles was performed to identify any additional eligible studies. A systematic search was conducted across all databases from September to October 2024 to identify articles published between 2010 and 2024. All studies were published in English language.

A range of search terms was employed and applied consistently across all databases. These terms included: ("Parkinson's disease") AND ("dysphagia" OR "swallowing impairment") AND ("speech therapy") AND ("music therapy") AND ("swallowing").

Inclusion and exclusion criteria

Table 1 shows the Participant, Intervention, Comparison, Outcome, and Study design (PICOs) frameworks that were used for selection criteria.

Table 1. Inclusion and exclusion criteria.

PICOS	Inclusion criteria	Exclusion criteria
Population	Patients who were diagnosed with Parkinson's disease at any stage of disease progression and severity and patients who had dysphagia of any severity.	Patients who were with non-parkinsonian conditions and without dysphagia.
Intervention	Speech therapy integrated with music therapy	Speech therapy or music therapy alone as a primary intervention.
Comparison	Studies with and without a comparative group were considered, such as music therapy or speech therapy alone or without intervention.	N/A
Outcome	Swallowing-related outcomes, as evaluated by instrumental assessment and/or functional scores.	Swallowing-related outcomes were not considered as primary outcomes.
Study design	Randomised and non-randomised studies.	Editorials, opinions and commentaries, and qualitative studies

Data extraction

Two independent authors (NW and VB) screened titles and abstracts of the records identified through the searches described above. The results from the searches were exported into the industry standard bibliographic tool Endnote™ to merge search results from different databases, and duplicates were removed. Once the initial results list was finalized, two authors (NW and VB) obtained full-text copies of all the remaining studies that fulfilled the listed inclusion criteria. They independently assessed the studies based on the inclusion criteria and decided whether to include or exclude studies. We resolved any disagreements through discussion, but when consensus could not be reached, a third reviewer (ES) was consulted, and her decision was final.

The main characteristics of the included studies were extracted by the two independent authors (NW and VB) using a designed template. We resolved any disagreements through discussion, but when consensus could not be reached, a third reviewer (ES) was consulted, and her decision is final. Data were extracted on study characteristics (author, year, country, type of study, and sample size), patient characteristics (diagnostic, age, and gender), intervention (type of intervention and comparator groups), and outcome measures (measurements and all swallow related outcomes).

Quality assessment

The quality of the studies included in this review were evaluated using the JBI critical appraisal tools according to research type. The two independent authors (NW and VB) reviewed the methodological quality of the included studies. Any disagreement was resolved by consensus, and whenever this was not possible, a third reviewer (ES) was consulted, whose decision was final.

The JBI quasi-experimental studies checklist is made up of nine items: certainty of cause and effect, pre-homogeneity verification, exposure to the same environment outside of the intervention, presence or absence of a control group, pre- and post-intervention effect measures, description of dropouts, equivalence of outcome measures, appropriateness of outcome variable measures, and statistical analysis methods. For each item, the checklist assigned a score of 1 to "yes" and 0 to "unclear," "no," and "not applicable".[24]

Table 2 shows the quasi-experimental study that were appraised using the JBI tool.[24] The quality assessment score for quasi-experimental study was 4 out of 9 (high risk). The quasi-experimental design of this study is vulnerable to selection bias, owing to the lack of a control group.

Table 2. The JBI critical appraisal checklist for quasi-experimental study.

JBI critical appraisal checklist for quasi-experimental study	Funderburke[23]
Is it clear in the study what is the "cause" and what is the "effect" (i.e., there is no confusion about which variable comes first)?	Y
Was there a control group?	N
Were the participants included in any comparisons similar?	NA
Were the participants included in any comparisons receiving similar treatment/care, other than the exposure or intervention of interest?	NA
Were there multiple measurements of the outcome both pre- and post-intervention/ exposure?	Y
Were the outcomes of participants included in any comparisons measured in the same way?	NA
Were outcomes measured in a reliable way?	Y
Was follow-up complete and if not, were differences between groups in terms of their follow-up adequately described and analyzed?	N
Was appropriate statistical analysis used?	Y
Overall appraisal	**4/9** (High risk)

Note: Y: yes, N: no, NA: not applicable.

The JBI RCT studies checklist is made up of 13 items: randomization, allocation concealment, pre-homogeneity verification, blinding (participants, interventors, and assessors), identical conditions other than experimental treatment, description of dropouts, analysis based on randomization, equivalence of outcome measures, appropriateness of outcome variable measures and statistical analysis methods, and appropriateness of the study design. For each item, the checklist assigned a score of 1 to "yes" and 0 to "unclear," "no," and "not applicable".[25]

Table 3 shows the critical appraisal for the randomized clinical trials that were assessed using the JBI tool.[25] The quality assessment score for RCT study was 11 out of 13 (low risk).

Table 3. The JBI critical appraisal checklist for randomized controlled trials.

JBI critical appraisal checklist for randomized controlled trials	Mohseni et al.[25]
Was true randomization used for assignment of participants to treatment groups?	Y
Was allocation to treatment groups concealed?	Y
Were treatment groups similar at the baseline?	Y
Were participants blind to treatment assignment?	Y
Were those delivering treatment blind to treatment assignment?	N
Were outcomes assessors blind to treatment assignment?	N
Were treatment groups treated identically other than the intervention of interest?	Y
Was follow up complete and if not, were differences between groups in terms of their follow up adequately described and analyzed?	Y
Were participants analyzed in the groups to which they were randomized?	Y
Were outcomes measured in the same way for treatment groups?	Y
Were outcomes measured in a reliable way?	Y
Was appropriate statistical analysis used?	Y
Was the trial design appropriate, and any deviations from the standard RCT design (individual randomization, parallel groups) accounted for in the conduct and analysis of the trial?	Y
Overall appraisal	**11** (Low risk)

Data analysis

A descriptive analysis of the data was conducted. The researchers did not perform analysis of subgroups or subsets. Meta-analysis of the studies was not undertaken due to the heterogeneity of the interventions and outcome measures employed in the studies. This makes it challenging to integrate the data in a meaningful way and perform a meta-analysis.

Results

Search results

Figure 1 shows the PRISMA flowchart of the study. A comprehensive literature search in electronic databases (N=34,128) and a hand search (N=13) yielded 34,141 studies. After duplicates were removed, 14,907 studies remained and were kept for title and abstract review. Then, 14,894 studies were excluded after screening due to factors such as unrelated studies (N=11,778), non-English language (N=655), and other considerations (N=2,461). Thereafter, 13 studies were assessed for full-text articles for eligibility, and 11 studies were excluded for reasons such as not integrating speech and music (N=9) or not focusing on swallowing in Parkinson's disease patients (N=2). Consequently, the review included the remaining two studies.

Table 4 shows study characteristics included in the study. Two of the studies included in this review are clinical trials, one was a quasi-experimental study,[21] and one was a randomized study.[22] Studies were published, one in 2012[21] and another one in 2023.[22] Of the total studies, one was conducted in the USA,[21] and one in Iran, West Asia.[22]

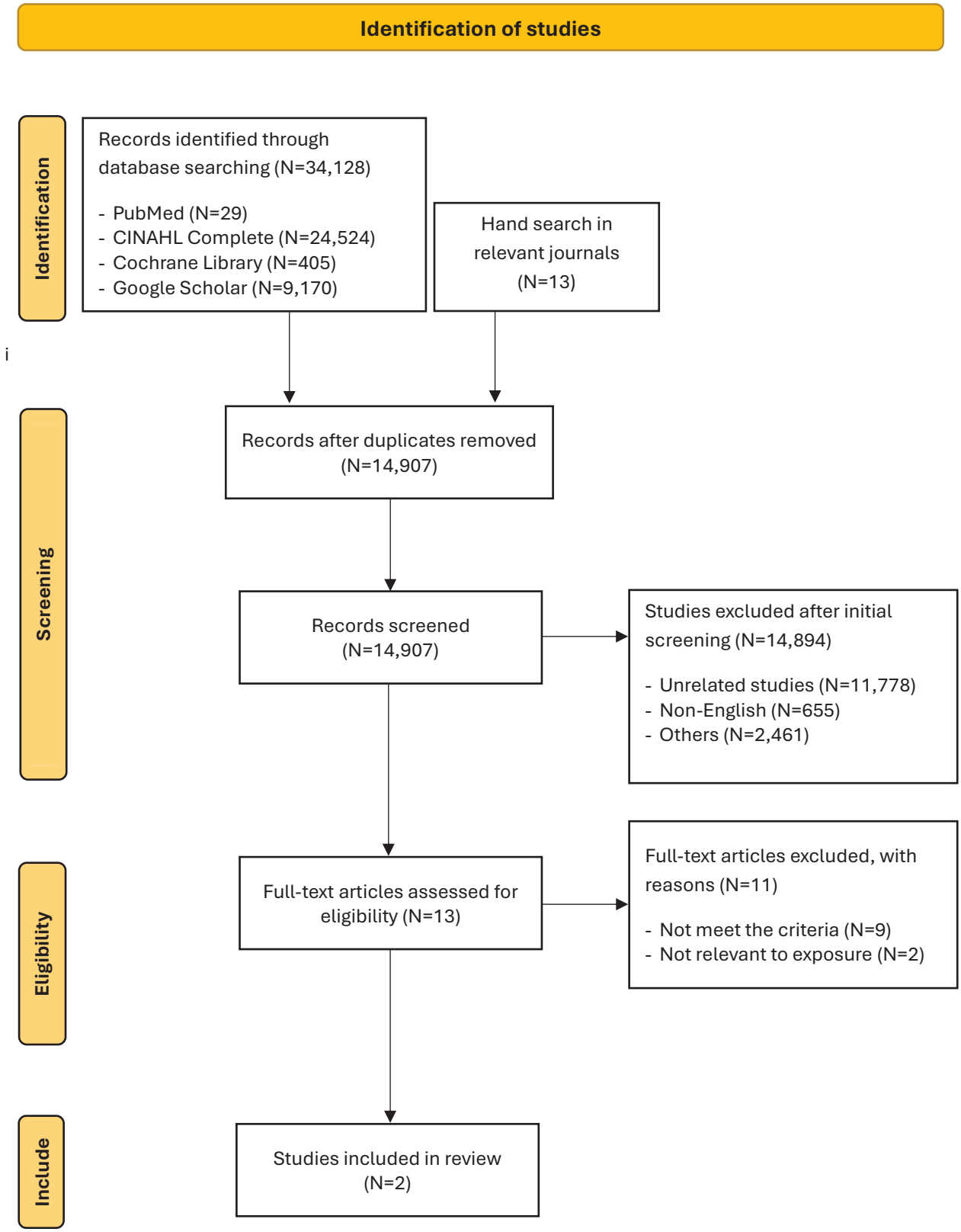


Figure 1. *Methodology flowchart.*

Table 4. Included studies.

Study/ Year/ Country	Study design	Participant	Interventions	Control	Duration	Interventionist	Measurements on swallowing function	Outcomes
Funder burke[21] / 2012 / USA	Quasi-experimental design	- Three male patients with idiopathic Parkinson's disease - Ages: 67 and 85 years - Patients with Parkinson's disease being at stages 2.5 or better based on the H&Y scale. - Patients had been on a stable regimen of antiparkinsonian and psychotropic medications for 30 days prior to participation.	ICST	No	8 sessions for 8 weeks (2.5-3 hours per session)	Speech therapist	- Video fluoroscopy, utilizing the modified barium swallow study (MBS) - Swallowing quality of life questionnaire (SWAL-QOL)	- A choral singing program, integrating speech therapy and music therapy, may lead to positive changes in swallowing measures for individuals with Parkinson's disease. - Improvements were noted in swallowing evaluations, with increases in angle of elevation and magnitude displacement for participants. - Therapeutic benefits can occur even with non-swallowing specific activities, highlighting the importance of context in motor movement rehabilitation.

Table 4. Included studies. (Continue)

Study/ Year/ Country	Study design	Participant	Interventions	Control	Duration	Interventionist	Measurements on swallowing function	Outcomes
Mohseni *et al.*[22] / 2023 / Iran	RCT	- Thirty-three patients with idiopathic Parkinson's disease - 25 men and 8 women with complaints of swallowing impairment - Mean age, 58.88 years - Patients with Parkinson's disease being at stages 1 to 2 (mild) based on the H&Y scale. - Patients were in the "on-state" of the medication cycle, which typically occurs one hour after medication.	Combination therapy (speech therapy + music therapy)	- Speech therapy - Music therapy	12 tele-rehabilitation sessions for 4 weeks (1 hour per session)	Speech therapist (to perform speech exercises) and a music and singing specialist (to perform singing exercises)	- Dysphagia handicap index (DHI) - Swallowing disturbance questionnaire (SDQ)	- Significant improvements in swallowing function were observed across all treatment groups, with the combination therapy group demonstrating the greatest benefits. - The combination therapy group achieved superior outcomes compared with both conventional speech therapy and music therapy groups on the SDQ and DHI scores. - These improvements were sustained at a 3-month follow-up, indicating lasting therapeutic effects. - Overall, the findings suggest that both speech therapy and combination therapy are effective in reducing dysphagia symptoms in patients with Parkinson's disease, with combination therapy offering the most pronounced benefits.

Thirty-six participants included in the study were patients with idiopathic Parkinson's disease who were clinically evaluated and ultimately diagnosed by a neurologist.[21,22] The number of participants with Parkinson's disease included in each study was 3 in one study[21] and 33 in another one.[22] The total number of patients included in the study was 36, comprising 28 males and 8 females.[21,22]

Disease severity was characterized using the Hoehn and Yahr (H&Y) scale, which included subjects with a score of 1-2[22] and 2.5,[21] reflecting the mild functional disability. Both studies recorded information about the participants' medication cycle during assessment or treatment. In one study, the patients were in the "on-state" of the medication cycle, which typically occurs one hour after medication.[22] In another study, patients had been on a stable regimen of antiparkinsonian and psychotropic medications for 30 days prior to participation.[21]

A total of 2 studies reported the effect of speech therapy in combination with music therapy.[21,22] Interventions were heterogeneous across the included studies. One study was interdisciplinary choral singing therapy (ICST), including vocal exercise and choral singing.[21] Another study was conventional speech therapy (LSVT therapy and vocal exercises) combined with music therapy (motor, breathing, vocal, and singing exercises).[22] In one study, speech therapy in combination with music therapy was compared to speech therapy or music therapy alone,[22] while the remaining study was without a comparative group.[21]

The intervention duration in one study was 8 weeks and was conducted by a speech therapist.[21] Another study was 4 weeks, and a speech therapist contributed to the speech intervention, while a music therapist contributed to the music intervention.[22]

The included studies assessed different assessments but focused on domains directly related to swallowing outcomes. One study used videofluoroscopy, utilizing the MBS, and self-reported outcomes were measured using the SWAL-QOL.[21] Another study used DHI (functional, physical, and emotional sub-tests, and as total score) and SDQ scores.[22]

A study by Funderburke[21] demonstrated that a choral singing program, integrating speech therapy and music therapy, may lead to positive changes in swallowing measures for individuals with Parkinson's disease. The findings showed that swallowing function increased the angle of elevation and amplitude of hyoid bone displacement during three consecutive swallows of 10 ml of thin liquid and pudding-thick barium. Although no statistically significant improvements were observed in SWAL-QOL scores, slight numerical increases were noted in selected subscales such as communication and fear of eating. These subjective trends, while limited by the small sample size and short training duration, were directionally consistent with the physiological gains observed. Therapeutic benefits can therefore occur even with non-swallowing-specific activities, highlighting the importance of context in motor movement rehabilitation.

A study by Mohseni *et al.*[22] demonstrated that a telerehabilitation program combining speech therapy and music therapy significantly improved swallowing functions in patients with Parkinson's disease. The findings showed that the combination therapy group did better on the SDQ and all its sub-tests (functional, physical, and emotional), as well as the total score of the DHI. They also did better on the SDQ score than the two other groups that only received speech therapy or music therapy. This suggests that the combined approach is more effective in addressing the multifaceted nature of dysphagia in Parkinson's disease. The study highlighted that the combination therapy group outperformed the other groups in all sub-tests and the total score of DHI, indicating a comprehensive improvement in swallowing function, including functional, physical, and emotional aspects. The interventions maintained their improvements during a 3-month follow-up, demonstrating their lasting effects.

Discussion

To the best of our knowledge, this was the first systematic review to investigate the effectiveness of integrating speech therapy with music therapy on swallowing outcomes in patients with Parkinson's disease.

Based on our findings, it is not possible to determine with certainty which types of speech therapy and music therapy are most effective for improving swallowing function in patients with Parkinson's disease, largely due to the heterogeneity of the included studies. There have been only two studies that have focused on the integration of speech therapy and music therapy to improve swallowing function in patients with Parkinson's disease.[21,22] Nevertheless, our review indicates that the combined use of speech therapy and music therapy, as a non-pharmacological approach, provides meaningful benefits for swallowing outcomes. These findings are generally consistent with previous evidence supporting the effectiveness of speech therapy alone[14-17] and music therapy alone[18] in improving swallowing function in patients with Parkinson's disease.

Both included studies provide evidence that speech therapy even when primarily targeting communication can positively influence swallowing mechanisms in individuals with Parkinson's disease. In Funderburke's study,[21] structured vocal exercises such as lip buzzing, glissandos, messa di voce, and articulation practice were incorporated to improve respiratory–phonatory coordination, vocal intensity, and articulatory precision. These tasks engaged the tongue, lips, jaw, and submental muscles, contributing

to oral bolus formation and propulsion, and ultimately supporting swallowing physiology. Similarly, Mohseni *et al*.[22] reported that conventional speech therapy adapted from LSVT principles including sustained vowels, pitch variation, and loudness training-significantly improved swallowing outcomes when delivered alone or in combination with music therapy. The program emphasized high effort phonatory tasks, which reinforced respiratory drive, glottal closure, and submental muscle activation, thereby supporting oral bolus propulsion and airway protection.

These findings align with the broader literature on intensive speech therapy, particularly Lee Silverman Voice Treatment (LSVT LOUD®), which has been shown to improve both speech and swallowing outcomes in individuals with Parkinson's disease. Unlike traditional multifocus therapies, LSVT emphasizes high-effort phonation and motor learning principles, resulting in increased respiratory drive, glottal closure, and hyolaryngeal excursion.[13-16] Importantly, LSVT's efficacy has been attributed to principles of motor learning and activity-dependent neuroplasticity, which suggest that intensive vocal training can induce long-term adaptive changes in shared neural pathways involved in both speech and swallowing.[26-28]

Although neither of the included studies directly applied the LSVT protocol, both incorporated interventions grounded in the same principle of intensive, high-effort vocalization. In Funderburke's study,[21] this was delivered through ICST, while in Mohseni *et al*.[22] it was implemented as an LSVT-adapted program. These approaches share a common mechanism that provides a biological rationale for the observed cross-system effects, supporting the view that intensive speech therapy whether delivered through structured vocal drills, ICST, or other high-effort phonatory formats can indirectly enhance swallowing safety.

Beyond speech therapy, singing represents a novel approach that leverages the principles of neuroplasticity, suggesting that musical stimuli can activate neural processes to influence non-musical behaviors such as swallowing. The notion that singing may enhance swallowing function given that respiration, phonation, and articulation engage overlapping neuroanatomical substrates-supports the theoretical basis for music therapy in this context. The shared neural networks involved in singing and speaking further reinforce the auditory–motor feedback loop that singing specifically activates in the brain.[29] Swallowing and vocalization both require coordinated activation of muscles in the upper airway, including the oral, pharyngeal, laryngeal, and respiratory areas.[30] This physiological rationale is supported by Funderburke's study,[21] in which the choral singing component of ICST produced measurable improvements in hyolaryngeal excursion on videofluoroscopy, indicating a direct physiological benefit of singing-based training on swallowing safety.

Remarkably, music may also enhance enjoyment. The participatory and pleasurable aspects of music therapy have been shown to improve patient motivation and adherence to rehabilitation programs.[22] This is particularly critical in managing chronic illnesses such as Parkinson's disease, where long-term benefits depend on sustained therapy participation.[31] Specifically, choral singing has been incorporated into therapeutic programs to help individuals with Parkinson's disease manage consequences of the condition, including social isolation, low mood, and communication difficulties.[32] Music therapy in Parkinson's disease has been shown to improve not only voice and swallowing, but also emotional and social functioning. However, extended session durations may result in fatigue, reduced concentration, and decreased motivation.[21] Therefore, collaboration between speech therapists and music therapists is essential to determine appropriate session lengths tailored to the needs of individuals with Parkinson's disease.

In summary, our findings support the notion that, although speech therapy and music therapy are not primarily designed to rehabilitate swallowing, their training paradigms can indirectly enhance swallowing physiology. These results underscore the potential value of implementing integrated speech and music therapy programs more broadly in future rehabilitation efforts for individuals with Parkinson's disease.

Practical implications

Collaboration between speech therapists and music therapists is vital, as their complementary expertise optimizes patient care. In one study, the integrated intervention was delivered solely by a speech therapist,[21] while in another, the speech therapist provided the speech therapy component, and the music therapist delivered the music therapy component.[22] Their joint involvement is crucial for promoting safe swallowing practices and improving quality of life in patients with Parkinson's disease. We further suggest that the development of standardized speech therapy–music therapy protocols is needed to enhance replication and offer clinicians more precise practice guidance. Since both approaches are safe, effective, and non-pharmacological, they should receive more therapeutic recognition and promotion.

Limitations

This systematic review has several limitations that should be acknowledged. Firstly, the final review included only two studies with a combined sample of 36 participants, which limited the strength of the findings. In addition, the quasi-experimental study carried a high risk of selection bias due to the absence of a control group. Although the results are encouraging, further research employing RCT designs is needed to identify effective intervention types that can support the clinical integration of speech therapy and music therapy for patients with Parkinson's disease. Furthermore,

although integrating speech therapy with music therapy shows promising effects on swallowing in individuals with Parkinson's disease, the current evidence is too limited to draw firm conclusions, highlighting the need for further investigation.

Conclusion

In conclusion, based on our findings, it is not yet possible to determine the specific types of speech therapy interventions that are integrated with music therapy for patients with Parkinson's disease, largely due to the heterogeneity of the included studies. Nevertheless, the evidence indicates that combining speech therapy with music therapy shows promise for improving swallowing function in this population. Such an interdisciplinary approach has the potential to enhance rehabilitation outcomes and improve quality of life. Further research is required to establish standardized speech therapy–music therapy protocols that can strengthen the evidence base and support broader clinical implementation.

Ethical approval

This study had received ethical approval from The Ethics Committee, Faculty of Associated Medical Sciences, Chiang Mai University (CMU), Thailand (AMSEC-67EM-038**).**

Funding

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

Conflicts of Interest

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

CRediT authorship contribution statement

Natwipa Wanicharoen: conceptualization, data curation, formal analysis, investigation, methodology, project administration, supervision, visualization, writing: original draft, writing: review and editing; **Vich Boonrod**: conceptualization, data curation, formal analysis, investigation, methodology, writing: original draft; **Erin Songdechagraiwud**: data curation, formal analysis; **Palita Yaemsuan**: visualization, writing: review and editing, correspondence.

References

[1] Marras C, Beck JC, Bower JH, et al. Prevalence of Parkinson's disease across North America. NPJ Parkinsons Dis. 2018; 4: 21. doi:10.1038/s41531-018-0058-0.

[2] Parkinson's disease: Causes, symptoms, and treatments. National Institute of Aging. 2022. https://www.nia.nih.gov/health/parkinsons-disease/parkinsons-disease-causes-symptoms-and-treatments. (accessed Dec 23, 2024).

[3] Durcan R, Wiblin L, Lawson RA, et al. Prevalence and duration of non-motor symptoms in prodromal Parkinson's disease. Eur J Neurol. 2019; 26(7): 979-85. doi:10.1111/ene.13919.

[4] Sung HY, Kim JS, Lee KS, et al. The prevalence and patterns of pharyngoesophageal dysmotility in patients with early-stage Parkinson's disease. Mov Disord. 2010; 25(14): 2361-8. doi:10.1002/mds.23290.

[5] Ertekin C, Tarlaci S, Aydogdu I, et al. Electrophysiological evaluation of pharyngeal phase of swallowing in patients with Parkinson's disease. Mov Disord. 2002; 17(5): 942-9. doi:10.1002/mds.10240.

[6] Tjaden K. Speech and swallowing in Parkinson's disease. Top Geriatr Rehabil. 2008; 24(2): 115-26. doi:10.1097/01.TGR.0000318899.87690.44.

[7] Patel B, Legacy J, Hegland KW, Okun MS, Herndon NE. A comprehensive review of the diagnosis and treatment of Parkinson's disease dysphagia and aspiration. Expert Rev Gastroenterol Hepatol. 2020; 14(6): 411-24. doi:10.1080/17474124.2020.1769475.

[8] Kalf JG, de Swart BJ, Bloem BR, Munneke M. Prevalence of oropharyngeal dysphagia in Parkinson's disease: a meta-analysis. Parkinsonism Relat Disord. 2012; 18(4): 311-5. doi:10.1016/j.parkreldis.2011.11.006.

[9] Gong S, Gao Y, Liu J, et al. The prevalence and associated factors of dysphagia in Parkinson's disease: A systematic review and meta-analysis. Front Neurol. 2022; 13: 1000527. doi:10.3389/fneur.2022.1000527.

[10] Carneiro D, das Graças Wanderley de Sales Coriolano M, Belo LR, de Marcos Rabelo AR, Asano AG, Lins OG. Quality of life related to swallowing in Parkinson's disease. Dysphagia. 2014; 29(5): 578-82. doi:10.1007/s00455-014-9548-3.

[11] Chan HF, Ng ML, Kim H, Kim DY. Swallowing-related quality of life among oral-feeding Chinese patients with Parkinson's disease - a preliminary study using Chinese SWAL-QOL. Disabil Rehabil. 2022; 44(7): 1077-83. doi:10.1080/09638288.2020.1791979.

[12] Suttrup I, Warnecke T. Dysphagia in Parkinson's disease. Dysphagia. 2016; 31(1): 24-32. doi:10.1007/s00455-015-9671-9.

[13] Robertson C, Thomson F. Speech therapy in Parkinson's disease: a study of the efficacy and long-term effects of intensive treatment. Br J Disord Commun. 1984; 19(3): 213-24. doi:10.3109/13682828409029837.

[14] El Sharkawi A, Ramig L, Logemann JA, Pauloski BR, Rademaker AW, Smith CH, et al. Swallowing and voice effects of Lee Silverman Voice Treatment (LSVT®): a pilot study. J Neurol Neurosurg Psychiatry. 2002; 72(1): 31-6.

[15] Russell JA, Ciucci MR, Connor NP, Schallert T. Targeted exercise therapy for voice and swallow in persons with Parkinson's disease. Brain Res. 2010; 1341: 3-11. doi:10.1016/j.brainres.2010.03.029.
[16] Miles A, Jardine M, Johnston F, de Lisle M, Friary P, Allen J. Effect of Lee Silverman Voice Treatment (LSVT LOUD®) on swallowing and cough in Parkinson's disease: A pilot study. J Neurol Sci. 2017; 383: 180-7. doi:10.1016/j.jns.2017.11.015.
[17] Baijens LW, Speyer R. Effects of therapy for dysphagia in Parkinson's disease: systematic review. Dysphagia. 2009; 24(1): 91-102. doi:10.1007/s00455-008-9180-1.
[18] Stegemöller EL, Hibbing P, Radig H, Wingate J. Therapeutic singing as an early intervention for swallowing in persons with Parkinson's disease. Complement Ther Med. 2017; 31: 127-33. doi:10.1016/j.ctim.2017.03.002.
[19] Yeo MS, Hwang J, Lee HK, Kim SJ, Cho SR. Therapeutic singing-induced swallowing exercise for dysphagia in advanced-stage Parkinson's disease. Front Neurol. 2024; 15: 1323703. doi:10.3389/fneur.2024.1323703.
[20] Machado Sotomayor MJ, Arufe-Giráldez V, Ruíz-Rico G, Navarro-Patón R. Music therapy and Parkinson's disease: A systematic review from 2015-2020. Int J Environ Res Public Health. 2021; 18(21): 11618. doi: 10.3390/ijerph182111618.
[21] Funderburke SK. Effects of interdisciplinary choral singing on the voice and swallowing function of individuals with Parkinson's disease [master's thesis]. Gainesville: University of Florida; 2012.
[22] Mohseni Z, Saffarian A, Mohamadi R, Abolghasemi J, Habibi SA. Effect of conventional speech therapy combined with music therapy on swallowing in patients with Parkinson's disease (Telerehabilitation): a randomized-controlled trial. k. Middle East J Rehabil Health Stud 2023;10(1): e131572. doi:10.5812/mejrh-131572.
[23] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021; 372: n71. doi:10.1136/bmj.n71.
[24] Barker TH, Habibi N, Aromataris E, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for quasi-experimental studies. JBI Evid Synth. 2024; 22(3): 378-88. doi:10.11124/JBIES-23-00268.
[25] Barker TH, Stone JC, Sears K, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for randomized controlled trials. JBI Evid Synth. 2023; 21(3): 494-506. doi:10.11124/JBIES-22-00430.
[26] Kleim JA, Jones TA. Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. J Speech Lang Hear Res. 2008; 51(1): S225-S239. doi:10.1044/1092-4388(2008/018).
[27] Ramig L, Halpern A, Spielman J, Fox C, Freeman K. Speech treatment in Parkinson's disease: Randomized controlled trial (RCT). Mov Disord. 2018; 33(11): 1777-91. doi:10.1002/mds.27460.
[28] Ramig LO, Sapir S, Countryman S, et al. Intensive voice treatment (LSVT) for patients with Parkinson's disease: a 2 year follow up. J Neurol Neurosurg Psychiatry. 2001; 71(4): 493-8. doi:10.1136/jnnp.71.4.493.
[29] Ozdemir E, Norton A, Schlaug G. Shared and distinct neural correlates of singing and speaking. Neuroimage. 2006; 33(2): 628-35. doi:10.1016/j.neuroimage.2006.07.013.
[30] Jacob P, Kahrilas PJ, Logemann JA, Shah V, Ha T. Upper esophageal sphincter opening and modulation during swallowing. Gastroenterology. 1989; 97(6): 1469-78. doi:10.1016/0016-5085(89)90391-0.
[31] Winiker K, Kertscher B. Behavioural interventions for swallowing in subjects with Parkinson's disease: A mixed methods systematic review. Int J Lang Commun Disord. 2023; 58(4): 1375-1404. doi:10.1111/1460-6984.12865.
[32] Fogg-Rogers L, Buetow S, Talmage A, et al. Choral singing therapy following stroke or Parkinson's disease: an exploration of participants' experiences. Disabil Rehabil. 2016; 38(10): 952-62. doi:10.3109/09638288.2015.1068875.

Journal of Associated Medical Sciences 2026; 59(1): 121-128

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 stress, anxiety, depression, and reported self-esteem among mothers of children with autism spectrum disorder

Punitha Poyyaamozhi[1*], Leonard Joseph[2], Shanthi Nambi[3], K Raman[4], Inbadurai Palaniswami[5], V Rewanthkumar[6], Christine Muya[7]

[1]*Saveetha college of occupational therapy, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India.*
[2]*Centre of Regenerative Medicine and Devices, School of Education, Sports and Health Sciences, University of Brighton, East Sussex, UK.*
[3]*Department of Psychiatry, Saveetha Medical College and Hospital, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India.*
[4]*Department of Psychiatry, Saveetha Medical College and Hospital, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India.*
[5]*Rashid hospital, Dubai Health, United Arab Emirates.*
[6]*Lumos Child Development Centre, Chennai, Tamil Nadu, India.*
[7]*Department of Rehabilitation Sciences, Amref International University, Nairobi, Kenya.*

ARTICLE INFO

Article history:
Received 22 April 2025
Accepted as revised 9 October 2025
Available online 16 October 2025

Keywords:
Mothers, autism spectrum disorder, stress, anxiety, depression, self-esteem.

ABSTRACT

Background: Mothers of children with autism spectrum disorder (ASD) have an increased risk of depression, anxiety, stress, and diminished self-esteem due to the numerous challenges involved in raising children with socio-behavioural difficulties.

Objectives: This study investigated the prevalence of depression, anxiety, stress, and low self-esteem among mothers of children with ASD.

Materials and methods: A cross-sectional descriptive study was conducted among 246 mothers of children with ASD. The prevalence of depression, anxiety, and stress was assessed using the Depression, Anxiety and Stress Scale-21 Items, and self-esteem was measured using Rosenberg's Self-Esteem Scale. Demographic data for both mothers and children were collected using a pro forma. Data are summarised as frequencies, means (M), standard deviations (SD), and percentages.

Results: The prevalence rates among mothers were as follows: depression, 84.14% (N=207); anxiety, 77.2% (N=190); stress, 74% (N=167); and diminished self-esteem, 67.9% (N=167). Mean severity scores indicated moderate levels of depression (M=17.26, SD=7.7), anxiety (M=12.7, SD=7.7), and stress (M=19.49, SD=8.0), and low self-esteem (M=23.25, SD=9.5).

Conclusion: Mothers of children with ASD had high prevalence rates of depression, anxiety, stress, and low self-esteem. These findings highlight the psychological challenges faced by mothers caring for children with ASD. Therapeutic interventions should therefore address the mental health needs of these mothers alongside the rehabilitation of their children.

* **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: 10.12982/JAMS.2026.015
E-ISSN: 2539-6056

Introduction

Parents of children with autism spectrum disorder (ASD) face numerous challenges, including supporting other family members, meeting the needs of siblings, and caring for the child with ASD. Additional difficulties include accessing appropriate health and childcare services, securing suitable educational opportunities

for their children with ASD, and managing the social impact on the family.[1] Mothers of children with ASD often experience greater challenges than fathers in raising their children, irrespective of the children's maladaptive behaviours, communication deficits, or social skill limitations, and have higher levels of stress, anxiety, and depression than the child's father.[2]

A meta-analysis on parenting stress indicated that parents of children with ASD experience higher levels of depression, anxiety, and stress than the parents of children with other developmental disorders.[3] Mothers raising children with ASD are often physically exhausted due to the constant supervision their children require.[4] They also report emotional difficulties such as denial, distress, sadness, and worry, while facing social challenges including criticism, isolation, and ignorance from family members and society in response to their child's atypical behaviour.[5] Furthermore, many parents are compelled to leave their jobs because of the demands of caring for a child with ASD. When combined with high healthcare costs, this loss of income creates significant economic challenges for families, further contributing to parental stress.[6]

A study conducted in Ghana reported that parents of children with ASD faced difficulties in holding meaningful conversations with their children, financial burdens related to their care, stigmatisation associated with autism, and the demands of their children.[7] Social stigma was particularly common among mothers of children with ASD, influenced by cultural factors, especially in Eastern and Middle Eastern countries.[8] Parental stress was primarily attributed to the burden of childcare, challenges in managing family relationships, limited support systems, financial pressures, and the social stigma surrounding childhood disability.[9] The high stress levels experienced by mothers also affected their physical health, psychological well-being, personal relationships, interactions with society, and financial circumstances.[10] Psychological well-being is a vital component of overall health for these mothers, encompassing dimensions such as stress, depression, anxiety, and self-esteem, which collectively affect an individual's mental health and quality of life.[11]

In the Indian subcontinent, many parents have reported high levels of post-diagnosis stress, poorer parental functioning, increased marital conflict, difficulties with family socialisation, and feelings of inadequacy.[12] Among mothers of children with ASD, stress levels tend to rise as the caregiver's economic condition worsens.[13] In addition to financial pressures, most caregivers and mothers have expressed concern about a lack of family support.[14] A study conducted in India reported that mothers of children with ASD have an increased risk of psychological distress, which is often overlooked and insufficiently addressed when planning long-term child management.[15]

In summary, various factors such as psychological distress, financial concerns, lack of social and family support along with feeling of isolation increases the care burden of the mothers contributing to concerns on their health and wellbeing. There is, therefore, a pressing need to develop health and well-being programmes for mothers of autistic children. However, before any local intervention can be designed, it is essential to understand the scope and extent of these problems in the local cultural context. This study determined the prevalence of depression, anxiety, stress, and low self-esteem among mothers of children diagnosed with ASD in an urban community in India. The findings of this study can provide clear and locally relevant data on their psychological well-being and can guide the creation of targeted strategies to support these mothers' mental health while also contributing to the effective rehabilitation of their children with ASD.

Materials and methods

Study design and setting

This cross-sectional descriptive study was conducted in and around Chennai, Tamil Nadu, India, between 2023 and 2024. A simple random sampling method was used to select 14 paediatric therapy centres and 5 special schools from a sampling frame of 45 major paediatric therapy centres and 7 government-aided special schools. The study information sheet and details of the procedures were provided to the managing directors of the centres and principals of the government-aided special schools to facilitate participant recruitment. Mothers who provided informed consent and agreed to participate were included in the study. Written consent was obtained from all participants before data collection. This study was approved by the university's institutional ethics committee (approval number: 004/06/2023/IEC/SMCH).

Participant characteristics

This study recruited 246 mothers of children with ASD who met the specified inclusion and exclusion criteria. The inclusion criteria were as follows: 1) mothers of children with ASD, 2) age between 25 and 45 years for mothers, 3) age between 3 and 10 years for children with ASD, and 4) children with ASD who received occupational therapy and other therapies, and (5) Residing in and around Chennai. Mothers were excluded if they had severe physical health issues, a history of psychological disorders and were undergoing medication, were not accompanying their children with ASD, had discontinued therapy for their children, or were pregnant at the time of the survey. It was agreed that the final study report and a summary of the findings would be shared with the managing directors, principals, and participants because they would directly benefit from the results. The research findings would also be disseminated to all participating mothers and the occupational therapists providing therapy to the children through the creation of a closed community social media group.

Measurement

A study proforma was developed to collect the demographic characteristics of the mothers and their children with ASD. For the children, information gathered included age, sex, education, number of siblings, duration of therapy sessions, time since ASD diagnosis, communication abilities, and level of dependence in activities of daily living (ADL). The severity of autism was assessed using the Indian Scale for Assessment of Autism (ISAA).[16] For the mothers, data collected included age, educational qualification, family structure, economic status based on the Standard of Living Index (SLI),[17] employment status, marital status, place of residence, exposure to social media, support from the husband and other family members, presence of trustworthy friends, and availability of personal time.

Depression, Anxiety and Stress Scale-21 Items

The Depression, Anxiety, and Stress Scale–21 Items (DASS-21) was used to assess the levels of depression, anxiety, and stress experienced by mothers caring for their children with ASD. The DASS-21 consists of three self-report subscales, each comprising seven items, designed to measure depression, anxiety, and stress. The depression subscale evaluates constructs such as dysphoria, hopelessness, devaluation of life, self-deprecation, reduced interest or engagement, anhedonia, and inertia. The anxiety subscale assesses aspects including autonomic arousal, skeletal muscle tension, situational anxiety, and the subjective experience of anxious affect. The stress subscale measures chronic, non-specific arousal, including difficulty relaxing, nervous excitability, and a tendency to become easily distressed, agitated, irritable, overly reactive, or impatient. Scores for depression, anxiety, and stress are calculated by summing the relevant item scores, with results interpreted as normal, mild, moderate, or severe according to the established cutoffs.[18]

Rosenberg Self-Esteem Scale

Self-esteem is defined as a positive or negative attitude towards oneself[19] and can also be described as an individual's sense of self-worth.[20] The Rosenberg Self-Esteem Scale (RSS), developed by Rosenberg, is used to assess an individual's level of self-esteem.[21] The RSS comprises 10 items, five expressed as positive statements and five as negative statements. Participants respond by selecting one of four options: strongly disagree, disagree, agree, or strongly agree. Scores range from a minimum of 10 to a maximum of 40, with self-esteem levels categorised as low (10-25), medium (26-29), or high (30-40).[22]

Statistical analysis

The sample size for the study was calculated using the Krejcie and Morgan table.[23] The sampling frame comprised approximately 700 mothers drawn from paediatric centres and special schools. Based on the Krejcie and Morgan table, a sample size of 248 was required. Data were analysed using SPSS version 26.0 (IBM, New York, USA). The Kolmogorov–Smirnov test confirmed the normality of the data distribution. Demographic data were summarised using frequencies and percentages. Descriptive statistics were applied to present the prevalence of depression, anxiety, stress, and self-esteem among mothers of children with ASD. Data on these variables were summarised in terms of frequency, mean (M), standard deviation (SD), and percentage.

Results

Table 1. presents the demographic characteristics of the 246 mothers of children diagnosed with ASD. Of these, 116 (47%) were aged between 31 and 35 years, making this the largest age group. The majority, 194 (78.9%), held a degree-level qualification. In terms of family structure, 187 (76%) lived in a nuclear family. Economic status data revealed that 163 (66.3%) were classified as having a medium economic standing. Employment information revealed that 133 (54.1%) were homemakers, outnumbering those in active employment or those who had withdrawn from the workforce. Marital status data indicated that 237 (96.3%) were married, whereas nine were separated. Geographically, 203 (82.5%) lived in urban areas. Regarding social media use, 110 (44.7%) were passive observers, 25.6% were active participants, and 29.7% were inactive. Support networks appeared strong, with 211 (85.8%) receiving spousal support and 198 (80.5%) receiving additional family support. In terms of social connections, 175 (71.1%) reported having trustworthy friends. Personal time availability varied, with 156 (63.4%) stating they had time for themselves, whereas 90 (36.6%) reported having none. Awareness of their child's condition was present in 174 (70.7%) of the mothers.

Table 1. Demographic characteristics of Mothers of Children with ASD

Total Number of participants (mothers), N=246			
		Nos	**Percentage**
Age of mother	25-30	88	36%
	31-35	116	47%
	36-40	42	17%
Educational qualification	School	52	21.1%
	Degree	194	78.9%
Family status	Joint family	59	24%
	Nuclear family	187	76%
Economic status (based on Standard of Living Index scale) (SLI)	Low	32	13%
	Medium	163	66.3%
	High	51	20.7%
Employment status	Working	71	28.9%
	Withdrawal	42	17.1%
	Home maker	133	54.1%
Marital status	Married	237	96.3%
	Separated	9	3.7%
Place of living	Rural	43	17.5%
	Urban	203	82.5%
Exposure to social media	Active	63	25.6%
	Inactive	73	29.7%
	Silent observer	110	44.7%
Husband's support to mother	Yes	211	85.8%
	No	35	14.2%
Other family member's Support	Yes	198	80.5%
	No	48	19.5%
Mother with trust worthy friends	Yes	175	71.1%
	No	71	28.9%
Mothers personal time	No time	90	36.6%
	Yes	156	63.4%
Mother's knowledge of child's condition	Known	174	70.7%
	Unknown	72	29.3%

Table 2. presents the demographic characteristics of the children diagnosed with ASD whose mothers participated in this study. The majority, 156 (63.4%), were aged between 3 and 5 years. Of the 246 children, 182 (74.0%) were boys. In terms of severity, 161 (65.5%) were classified as having a moderate level of autism. Educational settings varied widely, ranging from no formal schooling to inclusive education. A total of 206 (82.7%) children were receiving regular therapeutic interventions. Among the cohort, 104 (42.3%) had been diagnosed with ASD within the past year. Communication abilities were almost evenly distributed, with 122 (49.6%) exhibiting verbal communication and 125 (50.4%) demonstrating non-verbal communication. Furthermore, 188 children were reported to be the only child in their family. Regarding functional independence, 130 (52.8%) were fully dependent on their caregivers for ADL, whereas 116 (47.2%) were partially dependent. Figure 1 illustrates the prevalence rates of depression, anxiety, stress, and self-esteem among mothers of children with ASD. The rates observed were 84.14% (N=207) for depression, 77.2% (N=190) for anxiety, 74% (N=167) for stress, and 67.9% (N=167) for low self-esteem. Figure 2 presents the severity levels of depression, anxiety, stress, and self-esteem among the participants. Based on the DASS-21, the mothers demonstrated moderate levels of depression (M=17.26, SD=7.72), anxiety (M=12.70, SD=7.71), and stress (M=19.49, SD=8.09). According to the RSS, they exhibited decreased self-esteem (M=23.25, SD=9.59).

Table 2. Demographic characteristics of children with ASD of participating mothers

		Number	Percentage
Child's age	3 to 5	156	63.4%
	5 to 7	90	36.6%
Child's Gender	Male	182	74%
	Female	64	26%
Severity of Autism (ISAA)	Mild	48	19.5%
	Moderate	161	65.5%
	Severe	37	15%
Child's education mode	Inclusive	68	27.6%
	No schooling	28	11.4%
	Regular stream	93	37.8%
	Special	57	23.2%
Therapy participation	Regular	206	83.7%
	Irregular	40	16.3%
Duration of Condition from Diagnosis	1 year	104	42.3%
	2 years	84	34.1%
	3 Years	41	16.7%
	4 Years	17	6.9%
Communication Level of the Child	Verbal	122	49.6%
	Non verbal	124	50.4%
Siblings for child	Any sibling	7	2.8%
	First child	51	20.7%
	Only child	188	76.5%
Child's dependency in ADLs on mother	Fully dependent	130	52.8%
	Partially ependent	116	47.2%

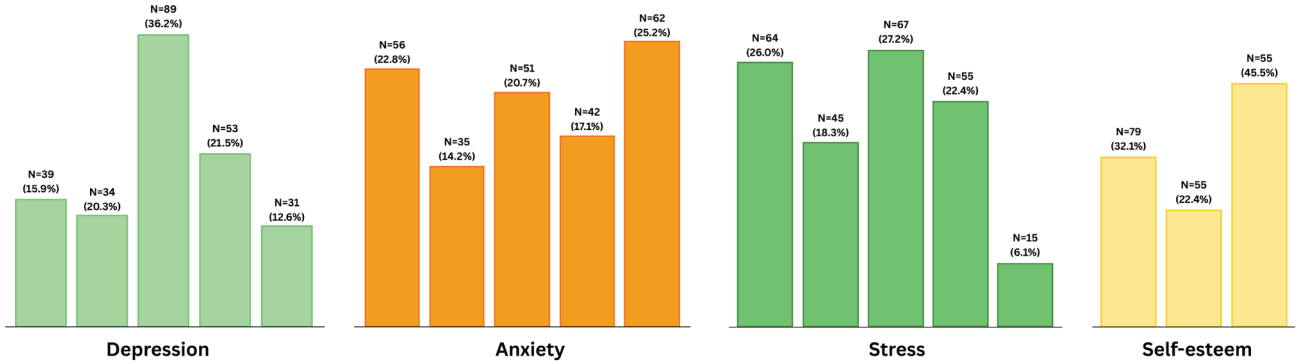


Figure 1. *illustrates the prevalence rates of depression, anxiety, stress, and self-esteem among mothers of children with ASD.*

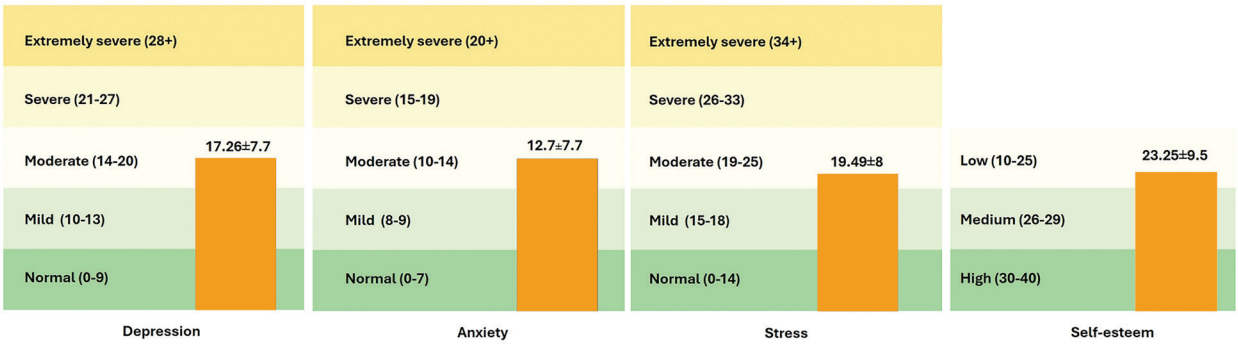


Figure 2. *presents the severity levels of depression, anxiety, stress, and self-esteem among the participants.*

Discussion

This cross-sectional study determined the prevalence of depression, anxiety, stress, and self-esteem among mothers of children diagnosed with autism spectrum disorder (ASD) and revealed substantial impacts on their psychological well-being. Notably, 84.14% (N=207) of participants experienced depression of varying severity, 77.2% (N=190) reported symptoms of anxiety, and 74% (N=182) experienced stress, as measured using the DASS-21. In addition, 67.9% (N=167) had self-esteem levels ranging from low to medium. Moderate to severe depression, anxiety, and stress were observed in 64%, 63%, and 55% of the cohort, respectively. These findings indicate the significant mental health challenges faced by mothers of children with ASD and highlight the complex interplay between caregiving responsibilities and psychological health.

The high prevalence of depression observed in this study, 84.14% (N=207) with a mean severity score of 17.26 (moderate level), is consistent with previous research. For example, Bramhereported moderate to severe depression among mothers of children with ASD, attributing it to the chronic emotional strain of caregiving.[24] Selvakumar *et al.*[15] similarly noted that the increased workload and stress associated with raising a child with ASD exacerbate depressive symptoms.[15] In addition, Secinti *et al.* identified internalised stigmatisation and the behavioural characteristics of children with ASD as significant contributors to caregiver depression.[25] Collectively, these findings suggest that depression in this population result from both internal factors (such as caregiving demands) and external pressures (such as stigma), indicating the need for a multifaceted approach to intervention.

In the present study, anxiety was identified as a particularly prominent issue, affecting 77.2% of the mothers, with a mean severity score of 12.70. This finding is consistent with previous research. This high prevalence may be attributable to the unpredictable nature and intensity of ASD symptomatology, compounded by limited social support.[26,27] Similarly, Falk *et al.* reported that decreased social support and greater severity of ASD symptoms in children significantly increase maternal anxiety levels.[27] The findings of this study support these observations, suggesting that anxiety is a primary psychological response to the ongoing uncertainty and social isolation faced by these mothers.

Stress was reported by 74% of participants, with a mean severity score of 19.49 (moderate level), highlighting the considerable demands of caregiving. Hill *et al.* observed that challenging behaviours in children with ASD, such as tantrums or aggression, significantly

increase parental stress,[28] a finding supported by Miranda *et al*. who reported an association of higher ASD symptom severity and behavioural difficulties with increased stress levels.[29] In this study, 55.7% of mothers experienced moderate to extreme stress, reflecting the cumulative effects of these factors and suggesting that stress in this population both precedes and results from depression and anxiety.

The finding that 67.9% of mothers had low to medium self-esteem, with a mean score of 23.25 (low level), adds an important dimension to this research. The ongoing demands of caregiving and the societal stigma associated with raising a child with ASD may erode maternal self-worth, particularly in the absence of adequate support or when faced with judgement. This observation is consistent with the wider literature on caregiver burden, which indicates that prolonged stress and social isolation frequently reduce self-esteem.[25] Similarly, Bawalesh reported decreased self-esteem among parents of children with ASD, with notably lower levels among older parents (aged 50-59 years) and those with highly educated children.[30] Cultural expectations can further complicate these dynamics; in many societies, children are regarded as essential to family completeness and a source of happiness. When a child is born with a disability or developmental challenge such as ASD, these expectations are disrupted, often leading to negative effects on family well-being.[31] In such circumstances, the home environment may become characterised by stress and dissatisfaction.[27] Thus, decreased self-esteem among parents may be an expected outcome of raising a child with ASD.[33] These findings are in line with previous studies, including those by Perumal *et al*. and Verte *et al*. both of which highlight the adverse effects of an ASD diagnosis on parents' psychological well-being and overall family functioning.[31,32]

This study has a few limitations. It relied exclusively on two self-report questionnaires to determine the psychological well-being of mothers of children with ASD. However, these instruments were valid and produced reliable findings on the psychological health of mothers caring for children with ASD. A key strength of the study is that it offers an accurate representation and locally contextualised evidence on the psychological health of mothers caring for children with ASD. Furthermore, it establishes baseline data that can inform the development of psychological health interventions for mothers. Future research could broaden the scope by incorporating a wider range of assessment tools, exploring mothers' experiences through qualitative methods, and extending the investigation to include fathers, who are also affected by their child's behaviours.

Conclusion

Mothers caring for children with ASD in an urban community in India exhibited high prevalence rates of stress, anxiety, depression, and low self-esteem. In the Indian context, multiple factors were identified as contributing to these psychological challenges. However, the underlying causes driving the observed increases in stress, anxiety, depression, and reduced self-esteem remain insufficiently explored. Further research may benefit from employing inferential statistical analyses to investigate potential relationships and identify significant predictors. Future research is needed to examine the risk factors that affect the psychological health and well-being of these mothers. Additional studies should also aim to generate evidence to strengthen their psychological well-being. Such work would support the development of tailored interventions to improve mental health, enabling mothers to lead more fulfilling lives and provide better support to their children with ASD.

Ethical approval

This study has been approved by the Institutional Ethics Committee, Saveetha Medical College Hospital, Saveetha Institute of Medical and Technical Sciences, number- 004/06/2023/IEC/SMCH.

Funding

No funding was received for this study.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

All authors contributed to the study design and interpretation of results, reviewed and approved the final manuscript; **Punitha Poyyaamozhi**: data collection, drafted of the manuscript and coordinated revisions.

Acknowledgements

We would like to thank Dr M. Arun Kumar, Principal of Saveetha College of Occupational Therapy, for his support and guidance. We are also grateful to the directors and principals of the schools and centres who agreed to our request to conduct this survey among their clients as well as to all the mothers of children with ASD who consented to complete the questionnaires. Our heartfelt thanks also go to our friends and family for their constant support and assistance throughout this process.

Reference

[1] Ooi KL, Ong YS, Jacob SA, Khan TM. A meta-synthesis on parenting a child with autism. Neuropsychiatr Dis Treat. 2016; 12: 745-62. doi: 10.2147/NDT.S100634.

[2] Johnson N, Frenn M, Feetham S, Simpson P. Autism spectrum disorder: parenting stress, family functioning and health-related quality of life. Fam Syst Health. 2011; 29(3): 232-52. doi: 10.1037/a0025341. PMID: 21928891.

[3] 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.

[4] Giallo R, Wood CE, Jellett R, Porter R. Fatigue, wellbeing and parental self-efficacy in mothers of children with an autism spectrum disorder. Autism. 2013; 17(4): 465-80. doi: 10.1177/1362361311416830.

[5] Acharya S, Sharma K. Lived experiences of mothers raising children with autism in Chitwan District, Nepal. Autism Research and Treatment. 2021; 6614490: 12 ages, 2021. doi: 10.1155/2021/6614490.

[6] Tathgur MK, Kang HK. Challenges of the caregivers in managing a child with autism spectrum disorder-A qualitative analysis. Indian J Psychol Med. 2021; 43(5): 416-21. doi: 10.1177/02537176211000769.

[7] Agyekum HA. Challenges and coping strategies for parents with autistic children. J Ment Disord Treat. 2018; 4: 166. doi:10.4172/2471- 271X.1000166.

[8] Thi LK, Mai TLA, Phương AH. Stigma among parents of children with autism: An integrative review. PRIJNR. 2023; 27(3): 530-48. Available from: https://he02.tci-thaijo.org/index.php/PRIJNR/article/view/261650.

[9] Malhi P, Shetty AR, Bharti B, Saini L. Parenting a child with autism spectrum disorder: A qualitative study. Indian J Public Health. 2022; 66(2): 121-7. doi: 10.4103/ijph.ijph_1779_21. PMID: 35859492.

[10] Shahbaz T, Khalid S, Ali A, Yaqoob S. Parenting stress among parents of children with autism spectrum disorder. JHRR. 2023; 3(2): 724-9. Available from: https://jhrlmc.com/index.php/home/article/view/214.

[11] World Health Organization. Mental health [Internet]. World Health Organization. World Health Organization; 2022. Available from: https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response.

[12] Jose P, Ravindranath R, Joseph LM, Rhodes EC, Ganapathi S, Harikrishnan S, et al. Patient, caregiver, and health care provider perspectives on barriers and facilitators to heart failure care in Kerala, India: A qualitative study. Wellcome Open Res. 2021; 5: 250. doi: 10.12688/wellcomeopenres.16365.2.

[13] Ramachandran A, Vyas N, Pothiyil DI. Stress among the caregivers of mentally disabled children visiting a rehabilitation centre in Chennai, Tamil Nadu – A cross-sectional study. Clin Epidemiol Glob Health. 2020; 8(4): 1155-7. doi:10.1016/j.cegh.2020.04.008.

[14] Minhas A, Vajaratkar V, Divan G, Hamdani SU, Leadbitter K, Taylor C, et al. Parents' perspectives on care of children with autistic spectrum disorder in South Asia - Views from Pakistan and India. Int Rev Psychiatry. 2015; 27(3): 247-56. doi: 10.3109/09540261.2015.

[15] Selvakumar N, Panicker AS. Stress and coping styles in mothers of children with autism spectrum disorder. Indian J Psychol Med. 2020; 42(3): 225-32. doi: 10.4103/IJPSYM.IJPSYM_333_19.

[16] Indian Scale for Assessment of Autism (ISAA): National Institute for the Mentally Handicapped. (n.d.). Indian Scale for Assessment of Autism (ISAA) Manual. An ISO 9001:2000 Institution. Retrieved from ISAA Test Manual.

[17] Pandey, A. K. and Dixit, A. (2008) "Standard of Living Index (SLI) and Inequalities in the Rural Areas of Uttar Pradesh", Journal of Global Economy, 4(1), pp. 15–27. doi: 10.1956/jge.v4i1.130.

[18] Lovibond PF, Lovibond SH. The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther. 1995; 33(3): 335-43. doi: 10.1016/0005-7967(94)00075-u.

[19] Çikrıkci Ö. The effect of internet use on well-being: Meta-analysis. Comput Human Behav. 2016; 65: 560-6. doi: 10.1016/j.chb.2016.09.021.

[20] Crocker J, Wolfe CT. Contingencies of self-worth. Psychol Rev. 2001; 108(3): 593-623. doi: 10.1037/0033-295x.108.3.593.

[21] Mann M, Hosman CM, Schaalma HP, de Vries NK. Self-esteem in a broad-spectrum approach for mental health promotion. Health Educ Res. 2004; 19(4): 357-72. doi: 10.1093/her/cyg041.

[22] García JA, Y Olmos FC, Matheu ML, Carreño TP. Self esteem levels vs global scores on the Rosenberg self-esteem scale. Heliyon. 2019; 5(3): e01378. doi: 10.1016/j.heliyon.2019.e01378.

[23] Krejcie RV, Morgan DW. Determining sample size for research activities. Educ Psychol Meas. 1970; 30(3): 607-10. doi: 10.1177/001316447003000308.

[24] Bramhe V. Depression and anxiety in mothers of children with autism spectrum disorders and intellectual disabilities. Indian J Ment Health. 2019; 6(3): 272-9. doi: 10.30877/IJMH.6.3.2019.272-279.

[25] Dilara DS, Diş D, Albayrak ZS, *et al*. Depression and parental distress among caregivers of autistic children: a serial mediator analysis in caregivers of autistic children. BMC Psychol. 2024; 12: 339. doi: 10.1186/s40359-024-01704-x.

[26] Sinha R, Rajpoot K, Uniyal A, Syed VS. Perceived mental health and quality of life in mothers of children with autism spectrum disorder in armed forces in a tertiary care center of North India: A Cross-sectional study. Ann Child Neurol. 2024; 32(4): 245-9. doi.org/10.26815/acn.2024.00612

[27] Altourah AJ, Al Ansari AM, Jahrami HA. Depression, anxiety and stress among mothers of children with autism spectrum disorder. Bahrain Med Bull. 2020; 42(2): 125-8. https://www.researchgate.net/

publication/341592730.

[28] Falk NH, Norris K, Quinn MG. The factors predicting stress, anxiety and depression in the parents of children with autism. J Autism Dev Disord. 2014; 44(12):3185-203. doi: 10.1007/s10803-014-2189-4. PMID: 25022253.

[29] Hill AP, Zuckerman KE, Hagen AD, Kriz DJ, Duvall SW, van Santen J, et al. Aggressive behaviour problems in children with autism spectrum disorders: Prevalence and correlates in a large clinical sample. Res Autism Spectr Disord. 2014; 8(9): 1121-33. doi: 10.1016/j.rasd.2014.05.006.

[30] Miranda A, Mira A, Berenguer C, Rosello B, Baixauli I. Parenting stress in mothers of children with autism without intellectual disability. Mediation of behavioral problems and coping strategies. Front Psychol. 2019; 10: 464. doi: 10.3389/fpsyg.2019.00464.

[31] Bawalsah J. Self-esteem in parents of children diagnosed with autism spectrum disorder. CIER. 2015; 8(3): 215-29. doi: 10.19030/cier.v8i3.9415.

[32] Perumal V, Veeraraghavan V, Lekhra OP. Quality of life in families of children with autism spectrum disorder in India. J Pharm Res. 2014; 8: 791-7. doi: 10.13140/2.1.4813.6320.

[33] Verté S, Roeyers H, Buysse A. Behavioural problems, social competence and self-concept in siblings of children with autism. Child Care Health Dev. 2003; 29(3): 193-205. doi: 10.1046/j.1365-2214.2003.00331.x.

Journal of Associated Medical Sciences 2026; 59(1): 129-136

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 on the prevalence of unexpected alloantibodies in Thai patients across Thailand

Udom Lao-on[1,2*] and Kunwadee Lao-on[1]

[1]*Department of Medical Technology, Faculty of Allied Health Sciences, Thaksin University, Phatthalung Province, Thailand.*
[2]*Department of Medical Technology, School of Allied Health Sciences, Walailak University, Nakhon Si Thammarat Province, Thailand.*

ARTICLE INFO

Article history:
Received 23 July 2024
Accepted as revised 30 September 2025
Available online 22 October 2025

Keywords:
Alloantibody, alloimmunization, antibody identification, hemolytic transfusion reaction.

ABSTRACT

Background: Alloantibodies are immune responses produced by exposure to foreign red blood cell antigens. However, sensitized alloantibodies can lead to hemolytic transfusion reactions and complicate cross-matching procedures.

Objectives: This review article aims to demonstrate the prevalence of unexpected alloantibodies across Thailand's four major geographical regions.

Materials and methods: A comprehensive literature search was conducted across PubMed, Google Scholar, and ThaiJo databases using keywords related to unexpected alloantibodies in Thailand, yielding 17 articles categorized by geographical region. Studies were filtered using strict criteria (published 2013-2025, sample size > 100, reporting > 2 alloantibody types) for final data analysis.

Results: It is revealed that anti-Mi^a is the highest unexpected alloantibody, with a prevalence rate of 25.9-37.2% across all four regions, followed by anti-E (15.6-20.7%), anti-Le^a (4.7-14.6%), and anti-Le^b (4.6-8.0%), respectively. Interestingly, anti-P1 is remarkably increased in the Northern region (15.1%), whereas anti-Jk^a (4.7%) and anti-Jk^b (2.5%) are specifically higher in the Northeastern region.

Conclusion: This review highlights the importance of practicing antibody screening protocols, choosing compatible antigen status, and establishing rare donor registries in Thailand to avoid unexpected alloimmunization, thereby ensuring optimal transfusion safety.

Introduction

Unexpected alloantibodies in blood transfusion

Blood transfusion is a vital component of modern medical care, used to manage acute hemorrhage, support patients undergoing major surgery, and treat chronic hematologic conditions such as thalassemia and hemophilia.[1,2] While the compatibility of ABO and Rh(D) blood groups is routinely assessed prior to transfusion, a broader spectrum of antigens from various blood group systems can also provoke immune responses. The presence of unexpected alloantibodies, also known as irregular antibodies, is an important obstacle to safe transfusion practices, especially in patients who require regular transfusions.[3]

Unexpected alloantibodies are antibodies that form in response to exposure to foreign red cell antigens that are not naturally present in the individual's own red blood cells.[4-6] These antibodies are termed "unexpected" because they are not ABO (non-red

* **Corresponding contributor.**
Author's Address: *Department of Medical Technology, School of Allied Health Sciences, Walailak University, Nakhon Si Thammarat Province, Thailand.*
E-mail address: udom.la@mail.wu.ac.th
doi: 10.12982/JAMS.2026.016
E-ISSN: 2539-6056

cell stimulating) antibodies and typically arise only after sensitization through transfusion, pregnancy, or occasionally organ transplantation.[7] Unlike non-red cell stimulating antibodies (like anti-A, B), alloantibodies are produced when a person has been exposed to foreign red cell antigens from systems such as Rh (like C, c, E, e), Kell, Duffy, Kidd, MNS, and others.[8]

The mechanism of alloantibody development is immunologic. When an individual is exposed to red cell antigens that their immune system recognizes as foreign, an immune response is mounted. Antigen-presenting cells process the foreign antigen and activate helper T cells, which in turn stimulate B cells to produce specific antibodies.[9,10] These antibodies, primarily of the IgG type, can traverse the placenta or interact with transfused red cells, thereby causing hemolysis. Importantly, once sensitized, a patient can mount an anamnestic response upon re-exposure to the same antigen, leading to a delayed hemolytic transfusion reaction (DHTR).[11]

The clinical implications of unexpected alloantibodies are considerable. They can cause acute or delayed hemolytic transfusion reactions, complicate cross-matching procedures, and limit the availability of compatible blood units.[12] In multi-transfused patients, such as those with thalassemia major, hemophilia, hematologic malignancies, or chronic renal failure, the risk of alloimmunization increases with each exposure to antigenically diverse donor red blood cells.[13] Alloantibodies also pose a risk in pregnant women, potentially leading to hemolytic disease of the fetus and newborn (HDFN) if maternal antibodies cross the placenta and attack fetal red cells.[14]

Antibody screening and identification methods according to Thai National Blood Center Standards

The detection and proper identification of unexpected alloantibodies are essential for ensuring transfusion safety. Alloimmunized patients are at increased risk of hemolytic transfusion reactions (HTRs) if incompatible red cell units are inadvertently transfused. Even clinically "weak" antibodies can lead to serious or delayed complications. Therefore, pre-transfusion antibody screening, particularly in patients with a history of transfusion, pregnancy, or autoimmune disease, is vital to avoid incompatible transfusions.[15] The significance is even greater in chronically transfused populations, such as patients with thalassemia major or myelodysplastic syndromes, where repeated antigen exposure amplifies the risk of antibody formation and transfusion complications.[6]

Moreover, the presence of unexpected antibodies can significantly delay the provision of compatible blood, especially when rare antibodies or multiple alloantibodies are involved. This delay can be life-threatening in emergencies where rapid transfusion is necessary.[16]

Antibody screening and identification procedures are critical components of safe blood transfusion practices, designed to detect unexpected antibodies that may cause hemolytic transfusion reactions. The National Blood Centre, Thai Red Cross Society, has established standardized protocols based on the indirect antiglobulin test (IAT) to ensure consistent and reliable results across all blood banking facilities,[17,18] as illustrated in Figure 1. The antibody screening process involves testing patient or donor serum against a panel of 2-3 group O screening cells with known antigen profiles using a three-phase approach: room temperature (immediate spin), 37 °C incubation (15-30 minutes), and the antiglobulin phase using anti-human globulin (AHG).[19] Each phase is followed by centrifugation at 1000 x g for 15 seconds, with careful washing steps using normal saline before the AHG phase. Results are graded from 0 to 4+ based on the degree of agglutination observed,[20] with quality control measures including the use of Coombs' control cells (IgG-sensitized red cells) to validate negative results.

When antibody screening yields positive results (≥1+), antibody identification must be performed using an extended panel of 8-16 group O cells with comprehensive antigen profiles covering common and uncommon blood group antigens.[19] The identification process follows the three-phase IAT or cooperates with additional techniques such as enzyme treatment (papain or ficin) employed when necessary to enhance certain antibody reactions or distinguish between different antibody specificities. Analysis follows the "rule of three" principle, requiring at least three antigen-positive cells to react positively, while antigen-negative cells must remain non-reactive for definitive antibody identification. The procedure may be extended with selected cells, DTT treatment to differentiate IgM from IgG antibodies, and auto-control testing to rule out autoantibodies.[21] These standardized protocols ensure the safety and efficacy of blood transfusion practices throughout Thailand's healthcare system.

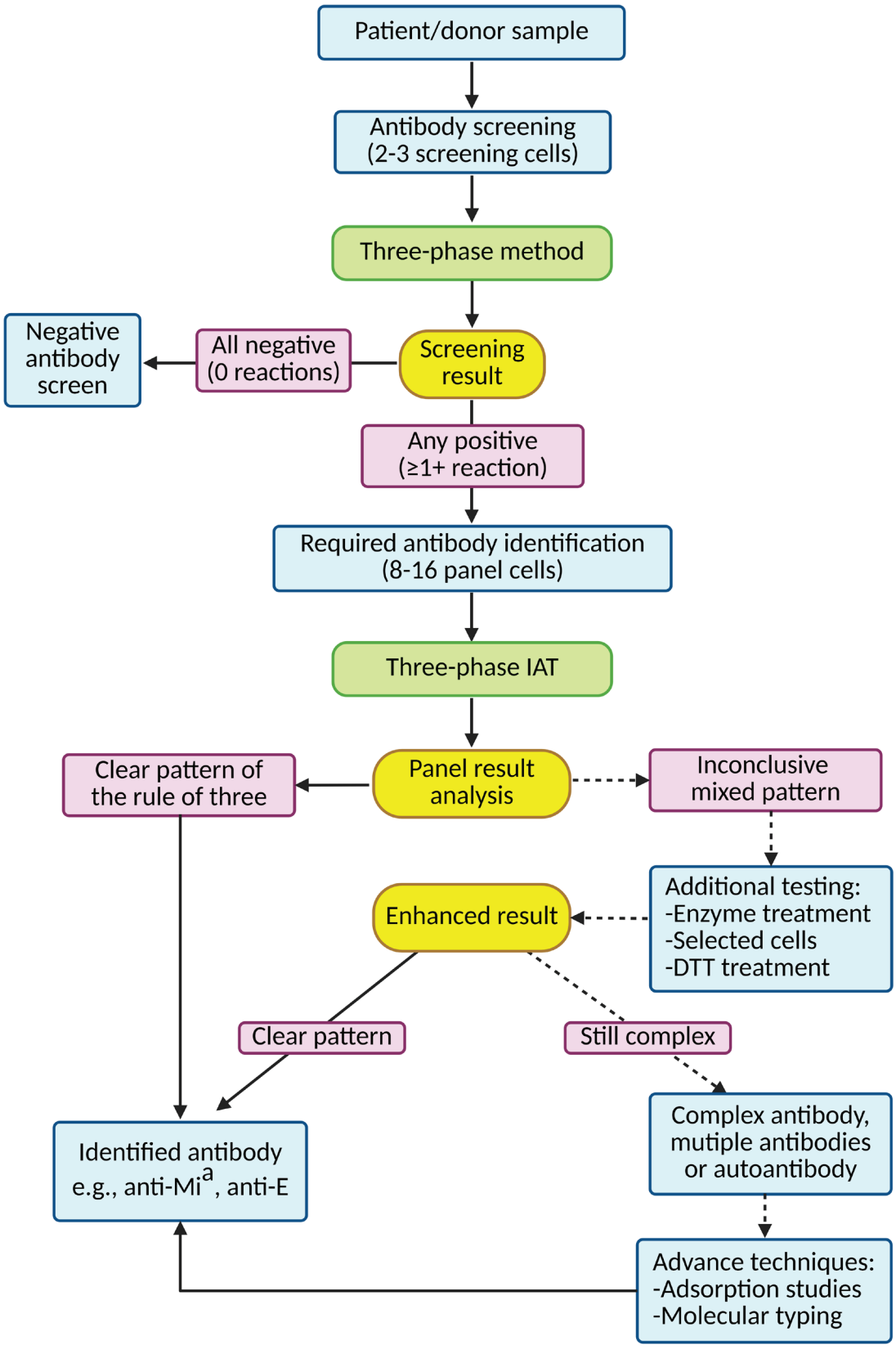


Figure 1. *Antibody screening and identification protocols in Thailand.*

Materials and methods

Literature search

A comprehensive literature search was conducted using multiple databases, including PubMed, Google Scholar, and ThaiJo. Relevant keywords and search terms were determined based on the research topic, specifically "Prevalence," "Unexpected alloantibody," "Red blood cell alloimmunization," and "Thailand." The titles and abstracts of previous publications were carefully screened to identify studies that were directly relevant to unexpected alloantibodies in Thailand. A total of 17 articles related to the prevalence of unexpected alloantibodies in Thailand were compiled and categorized according to geographical regions: Northern region (3 articles), Northeastern region (5 articles), Central region (7 articles), and Southern region (2 articles).

Subsequently, exclusion criteria were established to ensure data quality and relevance. Only studies published between 2013 and 2025 were included to ensure the most current data. Research conducted in patient populations exceeding 100 subjects (N>100) was selected to ensure statistical reliability, and studies reporting the identification of more than two types of alloantibodies were included to achieve appropriate data distribution. Following the application of these criteria, the prevalence of unexpected alloantibodies in each region was synthesized from representative high-quality articles for data extraction and analysis.

Results

The prevalence of unexpected alloantibodies in Thailand

In Thailand, several studies have investigated the prevalence of unexpected alloantibodies among transfusion recipients, especially in multi-transfused populations. Prevalence rate studies conducted across major Thai medical centers reveal that approximately 0.5-8.0% of hospitalized patients demonstrate detectable irregular antibodies.[22-28] The most common unexpected alloantibodies found in Thai patients are anti-Mi^a, anti-E, anti-Le^a, and anti-Le^b, as shown in Figure 2.

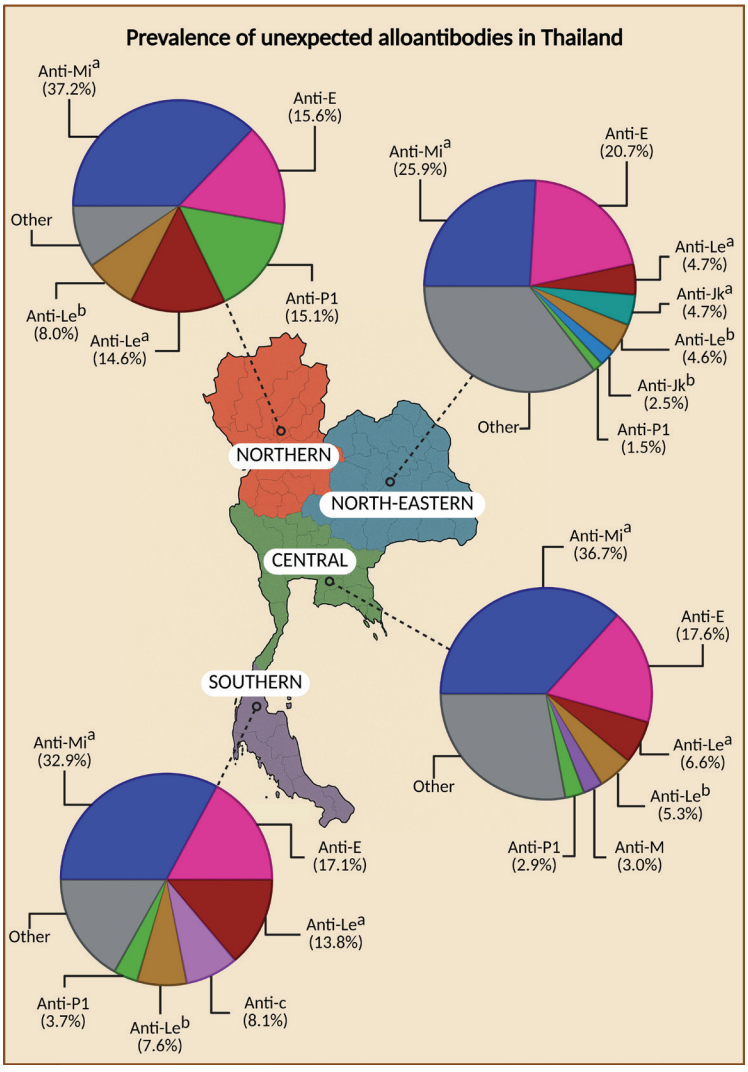


Figure 2. *The prevalence of alloantibodies in Thailand.*

A study of antibody identification across Thailand's four major geographical regions demonstrates significant variations in alloantibody distribution patterns, with the Northern region showing anti-Mi^a (37.2%), anti-E (15.6%), anti-P1 (15.1%), anti-Le^a (14.6%), and anti-Le^b (8.0%)[24], the Northeastern region exhibiting anti-Mi^a (25.9%), anti-E (20.7%), anti-Le^a (4.7%), anti-Le^b (4.6%), anti-Jk^a (4.7%), anti-Jk^b (2.5%), and anti-P1 (1.5%)[22], the Southern region demonstrating anti-Mi^a (32.9%), anti-E (17.1%), anti-Le^a (13.8%), anti-c (8.1%), anti-Le^b (7.6%), and anti-P1 (3.7%)[25]; and the Central region presenting anti-Mi^a (36.7%), anti-E (17.6%), anti-Le^a (6.6%), anti-Le^b (5.3%), anti-M (3.0%), and anti-P1 (2.9%).[27] Investigation of these prevalence patterns reveals that anti-Mi^a consistently demonstrated the highest prevalence (25.9-37.2%) across all four regions, followed by anti-E (15.6-20.7%), while anti-Le^a and anti-Le^b maintained relatively high prevalence rates of 4.7-14.6% and 4.6-8.0%, respectively, distributed throughout all regions.

Notably, anti- P1 showed consistently low prevalence (1.5-3.7%) across most regions with a remarkable exception in the Northern region (15.1%)[24], and anti-Kidd antibodies were rarely reported in most areas, but anti-Jk^a and anti-Jk^b demonstrated unexpectedly high prevalence (4.7 and 2.5%) specifically in the Northeastern region,[22] suggesting potential genetic and ethnic factors influencing alloantibody distribution patterns within the Thai population that warrant further investigation into underlying demographic and genetic determinants.

The high prevalence of anti-Mi^a is attributed to the high frequency of Mi^a-negative individuals (80-90%) in the Thai population,[27,29,30] This negative antigen expression makes Thai patients particularly susceptible to anti-Mi^a formation when Mi^a-negative recipients receive blood from Mi^a-positive donors (10-20%).[27,29,30] Anti-Mi^a is characterized as an IgM and/or IgG types. If anti-Mi^a is produced as an IgG type, it represents the clinical significance, capable of causing both hemolytic transfusion reactions and hemolytic disease in the fetus and newborn.[31,32] This antibody responds well to enzyme treatment during laboratory testing and is frequently found in association with other MNS system antibodies such as anti-M or anti-N.[33] The formation of anti-Mi^a is particularly problematic in patients with chronic diseases requiring regular transfusions, such as thalassemia major and hemophilia,[32] where the lack of routine Mi^a antigen screening during blood matching significantly increases the risk of alloimmunization.

Anti-E formation represents the second major concern in Thai transfusion medicine, with approximately 60-65% of the Thai population being E antigen-negative,[29,30] creating substantial opportunities for E antigen exposure during transfusion from 27-32% of E antigen-positive donors.[29,30] This antibody is predominantly responsible for hemolytic disease in fetuses and newborns in Thailand and can cause severe hemolytic transfusion reactions.[34,35] The formation of anti-E typically occurs when E-negative recipients receive blood from E-positive donors or when E-negative mothers carry E-positive fetuses during pregnancy.[36] The antibody is an IgG immunoglobulin that is best detected during the antiglobulin phase and represents one of the most clinically significant antibodies requiring antigen-negative blood selection for affected patients.[37]

Anti-Le^a and anti-Le^b pose difficulties for the Thai population due to the complex relationship between Lewis antigen expression and secretor status.[38] Anti-Le^a and anti-Le^b formation typically occur in Le(a-b-) individuals, who comprise approximately 20-24% of the Thai population, when exposed to blood from 61.0% Le(a-b+), 7.6% Le(a+b-), or 11.3% Le(a+b+) donors.[39,40] Both antibodies are predominantly IgM immunoglobulins with generally low clinical significance, though they can cause crossmatch incompatibilities and procedural delays.[38,41] An important characteristic of Lewis antibodies is their potential for spontaneous disappearance over time and their absence on neonatal red blood cells, which affects their clinical management strategies.[42]

These epidemiological findings have profound implications for blood banking practices in Thailand, necessitating population-specific screening strategies, targeted donor recruitment programs, and modified antibody identification protocols that account for the unique spectrum of alloantibodies encountered in Thai patients, thereby ensuring optimal transfusion safety and compatibility in this genetically distinct population.

Discussion

Risk factors for alloantibody development

Besides the frequency of red cell antigen expression that is attributed to genetic and ethnic differences,[22,24,25,27] there are multiple factors that induce alloantibody development and increase opportunities for antigen-antibody mismatches during transfusion in Thai patients. The immunocompetent status of recipients, including their HLA haplotypes and previous exposure to immunosuppressive medications, significantly influences the likelihood of antibody formation.[43] Demographic factors such as age and gender are equally important, with women of reproductive age facing higher risks due to pregnancy-related antigen exposure,[44] elderly patients requiring frequent transfusions,[45] and children with chronic diseases necessitating regular blood support.[46]

Antigen exposure patterns constitute another critical category of risk factors. The frequency and timing of blood transfusions directly correlate with alloimmunization risk, with patients receiving more than three units showing significantly increased antibody formation rates.[47] The type of blood product administered also influences risk, with red blood cell transfusions carrying higher immunization potential

compared to platelets.[48] Pregnancy-related factors, including the number of pregnancies, history of miscarriage or preterm delivery, and the occurrence of fetomaternal hemorrhage, create additional opportunities for foreign antigen exposure and subsequent antibody development.[49]

Management strategies and future directions

The management of alloantibody formation in Thai patients requires comprehensive strategies addressing both immediate clinical needs and long-term systematic improvements. Current challenges include aging populations requiring increased blood utilization, improved survival rates of patients with chronic diseases, expanding use of blood products in medical treatments, and limited implementation of blood conservation strategies.

For transfusion-dependent patients, such as those with thalassemia, it may be advisable to test for clinically significant antigens, including C, c, E, e, Mi^a, and the Kidd, Duffy, Kell, MNS, Lewis, and P1PK systems, prior to their first transfusion. For patients who have already received transfusions, red cell genotyping could be recommended. Moreover, it is suggested that every blood unit undergo an AHG crossmatch, with antibody screening performed alongside each crossmatch. In cases where the antibody screening is positive, the specific antibody should be identified, and antigen typing performed on both the patient's sample and the donor blood. Even if the crossmatch is compatible, selecting blood units that lack the corresponding antigen to the detected antibody is advised to maximize patient safety.

It may be advisable to issue a medical alert card specifying the type of alloantibody to assist healthcare personnel in careful monitoring, especially for antibodies demonstrating a dosage effect. Technological advances offer promising solutions through molecular typing replacing serological methods,[50] digital database systems for donor and patient matching,[51] development of point-of-care testing capabilities,[52] and research into artificial blood substitutes,[53]

Policy and guideline improvements should focus on developing national blood phenotyping programs, expanding rare donor registries, updating blood banking standards, and enhancing education and training programs for healthcare professionals.[54,55] Research priorities should include comprehensive studies of alloimmunization patterns specific to the Thai population, cost-effectiveness analyses of prevention strategies, development of risk prediction models, and investigation of novel therapeutic approaches for heavily alloimmunized patients. The implementation of these recommendations requires coordinated efforts between blood centers, hospitals, regulatory bodies, and research institutions to create a comprehensive approach to alloantibody prevention and management that addresses the unique needs of Thai patients while maintaining the highest standards of transfusion safety.

Conclusion

This review demonstrates the entire prevalence of unexpected alloantibodies throughout Thailand, revealing that anti-Mi^a is the most common alloantibody, followed by anti-E, anti-Le^a, and anti-Le^b. This review highlights the importance of practicing antibody screening protocols, choosing matched antigen status, and establishing rare donor registries in Thailand to mitigate the risks posed by unexpected alloantibodies.

Funding

This research received no external funding.

Conflict of interest

The author declares no conflict of interest.

CRediT authorship contribution statement

Udom Lao-on: conceptualization, investigation, visualization, writing: original draft preparation, reviewing and editing; **Kunwadee Lao-on**: conceptualization, visualization, writing: reviewing and editing.

Acknowledgements

We acknowledge the support of the Research Unit, School of Allied Health Sciences, Walailak University, which provided a BioRender subscription for this study.

Reference

[1] Talukdar B, Bhattacharya P. Novel aspects in blood transfusion – From donor to patient. J Hematol Allied Sci. 2025; 5(1): 18-25. doi:10.25259/JHAS_66_2024.

[2] Tenorio GC, Gupte SC, Munker R. Transfusion Medicine and Immunohematology. In: Munker R, Hiller E, Glass J, Paquette R, Editors. Modern Hematology: Biology and Clinical Management. New Jersey: Humana Press; 2007: pp 401-32. doi: 10.1007/978-1-59745-149-9_22.

[3] Daniels G, Poole J, De Silva M, Callaghan T, MacLennan S, Smith N. The clinical significance of blood group antibodies. Transfus Med. 2002; 12(5): 287-95. doi: 10.1046/j.1365-3148.2002.00399.x.

[4] Pham BN, Le Pennec PY, Rouger P. Allo-immunisation anti-érythrocytaire. Transfus Clin Biol. 2012; 19(6): 321-32. doi: 10.1016/j.tracli.2012.06.007.

[5] Schonewille H, Van De Watering LMG, Loomans DSE, Brand A. Red blood cell alloantibodies after transfusion: factors influencing incidence and specificity. Transfusion. 2006; 46(2): 250-6. doi: 10.1111/j.1537-2995.2006.00708.x.

[6] Tormey CA, Hendrickson JE. Transfusion-related red blood cell alloantibodies: induction and consequences. Blood. 2019; 133(17): 1821-30. doi: 10.1182/blood-2018-08-833962.

[7] Yang S, Yang HJ, Park HJ, Kwon SW. The Frequency

and distribution of unexpected red Cell Antibodies and Analysis of Antigen Exposure. Korean J Clin Lab Sci. 2014; 45: 164-9.

[8] Kahar MA, Patel RD. Phenotype frequencies of blood group systems (Rh, Kell, Kidd, Duffy, MNS, P, Lewis, and Lutheran) in blood donors of south Gujarat, India. Asian J Transfus Sci. 2014; 8(1): 51-5. doi: 10.4103/0973-6247.126693.

[9] Hickey MJ, Valenzuela NM, Reed EF. Alloantibody Generation and Effector Function Following Sensitization to Human Leukocyte Antigen. Front Immunol. 2016; 7(2016): 00030. doi: 10.3389/fimmu.2016.00030.

[10] Tinckam KJ, Chandraker A. Mechanisms and role of HLA and non-HLA alloantibodies. Clin J Am Soc Nephrol. 2006; 1(3): 404-14. doi: 10.2215/CJN.00270106.

[11] Zerra PE, Josephson CD. Chapter 73 - Delayed Hemolytic Transfusion Reactions. In: Shaz BH, Hillyer CD, Schwartz J, Gil MR, editors. Transfusion Medicine and Hemostasis. 4th ed. New York: Elsevier; 2025: pp 343-5. doi: 10.1016/B978-0-323-96014-4.00087-2.

[12] Panch SR, Montemayor-Garcia C, Klein HG. Hemolytic Transfusion Reactions. N Engl J Med. 2019; 381(2): 150-62. doi: 10.1056/NEJMra1802338.

[13] Bhuva DK, Vachhani JH. Red cell alloimmunization in repeatedly transfused patients. Asian J Transfus Sci. 2017; 11(2): 115-20. doi: 10.4103/0973-6247.214347.

[14] Delaney M, Matthews DC. Hemolytic disease of the fetus and newborn: managing the mother, fetus, and newborn. Hematology Am Soc Hematol Educ Program. 2015; 2015(1): 146-51. doi: 10.1182/asheducation-2015.1.146

[15] Arinsburg SA. Chapter 21 - Pretransfusion Testing. In: Shaz BH, Hillyer CD, Reyes Gil M, editors. Transfusion Medicine and Hemostasis. 3rd ed. New York: Elsevier; 2019: pp 107-16. doi: 10.1016/B978-0-12-813726-0.00021-0.

[16] Saviano A, Perotti C, Zanza C, Longhitano Y, Ojetti V, Franceschi F, *et al*. Blood Transfusion for Major Trauma in Emergency Department. Diagnostics. 2024; 14(7): 708. doi: 10.3390/diagnostics14070708.

[17] Thai Red Cross Society, National Blood Centre. Blood Administration and Monitoring. In: Pimol C, Sasitorn B, Aphisit T, editors. The Appropriate Use of Blood and Blood Components. 2nd ed. Bangkok: Pimnaphat; 2023: pp 40-1.

[18] Thai Red Cross Society, National Blood Center. Compatibility testing and selection of components. In: Pimol C, Sasitorn B, Aphisit T, editors. Standards for blood banks and transfusion service. 5th ed. Bangkok: Udomsuksa; 2024: pp 83-8.

[19] Kahar MA. Formulation of in-house screening and panel red cell for red cell antibody detection and identification in blood recipients: How to do it? J Hematol Allied Sci. 2024; 4(3): 94-102. doi: 10.25259/JHAS_43_2024.

[20] Chang Y-J, Ho C-Y, Zhou X-M, Yen H-R. Determination of degree of RBC agglutination for blood typing using a small quantity of blood sample in a microfluidic system. Biosens Bioelectron. 2018; 102: 234-41. doi: 10.1016/j.bios.2017.11.008.

[21] Johnson ST, Puca KE. Evaluating patients with autoimmune hemolytic anemia in the transfusion service and immunohematology reference laboratory: pretransfusion testing challenges and best transfusion-management strategies. Hematology Am Soc Hematol Educ Program. 2022; 2022(1): 96-104. doi: 10.1182/hematology.2022000406.

[22] Butryojantho C, Junta N, Pimphumee R, Srichai S, Darunikorn P, Puapairoj C, *et al*. Antibody screening and the prevalence of unexpected antibodies in the patients of Srinagarind Hospital. J Hematol Transfus Med. 2017; 27(1): 27-33 (in Thai).

[23] Daowsee K, Urwijitaroon Y. Red cell alloantibodies in transfused patients detected by standard tube test and column gel agglutination: A Study for Improving Transfusion Service. J Hematol Transfus Med. 2013; 23(4): 269-76 (in Thai).

[24] Po-ngernnak P, Sasjeenpong S, Chuesakul K, Pangwangthong K. The Prevalence of Red Blood Cell Alloantibodies in Lower Northern Thailand. J Med Assoc Thai. 2016; 99(12): 1337-43. PMID: 29952524.

[25] Promwong C, Siammai S, Hassarin S, Buakaew J, Yeela T, Soisangwan P, *et al*. Frequencies and specificities of red cell alloantibodies in the Southern Thai population. Asian J Transfus Sci. 2013; 7(1): 16-20. doi: 10.4103/0973-6247.106718.

[26] Rattakul A, Intharanut K, Nathalang O. The prevalence of red cell antibodies among patients and pregnant women in Pakchongnana Hospital. J Hematol Transfus Med. 2019; 29(4): 317-24 (in Thai).

[27] Rungroung T, Vejbaesya S, Rodpong P, Pimsamsee W, Chongkolwatana V. Red cell phenotyping in blood donors to provide safe blood transfusion for transfused patients at Siriraj Hospital. J Hematol Transfus Med. 2018; 28(4): 423-30 (in Thai).

[28] Thompat W, Boonpradit P, Wisetwut P, Oparatchatawat C, Srijinda S. The prevalence of unexpected antibodies in transfused cancer patients at National Cancer Institute. J Hematol Transfus Med. 2022; 32(4): 299-306 (in Thai).

[29] Romphruk A, Butryojantho C, Jirasakonpat B, Junta N, Srichai S, Puapairoj C, *et al*. Phenotype frequencies of Rh (C, c, E, e), M, Mi^a and Kidd blood group systems among ethnic Thai blood donors from the north-east of Thailand. Int J Immunogenet. 2019; 46(3): 160-5. doi: 10.1111/iji.12420.

[30] Toompokalang S, Peangchaiyaphoom N, Suwanwootichai P. C, E, c, e and Mi^a antigen test using Ortho Vision Max automated analyzer in

blood donors at the Regional Blood Centre III, Chonburi province. J Hematol Transfus Med. 2023; 33(1): 21-7 (in Thai).

[31] Agrawal S, Chowdhry M. A case report on anti-Mi^a antibody in a multi-transfused patient from India. Transfus Apher Sci. 2019; 58(5): 625-7. doi: 10.1016/j.transci.2019.08.027.

[32] Pahuja S, Sehgal S, Sharma G, Singh M, Yadav R. The anti-Mi^a antibody – Report of four cases in a tertiary care hospital with review of literature. Glob J Transfus Med. 2019; 4(1): 79-83. doi: 10.4103/gjtm.gjtm_2_19.

[33] Reid ME, Storry JR. Low-incidence MNS antigens associated with single amino acid changes and their susceptibility to enzyme treatment. Immunohematology. 2001; 17(3): 76-81. doi: 10.21307/immunohematology-2019-551.

[34] Cheepsattayakorn R, Fongsatitkul L, Chotinaruemol S, Mahawongtong M. Anti-E as a cause of hemolytic disease of the newborn. J Med Assoc Thai. 1997; 80 Suppl 1: S1-4. PMID: 9347637.

[35] Usman ASi, Mustaffa R, Ramli N, Diggi SA. Hemolytic disease of the fetus and newborn caused by anti-E. Asian J Transfus Sci. 2013; 7(1): 84-5. doi: 10.4103/0973-6247.106750.

[36] Joy SD, Rossi KQ, Krugh D, O'Shaughnessy RW. Management of pregnancies complicated by anti-E alloimmunization. Obstet Gynecol. 2005; 105(1): 24-8. doi: 10.1097/01.aog.0000149153.93417.66.

[37] Michalewska B, Ejduk A, Pniewska K. Acute haemolytic transfusion reaction apparently caused by the 'enzyme-only' anti-E. Vox Sang. 2005; 89(1) :61-1. doi: 10.1111/j.1423-0410.2005.00641.x.

[38] Combs MR. Lewis blood group system review. Immunohematology. 2009; 25(3): 112-8. doi: 10.21307/immunohematology-2019-243.

[39] Nanakorn N, Boonthongkhao S, Mitundee S, Tonwong N. Prevalence of Lewis blood group polymorphisms in Southern Thai blood donors. J Health Sci Med Res. 2022; 40(4): 10. doi: 10.31584/jhsmr.2021847.

[40] Nathalang O, Kuvanont S, Punyaprasiddhi P, Tasaniyanonda C, Sriphaisal T. A preliminary study of the distribution of blood group systems in Thai blood donors determined by the gel test. Southeast Asian J Trop Med Public Health. 2001; 32(1): 204-7. PMID: 11485086.

[41] Subramaniyan R. Serological characteristics of Lewis antibodies and their clinical significance – A case series. Hematol Transfus Cell Ther. 2023; 45(2): 159-64. doi: 10.1016/j.htct.2021.07.007.

[42] Kijkornpan S, Suebsaeng C, Chiewsilp P. Lewis antigens on newborn red cells. J Med Assoc Thai. 2000; 83(Suppl 1): S46-8. PMID: 10865406.

[43] Karnes JH, Miller MA, White KD, Konvinse KC, Pavlos RK, Redwood AJ, *et al*. Applications of Immunopharmacogenomics: Predicting, Preventing, and Understanding Immune-Mediated Adverse Drug Reactions. Annu Rev Pharmacol Toxicol. 2019; 59: 463-86. doi: 10.1146/annurev-pharmtox-010818-021818.

[44] Moncharmont P, Barday G, Meyer F. Red blood cell alloimmunisation in 18 to 50-year old transfused women: a 3-year study. Blood Transfus. 2015; 13(2): 345-6. doi: 10.2450/2014.0211-14.

[45] Moncharmont P, Barday G, Py JY, Meyer F.Acquired red blood cell alloantibodies in transfused patients of 80 years or over: a 2008-2013 national haemovigilance survey. Blood Transfus. 2017; 15(3): 254-8. doi: 10.2450/2016.0328-15.

[46] Dogra A, Sidhu M, Kapoor R, Kumar D. Study of red blood cell alloimmunization in multitransfused thalassemic children of Jammu region. Asian J Transfus Sci. 2015; 9(1): 78-81. doi: 10.4103/0973-6247.150958.

[47] Fluit CRMG, Kunst VAJM, Drenthe-Schonk AM. Incidence of red cell antibodies after multiple blood transfusion. Transfusion. 1990; 30(6): 532-5. doi: 10.1046/j.1537-2995.1990.30690333485.x.

[48] Park HS, Shin KS, Son BR. Incidence of red blood cell alloantibody formation after platelet concentrate transfusions. Korean J Blood Transfus. 2019; 30(1): 33-41. doi: 10.17945/kjbt.2019.30.1.33.

[49] Mor G, Cardenas I. The immune system in pregnancy: a unique complexity. Am J Reprod Immunol. 2010; 63(6): 425-33. doi: 10.1111/j.1600-0897.2010.00836.x.

[50] Westhoff CM. Molecular testing for transfusion medicine. Curr Opin Hematol. 2006; 13(6): 471-5. doi: 10.1097/01.moh.0000245695.77758.3d.

[51] Shah A, Shah D, Shah D, Chordiya D, Doshi N, Dwivedi R. Blood bank management and inventory control database management system. Procedia Comput Sci. 2022; 198: 404-9. doi: 10.1016/j.procs.2021.12.261.

[52] Raja ML, Musi R, Fattorini M, Piva E, Putoto G. Point of care Testing and transfusion safety in resource limited settings: A review. J Blood Disorders Transf. 2015; 6(2): 1000269. doi: 10.4172/2155-9864.1000269.

[53] Sarkar S. Artificial blood. Indian J Crit Care Med. 2008; 12(3): 140-4. doi: 10.4103/0972-5229.43685.

[54] Gammon RR, Dubey R, Gupta GK, Hinrichsen C, Jindal A, Lamba DS, *et al*. Patient Blood Management and Its Role in Supporting Blood Supply. J Blood Med. 2023; 14: 595-611. doi: 10.2147/jbm.S387322.

[55] Thirunavukkarasu J, Subinesh KJ, Dharshan P, Boobesh S, Editors. A smart blood donor locator and component matching system for rapid disaster response. 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES); 2023: 1-7. doi: 10.1109/ICSES60034.2023.10465383.

Journal of Associated Medical Sciences 2026; 59(1): 137-144

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 study of correlation between cholelithiasis and fatty liver from ultrasonography of patients in Vajira Hospital

Patamaporn Molee[1*], Wanisacha Kuisakorn[2], Yanutta Laosuksanttiwong[3], Natchaya Kerdpon[4], Mananya Pirom[5]

[1]Department of Radiological Technology, Faculty of Medicine Vajira Hospital, Navamindradhiraj University, Bangkok, Thailand.
[2]Vajira Center of Excellence In Cancer Care, Faculty of Medicine Vajira Hospital, Navamindradhiraj University, Bangkok, Thailand.
[3]Radiology Center, Krungthai General Hospital, Nonthaburi Province, Thailand.
[4]Magnetic Resonance Imaging Center, Udorn Medical Group, Pathum Thani Hospital, Pathum Thani Province, Thailand.
[5]Group of Out-Reach Health Services, Maha Sarakham Hospital, Maha Sarakham Province, Thailand.

ARTICLE INFO

Article history:
Received 25 August 2025
Accepted as revised 20 October 2025
Available online 29 October 2025

Keywords:
Cholelithiasis, fatty liver, ultrasonography, retrospective study, epidemiology, Thailand.

ABSTRACT

Background: The rising consumption of Western-style fast food—high in carbohydrates, sugar, and fat—together with a fast-paced lifestyle, occupational stress, and physical inactivity, represents a major risk factor for fatty liver disease, which can progress to hepatitis, cirrhosis, or hepatocellular carcinoma. Evidence linking fatty liver and cholelithiasis is growing globally, yet data from Thailand and Southeast Asia remain scarce, despite both conditions being increasingly prevalent. Cholelithiasis, a common gastrointestinal disorder affecting 5–10% of the Thai population, particularly females and older adults, may also be associated with fatty liver disease.

Objectives: This study investigated the correlation between cholelithiasis and fatty liver disease using ultrasonographic data from patients at Vajira Hospital, aiming to address this regional knowledge gap and inform preventive strategies.

Materials and methods: A retrospective analysis was performed on 311 patients aged ≥50 years who underwent upper abdominal ultrasonography at Vajira Hospital between 2023 and 2024. Collected data included gender, age, gallstone size, and diagnostic findings. Patients were classified into two groups: 132 with cholelithiasis and 179 without. Statistical analyses included descriptive statistics (mean, SD, frequency, percentage) and inferential tests (t-test, chi-square test, and logistic regression).

Results: Of the 311 patients, 162 (52.1%) had fatty liver, with a mean age of 62±8.46 years; 62.3% were female and 37.7% were male. Among these, 73 patients (45.1%) had concurrent cholelithiasis (66% female, 34% male), with a mean gallstone size of 1.09±0.53 cm. The remaining 89 patients (54.9%) had fatty liver disease without cholelithiasis (60% female, 40% male). A significant inverse association was observed between age and fatty liver disease (p= 0.003), indicating that younger patients within this ≥50-year cohort were more likely to be affected.

Conclusion: Cholelithiasis was not significantly associated with fatty liver disease. However, fatty liver showed a novel inverse correlation with age, with higher prevalence among younger adults within the ≥50-year cohort. This finding, which contrasts with previous reports of increasing prevalence with age, may reflect methodological factors and lifestyle influences in Thailand's urban population. These results emphasize the need for early screening and targeted public health interventions for working-age adults, while acknowledging limitations related to study design, absence of key confounders, and restricted age range.

***Corresponding contributor.**
Author's Address: *Department of Radiological Technology, Faculty of Medicine Vajira Hospital, Navamindradhiraj University, Bangkok, Thailand.*
E-mail address: patamaporn.m@nmu.ac.th
doi: 10.12982/JAMS.2026.017
E-ISSN: 2539-6056

Introduction

Fatty liver disease (FLD) is a condition characterized by excessive accumulation of triglycerides in hepatic tissue. In Thailand, the prevalence of FLD is approximately 45.2%, with a higher occurrence in males than females, and the highest frequency observed at a median age of 56 years.[1] Key risk factors include obesity, diabetes mellitus, metabolic syndrome, hyperlipidemia, and unhealthy lifestyles, particularly among urban populations adopting high-fat, high-sugar diets, experiencing work-related stress, and lacking physical activity.[2,3]

FLD can progress in two forms: non-inflammatory, which accounts for around 75% of cases and poses a lower risk of liver fibrosis and cancer, and inflammatory, which accounts for the remaining 25% and is more likely to lead to cirrhosis and hepatocellular carcinoma.[4] A growing concern is the increasing incidence of hepatocellular carcinoma among patients with non-alcoholic fatty liver disease (NAFLD), as reported in recent studies, with incidence reaching 0.44 per 1,000 persons per year.[2] This burden makes NAFLD a major public health challenge in Thailand and across Asia.

Several studies have suggested a potential association between FLD and gallstone disease (cholelithiasis), which affects 5–10% of the Thai population, with higher prevalence in females and older adults.[5] Both conditions share common risk factors such as obesity, insulin resistance, and metabolic disturbances.[6] Recent systematic reviews have synthesized the evidence linking gallstone disease and non-alcoholic fatty liver disease (NAFLD). Slouha *et al*. concluded that the relationship is likely bidirectional, although substantial heterogeneity across populations and study designs limits definitive inference.[7] In parallel, Wong *et al*. reported that the global burden of NAFLD is increasing, with particularly rapid growth in Asian populations, underscoring the importance of regional data.[3] Taken together, these reviews highlight both the plausibility of a shared pathophysiological basis between gallstones and NAFLD and the urgent need for context-specific studies in Thailand to address this knowledge gap.[3] However, evidence directly examining this association in Thai or Southeast Asian populations remains scarce. Clarifying this relationship is essential for clinical risk stratification and for designing targeted public health interventions in settings where both conditions are becoming increasingly prevalent.

Ultrasonography is a reliable and non-invasive imaging modality for detecting both conditions, offering sensitivity of 84.8% and specificity of 93.6% for detecting moderate to severe histologically defined fatty liver.[8] Beyond its diagnostic utility, ultrasonography is widely accessible in Thailand, making it particularly suitable for hospital-based and population-level screening. Despite the growing body of international evidence, few studies have evaluated the association between FLD and cholelithiasis in Thai or Southeast Asian populations, where the burden of these conditions is rapidly increasing. Addressing this gap in the local context is essential for clinical risk stratification and for guiding targeted public health strategies. Therefore, this study investigates the association between FLD and cholelithiasis using ultrasonographic findings in patients at Vajira Hospital, aiming to inform prevention and early detection efforts.

Materials and methods

Study design

This study employed a retrospective design based on patients who underwent upper abdominal ultrasonography at the Diagnostic Radiology Department of Vajira Hospital between 2023 and 2024. A total of 311 patients aged 50 years and above were included. This age cutoff was selected because both fatty liver and cholelithiasis are more prevalent in older adults. Previous Thai studies, including a hospital-based study in Maha Sarakham and a community-based cohort in Chiang Rai, demonstrated higher prevalence rates among adults aged ≥40-50 years, supporting the relevance of this criterion for identifying high-risk populations.[9, 10]

Patient data were retrieved from the hospital's medical records and ultrasound imaging system. The variables collected included gender, age, gallstone size, radiologist diagnosis, and corresponding ultrasonographic images. Patients were classified into two groups: 132 with gallstones (cholelithiasis) and 179 without. Data on key confounding variables, including body mass index (BMI), diabetes mellitus, and alcohol intake, were unavailable in the retrospective dataset and therefore were not incorporated into the statistical models. The absence of these key confounders may have reduced the robustness of the regression analysis, a limitation acknowledged in this study.

Descriptive statistics, including mean and standard deviation, were used to summarize continuous variables, while frequencies and percentages were used for categorical variables. Comparative analyses between groups were performed using independent samples t-test and chi-square test. Additionally, logistic regression analysis was conducted to assess the association between selected variables and the presence of fatty liver disease. Model assumptions, including multicollinearity and overall model fit, were assessed, although the limited availability of confounding variables likely reduced statistical power.

Study procedure

The data collection process for this study was conducted at Vajira Hospital using the EVInsite® radiologic imaging and reporting system. Retrospective data were obtained from ultrasonographic examinations performed between January 2023 and December 2024.

The procedure was carried out in the following steps:

1. **Patient selection**

   Patients who underwent upper abdominal ultrasonography during the study period were

identified. Records were screened to include only patients aged 50 years and above who met the inclusion criteria.

2. **Data retrieval**
   Eligible patient records were retrieved from the EVInsite® system, including both the ultrasound images (Figure 1 and Figure 2) and the official diagnostic reports interpreted by radiologists.
3. **Data extraction**
   Variables extracted included gender, age, gallstone size, and the presence or absence of fatty liver and cholelithiasis.
4. **Data entry and analysis preparation**
   All data were systematically recorded in a structured format and entered into IBM SPSS Statistics version 28.0.0.0. for analysis.

This structured and stepwise approach ensured accurate and ethical handling of patient data while maintaining the scientific integrity of the research process.

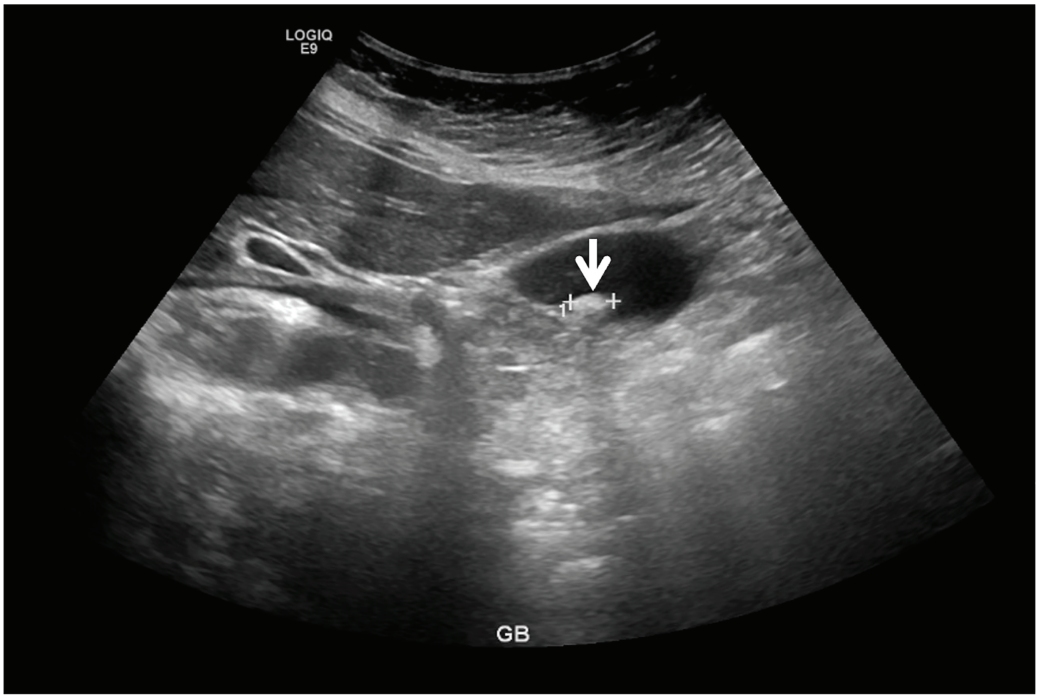


Figure 1. *Ultrasonographic image of the gallbladder demonstrating a hyperechoic calculus (arrow) with posterior acoustic shadowing, consistent with gallstones.*

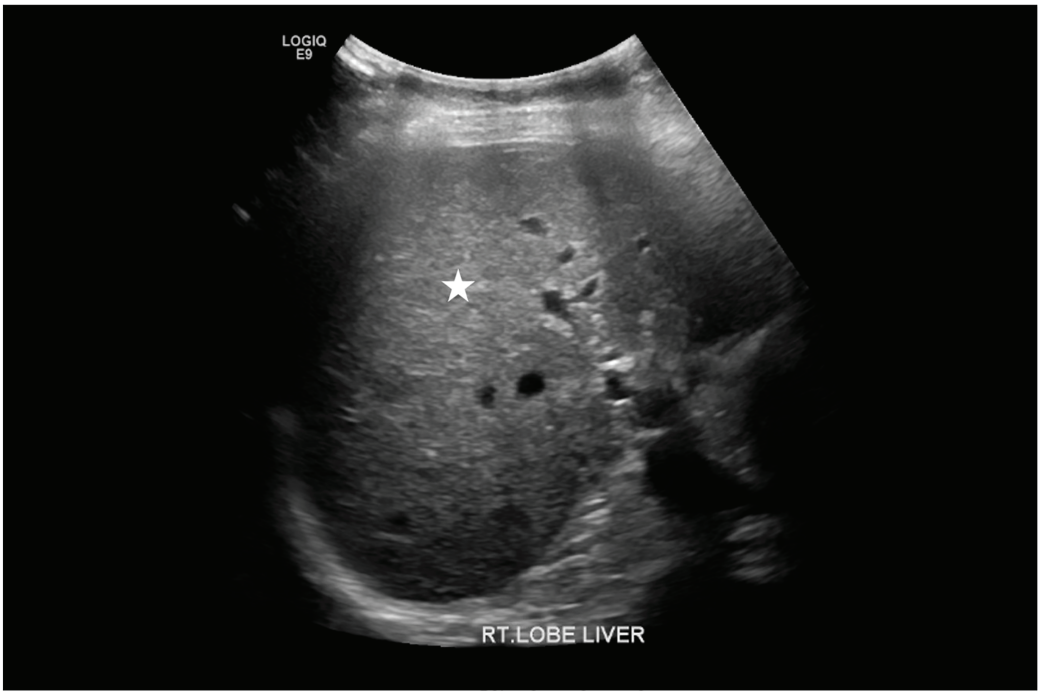


Figure 2. *Ultrasonographic image of the right hepatic lobe demonstrating increased parenchymal echogenicity (star) and blurring of vascular margins, consistent with fatty liver disease.*

Data analysis

All analyses were performed using IBM SPSS Statistics software, version 28.0.0.0. Both descriptive and inferential statistical methods were applied.

Descriptive statistics summarized the characteristics of the study population: quantitative variables (age, gallstone size) were expressed as mean±SD, while qualitative variables (gender, diagnosis of gallstone disease, and diagnosis of fatty liver disease) were presented as frequencies and percentages.

Inferential analyses tested study hypotheses and assessed associations between variables. examined relationships between categorical variables, and the independent-samples t-test compared mean values of continuous variables between groups. Logistic regression analysis was performed to evaluate the association between cholelithiasis and fatty liver disease. However, data on important confounders (BMI, diabetes, alcohol intake) were not available in the retrospective dataset and thus could not be incorporated into the regression model. As a result, the model may have been underpowered, and the findings should be interpreted with caution. Model adequacy was assessed using the Hosmer–Lemeshow goodness-of-fit test and pseudo R^2 values, which confirmed acceptable fit.

All statistical tests were two-tailed and performed at a 95% confidence level, with $p<0.05$ considered statistically significant.

Results

General characteristics of the study population

The demographic and clinical characteristics of patients with and without cholelithiasis are summarized in Table 1. Overall, females were more prevalent across all subgroups. However, the relative proportions of males and females were similar, indicating no substantial gender differences in the co-occurrence of fatty liver disease and cholelithiasis.

Table 1. Percentage distribution of patients with and without fatty liver in combination with cholelithiasis, by gender.

	Female N (%)	Male N (%)
Fatty liver with cholelithiasis (N=73)	48 (66)	25 (34)
Without fatty liver, with cholelithiasis (N=59)	37 (63)	22 (37)
Fatty liver without cholelithiasis (N=89)	53 (60)	36 (40)
Neither fatty liver nor cholelithiasis (N=90)	58 (64)	32 (36)

Age Distribution

Among patients with cholelithiasis, FLD was more common in those aged 50-70 years, whereas its prevalence declined in older age groups (Table 2). A similar trend was observed among patients without cholelithiasis, with the highest frequency in the 50-60-year age group, followed by a progressive decrease in older age categories (Table 3).

Table 2. Percentage distribution of patients with and without fatty liver among those with cholelithiasis, by age group.

Age group (years)	Fatty liver, N (%)	Without fatty liver, N (%)
50-60	41 (31.06)	24 (18.18)
61-70	23 (17.42)	18 (13.64)
71-80	5 (3.79)	8 (6.06)
81-90	4 (3.03)	8 (6.06)
91-100	0	1 (0.76)

Table 3. Percentage distribution of patients with and without fatty liver among those without cholelithiasis, by age group.

Age group (years)	Fatty liver, N (%)	Without fatty liver, N (%)
50-60	43 (24.02)	34 (18.99)
61-70	28 (15.64)	30 (16.76)
71-80	15 (8.38)	18 (10.06)
81-90	3 (1.68)	8 (4.47)
91-100	0	0

Gallstone size

The distribution of fatty liver cases by gallstone size is presented in Table 4. Fatty liver was most frequently observed in patients with gallstones ≤1 cm in diameter, whereas larger stones were less common in both groups.

Comparative and inferential analyses

Comparative analyses are presented in Table 5. No significant gender differences were observed between patients with and without fatty liver (p=0.330). Patients with fatty liver were significantly younger than those without (62.0±8.46 vs. 65.0±10.24 years, p=0.003). Gallstone size showed no significant difference between groups (p=0.373). Similarly, the prevalence of fatty liver was comparable between patients with and without cholelithiasis (p=0.797).

Logistic regression analysis

Multivariate logistic regression results are shown in Table 6. Age was the only variable independently associated with fatty liver disease (OR=0.967, 95% CI: 0.944-0.991, p=0.007), demonstrating an inverse association. Neither gender nor gallstone status was significantly associated with fatty liver.

Table 4. Percentage distribution of patients with and without fatty liver, by gallstone size.

Gallstone size (cm)	Fatty liver, N (%)	Without fatty liver, N (%)
0.00-1.00	42 (31.82)	31 (23.48)
1.01-2.00	27 (20.54)	26 (19.70)
2.01-3.00	4 (3.03)	2 (1.52)

Table 5. Baseline characteristics and comparative analysis of patients with and without fatty liver.

Variables	Fatty liver (N=162)	Without fatty liver (N=149)	*p* value
Gender			0.330
Male, N (%)	61 (37.7)	54 (36.2)	
Female, N (%)	101 (62.3)	95 (63.8)	
Age (years)	62±8.46	65±10.24	0.003*
Gallstone size (cm)	1.09±0.53	1.12±0.50	0.373
Gallstone, N (%)	73 (45.1)	59 (39.6)	0.797
Non-gallstone, N (%)	89 (54.9)	90 (60.4)	

***Note:** *statistically significant at p<0.05.*

Table 6. Multivariate logistic regression analysis of gender, age, and gallstone status in relation to fatty liver.

Variables	*p* value	OR (95%CI)
Gender	0.42	0.952 (0.597-1.520)
Age (years)	0.007*	0.967 (0.944-0.991)
Gallstone	0.343	1.247 (0.790-1.970)

***Note:** *statistically significant at p<0.05.*

Discussion

Gender

In this study, most patients with concurrent cholelithiasis and fatty liver were female (66%), consistent with the reports of Ashraf A *et al*. and Kichloo A. *et al*.[6,11] These studies indicate that physiological changes in females, particularly during menopause, contribute to insulin resistance, central obesity, and increased fat deposition, potentially leading to fatty liver disease due to fluctuations in estrogen levels. Beyond hormonal shifts, alterations in cholesterol metabolism and bile acid composition in women have also been implicated, which may increase the likelihood of gallstone formation while simultaneously promoting hepatic lipid accumulation.[12,13] Estrogen can enhance cholesterol secretion into bile, thereby raising gallstone risk, and at the same time influence hepatic lipid homeostasis, contributing to steatosis.[12,14] In addition, insulin resistance—a common feature in postmenopausal women—further exacerbates both hepatic triglyceride accumulation and gallstone risk.[15,16] Lifestyle factors prevalent among Thai urban women, including reduced physical activity and high-fat dietary intake, may amplify these biological susceptibilities (e.g., in Thai patients NAFLD, combined exercise and dietary modification trials have

demonstrated reduction in hepatic fat),[17] and meta-analyses show diet/exercise interventions consistently improve fatty liver outcomes across populations.[18] Taken together, these mechanisms may explain why women, particularly those aged ≥50 years, are more likely than men to present with both cholelithiasis and fatty liver disease.

Age

In our study, patients aged 50-60 years with cholelithiasis were more likely to have fatty liver disease than those without, consistent with the findings of Rattanangamkul, who reported the highest prevalence in this age group.[19] The mean age of patients with fatty liver disease was 62 years, significantly lower than that of patients without the condition (p=0.003). However, this result contrasts with studies from Chinese cohorts which reported an increasing prevalence of concurrent cholelithiasis and fatty liver disease with advancing age.[11,20,21] This discrepancy may partly reflect our study's focus on patients aged ≥50 years, consistent with Thai data indicating that the burden of NAFLD is concentrated in older adults.[9,10] Nevertheless, the restricted age range and limited sample size across certain strata may have reduced our ability to detect broader age-related trends and limited generalizability to younger populations.

Systematic reviews have also reported age-dependent patterns and considerable between-study variability, with stronger associations often observed in metabolically high-risk or younger cohorts. In contrast, our study focused on adults aged ≥50 years and identified an inverse relationship between age and fatty liver within this band, which may attenuate any crude association with gallstone disease. This is consistent with review-level observations that population structure and unmeasured confounding (e.g., obesity, diabetes, alcohol use) can influence pooled estimates.[7]

Beyond methodological considerations, lifestyle and dietary transitions among Thailand's urban, working-age adults may partly explain the observed higher prevalence of fatty liver disease in younger patients within this ≥50-year cohort. Increased consumption of Western-style fast food, reduced physical activity, and occupational stress have been identified as key contributors to NAFLD risk in Asian populations.[22,23] Long working hours have also been associated with metabolic derangements and lean NAFLD,[24] while adherence to healthier lifestyle patterns is protective.[25] Taken together, these findings support the hypothesis that younger individuals in older Thai cohorts may face greater exposure to adverse lifestyle factors, thereby elevating their susceptibility to fatty liver disease despite being relatively younger within the study population.

Gallstone size

Gallstone size ranged from 0.00 to 3.00 cm and was classified into three size categories. The mean gallstone size was 1.09 cm in patients with fatty liver disease and 1.12 cm in those without the condition, demonstrated no significant difference (p=0.373). This finding differs from that of Rehman T. *et al.*, who reported that larger gallstone were associated with increased risk and severity of fatty liver disease.[26] The difference may be explained by methodological differences: our study categorized gallstones into predefined size ranges and did not assess fatty liver severity, whereas the previous study analyzed severity without stratifying by gallstone size.

Reviews comparing diagnostic modalities have noted that ultrasonography—while practical—has reduced sensitivity for mild steatosis compared with biochemical indices or advanced imaging, potentially biasing associations toward the null in hospital-based cohorts. The reliance on ultrasonography in our study may therefore partly explain the absence of a size–steatosis gradient.[8]

Association between cholelithiasis and fatty liver disease

In this analysis, 45.1% of patients with cholelithiasis also had fatty liver, compared with 54.9% of those without cholelithiasis. This association was not statistically significant (p=0.797), consistent with Lu Y *et al*. who also reported no significant association between cholelithiasis and non-alcoholic fatty liver disease (NAFLD) (p=0.15).[27] In contrast, Kim YK *et al.*, and Li X and Gao P found a significant association.[20,21] Their studies, however, were conducted in younger Chinese populations and limited to NAFLD cases, with a lower overall prevalence of cholelithiasis (8.8% in 897 participants) than in our cohort—likely reflecting differences in dietary patterns and population characteristics.

Recent evidence syntheses corroborate a link between gallstone and FLD. Slouha *et al*. reviewed observational studies and concluded that the association is likely bidirectional, while emphasizing substantial heterogeneity by population and study design.[7] Wong *et al*. highlighted the rapidly rising burden of fatty liver disease across Asian settings, underscoring the need for locally generated data.[3] Interpreted against this backdrop, our finding of no significant association between cholelithiasis and fatty liver in an older Thai cohort suggests that demographic structure, risk profiles, and setting-specific exposures may partly account for between-study differences. These considerations align with prior reviews noting effect modification by age, metabolic status, and diagnostic methods.

Taken together with recent reviews, our results suggest that context and case mix matter: in older Thai adults undergoing routine ultrasonography, cholelithiasis does not appear to be independently associated with fatty liver, and younger age within an older cohort carries a higher likelihood of steatosis. These insights support targeted screening and preventive strategies that

prioritize working-age groups and address metabolic risk in the Thai setting.[3]

From a clinical perspective, ultrasonography could be incorporated into routine health check-ups as a cost-effective screening tool for early detection of fatty liver disease, particularly among working-age adults. At the population level, these findings underscore the need for public health strategies in Thailand, including dietary education, community-based lifestyle interventions, and workplace wellness programs targeting urban populations. Such approaches could mitigate the rising burden of fatty liver disease and its complications.

Limitations

This study has several limitations. First, the retrospective cross-sectional design restricted data collection to information available in hospital medical records, which precluded the inclusion of several important confounders such as body mass index (BMI), diabetes, lipid profiles, and alcohol consumption. Second, the diagnosis of fatty liver disease relied exclusively on ultrasonography, which is operator-dependent and less sensitive for detecting mild steatosis compared with biochemical or histopathological confirmation. An additional limitation is the restriction of the study population to individuals aged ≥50 years. This age threshold was selected because of the higher prevalence of fatty liver disease and gallstone disease among older Thai adults.[9,10] Nevertheless, this design reduces the generalizability of our findings to younger populations who may also be at risk. Finally, the relatively modest sample size limited subgroup analyses and adjustments for unmeasured confounders. Future studies should therefore include broader age ranges, incorporate validated risk factor assessments, and employ multimodal diagnostic approaches to enhance external validity and diagnostic accuracy.

Conclusion

In conclusion, fatty liver disease was more common among younger adults in this hospital-based cohort, while no significant association was observed with gender or gallstone status. This novel inverse relationship with age, contrasting with previous studies that reported increasing prevalence in older adults, represents an important contribution to understanding the Thai context. The findings underscore the need for early detection and lifestyle modification programs to reduce the future burden of disease. Nevertheless, several limitations must be acknowledged, including retrospective design, reliance on ultrasonography alone, and the absence of key confounding variables such as body mass index, diabetes, and alcohol use, which limit generalizability. Despite these constraints, ultrasonography remains a practical screening tool in clinical practice. At the population level, targeted public health initiatives—including workplace wellness programs, dietary education, and community-based strategies—are urgently needed to mitigate the rising burden of fatty liver disease and its complications in Thailand.

Ethical approval

This study was approved by the Research Ethics Committee of the Faculty of Medicine Vajira Hospital, Navamindradhiraj University, Bangkok, Thailand (COA 172/2566).

Funding

This study did not receive funding from any agency and has no conflicts of interest with any funding source.

Conflict of interest

The authors declare no conflicts of interest in this study.

CRedit authorship contribution statement

Patamaporn Molee: conceptualization, methodology, supervision, writing: original draft preparation, review and edit, ethics approval and research consultation; **Wanisacha Kuisakorn**: data collection and analysis, results interpretation; **Yanutta Laosuksanttiwong**: data collection and analysis, results interpretation; **Natchaya Kerdpon**: data collection and analysis, results interpretation; **Mananya Pirom**: clinical interpretation consultation.

Acknowledgements

Faculty of Medicine Vajira Hospital, Navamindradhiraj University

References

[1] Sono S, Kaewdech A, Wongsawanon C, Varanuntakul T, Jarupanich T, Sritanabutr N, et al. Metabolic dysfunction-associated fatty liver disease (MAFLD) in the adult population attending a health check-up program in Thailand: prevalence and fibrosis status. Sci Rep. 2025; 15(1): 21429. doi: 10.1038/s41598-025-06874-1.

[2] Huh Y, Cho YJ, Nam GE. Recent Epidemiology and Risk Factors of Nonalcoholic Fatty Liver Disease. J Obes Metab Syndr. 2022; 31(1): 17-27. doi: 10.7570/jomes22021.

[3] Wong VW-S, Ekstedt M, Wong GL-H, Hagström H. Changing epidemiology, global trends and implications for outcomes of NAFLD. J Hepatol. 2023; 79(3): 842-52. doi: 10.1016/ j.jhep.2023.04.036.

[4] Sepehrimanesh M, Niknam R, Ejtehadi F, Fattahi MR, Safarpour A. Association between non-alcoholic fatty liver disease and metabolic syndrome with gallstone disease, South Iran: A population-based study. Diabetes Metab Syndr Obes. 2020; 13: 1449-58. doi: 10.2147/dmso.S246949.

[5] Srisubat A, Thaiyakul A, Kwanjaroensub S, Sa-nga N, Paiprarert S, Pinyoteppratarn R, et al. Economic evaluation of laparoscopic cholecystectomy versus open cholecystectomy. J Dept Med Serv.

2023; 48(1): 93-103. doi: 10.14456/jdms.2023.12.
[6] Kichloo A, Solanki S, Haq KF, Dahiya D, Bailey B, Solanki D, et al. Association of non-alcoholic fatty liver disease with gallstone disease in the United States hospitalized patient population. World J Gastrointest Pathophysiol. 2021; 12(2): 14-24. doi: 10.4291/wjgp.v12.i2.14.
[7] Slouha E, Biput SJ, Kuteyi A, Kalloo AE, Gorantla VR. Non-alcoholic fatty liver disease and gallstones: A systematic review. Cureus. 2023; 15(9): e45027. doi: 10.7759/cureus.45027.
[8] Hernaez R, Lazo M, Bonekamp S, Kamel I, Brancati FL, Guallar E, et al. Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: A meta-analysis. Hepatology. 2011; 54(3): 1082-90. doi: 10.1002/hep.24452.
[9] Pitug B. Prevalence of nonalcoholic fatty liver disease (NAFLD) and correlation between revalence of NAFLD and obesity in people living in borabue district, Mahasarakham province. Mahasarakham Hosp J. 2019; 14(2): 22-8. Available from: https://he02.tci-thaijo.org/index.php/MKHJ/article/view/204616 [cited 2025 Sep 11].
[10] Chairattanasakun J, Porruan R. The prevalence and associated factors of fatty liver among participants in Cholangiocarcinoma Screening and Care Program (CASCAP) of Mueang Chiang Rai District. Chiangrai Med J. 2024; 16(2): 1-9. Available from: https://he01.tci-thaijo.org/index.php/crm journal/article/view/261988 [cited 2025 Sep 11].
[11] Ashraf A, Yousaf F, Omer M, Bacha DR, Waseem M, Ullah K, et al. Sonographic association between fatty liver and gall bladder stones among all adult patients visiting private clinics of Lahore City. Health Med Sci. 2019; 2: 21-6. doi: 10.31014/aior.1994.02.01.15.
[12] Wang HH, Liu M, Clegg DJ, Portincasa P, Wang DQ. New insights into the molecular mechanisms underlying effects of estrogen on cholesterol gallstone formation. Biochim Biophys Acta. 2009; 1791(11): 1037-47. doi: 10.1016/j.bbalip.2009.06.006.
[13] Sun H, Warren J, Yip J, Ji Y, Hao S, Han W, et al. Factors influencing gallstone formation: A review of the literature. Biomolecules. 2022;12(4): 550. doi: 10.3390/biom12040550.
[14] Dong J, Dennis KMJH, Venkatakrishnan R, Hodson L, Tomlinson JW. The impact of estrogen deficiency on liver metabolism: Implications for hormone replacement therapy. Endocr Rev. 2025; Epub ahead of print. doi: 10.1210/endrev/bnaf018.
[15] Kim SE, Min JS, Lee S, Lee DY, Choi D. Different effects of menopausal hormone therapy on non-alcoholic fatty liver disease based on the route of estrogen administration. Sci Rep. 2023; 13(1): 15461. doi: 10.1038/s41598-023-42788-6.
[16] Kasarinaite A, Sinton M, Saunders PTK, Hay DC. The influence of sex hormones in liver function and disease. Cells. 2023; 12(12): 1604. doi: 10.3390/cells12121604.
[17] Charatcharoenwitthaya P, Kuljiratitikal K, Aksornchanya O, Chaiyasoot K, Bandidniyamanon W, Charatcharoenwitthaya N. Moderate-intensity aerobic vs resistance exercise and dietary modification in patients with nonalcoholic fatty liver disease: A randomized clinical trial. Clin Transl Gastroenterol. 2021; 12(3): e00316. doi: 10.14309/ctg.0000000000000316.
[18] Katsagoni CN, Georgoulis M, Papatheodoridis GV, Panagiotakos DB, Kontogianni MD. Effects of lifestyle interventions on clinical characteristics of patients with non-alcoholic fatty liver disease: A meta-analysis. Metabolism. 2017; 68: 119-32. doi: 10.1016/j.metabol.2016.12.006.
[19] Rattanangamkul M. Relation of non-alcoholic fatty liver disease (NAFLD) and metabolic syndrome in Thais.: College of Integrative Medicine, Dhurakij Pundit University; 2017.
[20] Kim YK, Kwon OS, Her KH. The grade of non-alcoholic fatty liver disease is an independent risk factor for gallstone disease: An observational Study. Medicine (Baltimore). 2019; 98(27): e16018. doi: 10.1097/md.0000000000016018.
[21] Li X, Gao P. Fatty liver increases gallstone disease risk in younger Chinese patients. Medicine (Baltimore). 2019; 98(22): e15940. doi: 10.1097/md.0000000000015940.
[22] Hallsworth K, Adams LA. Lifestyle modification in NAFLD/NASH: Facts and figures. JHEP Rep. 2019; 1(6): 468-79. doi: 10.1016/j.jhepr.2019.10.008.
[23] Jindarattanaporn N, Phulkerd S, Chamratrithirong A, Soottipong Gray R, Pattaravanich U, Loyfah N, et al. How an agreement with restriction of unhealthy food marketing and sodium taxation influenced high fat, salt or sugar (HFSS) food consumption. BMC Public Health. 2024; 24(1): 586. doi: 10.1186/s12889-024-18069-w.
[24] Lim GY, Chang Y, Kim I, Ryu S, Kwon R, Song J. Long working hours and increased risks of lean non-alcoholic fatty liver disease among Korean men and women. Sci Rep. 2023; 13(1): 12230. doi: 10.1038/s41598-023-39154-x.
[25] Ling Z, Zhang C, He J, Ouyang F, Qiu D, Li L, et al. Association of healthy lifestyles with non-alcoholic fatty liver disease: A prospective cohort study in Chinese government employees. Nutrients. 2023; 15(3): 604. doi: 10.3390/nu15030604.
[26] Rehman T, Shafi S, Amwar M, Nazeer M, Hanjra H, Hassan S, et al. Sonographic evaluation of cholelithiasis and its correlation with normal/fatty liver. Int J Health Med Nurs Pract. 2021; 91: 51-60. doi: 10.7176/JHMN/91-07.
[27] Lu Y, Hu L, Song J, Wan J, Chen H, Yin J. Gallstone disease and nonalcoholic fatty liver disease in patients with type 2 diabetes: a cross-sectional study. BMC Endocr Disord. 2021; 21(1): 231. doi: 10.1186/s12902-021-00899-z.

Journal of Associated Medical Sciences 2026; 59(1): 145-152

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

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.

ARTICLE INFO

Article history:
Received 24 July 2025
Accepted as revised 27 October 2025
Available online 30 October 2025

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 seven subskills. 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 identitiescore 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 participation in daily activities, social relationships, and rehabilitation.[1-2] Performance-based assessments, such as the Executive Function Performance Test (EFPT), have shown utility in detecting subtle executive dysfunction among marginalized individuals, even when standard cognitive screenings do not indicate impairments.[3]

* **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: 10.12982/JAMS.2026.018
E-ISSN: 2539-6056

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.[9-11] However, its use in Southeast Asian institutional settings remains underexplored, despite systematic reviews highlighting the importance of structured functional assessments for individuals transitioning from homelessness.[12-13] Moreover, psychological factors such as emotional regulation, motivation, and purpose are often overlooked in standard psychiatric or cognitive evaluations, particularly among older or institutionalized populations.[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 provides 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 a government-operated protective shelter in Thailand. The study adhered to the STROBE guidelines for observational research. Ethical approval was granted by (removed for anonymity). 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 mental disorder diagnosis 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 by shelter psychologists, social workers, and occupational therapists to ensure capacity for functional work-like activity. The Brief Psychiatric Rating Scale (BPRS) was selected to assess the severity of psychiatric symptoms because it provides a dimensional measure across a broad range of symptom domains, rather than definitive diagnostic cut-off points. Although participants exhibited psychiatric symptomatology, only those without acute psychotic symptoms were included, as verified through psychiatric records and confirmation by clinical staff. This inclusion criterion ensured that participants were clinically stable and able to engage meaningfully in occupational tasks, while still representing the spectrum of chronic psychiatric conditions commonly observed in long-term institutional settings.[23]

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. To ensure suitability, the adapted version was reviewed by three occupational therapy experts, yielding a content validity index (CVI) of 1.00. In this study, the adapted tool also demonstrated acceptable internal consistency (Cronbach's α=0.68) and test–retest reliability (intraclass correlation coefficient, ICC=0.77). These psychometric properties support the use of the adapted AWP in this context.

The AWP measures motor, process, and communication-interaction skills on a 4-point ordinal scale. This structured, performance-based approach is increasingly used to understand functional needs among individuals in institutional or homeless settings.[12–13,15] For this study, seven AWP subskills were prioritized based on their applicability and reliability in the shelter context. For interpretation, domain and subskill scores were categorized into three levels based on the percentage of the maximum possible score: high (≥80%), moderate (50-79%), and low (<50%). Each participant completed one structured task 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.[24-26] 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 the screening and ethical readiness of participants. 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 the RUDAS scores ranging from 20 to 30 (mean=26.77, SD=2.06). The BPRS scores indicated varying psychiatric symptom severity, ranging from 18 to 39 (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,Λ1) [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 20-30]

Table 1. Demographic and clinical characteristics of participants (service users)(n=60)(Continue)

Characteristic	Mean (SD) or n (%)
BPRS score	24.87 (5.54) [range 18-39]
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).

Descriptive occupational performance

Total scores on the adapted AWP indicated substantial variation in occupational performance (mean=32.4, SD=5.8). Participants generally demonstrated higher scores in motor domains, while process and communication-interaction skills were comparatively lower. This pattern reflects functional heterogeneity in this institutionalized population. Domain and subskill scores are detailed in Table 2. Total scores on the adapted AWP demonstrated substantial variation across participants. Based on the established cut-off values, 35% of participants achieved high performance (≥80%), 50% showed moderate performance (50-79%), and 15% scored in the low performance range (<50%).

Table 2. AWP domain and subskill scores (n = 60)

AWP Domain	Subskill	Mean± SD	Range	Classification (n, %)
Motor	Total subskill score	19.58±0.83	16-20	High = 60 (100%)
	Posture	3.93±0.31		Moderate = 0
	Mobility	3.88±0.32		Low = 0
	Coordination	3.98±0.13		
	Strength	3.93±0.25		
	Physical Energy	3.85±0.36		
Process	Total subskill score	18.57±2.16	9-20	High = 57 (95.00%)
	Mental Energy	3.87±0.47		Moderate = 1 (1.67%)
	Knowledge	3.52±0.72		Low = 2 (3.33%)
	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	High = 50 (83.33%)
-Interaction	Physicality	3.70±0.53		Moderate = 10 (16.67%)
	Language	3.67±0.48		Low = 0
	Relations	3.38±0.83		
	Informations & Exchange	3.75±0.54		
Total AWP domian score		52.65±3.50	39-56	High = 58 (96.67%)
				Moderate = 2 (3.33%)
				Low = 0

Note: AWP = Assessment of Work Performance. Subskills grouped under each domain as evaluated in this sample.

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	.830	
Relations	.787	
Task organization	.694	
Organizations of space & objects	.689	
Knowledge	.668	
Mental energy	.630	
Informations & exchange		.758
Language		.745
Physicality		.389

Note: Extraction Method: 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, loss of roles, 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, greater task completion, 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, supporting previous findings that better cognitive status and fewer psychiatric symptoms correlate 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 evaluation 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 was approved 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.

Funding

This project was supported by the New Researcher Grant, fiscal year 2024, awarded by the 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 & editing,project administration, and funding acquisition; **Watthanaree Ammawat:** conceptualization, methodology, formal analysis, investigation, and writing-review & editing; **Maliwan Rueankam:** conceptualization, methodology, investigation, data curation, writing-review & editing, and supervision; **Winai Chatthong:** methodology, investigation, writing-review & editing; **Supalak Khemthong:** conceptualization, methodology, writing-original draft, review & editing.

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 homeless 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 J. Development and evaluation of validity and utility of the instrument Assessment of Work Performance (AWP) [dissertation]. Linköping: Linköping University; 2007. Available from: https://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-10577

[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] Yang Y, Hu L, Ng SM. Enhancing occupational identity and self-regulation through narrative-based OT in cancer care. Hong Kong J Occup Ther. 2020; 33(1): 19–26. doi.org/10.1177/1569186120917893.

[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:10.1017S1041610204000043.

[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. Available from: https://psycnet.apa.org/record/1994-27973-001

[24] Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis. Struct Equ Modeling. 1999;6(1):1–55.doi.org/10.1080/10705519909540118.

[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. Available from: 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(1): 153-161

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Assessment of natural radioactivity concentrations in soil samples from Mae Chaem District, Chiang Mai Province, Thailand using a HPGe detector

Siriprapa Somboon, Jirachaya Chaijinda, Kewalin Ruktinnakorn, Sorawipat Intamoon, Bharinee Promprakob, Tarika Thumvijit, Sompong Sriburee*

Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 16 September 2025
Accepted as revised 22 October 2025
Available online 31 October 2025

Keywords:
Natural radioactivity, health, soil, gamma spectrometry, HPGe.

ABSTRACT

Background: Understanding the existence of natural radioactivity in different soil origins is essential for evaluating associated radiological risks.

Objectives: The purpose of this study was to determine the natural radioactivity concentrations in soil samples from Mae Chaem district using a HPGe detector.

Materials and methods: Thirty-five soil samples were randomly collected for laboratory analysis from four geological regions, ten from Triassic granite (Trgr), Silurian Devonian-Carboniferous (SDCtp), and Ordovician (O) locations and five from the Quaternary sediments, namely the Terrace deposits (Qt) area. The radioactivity of ^{226}Ra, ^{232}Th, and ^{40}K was measured using a gamma spectrometry with a high-purity germanium detector. The three radiation hazard indices for soil samples, including absorbed dose rate in outdoor air (D_{air}), the annual effective dose equivalent (AED_{out}), and the excess lifetime cancer risk outdoor ($ELCR_{out}$) were determined.

Results: The concentration of natural radioactivity for ^{226}Ra of Trgr, SDCtp, O, and Qt were found in the range of 50 to 172, 46 to 179, 59 to 285, and 46 to 77 Bq kg^{-1}, respectively. The activity concentration of ^{232}Th varied from 69 to 148, 53 to 239, 60 to 273, and 51 to 120 Bq kg^{-1}, whereas the activity concentration of ^{40}K ranged from 351 to 1100, 299 to 1356, 190 to 1364, and 779 to 1188 Bq kg^{-1}, respectively. The D_{air} varied from 88 to 207, 68 to 275, 94 to 345, and 85 to 146 nGy h^{-1}, with average values 139, 131, 170, and 110 nGy h^{-1} for Trgr, SDCtp, O and Qt samples, respectively. All sample values were higher than the worldwide population-weighted value of 59 nGy h^{-1} reported by the United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR, 2000). The AED_{out} and the $ELCR_{out}$ values also were higher than global average values by UNSCEAR.

Conclusion: In soil samples collected from four specific regions in the Mae Chaem district, the average natural activity concentrations of ^{226}Ra, ^{232}Th, and ^{40}K were found to be above the worldwide population-weighted value. The D_{air}, AED_{out} in external terrestrial radiation from outdoor sources and $ELCR_{out}$ values were higher than reported by UNSCEAR. These results indicate a high radiological health risk associated with prolonged exposure to natural radionuclides in soils from the Mae Chaem district.

Introduction

The public is exposed to ionizing radiation emitted from both natural and man-made sources. Approximately 82% of the human absorbed dose is from two main natural sources which cannot be

* **Corresponding contributor.**
Author's Address: *Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: sompong.sriburee@cmu.ac.th
doi: 10.12982/JAMS.2026.019
E-ISSN: 2539-6056

avoided. The two main sources are cosmic rays acting on the atmosphere and radioactive nuclides that originated in the Earth's crust.[1] Soil is a major source of natural radioactivity, with soil nature influenced by the climate, type of parent material, surface slope, and decomposition time.[2] The parent material of soil mainly consists of natural weathered rock, including sedimentary rocks, igneous rocks, or geologically young sediments that overlie the bedrock. Soil type depends on the parent material from which it was derived. The geological and geographical conditions of a region significantly impact the natural radioactivity.[3] Additional, radiation levels are associated with the type of rock, which higher radiation levels are found in igneous rocks such as granite. In contrast, sedimentary rocks have lower radionuclide content. Nonetheless, radiation levels are modified by human practices. The main source of external exposure is emitted from primordial radionuclides in the ^{238}U series, ^{232}Th series, and ^{40}K. The emission of energy from these radionuclides poses human health hazards.[4] Moreover, radiation has detrimental effects causing genetic disorders, leukemia, and various forms of cancer.[5-7]

Hence, understanding the distribution of natural radioactivity in each location is important. This knowledge makes it possible to accurately assess radiation health hazard indices and estimate the risk associated with prolonged potential exposure to background radiation.[8-10]

Several previous studies have reported on the level of natural radioactivity in other countries, yet limited assessments on natural radioactivity have been conducted in Thailand. Most reports originate from the southern region of Thailand,[11-14] a region characterized by various types of rocks and mining activities. Mae Chaem district is located 80 km southwest of Chiang Mai province between latitude 18° 11' to 18° 21' N and longitude 98° 18' to 98° 30' E. The elevation of the district ranges between 400 and 2,565 meters above sea level. More than 70% of the district area is mountainous and there is a small amount of arable land. The geology of Mae Chaem district is characterized by a complex stratigraphic sequence of sedimentary and metamorphic rocks. The basement rocks in this region mainly consist of Precambrian gneiss, which is widely distributed and is predominant around the eastern watershed of the Mae Chaem River. This area also contains various rock types and quaternary sediments. Several factors complicate the stratigraphic sequence of layered sedimentary and metamorphic rocks in Mae Chaem, including variable metamorphic effects as well as erosion, daily human soil activity and thrust faulting.[15] Therefore, evaluating natural radioactivity in soils of Mae Chaem basin's complex geology is essential. This study determined the specific activity of ^{226}Ra, ^{232}Th, and ^{40}K in soil samples from four regions of Mae Chaem district to provide a baseline for background radiation levels.

Materials and methods

Location of the study area

The study area for this research was Mae Chaem district in Chiang Mai Province, Northern Thailand. Figure 1. presents a geological map of the area. The district consists of igneous rock in the form of Triassic granite (Trgr), various sedimentary rocks (sandstone, siltstone, shale, and limestone) and metamorphic rocks (schist, gneiss, and calc-silicate). The sedimentary and metamorphic rocks in Mae Chaem district consist of Quaternary sediments including Terrace deposits (Qt), Silurian-Devonian-Carboniferous (SDCtp), Precambrian (PE), Cambrian (C), Ordovician (O), and Tertiary Miocene (Tm). The population of Mae Chaem is primarily located in the lowland area below 600 meters above sea level (m.a.s.l.). Some small villages are located within the mid-elevation zone, which ranges from 600 to 1,000 m.a.s.l., while some communities are found at higher elevations. Most ethnic northern Thai are clustered near the Chaem river basin.[15] The study area was therefore confined to the lowland region adjacent to the Chaem River. In north of Mae Chaem district is difficult to navigate and access because it is highly forested with limited settlement, while the steep hills and mountains restrict global positioning system (GPS) connectivity. Consequently, only four specific geological formations were selected, namely Trgr, SDCtp, O, and Qt region. Soil samples were collected randomly along the route in the Chaem basin. GPS and a geographic information system (GIS) were used to randomly select soil sampling locations, with a minimum distance of 1 kilometer between samples.

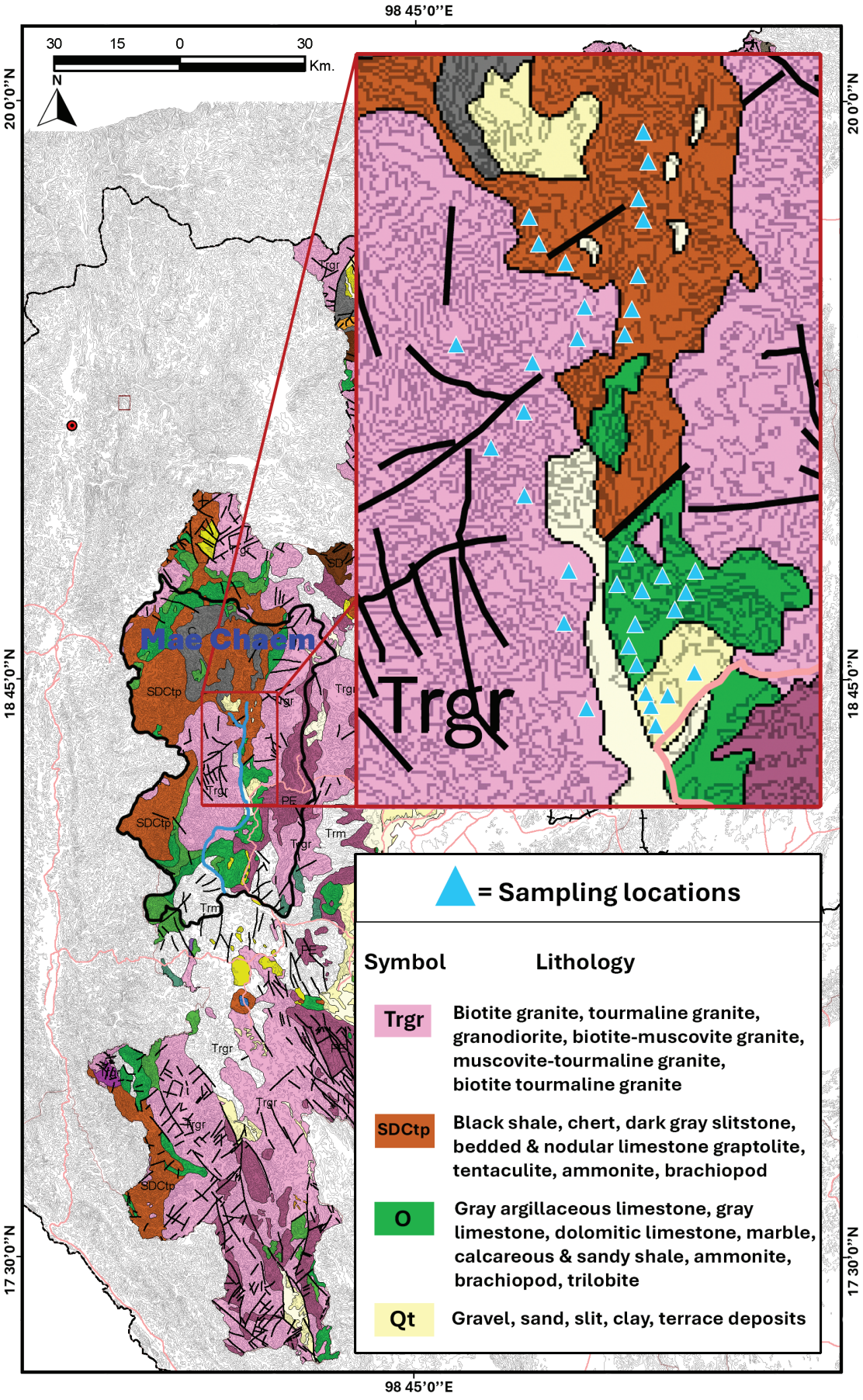


Figure 1. *Sampling locations for the four geological formations.*

Sampling and sample preparation

Thirty-five soil samples from four different geological areas were collected for laboratory analysis. These included ten sample each from the Trgr, SDCtp, and O areas, and five from the Qt area of the alluvial plain. Before soil collection, surface leaves, grass, and other organic material were removed. One kilogram of soil was taken at a depth of 0 to 20 cm and packed in a plastic bag for transport to the laboratory. The samples were dried in an oven at 110°C until a consistent weight was achieved. After drying, the samples were crushed into a fine powder and sieved to a particle size of less than 2 mm. The samples were deposited into 100 ml plastic containers and sealed tightly. Each sample was stored for a minimum of 4 weeks before gamma spectrometry measurements to ensure radioactivity equilibrium between ^{226}Ra, ^{232}Th, and their decay products.[16]

Radionuclides measurements

Specific activities of the radionuclides ^{226}Ra, ^{232}Th, and ^{40}K were measured using a high-purity germanium (HPGe) coaxial detector (Canberra, Model, USA). Data collection and analysis were conducted using a Genie2000 software package (Canberra, USA). The detector's energy calibration was conducted using mixed radionuclide point sources of ^{133}Ba, ^{137}Cs, ^{57}Co, and ^{60}Co. An energy resolution of approximately 1.8 keV FWHM at 1332 keV and an efficiency of approximately 33% were utilized. The container's geometric efficiency was determined by the IAEA-478 reference materials (International Atomic Energy Agency, Vienna, Austria). Each sample was measured with an acquisition time of 24 hours. The specific activities of radionuclides ^{226}Ra, ^{232}Th, and ^{40}K were analyzed from gamma ray peaks at various energies. The activity of ^{226}Ra radionuclides were determined using gamma rays from ^{214}Pb at 351.9 keV and ^{214}Bi at 609.3 keV. The gamma ray peaks from ^{208}Tl at 583.2 keV and ^{228}Ac at 911.2 keV were used to determine ^{232}Th, while the gamma peak energy of 1460.8 keV was used to evaluate ^{40}K.[17]

Calculation of radiation hazard indices for soil samples

Three parameters were determined for the soil samples including absorbed dose rate in outdoor air (D_{air}; nGy h^{-1}), the annual effective dose equivalent (AED_{out};mSv y^{-1}), and the excess lifetime cancer risk outdoor ($ELCR_{out}$) using the following equations:

$$D_{air}\ (nGy\ h^{-1}) = 0.462A_{Ra} + 0.604A_{Th} + 0.0417A_{K}$$

$$AED_{out}\ (mSv\ y^{-1}) = D \times 8760 \times 0.2 \times 0.7 \times 10^{-6}$$

$$ELCR_{out} = AED_{out} \times LF \times RF$$

Where A is a specific activity of radionuclides, LF and RF are the life expectancy of Thai people (75 years)[18] and risk factor by International Commission on Radiological Protection[19] 103 recommends a threshold of 0.05 Sv^{-1} for general population, leading to stochastic effects in low bakground radiation, respectively.

Results

Natural radioactivity concentration

The activity concentrations of natural radionuclides ^{226}Ra, ^{232}Th, and ^{40}K in all samples were compared with the data from UNSCEAR and the average values for southern Thailand[12] which have previously been research, as shown in Figure 2. The specific activity concentrations for ^{226}Ra of soil ranged from 50 to 172 (Trgr), 46 to 179 (SDCtp), 59 to 285 (O), and 46 to 77 Bq kg^{-1}(Qt). The average values for ^{226}Ra were 102±43 (Trgr), 85±48 (SDCtp), 128 ± 68 (O), and 56±12 Bq kg^{-1} (Qt). These values exceed the worldwide population-weighted value of 33 Bq kg^{-1}. For ^{232}Th, the activity concentrations ranged from 69 to 148 (Trgr), 53 to 239 (SDCtp), 60 to 273 (O), and 51 to 120 Bq kg^{-1} (Qt). The average concentrations were 102±25, 110±65, 132±63, and 72±29 for Trgr, SDCtp, O, and Qt, respectively. These values are also higher than the worldwide population-weighted value of 45 Bq kg^{-1}. The specific activity concentrations of ^{40}K ranged from 351 to 1100 (Trgr), 299 to 1356 (SDCtp), 190 to 1364 (O), and 779 to 1188 Bq kg^{-1} (Qt). The radioactivity concentrations were 685±280 (Trgr), 560±323 (SDCtp), 702±356 (O), and 969±146 (Qt). These values are also above the worldwide population-weighted value of 420 Bq kg^{-1}. The activity concentrations of natural radionuclides ^{226}Ra, ^{232}Th, and ^{40}K varied across different regions. However, a one-way ANOVA performed with SPSS version 17 (FB7E105EFD8A514130CC) indicated no statistically significant differences between the geological groups. The p-values obtained were 0.074, 0.216, and 0.133, respectively.

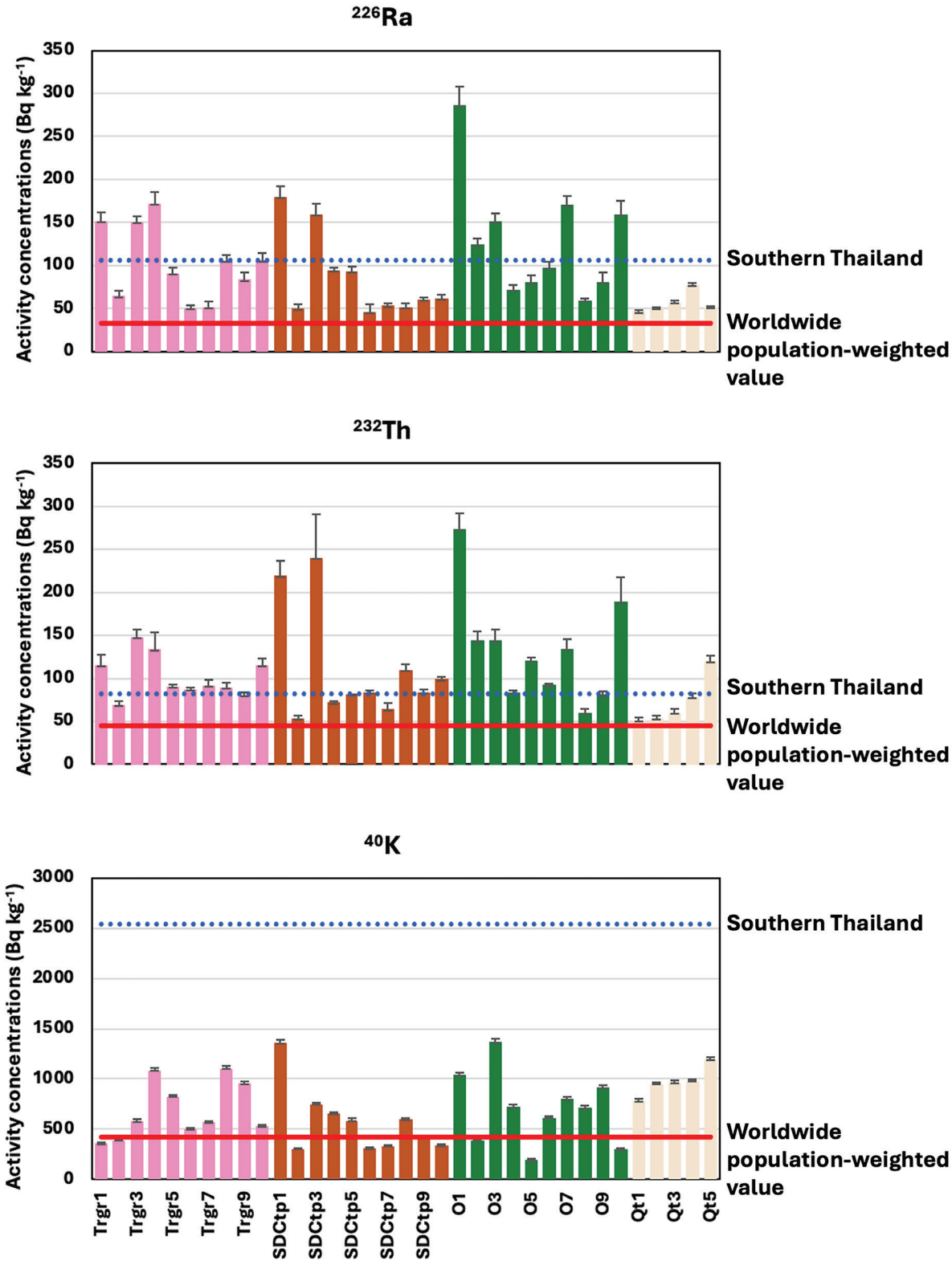


Figure 2. *Comparison of activity concentrations of natural radionuclides ^{226}Ra, ^{232}Th, and ^{40}K in all samples to UNSCEAR 2000 and averages data from southern Thailand.*

Radiological hazard indices

Figure 3. presents the three radiological hazard indexes. The absorbed dose rates were converted into an annual effective dose rate by adopting a conversion factor of 0.7 Sv Gy^{-1} and 0.2 as the outdoor occupancy factor, considering that people spend on average 20% of their time outdoors. The absorbed dose rates caused by terrestrial gamma rays at 1 meter above ground ranged from 88 to 207 (Trgr), 68 to 275 (SDCtp), 94 to 345 (O), and 85 to 146 nGy h^{-1}(Qt). The average values for these samples were 139 (Trgr), 131 (SDCtp), 170 (O), and 110 nGy h^{-1} (Qt). All sample values were higher than the value of 59 nGy h^{-1} reported by UNSCEAR 2000. The AED_{out} in all areas exceeded the external terrestrial radiation from outdoor sources, which is valued at 0.07 mSv y^{-1}. For $ELCR_{out}$, all samples exceed the worldwide average value of 0.25×10^{-3} reported by UNSCEAR.

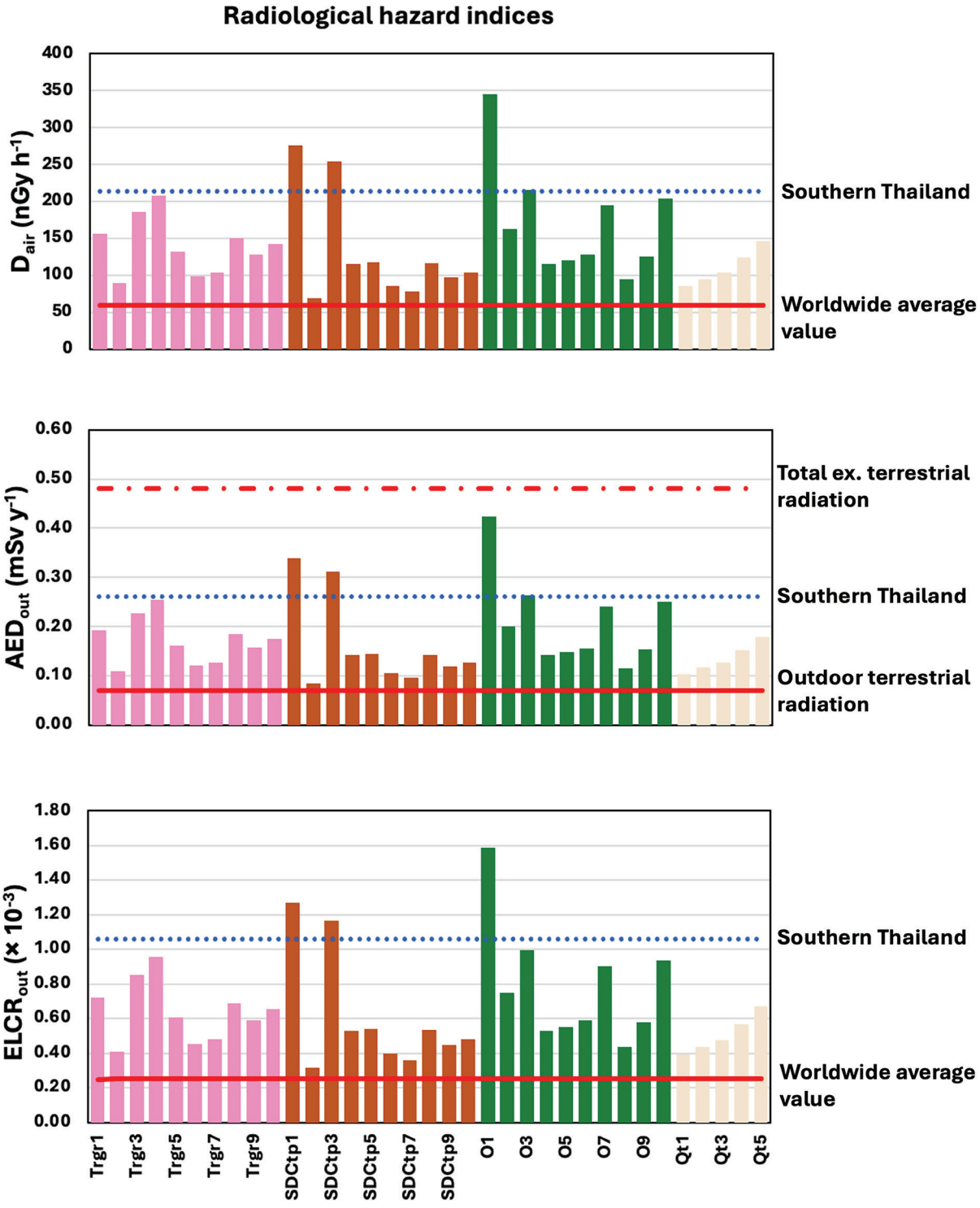


Figure 3. Comparison of three radiological hazard indices with UNSCEAR 2000 and data from southern Thailand.

Discussion

The average concentrations of natural radionuclides in Mae Chaem were found to be higher than the worldwide population-weighted value in the UNSCEAR 2000 report. The concentrations of elements in the geological formations caused variations in the levels of natural radionuclide activity at the different sampling sites. Typically, higher radiation levels are associated with granite from igneous rocks.[20] Autsavapromporn *et al*.[21] reported the amount of natural radionuclides ^{226}Ra, ^{232}Th, and ^{40}K in 48 soil samples at Kong Khaek subdistrict, Mae Chaem to be 47±20.9, 77.9±29.7 and 700.1±233 Bq kg^{-1}, respectively. That sample site is situated south of Mae Chaem district, and its geology is classified as type O. The natural radionuclides ^{226}Ra and ^{232}Th of these results were found to be twice as high as reported by Autsavapromporn, whereas the level of ^{40}K was approximately equal. Pisapak *et al*.[14] found that the Trgr geological area had a higher concentration of ^{226}Ra and ^{232}Th than the Qt area in Songkhla Province. In this study, activity concentrations of ^{226}Ra and ^{232}Th were high in O, Trgr, SDCtp, and Qt areas, respectively. Variations were caused by the deposited radionuclides in the soil or non-uniform distribution from variable sedimentary transport. Metamorphic and sedimentary rocks deriving from felsic materials inherit their parent rocks. A significant increase in sedimentary rocks can be achieved through the density driven accumulation of these minerals, typical of placer deposits, as well as by absorption and/or adsorption of radionuclide in organic matter. Some geographic areas with phosphate limestone in sedimentary rocks may contain radioactive minerals in their rock matrix. Arafa et al.[22] reported that both ^{226}Ra and ^{232}Th were found in higher amounts in sediment rock samples. The activity concentration of ^{40}K in soil is often associated with the presence of potassium-rich feldspars and micas, common in granitic and metamorphic rocks. However, in this study, the high ^{40}K levels in the O area are likely enhanced by agricultural activity. Approximately 82.08% of the Mae Chaem district is covered by forests. The largest land use in the district is agriculture at 16.92% of the total area.[23] Corn is the primary crop grown in Mae Chaem, supplemented by a few leafy greens, potatoes and other vegetables. Chemical fertilizers are employed by farmers in agricultural fields to enhance agricultural productivity. Chemical fertilizers mostly consist of phosphate, nitrate, ammonium and potassium salts. The application of these chemical fertilizers can increase the concentrations of natural radioactive nuclides, especially when potassium sulfate is used.[24] Extensive use of potassium-based fertilizers is a well-known source of elevated soil ^{40}K concentrations.[25] Human exposure to terrestrial radiation is influenced by the radionuclide content in the soil and other environmental factors but also depends on anthropogenic activities.[26, 27] Most previous studies have reported the terrestrial radiation dose in the southern region of Thailand, where various rocks and quaternary deposits are found. Additionally, several mineral resources have been discovered in this region. The southern region therefore has elevated natural radioactivity levels. The present study indicates that the average radioactivity of ^{232}Th in the study area was higher than levels reported in southern Thailand, while average the radioactivity levels of ^{226}Ra and ^{40}K were lower. This could be attributed to geological features. For instance, compared to the southern region, Mae Chaem district has a higher ratio of igneous rocks which results in thorium having a higher radioactivity than uranium.[20]

The absorbed dose rates were within range of external exposure rates from terrestrial gamma radiation in the Asia region report by UNSCEAR in 2008.[28] The high absorbed dose rates in air throughout the world are associated with thorium-bearing and uranium-bearing minerals in the soil. The outdoor absorbed dose rate in air, derived from soil radionuclide concentrations and compared with direct measurements in Thailand, was 0.8. Therefore, the absorbed dose rates in air will exceed the estimated values. The ambient background gamma dose rate in air at any specific location is not constant in time, in addition to variations from place to place. Consequently, ongoing surveillance is advised. The AED_{out} in all area was higher than the external terrestrial radiation from outdoor sources (0.07 mSv y^{-1}). However, the AED_{out} values were lower when compared to the total external terrestrial radiation reported by UNSCEAR (0.48 mSv y^{-1}) and the recommendations (1 mSv y^{-1}) from the International Commission on Radiological Protection (ICRP).[19] Chiang Mai[29] and southern Thailand have higher outdoor terrestrial radiation exposure than the global average value. The activity concentrations of ^{226}Ra, ^{232}Th, and ^{40}K, the absorbed dose rates, and the AED_{out} in Asian countries such as Vietnam[30] and China[31], are also higher than the worldwide average value. Additionally, indoor exposure to gamma rays from construction materials necessitates consideration of the total external terrestrial exposure. In Thailand, wood-frame construction is common, and the indoor and outdoor ratio was 0.6. Therefore, indoor exposure may have a low effect on total external terrestrial exposure compared with a high-ratio area. For $ELCR_{out}$, all samples exceed the worldwide average value reported by UNSCEAR (0.25×10^{-3}). The $ELCR_{out}$ refers to the likelihood of an individual developing cancer over their lifetime because of exposure to ionizing radiation. The investigated areas in this result have higher radiological health risks due to extended exposure to natural radionuclides from soils. Moreover, the radioactive series of ^{238}U and ^{232}Th both produce radon isotopes during their decay chains. The radon isotopes are gases that emit alpha particles during decay. These particles pose health hazards predominantly through the inhalation. Nevertheless, this study did not evaluate the internal radiation exposure from radon, which is a limitation of the investigation.

Conclusion

The natural activity concentrations of ^{226}Ra, ^{232}Th, and ^{40}K in soil from the Trgr, SDCtp, O, and Qt geological areas in the Mae Chaem district were measured. The radioactivity concentrations were found to exceed the average worldwide population-weighted value reported by UNSCEAR 2000. The absorbed dose rates in outdoor air, the AED_{out}, and the $ELCR_{out}$ were higher than the worldwide average values by UNSCEAR. These results indicate that long-term exposure to natural radioactivity in the soils from the Mae Chaem district poses a substantial radiological health risk. The results of this study can be utilized as a database for monitoring and comparing future natural radioactivity changes.

Ethical approval

The study did not include human or animal subjects and did not necessitate ethical approval.

Funding

None.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Siriprapa Somboon: Conceptualization, Methodology, Formal analysis, Investigation, and Writing – review & editing. **Jirachaya Chaijinda:** Methodology, Formal analysis, and Investigation. **Kewalin Ruktinnakorn:** Methodology and Investigation, **Sorawipat Intamoon:** Methodology, Formal analysis, and Investigation, **Bharinee Promprakob:** Methodology, Formal analysis, and Investigation, **Tarika Thumvijit:** Conceptualization, Methodology, and supervision, **Sompong Sriburee:** Conceptualization, Methodology, and supervision.

Acknowledgements

The author would like to thank Miss Ampika Jawana for her help in data analysis.

References

[1] UNSCEAR. Sources and effects of ionizing radiation. United Nations Scientific Committee on the Effects of Atomic Radiation, UNSCEAR 2000 Report to the general assembly, with scientific annexes, Volume I: Sources New York: United Nations; 2000. doi: 10.1016/j.chemosphere.2020.128459.

[2] Khan R, Islam HMT, Islam A. Mechanism of elevated radioactivity in Teesta river basin from Bangladesh: Radiochemical characterization, provenance and associated hazards. Chemosphere. 2021; 264(Pt 1): 128459. doi: 10.1016/j.chemosphere.2020.128459.

[3] Merdanoglu B, Altinsoy N. Radioactivity concentrations and dose assessment for soil samples from Kestanbol granite area, Turkey. Radiat Prot Dosim. 2006; 121(4): 399-405. doi: 10.1093/rpd/ncl055

[4] Kreuzer M, Sobotzki C, Fenske N, Marsh JW, Schnelzer M. Leukaemia mortality and low-dose ionising radiation in the WISMUT uranium miner cohort (1946-2013). Occup Environ Med. 2017; 74(4): 252-8. doi: 10.1136/oemed-2016-103795

[5] Hsieh WH, Lin IF, Ho JC, Chang PW. 30 years follow-up and increased risks of breast cancer and leukaemia after long-term low-dose-rate radiation exposure. Br J Cancer. 2017; 117(12): 1883-7. doi: 10.1038/bjc.2017.350.

[6] Leuraud K, Laurier D, Gillies M, Haylock R, Kelly-Reif K, Bertke S, et al. Leukaemia, lymphoma, and multiple myeloma mortality after low-level exposure to ionising radiation in nuclear workers (INWORKS): updated findings from an international cohort study. Lancet Haematol. 2024; 11(10): e761-e9. doi: 10.1016/S2352-3026(24)00240-0.

[7] Little MP, Wakeford R, Zablotska LB, Borrego D, Griffin KT, Allodji RS, et al. Radiation exposure and leukaemia risk among cohorts of persons exposed to low and moderate doses of external ionising radiation in childhood. Br J Cancer. 2023; 129(7): 1152-65. doi: 10.1038/s41416-023-02387-8.

[8] ICRP. Protection of the public in situations of prolonged radiation exposure. The application of the Commission's system of radiological protection to controllable radiation exposure due to natural sources and long-lived radioactive residues. Ann ICRP. 1999; 29(1-2): 1-109.

[9] WHO Handbook on indoor radon: A public health perspective. WHO Guidelines approved by the Guidelines Review Committee. Geneva2009.

[10] Mortazavi SMJ, Mortazavi Gh, Mortazavi SAR, Paknahad M. Is induction of anomalies in lymphocytes of the residents of high background radiation areas associated with increased cancer risk? J Biomed Phys Eng. 2019; 9(3): 367-72. doi: 10.31661/jbpe.v9i3Jun.654.

[11] Bhongsuwan T, Auisui SA. A high natural radiation area in Khao-Than hot spring, Southern Thailand. Radiat Prot Dosim. 2015; 167(1-3): 284-8. doi: 10.1093/rpd/ncv263.

[12] Kessaratikoon P, Boonkrongcheepa R, Polthum S. Measurement of radionuclides in surface soil and beach sand samples from Rayong Province (Thailand) and the evaluation of excess lifetime cancer risk. Scienceasia. 2021; 47: 120-9. doi: 10.2306/scienceasia1513-1874.2021.S016.

[13] Kritsananuwat R, Arae H, Fukushi M, Sahoo SK, Chanyotha S. Natural radioactivity survey on soils originated from southern part of Thailand as potential sites for nuclear power plants from radiological viewpoint and risk assessment. J Radioanal Nucl Ch. 2015; 305(2): 487-99. doi: 10.1007/s10967-015-3994-8.

[14] Pisapak P, Todorovic N, Bhongsuwan T. Correlation between radon and radium concentrations in soil and estimation of natural radiation hazards in Namom district, Songkhla province (Southern Thailand). Environ Earth Sci. 2017; 76(4): 1-11.

doi:10.1007/s12665-017-6439-6.
[15] Thanapakpawin P, Richey J, Thomas D, Rodda S, Campbell B, Logsdon M. Effects of landuse change on the hydrologic regime of the Mae Chaem river basin, NW Thailand. J Hydrol. 2007; 334(1-2): 215-30. doi: 10.1016/j.jhydrol.2006.10.012.
[16] Sroor A, El-Bahi SM, Ahmed F, Abdel-Haleem AS. Natural radioactivity and radon exhalation rate of soil in southern Egypt. Appl Radiat Isot. 2001; 55(6): 873-9. doi: 10.1016/S0969-8043(01)00123-3.
[17] Kranrod C, Chanyotha S, Pengvanich P, Kritsananuwat R, Hosoda M, Tokonami S. A Comparative study of the outdoor absorbed dose rate in air by In-situ and soil-sampling-based measurement methods. Radiat Environ Med. 2020; 9: 98-104. doi: 10.51083/radiatenvironmed.9.2_98.
[18] World Health Organization. Life expectancy and Healthy life expecancy - Data by country [Internet]. Published online 2020 [cited 11 October 2025]. Available from: https://data.who.int/countries/764.
[19] ICRP. Recommendations of the International Commission on Radiological Protection. Publication 103. 2007.
[20] Alnour IA. Concentration of uranium and thorium in granite rock in Kadugli, Sudan using γ-ray spectroscopy. Phys Sci Int J. 2019; 22(4): 1-7. doi: 10.9734/PSIJ/2019/v22i430137.
[21] Autsavapromporn N, Kranrod C, Kritsananuwat R, Sola P, Klunklin P, Chitapanarux I, et al. Health Impacts of Natural Background Radiation in High Air Pollution Area of Thailand. Toxics. 2024; 12(6): 1-17. doi: 10.3390/toxics12060428.
[22] Arafa W, Mahmoud H, Yousf E, Ashry A, Elaassy I, Elsersy A. Assessment of radiological hazards of sedimentary, igneous and sediments natural rocks. World J Nucl Sci Technol. 2024; 14: 131-45. doi: 10.4236/wjnst.2024.142008.
[23] Supasri T, Itsubo N, Gheewala SH, Sampattagul S. Life cycle assessment of maize cultivation and biomass utilization in northern Thailand. Sci Rep. 2020; 10(1): 3516. doi: 10.1038/s41598-020-60532-2.
[24] Abdullah HM, Ahmed AH, Samad AI, Hussein MI. Radioactivity measurements of chemical and organic fertilizers applied to farms in the Kurdistan Region, Iraq. Isotopes Environ Health Stud. 2025; 26: 1-12. doi: 10.1080/10256016.2025.2556934.
[25] Alharbi W. Natural radioactivity and dose assessment for brands of chemical and organic fertilizers used in Saudi Arabia. J Mod Phys. 2013; 4(03): 344-8. doi: 10.4236/jmp.2013.43047
[26] Alzubaidi G, Hamid FB, Abdul Rahman I. Assessment of natural radioactivity levels and radiation hazards in agricultural and virgin soil in the State of Kedah, North of Malaysia. Sci World J. 2016; 2016: 6178103. doi: 10.1155/2016/6178103.
[27] Mwalongo DA, Haneklaus NH, Carvalho FP, Lisuma JB, Kivevele TT, Mtei KM. Influence of phosphate fertilizers on the radioactivity of agricultural soils and tobacco plants in Kenya, Tanzania, and Uganda. Environ Sci Pollut Res Int. 2023; 30(34): 83004-23. doi: 10.1007/s11356-023-27543-8.
[28] UNSCEAR. Sources and effects of ionizing radiation. United Nations Scientific Committee on the Effects of Atomic Radiation. UNSCEAR 2008 Report to the general assembly, with scientific annexes, Volume I. New York: United nations; 2010.
[29] Autsavapromporn N, Krandrod C, Klunklin P, Kritsananuwat R, Jaikang C, Kittidachanan K, et al. Health effects of natural environmental radiation during burning season in Chiang Mai, Thailand. Life (Basel). 2022; 12(6): 1-13. doi: 10.3390/life12060853.
[30] Duonga VH, Santhanabharathib B, Pradhoshinib KP, Trunga TC, Huong Ngod TT, Nguyene T, et al. Radiological hazard assessments of natural radionuclides in surface soil in Chau Binh, Vietnam. J Environ Pollut Manag. 2025;2:172-81. doi: 10.1016/j.epm.2025.07.002.
[31] Huang Y, Wen W, Liu J, Liang X, Yuan W, Ouyang Q, et al. Preliminary Screening of Soils Natural Radioactivity and Metal(loid) Content in a Decommissioned Rare Earth Elements Processing Plant, Guangdong, China. Int J Environ Res Public Health. 2022;19(21). doi: 10.3390/ijerph192114566.

Journal of Associated Medical Sciences 2026; 59(1): 162-167

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

Reliability of the Thai version of the Spinal Cord Independence Measure Self-Report (SCIM-SR-Thai)

Kodchakorn Mekavutikul and Siam Tongprasert*

Department of Rehabilitation Medicine, Faculty of Medicine, Chiang Mai University, Thailand.

ARTICLE INFO

Article history:
Received 9 September 2025
Accepted as revised 27 October 2025
Available online 3 November 2025

Keywords:
Spinal cord injuries, disability evaluation, self report, psychometrics.

ABSTRACT

Background: The SCIM-SR, a self-report instrument, offers a patient-centered perspective. To ensure its clinical and scientific utility in Thailand, it is essential to establish the reliability of a translated and culturally adapted Thai version.

Objectives: To investigate the reliability of the SCIM-SR-Thai in chronic persons with SCI using a test-retest approach after 7 days.

Materials and methods: Thai adults aged 20-80 years with chronic SCI for at least 1 year were recruited. Participants filled out the SCIM-SR-Thai in-office once and were given a second copy to be completed at 7 days at their leisure. Participants were instructed to mail the questionnaire back within 14 days or would be considered a drop-out.

Results: A total of 59 participants completed the study. The majority were male (62.7%). The mean (SD) age was 48.5 (15.4) with median (IQR) time since SCI of 12 (6-21) years. Most of the participants were diagnosed with T1-S3 AIS A, B, and C 38/59 (64.4%). The SCIM-SR-Thai demonstrated excellent internal consistency and test-retest reliability (Cronbach's alpha 0.99, ICC=0.97; 95% CI=0.95-0.98). Analyses of individual subscales yielded similar results.

Conclusion: SCIM-SR-Thai demonstrates good reliability and internal consistency. The authors encourage adoption of this questionnaire as a standardized tool in clinical practice to assess functional independence in people with SCI.

Introduction

Chronic spinal cord injury (SCI) poses devastating lifelong consequences for persons with SCI and their families. The psychological and economic burden are immense, as they are abruptly striped away of their productive life and become dependent upon their caregivers.[1,2] Being confined to their beds, persons with SCI are also prone to serious medical complications, including pressure injuries, urinary tract infections, deep venous thrombosis, and autonomic dysreflexia, that further limits their longevity.[1] Therefore, after the acute stabilization of neurological state, rehabilitation during the subacute phase is critical in maintaining functional independence, which in turn would improve quality of life and prevent long term complications.[2]

The health care providers-assessed Spinal Cord Independence Measure Version III (SCIM III) is a widely accepted outcome measure that assesses activities of

* **Corresponding contributor.**
Author's Address: *Department of Rehabilitation Medicine, Faculty of Medicine, Chiang Mai University, Thailand.*
E-mail address: siam.tongpr@cmu.ac.th
doi: 10.12982/JAMS.2026.020
E-ISSN: 2539-6056

daily living (ADLs) and rehabilitation outcomes in SCI.[3] The SCIM III scores patient out of 100 by evaluating three domains: Self-care (6 items, score 0-20), Respiratory and Sphincter Management (4 items, score 0-40), and Mobility (9 items, score 0-40).[3] To further enable continuous monitoring in the community, a self-report version of the SCIM III (SCIM-SR) has been developed by Fekete *et al*.[4] It has been cross-culturally adapted and translated into multiple languages.[5-12] The Thai version of the SCIM-SR was first translated from English by Wilartratsami S. *et al.*, but their prototype was developed using cohort of healthy participants.[9] Their final version when tested in Thai people with SCI demonstrated excellent internal validity (Cronbach's alpha=0.91) and concurrent validity when compared to the Thai version of the SCIM III (Pearson's coefficient=0.94).[9] However, temporal consistency was tested only at a 3-day interval (Intraclass correlation coefficient=0.95), and thus the effects of memory and learning cannot be excluded at such a short timeframe.[9] Terwee *et al*. noted a duration of 1-2 weeks between test-retest would be appropriate to minimize patient's recall.[13]

To overcome these limitations, the authors have developed an in-house version of the self-report (SCIM-SR-Thai), which has undergone meaningful cross-cultural adaptations to better fit persons with SCI's context. The questionnaire was tested in actual SCI cases, as opposed to healthy subjects. While the authors have demonstrated comparable internal consistency (Cronbach's alpha=0.96) and concurrent validity (Pearson's coefficient=0.97) to Wilartratsami S *et al*.[9,12] its reproducibility has yet been explored. Because long-term surveillance of patient's functional dependence is essential to rehabilitation outcomes, the authors aimed to investigate the reliability of the SCIM-SR-Thai in chronic spinal cord injury patients using a test-retest approach after 7 days.

Materials and methods

This study was conducted between December 2023 to January 2024 targeting chronic Thai people with SCI at an outpatient clinic of a university hospital in Northern Thailand.

Participants

The inclusion criteria included patients; 1) aged between 20 to 80 years, and 2) diagnosed with SCI either from traumatic or non-traumatic causes for over a year; 3) who are proficient in Thai and can understand and answer the self-report either with or without assistance in writing. Patients with other comorbidities including neurocognitive impairment, psychiatric conditions, traumatic brain injury, and severe medical illnesses were excluded.

Sample size calculation

Sample size was calculated using the Bonett's formula to ensure a sufficiently narrow confidence interval for calculating the intraclass correlation coefficient (ICC).[14] The anticipated value of the ICC was based on results from Wilartratsami *et al.* of 0.95.[9] According to this estimation (as shown below), 59 participants were required.

$$N= 1 + 8Z^2_{\alpha/2}[(1-p)(1+k(k-1)p)]^2/ k(k-1)w^2$$

where,

- p is anticipated value of the ICC=0.95
- W is the desired width of the confidence interval= 0.2
- Number of item or raters or repeats (K) = 2
- $Z_{\alpha/2}$ is value from the standard normal distribution for the desired confidence level= 1.96 (for a 95% confidence interval).

Research instrument

The Thai version of the Spinal Cord Independence Measure Self-Report (SCIM-SR-Thai) consists of 19 self-reported questions (appendix). The SCIM-SR-Thai measures functional independence in three subscales scored total of 100: Self-care (6 items, score 0-20), Respiratory and Sphincter Management (4 items, score 0-40), and Mobility (9 items, score 0-40). The total score ranges from 0 to 100 with higher scores indicating a higher level of functional independence.[12]

Data collection

Participants were given informed consent before study initiation. Baseline demographic and clinical data were collected in accordance with the International SCI Core Data Set (Version 3.0).[15] Participants were to respond to the SCIM-SR-Thai in-office. Subjects were then given a second copy of the SCIM-SR-Thai and were instructed to complete it at 7 days at their leisure. Caregiver assistance in reading or comprehending the questions was not permitted. Finally, participants were instructed to mail the results back to the researchers via online messaging platform. Failure to submit the self-report within 14 days was considered a drop-out.

Data analysis

Statistical analysis was performed using SPSS statistics 28.0.0.[16] The frequency and percentage were used to describe categorical variables, mean and standard deviation (SD) for normally distributed numerical variables, and median with interquartile range (IQR) for non-normally distributed numerical variables. Since our results are non-normally distributed, the median was reported. Internal consistency was assessed using Cronbach's alpha with an acceptable cut-off of 0.70.[13] Reliability assessment was determined by test–retest reliability using the intraclass correlation coefficient (ICC), in which an ICC of at least 0.70 was considered reliable.[13]

Results

A total of 65 participants were recruited and 59 completed the study (Table 1). The other 6 participants did not completely answer questions 6 and 7 on sphincter management. The majority were male (62.7%). The mean (SD) age was 48.5 (15.4) with median (IQR) time since SCI of 12 (6-21) years. Most of the participants were diagnosed with T1-S3 AIS A, B, and C 38/59 (64.4%).

The median for test and re-test total SCIM-SR-Thai scores were 57 (48-68) and 56 (39-66), respectively (Table 2). The SCIM-SR-Thai demonstrated excellent internal consistency (Cronbach's alpha 0.99) and test-retest reliability at 7 days with an ICC of 0.97 (95% CI: 0.95-0.98). Additionally, analyses of individual subscales, namely self-care, respiration & sphincter management, and mobility, yielded similar results with all achieving acceptable internal consistency and reproducibility (Table 3). Self-care and mobility subscales showed better reliability at 7 days than respiratory & sphincter management subscale, with ICC of 0.96 (95% CI: 0.94-0.98) and 0.97 (95% CI: 0.95-0.98) versus 0.86 (95% CI: 0.77-0.91).

Table 1. Baseline demographic and clinical characteristics of participants.

Characteristics	N=59
Sociodemographic	
Gender[a]	
Male	37 (62.7)
Female	22 (37.3)
Age (years)[b]	48.15 (15.4)
Age group[a]	
20-29 years	4 (6.8)
30-59 years	41 (69.5)
≥ 60 years	14 (23.7)
Lesion characteristics	
Time since injury (years)[c]	12 (6 - 21)
Time since injury group[a]	
1-4 years	11 (18.6)
5-9 years	11 (18.6)
10-14 years	13 (22.0)
≥15 years	24 (40.7)
Severity of SCI[a]	
C1-4 AIS A, B, and C	8 (13.6)
C5-8 AIS A, B, and C	9 (15.2)
T1-S3 AIS A, B, and C	38 (64.4)
AIS D at any injury level	4 (6.8)

Note: SCI: spinal cord injury, ASIA: American Spinal Injury Association, AIS: ASIA Impairment Scale, [a]Number (%), [b]Mean (SD), [c]Median (IQR).

Table 2. Test and re-test scores of SCIM-SR-Thai and its subdomains (self-care, respiratory, mobility).

	Total score (score 0-100)		Self-care (score 0-20)		Respiratory and sphincter management (score 0-40)		Mobility (score 0-40)	
	Test	Re-test	Test	Re-test	Test	Re-test	Test	Re-test
Median (IQR)	57 (48-68)	56 (39-66)	19 (14-20)	18 (12-20)	25 (15-31)	25 (16-30)	16 (7-18)	15 (6-18)

Table 3. Intraclass correlation coefficients and Cronbach's alpha of SCIM-SR-Thai (total score) and its subdomains.

	Intraclass correlation (95% CI)	Cronbach's Alpha
Total score	0.97 (0.95-0.98)	0.99
Self-care	0.96 (0.94-0.98)	0.98
Respiratory	0.86 (0.77-0.91)	0.92
Mobility	0.97 (0.95-0.98)	0.98

Discussion

Chronic SCI is a heavy diagnosis with both physical and psychological burdens, in which optimal rehabilitation programs help to maintain patient's dependence.[2] In doing so, tracking one's progress after discharge is essential and can be implemented using self-reported questionnaires.[4] This work provides a more comprehensive psychometric assessment of the SCIM-SR-Thai that was previously developed. Reproducibility is a crucial parameter of any clinically useful measurement tool, so that real changes may be distinguished from random and systemic variations. Like previous studies, the authors have shown that the SCIM-SR-Thai demonstrates excellent test-retest reliability ensuring temporal consistency (Table 4).[4,8,9-11] Therefore, SCIM-SR-Thai may be a promising instrument that awaits further translation into practice.

Table 4. Results compared between previous studies and the present study.

	Language	Number of participants	Internal consistency (Cronbach's alpha)	Validity (Pearson's correlation coefficient)	Test-retest reliability (Intraclass correlation coefficient)
Fekete et al.[4]	English	99		0.87	0.90
Aguilar-Rodríguez et al.[5]	Spanish	100		0.998	
Bonavita et al.[6]	Italian	116		0.934	
Takeuchi et al.[7]	Japanese	100	0.89	0.95	
Wang et al.[8]	Chinese	147	0.908	0.935	0.876
Wilartratsami et al.[9]	Thai	32	0.91	0.94	0.95
Jörgensen et al.[10]	Swedish	90	0.89		0.98
Khatri et al.[11]	Nepali	45	0.801	0.968	0.968
Tongprasert et al. [12]	Thai	61	0.98	0.97	
Present study	Thai	59	0.985		0.97

SCIM-SR has become a convenient alternative to SCIM III, and multiple translations have allowed adoption of the questionnaire to non-English speaking countries.[5-12] The challenge is ensuring these locally modified instruments are well validated and reliable despite subtle cultural differences, to allow across-the-board comparison internationally. Like Thai SCIM-SR by Wilartratsami *et al.*, the authors have shown that respiration and sphincter management subscales tend to have slightly lower reliability than self-care and mobility (ICC of 0.859, 0.960 and 0.961, respectively). [9] However, ICC coefficients of the respiration and sphincter management subscales still met the quality criterion (>0.7).[13] A likely cause for the slightly lower ICC in the sphincter management domain is greater day-to-day variability or inconsistency in self-management practices in the community setting. Bowel and bladder function are often influenced by factors such as infection, medication use, fluid intake or food intake which may lead to different self-reported scores on the test and re-test days. This fluctuation is less common for domains like mobility or self-care, which are generally stable in chronic SCI**.** Previous Thai studies have included participants with differing demographic and clinical characteristics. For example, Wilartratsami *et al*.[9] studied predominantly acute or subacute cases, whereas Tongprasert *et al*.[12] included a higher proportion of participants with AIS D lesions. These differences in injury phase and severity may influence self-reported functional ability and psychometric outcomes. In our study, we recruited chronic SCI participants with stable functional performance, which minimizes variability due to recovery and reinforces the temporal reliability of the instrument across time. Despite demographic and severity differences across studies, the consistency of internal consistency and ICC results supports the robustness of SCIM-SR-Thai across various SCI populations.

The strength of this study includes an appropriate time lapse between test and retest. In concordance to other SCIM-SR, the duration of 7 days was chosen to limit the effects of recall memory, while ensuring participant's function would not change.[13]

Limitations

In terms of limitation, participants were instructed to complete the retest at their leisure, thus testing conditions were not controlled at home. In addition, convenience sampling from a single rehabilitation center, which may restrict the generalizability of the findings to broader SCI populations or different care settings. Multicenter studies with more diverse participants are recommended to strengthen external validity. Further works may attempt to correlate SCIM-SR-Thai with quality-of-life questionnaires, such as the World Health Organization quality of life (WHOQOL) index, to further elucidate its real-world utility. Quality of life assessment provides insight into patients lived experience and participation beyond functional ability, reflecting the real-world utility of rehabilitation outcomes.[17] The WHOQOL-BREF was proposed due to its validated Thai version and comparability across populations, but future research may incorporate SCI-specific instruments such as the Quality-of-Life Index-SCI Version (QLI-SCI), SCIQL-23 or SCI-QOL Battery for a more nuanced evaluation of condition-specific well-being.

Conclusion

SCIM-SR-Thai demonstrates good reliability and internal consistency. The authors encourage its adoption as a standardized tool in clinical practice to assess functional dependence in Thai people with SCI, especially in the community settings.

Ethical approval

This study was approved by the Research Ethics Committee of the study site (EC294/2566).

Funding

This work was supported by the Faculty of Medicine, Chiang Mai University, grant number 012/2567.

Conflict of interest

The authors declare that there is no conflict of interest.

CRediT authorship contribution statement

KM was responsible for designing the research question, collecting and analyzing the data, drafting the manuscript, and writing the final version of the manuscript. ST was responsible for designing the research question, analyzing the data, and commenting on the final version of the manuscript.

Acknowledgements

With great appreciation, the authors would like to thank all participants for their valuable cooperation in this study. Additionally, gratitude is given to Faculty of Medicine, Chiang Mai University for the support and facilitation provided during the data collection process.

References

[1] Bickenbach J. International perspectives on spinal cord injury. [Internet]. World Health Organization; 2013 [cited 2024 Nove 23]. Avaliable from: https://www.who.int/publications/i/item/international-perspectives-on-spinal-cord-injury.

[2] Nas K, Yazmalar L, Şah V, Aydın A, Öneş K. Rehabilitation of spinal cord injuries. World J Orthop. 2015 ; 6(1): 8-16. doi: 10.5312/wjo.v6.i1.8.

[3] tzkovich M, Gelernter I, Biering-Sorensen F, Weeks C, Laramee MT, Craven BC, et al. The Spinal Cord Independence Measure (SCIM) version III: reliability and validity in a multi-center international study. Disabil Rehabil. 2007; 29(24): 1926-33. doi: 10.1080/09638280601046302.

[4] Fekete C, Eriks-Hoogland I, Baumberger M, Catz A, Itzkovich M, Lüthi H, et al. Development and validation of a self-report version of the Spinal Cord Independence Measure (SCIM III). Spinal Cord. 2013; 51(1): 40-7. doi: 10.1038/sc.2012.87.

[5] Aguilar-Rodríguez M, Peña-Pachés L, Grao-Castellote C, Torralba-Collados F, Hervás-Marín D, Giner-Pascual M. Adaptation and validation of the Spanish self-report version of the Spinal Cord Independence Measure (SCIM III). Spinal Cord. 2015; 53(6): 451-4. doi: 10.1038/sc.2014.225.

[6] Bonavita J, Torre M, China S, Bressi F, Bonatti E, Capirossi R, et al. Validation of the Italian version of the Spinal Cord Independence Measure (SCIM III) Self-Report. Spinal Cord. 2016; 54(7): 553-60. doi: 10.1038/sc.2015.187.

[7] Takeuchi S, Uemura O, Unai K, Liu M. Adaptation and validation of the Japanese version of the Spinal Cord Independence Measure (SCIM III) self-report. Spinal Cord. 2021; 59(10): 1096-103. doi: 10.1038/s41393-021-00633-5.

[8] Wang T, Tang J, Xie S, He X, Wang Y, Liu T, et al. Translation and validation of the Chinese version of the Spinal Cord Independence Measure (SCIM III) Self-Report. Spinal cord. 2021; 59(10): 1045-52. doi: 10.1038/s41393-020-00601-5.

[9] Wilartratsami S, Luksanapruksa P, Santipas B, Thanasomboonpan N, Kulprasutdilok P, Chavasiri S, *et al.* Cross-cultural adaptation and psychometric testing of the Thai version of the Spinal Cord Independence Measure III-Self Report. Spinal Cord. 2021; 59(3): 291-7. doi: 10.1038/s41393-020-00556-7.

[10] Jörgensen S , Forslund EB, Lundström U,Nilsson E, Levi R, Berndtsson E, *et al.* Sound psychometric properties of the Swedish version of the Spinal Cord Independence Measure Self-Report. J Rehabil Med. 2021; 53(5): jrm00197. doi: 10.2340/16501977-2839.

[11] Khatri P, Prasertsukdee S, Suttiwong J. Reliability of the Nepali Version of the Spinal Cord Independence Measure Self-Report. Rehabil Res Pract. 2022; 2022: 9983464. doi: 10.1155/2022/9983464.

[12] Tongprasert S, Wongpakaran T, Soonthornthum C. Validation of the Thai version of the Spinal Cord Independence Measure Self-Report (SCIM-SR-Thai). Spinal Cord. 2022; 60(4): 361-7. doi: 10.1038/s41393-022-00779-w.

[13] Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007; 60(1): 34-42. doi: 10.1016/j.jclinepi.2006.03.012.

[14] Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. Stat Med. 2002; 21(9): 1331-5. doi: 10.1002/sim.1108.

[15] Biering-Sørensen F, Charlifue S, Chen Y, New PW, Noonan V, Post MWM, et al. International Spinal Cord Injury Core Data Set (version 3.0)-including standardization of reporting. Spinal Cord. 2023; 61(1): 65-8. doi: 10.1038/s41393-022-00862-2.

[16] IBM Corp. IBM® SPSS® Statistics 28 [Computer program]. New York: IBM Corp; 2021.

[17] Post MW, van Leeuwen CM. Psychosocial issues in spinal cord injury: a review. Spinal Cord. 2012; 50(5): 382-9. doi:10.1038/sc.2011.182.

Journal of Associated Medical Sciences 2026; 59(1): 168-175

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 comparative study of diagnostic performance for tuberculosis and rifampicin-resistant tuberculosis between standard and pooled sputum methods using GeneXpert Ultra at Lampang Hospital, Thailand

Sittichok Chaiwong[1], Surasit Suwannasin [2], Jutharat Hmoteh[2*]

[1]Clinical Microbiology Unit Division of Medical Technology Lampang Hospital, Lampang province, Thailand.
[2]Department of Medical Technology, Faculty of Allied Health Science, Phatthalung Campus, Thaksin University, Phatthalung, Thailand.

ARTICLE INFO

Article history:
Received 15 August 2025
Accepted as revised 17 October 2025
Available online 4 November 2025

Keywords:
Tuberculosis, sputum pooling, GeneXpert Ultra, diagnostic performance, cost effectiveness.

ABSTRACT

Background: Molecular diagnosis using the Xpert MTB/RIF Ultra assay enables rapid detection of *Mycobacterium tuberculosis* and rifampicin resistance.

Objectives: This study evaluated the diagnostic performance and cost efficiency of a two-sample pooled testing approach compared with standard individual testing.

Materials and methods: A cross-sectional diagnostic study was conducted among 3,504 presumptive TB patients. Sputum specimens were tested by Xpert Ultra individually, with those results serving as the reference standard, and in two-sample pools. The pooling protocol utilized 2.0 mL from each specimen to maximize load volume. Deconvolution, which required retesting both individual specimens, was mandatory for all positive and invalid pooled results. Performance metrics, RIF resistance concordance, and cost-effectiveness modeled on cartridge consumption and direct cost 550 Thai Baht per unit were compared.

Results: The pooled method showed excellent concordance with the individual method, yielding 240 concordant positives and 3,249 concordant negatives, with no statistically significant difference in MTB detection (McNemar χ^2=0.267, p=0.605). The pooled approach achieved a sensitivity of 96.39% (95% CI, 93.25-98.33) and a specificity of 99.82% (95% CI, 99.60-99.93). The assay maintained reliable detection of RIF resistance, indicating that pooling did not compromise molecular accuracy. In practice, 1,752 pooled runs were performed, with 105 (5.7%) error results and 246 positive pools requiring deconvolution. Including repeats and deconvolution, total cartridge use was 2,349 compared with 3,504 for individual testing, corresponding to an actual cost reduction of 32.96% (approx. 180,000 Thai Baht saved per 1,000 tests).

Conclusion: Two-sample pooled Xpert Ultra testing demonstrated high diagnostic accuracy and doubled analytical throughput. Although deconvolution limited cost savings to approximately 33%, the strategy proved highly cost-effective and operationally feasible. This method offers a practical, scalable approach for optimizing molecular TB diagnostics and resource utilization, especially in low- to moderate-prevalence settings.

* **Corresponding contributor.**
Author's Address: *Department of Medical Technology, Faculty of Allied Health Science, Phatthalung Campus, Thaksin University, Phatthalung, Thailand.*
E-mail address: jutharat.h@tsu.ac.th
doi: 10.12982/JAMS.2026.021
E-ISSN: 2539-6056

Introduction

Tuberculosis (TB) remains one of the leading infectious diseases globally, with over 10 million people affected each year and approximately 1.3

million deaths reported annually among HIV-negative individuals.[1] Despite extensive global public health efforts, TB control remains challenging, particularly in low- and middle-income countries (LMICs) where access to rapid and accurate diagnostic tools is limited.[2] Thailand, classified as a high TB burden country, continues to face substantial challenges in TB detection, especially among vulnerable populations such as people living with HIV and those with extrapulmonary TB. Conventional TB diagnostics including sputum smear microscopy and culture are routinely used in many health facilities. Nevertheless, the methods have notable limitations. Smear microscopy lacks sensitivity, particularly in paucibacillary cases and among HIV-positive individuals, while culture methods, although more sensitive, are time-consuming and resource-intensive.[3] These constraints hinder early diagnosis and timely initiation of treatment, which are critical for TB control and prevention of transmission.

The introduction of molecular diagnostics has revolutionized TB detection. The Xpert MTB/RIF Ultra assay (Cepheid Inc., USA) is a cartridge-based nucleic acid amplification test endorsed by the World Health Organization (WHO) for the rapid detection of *Mycobacterium tuberculosis* and rifampicin resistance[1]. The assay detects *M. tuberculosis* complex DNA and simultaneously analyzes mutations within the *rpoB* gene, which encodes the β-subunit of RNA polymerase, a key target of rifampicin. Mutations in this gene, particularly within the rifampicin resistance determining region, are responsible for over 95% of rifampicin-resistant TB cases.[13]

This assay delivers results within approximately two hours and demonstrates superior sensitivity and specificity compared to conventional methods, particularly in smear-negative and HIV-positive patients.[4,5] Nevertheless, the high per-test cost, including cartridges and instrumentation, restricts its widespread adoption in resource-limited settings such as provincial hospitals in Thailand.

To overcome economic constraints, sputum pooling has been proposed as an innovative approach to enhance diagnostic capacity while reducing costs. Pooling involves combining sputum samples from multiple individuals into a single cartridge, thus decreasing reagent consumption and overall expenses. If the pooled sample tests negative, all individuals in the pool are considered TB-negative; if positive, each sample undergoes individual testing to identify the infected cases.[6]

Studies from high-burden countries such as Brazil and Lao PDR have demonstrated that pooling sputum samples maintains high diagnostic accuracy while achieving substantial cost-savings in cartridge use. For example, in Lao PDR, pooling resulted in 30% to 40% percent savings in Xpert cartridge costs without compromising test performance.[7] Despite promising evidence from international studies, data on sputum pooling for TB diagnosis in Southeast Asia, particularly Thailand, remain scarce. Most existing research focuses on laboratory validation or modeling rather than real-world clinical implementation in provincial hospitals. Given Thailand's high TB burden and resource constraints, empirical evaluation of pooled sputum testing using Xpert MTB/RIF Ultra under local healthcare conditions is urgently needed.

Therefore, this study aims to evaluate the diagnostic performance and cost-effectiveness of pooled sputum testing at Lampang Hospital, a large provincial hospital in northern Thailand. By comparing pooled testing to the standard individual sample approach, this study seeks to provide evidence to support policy and clinical practice changes aimed at improving TB detection coverage and optimizing resource utilization in high-burden, resource-limited settings.

Materials and methods

Study design and setting

A cross-sectional diagnostic study was conducted to compare pooled and individual sputum sample testing using the Xpert MTB/RIF Ultra assay, with the objective of evaluating diagnostic performance and potential resource savings in a high-burden setting.[1,5,7,9] The study took place at the Clinical Microbiology Laboratory of Lampang Hospital, a provincial referral center in northern Thailand. Between October 2023 and September 2024, a total of 3,504 patients with presumptive tuberculosis were enrolled, each contributing a single sputum specimen.[1,5,7] For the pooling strategy, sputum specimens from two different patients were combined to create a single pooled sample. This method was applied to assess the feasibility and efficiency of pooled testing in optimizing diagnostic resources and reducing overall costs. For the pooled testing approach, sputum specimens from two patients were combined into a single test cartridge and processed following the manufacturer's protocol.[4,5] The results obtained from the pooled samples were directly compared with the results from the corresponding individually tested samples, aiming to determine the concordance, sensitivity, specificity, and potential cost savings associated with pooled testing in a real-world diagnostic setting. The study aimed to compare the diagnostic performance and cost-effectiveness of conventional GeneXpert testing versus pooled sputum testing using the Xpert MTB/RIF Ultra assay.

Participants and sample collection

Sputum samples were collected from presumptive tuberculosis patients whose chest X-rays (CXR) were abnormal and suggestive of pulmonary TB.[1,14] Routine diagnostic testing was performed as part of standard care, and secondary data were retrieved from the laboratory database. All samples were anonymized prior to analysis. Sputum samples were collected

as part of routine diagnostic care. Secondary data, including Xpert MTB/RIF Ultra assay results and patient demographic information, were extracted from the hospital's laboratory information system. All records were anonymized prior to analysis to ensure compliance with ethical standards and patient confidentiality.

Procedures

Individual sample testing

Each sputum specimen was processed by adding 1.5 mL of sample reagent to 0.5 mL of sputum in a tube. The mixture was vortexed and left at room temperature for 10 minutes, then mixed again and allowed to sit for an additional 5 minutes. A 2.0 mL aliquot of the resulting supernatant was transferred into a dedicated GeneXpert cartridge for each patient. Cartridges were carefully sealed and loaded into the GeneXpert GX-IV 10C system (Cepheid, USA) for analysis. The assay run time was approximately 90 minutes.[4,5] Results included detection of *Mycobacterium tuberculosis* and *rifampicin resistance*. Internal quality control measures included the use of a Sample Processing Control (SPC) to verify the adequacy of DNA extraction and PCR amplification, along with routine checks for cartridge integrity and probe performance. Additionally, the laboratory participated in an annual external quality assessment (EQA) program administered by the Department of Disease Control, Ministry of Public Health, Thailand, to ensure continued accuracy and reliability of test results.[1,14] Invalid or error results (<6% of total runs) were repeated according to standard operating procedures until a valid result was obtained.

Pooled sample testing procedure

To evaluate diagnostic efficiency and cost-saving potential, paired sputum specimens from two individual patients were processed as described above. For pooled testing, sputum specimens from two different patients were combined to form a pooled sample. Specifically, 2.0 mL of processed sputum from each specimen (total 4.0 mL) were mixed together in a single tube. Subsequently, 8.0 mL of sample reagent were added to the pooled specimen according to the manufacturer's instructions, resulting in a 1:2 sample-to-reagent ratio. The mixture was gently inverted to ensure proper homogenization and then incubated as described for individual testing. From the resulting solution, 4.0 mL were transferred into a single GeneXpert cartridge for analysis. This final volume is within the cartridge's loading capacity, and the procedure followed the standard GeneXpert MTB/RIF Ultra protocol.[7,9,10,12,13] Results from pooled testing were directly compared with the results of the corresponding individual tests to assess diagnostic concordance, sensitivity, specificity, and potential cost reduction. The pre-analytical procedure, which included sample–reagent mixing and vertexing, required approximately 15 minutes to complete for both the standard and pooled sample methods. In the pooled sample approach, an additional 5 minutes were required for the pooling process prior to reagent addition. The analytical run time on the Xpert MTB/RIF Ultra assay was the same for both methods: approximately 90 minutes for MTB not detected results and up to 110 minutes for MTB detected results.

Deconvolution and retesting strategy

Results from pooled testing were directly compared with the results of the corresponding individual tests to assess diagnostic concordance, sensitivity, specificity, and potential cost reduction. For any pooled sample yielding an 'MTB Detected' result or an 'Invalid/Error' result, the two individual specimens comprising that pool were retrieved. Deconvolution testing was performed by retesting both individual specimens separately using new Xpert Ultra cartridges. All cartridges used for initial pooling, deconvolution, and error retesting were included in the final cost analysis model.

Reference standard and comparative performance assessment

The diagnostic outcome derived from the standard individual GeneXpert MTB/RIF Ultra assay was designated as the primary reference standard for assessing the performance of the pooled-testing strategy. This approach was specifically necessitated to facilitate a direct comparison of the technical efficiency and cost implications between the two distinct Xpert Ultra application methods within the operational parameters of the existing molecular diagnostic workflow. Mycobacterial culture, the established gold standard, was intentionally omitted from routine testing due to constraints in resources and time, thereby confining the study's scope to an operational evaluation of sputum pooling as a procedural modification to the current Xpert platform.

Data analysis

The analysis employed SPSS for statistical comparison using McNemar's test to assess paired differences in detection rates between pooled and individual testing methods. Microsoft Excel was used to calculate key diagnostic indices, including sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio. Exact Clopper–Pearson 95% confidence intervals were calculated for sensitivity, specificity, positive and negative predictive values, and accuracy, while confidence intervals for likelihood ratios were derived using the standard log-transformation method. Trace results reported by the Xpert MTB/RIF Ultra assay were interpreted as MTB detected (positive) in accordance with the manufacturer's guidelines and WHO recommendations, and were included in the overall diagnostic performance analysis. Cost analysis

was performed in Excel using a cost-minimization approach to compare resource utilization between testing strategies.

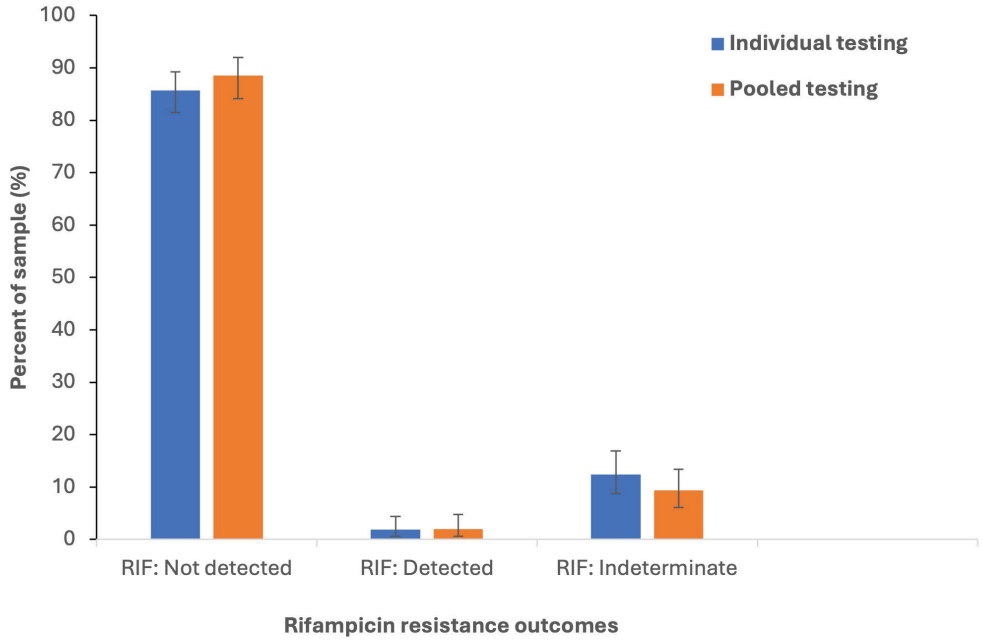


Figure 1. *Comparison of rifampicin-resistance status proportions obtained by the standard individual and pooled testing methods using the Xpert MTB/RIF Ultra assay. Rifampicin resistance was identified through rpoB gene mutation analysis. No statistically significant difference was observed between the two methods (Chi-square test, $p<0.05$), indicating equivalent diagnostic performance.*

Results

A total of 3,504 sputum specimens were simultaneously tested using both the individual method and a pooled approach, utilizing the Xpert MTB/RIF Ultra assay for molecular detection of *Mycobacterium tuberculosis* and rifampicin resistance. As shown in Table 1, there was a high degree of agreement between the two methods, with 240 concordant positive and 3,249 concordant negative results. Discordant results were minimal, with nine samples that were positive by the individual method but negative in the pooled testing. Comparison of the two methods revealed no statistically significant difference in the detection of *Mycobacterium tuberculosis*. The continuity-corrected McNemar chi-square was 0.267, and the associated *p*-value was 0.605. Conversely, there were six samples that were negative by the individual testing but positive in the pooled testing. The high concordance rate between the two methods, particularly for negative results, demonstrates the effectiveness of the pooling strategy for large-scale screening.

Table 1. Comparison of diagnostic results between pooled and individual Xpert MTB/RIF Ultra testing.

	Individual testing		
	Positive	**Negative**	**Total**
Pooled testing			
Positive	240	6	246
Negative	9	3,249	3,258
Total	249	3,255	3,504

The diagnostic performance of the pooled testing is summarized in Table 2. The assay demonstrated a sensitivity of 96.39%, with a 95% confidence interval ranging from 93.25 to 98.33, and a specificity of 99.82%, with a 95% confidence interval between 99.60 and 99.93. The overall accuracy reached 99.57%, with a 95% confidence interval of 99.29 to 99.76. The positive predictive value was 97.56%, and its 95% confidence interval extended from 94.77 to 99.10, while the negative predictive value was 99.72%, with a 95% confidence interval of 99.48 to 99.87. The high positive likelihood ratio of 523.0 demonstrates the strong discriminative ability of the pooled method in confirming true-positive cases, whereas the low negative likelihood ratio of 0.036 highlights its reliability in ruling out infection in negative samples.

Table 2. Diagnostic performance of pooled testing compared with individual testing for detection of *Mycobacterium tuberculosis* and rifampicin resistance Using Xpert MTB/RIF Ultra.

Diagnostic parameters	Values	(95% Confident Interval)
Sensitivity (%)	96.39	93.25-98.33
Specificity (%)	99.82	99.60-99.93
Accuracy (%)	99.57	99.29-99.76
Positive Predictive Value (PPV, %)	97.56	94.77-99.10
Negative Predictive Value (NPV, %)	99.72	99.48-99.87
Likelihood Ratio Positive (LR^+)	523.0	
Likelihood Ratio Negative (LR^-)	0.036	

In the Xpert MTB/RIF Ultra assay, detection of *Mycobacterium tuberculosis* complex DNA is accompanied by simultaneous analysis of mutations within the *rpoB* gene that confer rifampicin resistance. Figure 1. illustrates the distribution of rifampicin resistance outcomes obtained from individual and pooled Xpert MTB/RIF Ultra testing. Among MTB-positive samples, most were classified as RIF: Not detected, accounting for approximately 85-87% of all positive results. A small proportion (<2%) were RIF: Detected, indicating rifampicin resistance, while 8-10% were RIF: Indeterminate. The distributions of RIF resistance outcomes were comparable between the two testing approaches, and the Chi-square analysis confirmed that the difference was not statistically significant ($p>0.05$).

During pooled testing, cartridge errors occurred in 105 runs, corresponding to an overall error rate of approximately 5.7%. These errors primarily resulted from invalid or error signals reported by the Xpert MTB/RIF Ultra instrument. All affected pooled samples were retested individually for deconvolution, and the final results were based on the successful retests. The low error rate observed indicates that pooling two samples per cartridge did not substantially increase the instrument error rate compared with standard testing.

A comparison of testing capacity and cost between the standard and pooled methods is summarized in Table 3. The pooled strategy effectively doubled the testing capacity, allowing eight samples to be analyzed per run compared with four samples in the standard method, while the turnaround time per batch remained approximately two hours for both approaches. Under theoretical conditions where no repeats or follow-up testing were required, the cost per sample decreased by half from 550 THB to 275 THB while the total expenditure per run remained constant at 2,200 THB. In practice, cartridge consumption exceeded the theoretical minimum because of repeat runs and deconvolution of positive pools. Among 1,752 pooled runs, 105 (5.7%) produced invalid or error results, and 246 pools tested positive. These positive pools required 492 additional cartridges for deconvolution of both individual samples, together with 105 cartridges used for repeat or invalid runs, bringing the total cartridge use to 2,349. The corresponding total expenditure was 1,291,950 THB, representing an actual cost reduction of 32.96% compared with the standard individual method, which required 3,504 cartridges and 1,927,200 THB in total costs. Despite the additional cartridges required for deconvolution and repeat runs, the pooled testing approach maintained substantial savings while doubling analytical throughput.

Table 3. Comparison of testing capacity and cost between pooled testing compared with individual testing.

Parameters	Individual testing	Pooled testing	Fold change
Sample preparation time (min)	approx.15	approx. 20	~1.33
Samples per run	4	8	2.0
Analytical run time (min)	90-110	90-110	1.0
Total turnaround time per batch (hrs)	2	2	1.0
Cartridges used (no repeats)	3504	1752	0.5
Additional cartridges for deconvolution	-	492	-
Additional cartridges for errors	-	105	-
Cost per sample (THB)	550	275	0.5
Cost per run (THB)	2,200	2,200	1.0
Total cost (THB)	1,927,200	1,291,950	0.67
Actual cost reduction		32.96%	

Discussion

The high concordance observed between pooled and individual testing demonstrates that pooling two samples per cartridge can maintain diagnostic accuracy while substantially increasing throughput and conserving resources. These advantages are essential for large-scale screening programs and emergency responses. The pooled approach effectively identified true-negative samples, supporting efficient allocation of laboratory resources. However, pooled testing has an inherent limitation related to the dilution of samples with a low bacillary load. When a positive specimen is combined with negative specimens, the bacterial concentration may fall below the assay's limit of detection, leading to false-negative results. Similar findings have been reported in previous pooled-testing evaluations, which also showed only minimal loss of diagnostic performance in high-burden tuberculosis settings.[7,8] In comparison with previous studies, the results of this study were consistent with the general purpose of pooled testing, which aimed to maximize efficiency by accurately excluding the majority of negative samples while allowing a small number of discordant results that require follow-up retesting.[9,10] The observed sensitivity of 96.39% demonstrates a high probability of correctly identifying true positive samples. This level of performance was comparable to earlier studies in tuberculosis diagnostics that reported minimal loss in sensitivity relative to individual testing.[13] The combination of high positive and negative predictive values supported the application of pooled testing as a practical first-line screening method in areas of low to moderate tuberculosis prevalence. This strategy enabled the efficient use of molecular testing resources while maintaining reliable case detection, in agreement with the operational benefits documented in large-scale diagnostic programs.[14] The similar distribution of rifampicin resistance outcomes between the two testing methods indicates that pooling did not affect the assay's capacity to detect mutations in the *rpoB* gene. Because rifampicin resistance in *Mycobacterium tuberculosis* primarily arose from *rpoB* gene mutations, these comparable results confirmed that the pooled approach maintains molecular accuracy for both tuberculosis diagnosis and resistance surveillance, even when the bacterial load was shared between combined specimens.

From an operational and economic perspective, pooled testing provided substantial efficiency gains. The observed cost reduction of 32.96% represented an estimated saving of about 180,000 THB per 1,000 tests under actual laboratory conditions. These savings could be redirected toward expanding testing coverage, procuring reagents, or supporting other public health interventions.[15] Although the inclusion of deconvolution testing reduced the theoretical cost reduction from 50% to approximately 33%, the pooled testing approach remains economically advantageous and operationally feasible for tuberculosis screening in low- to moderate-prevalence settings. The scale of cost savings depended largely on underlying disease prevalence, as higher positivity rates required more deconvolution testing and reduced net savings. Even with these adjustments, pooled testing remained a practical and efficient diagnostic strategy that could expand testing capacity and reduced expenditure without additional capital investment.

Limitations

None

Conclusion

This study demonstrates that the two-sample pooled Xpert MTB/RIF Ultra assay provides reliable diagnostic performance while doubling testing capacity and reducing costs by 32.96% without extending

turnaround time. The approach performs closely to individual testing, offering a practical and scalable solution for tuberculosis screening in high-burden or resource-limited settings. Furthermore, the pooled Xpert MTB/RIF Ultra method maintained equivalent diagnostic performance for both *M. tuberculosis* and rifampicin resistance, underscoring its reliability for integrated TB diagnosis and resistance monitoring. By transforming routine diagnostic practice into research-based evidence, these findings support the adoption of pooled testing to strengthen screening programs and optimize the use of available resources.

Ethical approval

This study protocol was approved by the Ethics Committee of Lampang Hospital, Thailand (EC: 157/68).

Funding

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

Conflict of interest

The authors declare that there are no conflicts of interest related to this study.

CRediT authorship contribution statement

Chaiwong S: methodology, data curation, writing: review and edit, **Suwannasin S**: writing: review and edit, **Hmoteh J**: conceptualization, formal analysis, writing: original draft.

Acknowledgements

The authors thank Lampang Hospital for their support in providing laboratory resources and technical assistance, which were essential for the completion of this study.

References

[1] WHO consolidated guidelines on tuberculosis. Module 3: diagnosis–rapid diagnostics for tuberculosis detection: World Health Organization; 2024. Available from: https://www.who.int/publications/i/item/9789240089488.

[2] Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, et al. Tuberculosis. *Nat Rev Dis Primers*. 2016; 2(1): 16076. doi: 10.1038/nrdp.2016.76.

[3] Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database of Sys Rev. 2014; 2014(1): CD009593. doi: 10.1002/14651858.CD009593.pub3.

[4] Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med*. 2010; 363(11): 1005-15. doi: 10.1056/NEJMoa0907847.

[5] Dorman SE, Schumacher SG, Alland D, Nabeta P, Armstrong DT, King B, et al. Xpert MTB/RIF Ultra for detection of Mycobacterium tuberculosis and rifampicin resistance: a prospective multicentre diagnostic accuracy study. *Lancet Infect Dis*. 2018; 18(1): 76-84. doi: 10.1016/S1473-3099(17)30691-6.

[6] Abdalhamid B, Bilder CR, Garrett JL, Iwen PC. Cost effectiveness of sample pooling to test for SARS-CoV-2. Journal of infection in developing countries. 2020; 14(10): 1136. doi: doi.org/10.3855/jidc.13935.

[7] Iem V, Chittamany P, Suthepmany S, Siphanthong S, Somphavong S, Kontogianni K, et al. Pooling sputum for Xpert MTB/RIF and Xpert Ultra testing during the Covid-19 pandemic in Lao People's Democratic Republic. PLOS global public health. 2022; 2(4): e0000116. doi: 10.1371/journal.pgph.0000116.

[8] Zifodya JS, Kreniske JS, Schiller I, Kohli M, Dendukuri N, Schumacher SG, et al. Xpert Ultra versus Xpert MTB/RIF for pulmonary tuberculosis and rifampicin resistance in adults with presumptive pulmonary tuberculosis. Cochrane Database Syst Rev. 2021; 2(2): Cd009593. doi: 10.1002/14651858.CD009593.pub5.

[9] Vuchas C, Teyim P, Dang BF, Neh A, Keugni L, Che M, et al. Implementation of large-scale pooled testing to increase rapid molecular diagnostic test coverage for tuberculosis: a retrospective evaluation. Sci Rep. 2023; 13(1): 15358. doi: 10.1038/s41598-023-41904-w.

[10] Iem V, Bimba JS, Santos VS, Dominguez J, Creswell J, Somphavong S, et al. Pooling sputum testing to diagnose tuberculosis using xpert MTB/RIF and xpert ultra: a cost-effectiveness analysis. BMC Infect Dis. 2023; 23(1): 341. doi: 10.1186/s12879-023-08330-9.

[11] Iem V, Xangsayarath P, Chittamany P, Suthepmany S, Siphanthong S, Paboriboune P, et al. Pooling samples to increase testing capacity with Xpert Xpress SARS-CoV-2 during the Covid-19 pandemic in Lao People's Democratic Republic. PLoS One. 2022; 17(9): e0275294. doi: 10.1371/journal.pone.0275294.

[12] Zeng J, Huang H, Liu X, Huang Z, Liu W, Liu H, et al. Pooling sputum samples for the Xpert MTB/RIF assay: a practical screening strategy for highly infectious tuberculosis cases. BMC Infectious Diseases. 2024; 24(1): 122. doi: 10.1186/s12879-024-09020-w.

[13] Bimba J, Adekeye O, Iem V, Eliya T, Osagie I, Kontogianni K, et al. Pooling sputum samples for Xpert® MTB/RIF and Xpert® Ultra testing for TB diagnosis. Public Health Action. 2023; 13(1): 12-6. doi: 10.5588/pha.22.0052.

[14] World Health Organization consolidated guidelines on tuberculosis. Module 3: Diagnosis: World Health

Organization; 2025. Available from: https://www.who.int/publications/i/item/9789240107984.

[15] Bogere N, Bongomin F, Katende A, Ssebambulidde K, Ssengooba W, Ssenfuka H, et al. Performance and cost-effectiveness of a pooled testing strategy for SARS-CoV-2 using real-time polymerase chain reaction in Uganda. International Journal of Infectious Diseases. 2021; 113: 355-8. doi: 10.1016/j.ijid.2021.10.038.

[16] Yelin I, Aharony N, Tamar ES, Argoetti A, Messer E, Berenbaum D, et al. Evaluation of COVID-19 RT-qPCR Test in Multi sample Pools. Clinical Infectious Diseases. 2020; 71(16): 2073-8. doi: 10.1093/cid/ciaa531.

Journal of Associated Medical Sciences 2026; 59(1): 176-183

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Factors affecting hemoglobin levels of patients with end-stage kidney disease in the maintenance phase: A prospective observational study

Maryadi Maryadi[1*], Khudazi Aulawi[2], Uki Noviana[2]

[1]Department of Nursing, Faculty of Sport and Health, Universitas Negeri Gorontalo, Indonesia.
[2]Department of Nursing, Faculty of Medicine Public Health and Nursing, Universitas Gadjah Mada, Indonesia.

ARTICLE INFO

Article history:
Received 5 August 2025
Accepted as revised 23 November 2025
Available online 26 November 2025

Keywords:
End-stage kidney disease, hemodialysis, anemia, erythropoietin stimulating agent, vitamin B12 intake, duration of hemodialysis.

ABSTRACT

Background: Patients with end-stage kidney disease (ESKD) undergoing hemodialysis almost invariably develop anemia, which can significantly impair their quality of life. It is uncommon to find patients with consistently normal hemoglobin (Hb) levels; therefore, identifying the contributing factors is essential.

Objectives: This study aimed to identify factors affecting Hb levels in ESKD patients during maintenance phase.

Materials and methods: This prospective observational study was conducted at a regional hospital in Yogyakarta. The Hb cutoff point of 10 g/dL was used as the minimum standard for the maintenance phase, based on the Dialysis Consensus of the Indonesian Nephrology Association. The study included ESKD patients aged ≥18 years undergoing hemodialysis. The initial sample comprised 76 patients (38 with Hb >10 gm/dL and 38 with Hb ≤10 gm/dL); however, 65 patients were included in the final analysis after 11 were excluded due to dropout. Samples were matched based on age, sex, duration of hemodialysis, and comorbidities. Data were collected using patient medical record observation sheets, food record sheets, and the brief illness perception questionnaire (B-IPQ). Statistical analyses were performed using chi-square and logistic regression with a backward elimination method.

Results: A total of 65 participants were analyzed (33 patients with Hb >10 and 32 patients with Hb ≤10). Factors that showed significant associations with Hb levels were erythropoietin-stimulating agent (ESA) therapy odds ratio (OR)=7.162, 95% confidence interval [CI]=2.046-25.071, p=0.002), duration of hemodialysis (OR=4.909, 95% CI=1.299-18.551, p=0.019), and vitamin B12 intake (OR=4.238, 95% CI=1.183-15.176, p=0.027).

Conclusion: Erythropoietin stimulating agent therapy, duration of hemodialysis, and vitamin B12 intake were identified as factors influencing the improvement of Hb levels in patients with ESKD. Educational interventions are needed to enhance patients' understanding of the importance of maintaining balanced nutritional intake and adhering to optimal hemodialysis duration.

Introduction

End-stage kidney disease is characterized by a progressive decline in kidney function, which can have detrimental effects and potentially lead to death. End-stage kidney disease has become a major public health concern worldwide. The prevalence of ESKD requiring kidney transplantation continues to increase every year.[1] In Indonesia, the prevalence of active ESKD undergoing routine hemodialysis has increased by 58.9% over the past year.[2]

* **Corresponding contributor.**
Author's Address: *Department of Nursing, Faculty of Sport and Health, Universitas Negeri Gorontalo, Indonesia.*
E-mail address: maryadi@ung.ac.id
doi: 10.12982/JAMS.2026.022
E-ISSN: 2539-6056

Various physical complications arise due to the disease and hemodialysis procedures, one of which is fatigue and weakness caused by low Hb levels. Hemodialysis patients with normal Hb levels are rarely found, as anemia is a common complication among ESKD patients undergoing dialysis, resulting from either the underlying pathology or the dialysis process itself.[3] This condition is one of the major clinical issues in patients undergoing hemodialysis[1] and can have detrimental consequences, such as cognitive and psychological disorders, decreased cardiac function, angina, cardio-renal anemia syndrome, left ventricular hypertrophy, increased healthcare costs, reduced quality of life, worsening of clinical status, accelerated progression of heart disease, and increased risk of mortality.[4-6] It is also one of the major factors contributing to repeated hospitalizations among ESKD patients.[7]

Several studies conducted in different countries have examined the factors associated with anemia in ESKD patients, as reported by Shiferaw *et al*. in their systematic review. However, limited research has focused on the factors that play a critical role in maintaining and improving Hb levels among ESKD patients undergoing hemodialysis.[8] In Indonesia, 22% of ESKD patients receiving dialysis still have normal Hb levels.[2]

Several factors are thought to have a role in maintaining Hb stability in ESKD patients such as energy intake[9] which can increase the effectiveness of erythropoietin;[10] essential amino acids from protein intake, which support the formation of red blood cells;[11] iron, which can increase Hb levels and optimize erythropoietin therapy;[12] and folic acid and vitamin B12, which are essential in the production of healthy red blood cells.[13,14]

In addition, the duration of hemodialysis also affects Hb, as a longer duration of hemodialysis is associated with higher hemodialysis adequacy,[15] which may enhance erythrocyte glucose-6-phosphate dehydrogenase (G6PD) activity, thereby supporting proper red blood cell function.[16] Another factor is erythropoietin therapy which serves as the primary therapy for ESKD patients with anemia.[17] Additionally, patients' perception of their illness is thought to affect Hb levels, as those with a positive perception tend to maintain more stable emotional states and demonstrate greater adherence to recommended dietary and medication regimens.[18,19]

Several studies have analyzed the factors influencing Hb levels in patients with ESKD; however, the findings have been varied and inconsistent across different factors.[20-23] This study aims to identify patient-modifiable factors such as dietary intake, adherence to weekly hemodialysis regimens, and perceptions of their illness that may affect Hb levels. It is expected that the findings of this research will provide clinical implications for identifying factors that can help minimize the risk of anemia among ESKD patients.

Materials and methods

Study design

This prospective observational study was conducted and reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.[24] The research was conducted at Dr. Sardjito General Hospital, Yogyakarta, over a four-week period. The study subjects were determined based on Hb concentrations measured between June 16 and 18, 2022, and were followed for four weeks, starting from June 20 to July 19, 2022.

Sample

The sample size was calculated using the Lemeshow's formula, resulting in a minimum of 32 participants. Considering a 10% predicted dropout rate, the required minimum sample was 35. However, 38 patients in each group met the criteria, yielding a total sample of 76 participants (38 with Hb >10 gm/dL and 38 with Hb ≤10 gm/dL). However, 11 participants were excluded, resulting in 65 patients included in the analysis (33 with Hb >10 gm/dL and 32 with Hb ≤10 gm/dL). The Hb cutoff of 10 gm/dL was used as the minimum target for the maintenance phase in dialysis patients, based on the Dialysis Consensus of the Indonesian Nephrology Association.[25] The inclusion criteria were ESKD patients undergoing hemodialysis twice a week, aged ≥18 years, and having received hemodialysis for at least six months. The exclusion criteria were patients with human immunodeficiency virus (HIV) or acquired immune deficiency syndrome (AIDS), liver cirrhosis, spinal cord malignancy, renal carcinoma, a history of kidney transplantation, or prior chemotherapy. The dropout criteria were critical illness, failure to undergo hemodialysis twice a week, withdrawal from the study, additional iron therapy, or death.

Data collection

Data were obtained from food record sheets documenting calorie, protein, iron, folic acid, and vitamin B12 intake, collected three days per week over four weeks.[26] The records were entered into the Indonesian version of NutriSurvey software to analyze nutritional content consumed by each patient. The B-IPQ was used to assess patients' perceptions of illness,[27] and had been validated in ESKD patients, showing validity values of 0.328-0.844 and a reliability coefficient of 0.755.[28] Data on body mass index (BMI), duration of hemodialysis, adequacy of hemodialysis with clearance time volume (Kt/V), and ESA therapy were recorded using an observation sheet prepared by the researchers.

The accuracy of the data on food and beverages consumed depended on the ability and honesty of the participants in reporting. To minimize bias, the researchers reconfirmed participants' dietary information upon submission of the completed food record forms.

Data analysis

Data were analyzed using the statistical package for the social sciences (SPSS) version 22 (IBM Corp. Armonk, NY, USA). Bivariate tests were first conducted to determine candidate variables for logistic regression, defined as those with a $p<0.05$. Logistic regression analysis was then performed to identify factors influencing Hb levels, using the backward method to determine OR with 95% CI. The variables included in the logistic regression analysis were iron intake, vitamin B12 intake, hemodialysis duration, Kt/V, and ESA therapy (Table 1).

Table 1. Factors affecting normal Hb levels.

Variable	Coefficient	*p* value	OR	95% CI
Model 1				
Intake iron	0.532	0.442	1.702	0.439-6.596
Vitamin B12	1.054	0.148	2.870	0.689-11.953
Duration of hemodialysis	1.561	0.026*	4.761	1.206-18.801
Kt/V	0.794	0.296	2.212	0.499-9.800
ESA therapy	1.446	0.046*	4.245	1.023-17.617
Constant	-7.834	0.000	0.000	
Model 2				
Vitamin B12	1.066	0.140	2.905	0.705-11.976
Duration of Hemodialysis	1.484	0.031*	4.411	1.143-17.019
Kt/V	0.917	0.214	2.503	0.589-10.638
ESA therapy	1.580	0.025*	4.854	1.216-19.380
Constant	-7.324	0.000	0.001	
Model 3				
Vitamin B12	1.444	0.027*	4.238	1.183-15.176
Duration of Hemodialysis	1.591	0.019*	4.909	1.299-18.551
ESA therapy	1.969	0.002*	7.162	2.046-25.071
Constant	-7.301	0.000	0.001	

Note: Logistic regression backward method; *significant at $p<0.05$

Results

The recruitment and sample group determination process was conducted on June 18, 2022. A total of 191 patients underwent routine hemodialysis twice a week. In the first step, 26 were excluded (21 with hemodialysis for less than one month, 2 had HIV/AIDS, and 3 did not have the latest Hb test results). Of the 165 patients, 42 had Hb levels >10 gm/dL; however, 4 were excluded (3 declined to participate and 1 was scheduled for kidney transplantation), resulting in a group of 38 respondents with Hb levels >10 gm/dL, resulting in a group of 38 respondents with Hb levels >10 gm/dL. Another 38 patients with Hb levels ≤10 gm/dL were selected through matching on age, sex, length of hemodialysis, and comorbidities.

The number of samples analyzed was 65 (33 patients with Hb >10 gm/dL and 32 with Hb ≤10 gm/dL), since 11 participants were considered dropped out by the researchers because one participant resigned, one participant had to get special treatment in the Intensive care unit, one participant undergo hemodialysis less than two times a week because there were problems with vascular access, seven participants received additional iron intravenous therapy, and one participant died.

Table 2 shows that 58% of patients with Hb >10 gm/dL and Hb ≤10 gm/dL are aged 48.03±12.00 and 48.78±14.07 years, respectively. Among patients with Hb >10 gm/dL, 54% are male, whereas in the Hb ≤10 gm/dL group, 59% are female. The most recent education level in both groups was higher education. Regarding occupation, 33% of patients with Hb >10 work as government or private employees, while 31% of patients with Hb ≤10 gm/dL are housewives. Patients in both groups have been undergoing hemodialysis for 1-5 years. The most common eating disorder reported in both groups was decreased appetite. Hypertension was the most prevalent comorbidity among patients with Hb >10 gm/dL and Hb ≤10 gm/dL. Folic acid was the most frequently consumed vitamin or supplement. Patients with Hb >10 gm/dL were more likely to consume additional vitamins or supplements

such as Neurobion™, become C™, Albuvit™, and iron tablets, compared to patients with Hb ≤10. The results of regression analysis (Table 1) indicated that ESA therapy (OR=7.162, 95% CI=2.046-25.071, *p*=0.002), duration of hemodialysis (OR=4.909, 95% CI=1.299-18.551, *p*=0.019), and vitamin B12 intake (OR=4.238, 95% CI=1.183-15.176, *p*=0.027) were significantly associated with Hb levels in patients with ESKD.

Table 2. Characteristics of patients with ESKD.

Characteristic		**Hb >10 gm/dL N (%)**	**Hb ≤10 gm/dL N (%)**	***p* value**
Age[a]		48.03 ± 12.00	48.78 ± 14.07	0.646
Gender[b]	Male	18 (54)	13 (41)	0.261
	Female	15 (46)	19 (59)	
Education[c]	Primary school	4 (12)	7 (22)	0.078
	Secondary school	2 (6)	6 (19)	
	Higher education	15 (46)	12 (37)	
	Associate degree/bachelor	10 (30)	5 (16)	
	Master's degree	2 (6)	2 (6)	
Occupation[c]	Farmer/ laborer	2 (6)	2 (6)	0.641
	Housewife	7 (21)	10 (31)	
	Self-employed	7 (21)	6 (19)	
	Government/private	11 (33)	7 (21)	
	employee	1 (3)	1 (3)	
	Student	5 (16)	6 (19)	
	Retired			
Hemodialysis period[c]	6 - <12 months	2 (6)	1 (3)	0.120
	1 – 5 years	14 (42)	23 (71)	
	6 - 10 years	13 (39)	4 (13)	
	>10 years	4 (13)	4 (13)	
Eating disorders[c]	Nausea	2 (6)	4 (13)	0.630
	Anorexia	8 (25)	7 (21)	
	None	23 (69)	21 (66)	
Comorbidities[c]	Hypertension	31 (94)	31 (97)	0.407
	Diabetes	14 (42)	8 (25)	
	Gout	2 (6)	3 (9)	
Vitamins/supplements[c]	Used	25(76)	25 (75)	0.821
	Folic acid	25(76)	24 (75)	
	Neurobion™	8 (24)	4 (12)	
	Albuvit™	3 (9)	2 (6)	
	Becom C™	2 (6)	0 (0)	
	Iron tablets	1 (3)	0 (0)	
	None	8 (24)	7 (22)	
Intake Calories[b]	≥35 kkal/kg body weight/day	12 (36)	7 (22)	0.199
	<35 kkal/kg body weight/day	21 (64)	26 (78)	
Intake Protein[b]	≥1 g/kg body weight/day	19 (58)	16 (50)	0.540
	<1 g/kg body weight/day	14 (42)	16 (50)	
Intake Iron[b]	≥10 mg/day	21 (64)	11 (34)	0.018*
	<10 mg/day	12 (36)	21 (66)	
Intake Folic acid[b]	≥0,8 mg/day	25 (76)	24 (75)	0.943
	<0,8 mg/day	8 (24)	8 (25)	
Intake Vitamin B12[b]	≥3 µg/day	20 (61)	10 (31)	0.018*
	<3 µg/day	13 (39)	22 (69)	
BMI[b]	Normal	20 (62)	18 (56)	0.722
	Abnormal	13 (38)	15 (44)	

Table 2. Characteristics of patients with ESKD (continued).

Characteristic		Hb >10 gm/dL N (%)	Hb ≤10 gm/dL N (%)	*p* value
Adequation of hemodialysis (Kt/V)[b]	≥1.8	26 (79)	10 (31)	0.000*
	<1.8	7 (11)	22 (69)	
Duration of hemodialysis[b]	≥4.5 hours	28 (85)	15 (47)	0.001*
	<4.5 hours	5 (15)	17 (53)	
ESA therapy[b]	Appropriate	25 (76)	9 (28)	0.000*
	Inappropriate	8 (24)	23 (72)	
Perseption of disease[b]	Positive	21 (64)	15 (47)	0.174
	Negative	12 (36)	17 (53)	

Note: *significant at $p<0.05$, [a]mann whitney, [b]chi-square, [c]spearman rank.

Discussion

The factors influencing Hb levels in ESKD patients undergoing hemodialysis include ESA therapy, hemodialysis duration, vitamin B12 intake, and education. The dosage of alpha and beta ESA was categorized as 2000-5000 international units (IU) administered twice weekly or 80-120 IU/kg body weight per week. Continuous erythropoietin receptor activator (CERA) was administered at 0.6 µg/kg body weight or 50-75 µg every two weeks.[29] Erythropoiesis-stimulating agent therapy was identified as the dominant factor affecting Hb levels in patients with ESKD undergoing hemodialysis. Patients who received appropriate ESA therapy had a 7.162-fold higher likelihood of achieving Hb levels >10 gm/dL. In this study, the frequency of ESA therapy differed between the groups: patients with Hb >10 gm/dL received ESA therapy once weekly, whereas those with Hb ≤10 gm/dL received therapy twice weekly. Some patients with low Hb received ESA therapy cautiously due to post-dialysis increases in blood pressure, which could exacerbate hypertension.[30] In line with previous studies, regular administration of ESA therapy significantly increased Hb levels in patients with ESKD.[17]

Each hemodialysis session lasted 4-5 hours and was performed 2-3 times per week with the ideal total duration being 10-15 hours per week. However, most health insurance providers in Indonesia only cover an average of two sessions per week, and patients spending 5 hours per session are rare. Therefore, the maximum average duration per patient was set at 4.5 hours.[25] In this study, patients with an average hemodialysis duration of 4.5 hours had a 4.909 times chance of Hb levels compared to patients with an average hemodialysis duration of <4.5 hours. In line with the results described in Table 1, 85% of patients with a mean duration of hemodialysis of 4.5 hours had Hb levels of >10 gm/dL. Previous studies have also reported that patients undergoing hemodialysis for 12 hours per week had higher average Hb levels than those dialyzed for 8-10 hours per week.[31] Longer hemodialysis sessions improve dialysis adequacy, allowing for more effective removal of toxins.[15] Elevated blood urea nitrogen (BUN) levels have been shown to indirectly affect Hb levels; patients with high BUN have a 1.02-fold higher risk of developing anemia.[32]

Patients with adequate vitamin B12 intake had a 4.238-fold higher likelihood of achieving target Hb levels compared to those with lower intake, which aligns with previous research.[33] Vitamin B12 is a nutrient that functions as one of the main ingredients that play a role in the production of red blood cells.[34] Vitamin B12 is essential for red blood cell production and is a key marker in megaloblastic anemia, which arises from deficiencies in vitamin B12 and folic acid. Appropriate intake of vitamin B12 supports stable Hb levels in patients undergoing hemodialysis.[13] In this study, patients with Hb >10 gm/dL had higher average vitamin B12 intake than those with Hb ≤10 gm/dL; however, as shown in Table 1 , vitamin B12 intake was not a statistically significant predictor of Hb levels. Consistent findings have been reported in previous studies.[35] Only a small proportion of patients regularly consumed iron-rich foods, such as red meat, liver, green vegetables, and fish, and the variety of diet was limited. Adequate iron intake is crucial for increasing serum Hb levels and optimizing treatment response.[36]

This study found no significant effect of caloric, protein, or folic acid intake on Hb levels in ESKD patients undergoing hemodialysis. Both the Hb >10 gm/dL and Hb ≤10 gm/dL groups had average caloric intake within the lower category, which contrasts with previous studies on energy intake.[9] Patients must also restrict food intake to prevent excessive potassium and phosphate levels and to avoid fluid overload.[10] In line with previous research there was no significant relationship between protein intake and Hb levels.[35] In general, Essential amino acids from protein intake are generally necessary for red blood cell formation, as they play a crucial role in iron transport and utilization in the body.[37] Similarly, folic acid intake was not significantly associated with Hb levels in this study.[35] Most patients had taken daily folic acid as recommended, yet some still had Hb ≤10 gm/dL. Folic acid, a water-soluble vitamin, is essential for the production of healthy red blood cells.[14] Therefore, the intake of foods containing folic acid in patients with ESKD must adequate, as the human body cannot synthesize folic acid and must

obtain it through daily dietary intake.[38] This finding contrasts with previous studies on BMI.[39] In this study, less than half of the patients had comorbid diabetes mellitus. Patients with diabetes typically exhibit lower Hb levels than those without diabetes because blood glucose levels can influence Hb levels in individuals undergoing hemodialysis. On average, ESKD patients with controlled blood glucose levels maintained more stable Hb levels.[40]

Almost all patients in this study underwent hemodialysis twice weekly, and therefore, a Kt/V standard of 1.8 was applied. According to the dialysis consensus of the Indonesian Nephrology Association, the ideal Kt/V target is 1.2 for patients undergoing hemodialysis three times per week, whereas the target for those dialyzed twice weekly is 1.8.[25] In this study, hemodialysis adequacy was not significantly associated with Hb levels. Ginting *et a*l. also reported that patients achieving a Kt/V of 1.8 exhibited higher Hb levels than those with Kt/V <1.8.[41] Adequate hemodialysis is crucial for reducing anemia, morbidity, and mortality in patients with ESKD.[42] Additionally, another study reported that patients receiving adequate hemodialysis had higher erythrocyte G6PD activity.[16] Low G6PD activity is associated with reduced red blood cell survival, leading to decreased Hb levels.[43]

This study also examined patients' perceptions of illness, including both positive and negative beliefs, and their potential influence on health-related decision-making.[44,45] Most patients (64%) with Hb levels >10 gm/dL reported a positive perception of their disease compared to patients with Hb ≤10 gm/dL; however, no significant effect on Hb levels was observed. Research on the relationship between disease perception and Hb levels in hemodialysis patients remains limited. Patient perceptions in ESKD are often related to hopelessness, self-management, medication adherence, and dietary compliance. Patients with a positive perception tend to maintain more stable emotional status and follow self-management, medication, and dietary recommendations, which may have a favorable impact on Hb level.[18,19,46]

Limitations

This study has several limitations. First, including its cross-sectional design and relatively small sample size may limit the generalizability of the findings. Second, monitoring was conducted for only one month, with Hb measurements taken only twice, at baseline and at the end of the study. Third, other potentially influential factors, including infection, pure red cell aplasia (PRCA), dialysis access, smoking, drinking, and laboratory tests such as albumin levels, were not monitored. Fourth, the accuracy of dietary intake depended on the participants' ability and honesty in reporting, which was beyond the researchers' control. All participants were not willing to maintain food intake records for four weeks, as the types and portions consumed were not substantially different from previously recorded data. Consequently, dietary intake variables were analyzed using data recorded by patients for 2-3 weeks. Therefore, the results of this study should be interpreted with caution.

Conclusion

Erythropoietin stimulating agent therapy, duration of hemodialysis, and vitamin B12 intake were identified as factors influencing the improvement of hemoglobin levels in patients with ESKD. Enhanced education programs are needed to increase patient awareness of the importance of maintaining a balanced nutritional intake and completing hemodialysis sessions of adequate duration to sustain and enhance the hemoglobin levels. In this study, dietary intake data were obtained solely from patient-completed food records, which may have introduced variability due to individual perceptions of portion sizes. Additionally, hemodynamic variables were not assessed, highlighting the need for further research to validate and extend these findings.

Ethical approval

This study has approved by Medical and Health Research Ethics Committee of the Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia (KE/FK/0750/EC/2022).

Funding

None

Conflict of interest

The authors declare no conflict of interest in this study. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Maryadi M: conceptualization, methodology, data collections, data curation, writing original draft, review and editing, analyzed the data, and drafted the manuscript; **Khudazi Aulawi:** conceptualization, methodology, data curation; and supervision; **Uki Noviana:** conceptualization, methodology, and validation, and supervision.

Acknowledgements

We extend our gratitude to the respondents and the RSUP Dr. Ardito Yogyakarta for facilitating access to the respondents.

References

[1] CDC. Chronic Kidney Disease in the United States, 2021. Published 2021. https://www.cdc.gov/kidneydisease/publications-resources/CKD-national-facts.html%0AWhen.

[2] Indonesian Renal Registry. Report of Indonesian Renal Registry.; 2018. https://www.indonesianrenalregistry.org/data/IRR 2018.pdf.

[3] Cases A, Egocheaga MI, Tranche S, et al. Anemia

of chronic kidney disease: protocol of study, management and referral to nephrology. Nefrología. 2018; 38(1): 8-12. doi:10.1016/j.nefroe.2018.01.007.

[4] Mathias SD, Blum SI, Sikirica V, Johansen KL, Colwell HH, Okoro T. Symptoms and impacts in anemia of chronic kidney disease. J Patient-Reported Outcomes. 2020; 4(1): 64-74. doi:10.1186/s41687-020-00215-8.

[5] Offurum A, Wagner LA, Gooden T. Adverse safety events in patients with chronic kidney disease (CKD). Expert Opin Drug Saf. 2016; 15(12): 1597-607. doi:10.1080/14740338.2016.1236909.

[6] NICE. Guidance chronic kidney disease: Managing anaemia. National Institute for Health and Care Excellence. Published June 3, 2015. Accessed December 8, 2021. https://www.nice.org.uk/guidance/ng8.

[7] Wjayantii S, Yetti K, Masfuri. Analysis of the factors affecting the hospital readmission incidence of hemodialysis patients in Bandar Lampung, Indonesia. Enferm Clin. 2019; 29(2): 720-4. doi:10.1016/j.enfcli.2019.04.110.

[8] Shiferaw WS, Akalu TY, Aynalem YA. Risk factors for anemia in patients with chronic renal failure: a systematic review and meta-analysis. Ethiop J Health Sci. 2020; 30(5): 829-42. doi:10.4314/ejhs.v30i5.23.

[9] Obeagu EI, Babar Q, Obeagu GU. Megaloblastic anaemia - a review. Int J Curr Res Med Sci. 2021; 7(5): 17-24. doi:10.22192/ijcrms.2021.07.05.004.

[10] González-Ortiz A, Correa-Rotter R, Vázquez-Rangel A, Vega-Vega O, Espinosa-Cuevas Á. Relationship between protein-energy wasting in adults with chronic hemodialysis and the response to treatment with erythropoietin. BMC Nephrol. 2019; 20(1): 1-9. doi:10.1186/s12882-019-1457-0.

[11] Lopez MJ, Mohiuddin SS. Biochemistry, essential amino acids. StatPearls Publishing; 2022. Accessed August 2, 2022. https://www.ncbi.nlm.nih.gov/books/NBK557845/.

[12] Mikhail A, Brown C, Williams JA, et al. Renal association clinical practice guideline on anaemia of chronic kidney disease. BMC Nephrol. 2017; 18(1): 345-74. doi:10.1186/S12882-017-0688-1.

[13] Froese DS, Fowler B, Baumgartner MR. Vitamin B12, folate, and the methionine remethylation cycle—biochemistry, pathways, and regulation. J Inherit Metab Dis. 2019; 42(4): 673-85. doi:10.1002/jimd.12009.

[14] Hansen DK, Inselman AL. Folic acid. StatPearls Publishing; 2022. doi:10.1016/B978-0-12-386454-3.00731-4.

[15] Hekmat R. Hemodialysis adequacy and its impact on long-term patient survival in demographically, socially, and culturally homogeneous patients. Int J Nephrol. 2020; 2020. doi:10.1155/2020/9857123.

[16] Ayesh MH, Bataineh A, Elamin E, Khader Y, Alawneh K, Rababah M. Adequate Hemodialysis improves anemia by enhancing glucose-6-phosphate dehydrogenase activity in patients with end-stage renal disease. BMC Nephrol. 2014; 15(1): 1-5. doi:10.1186/1471-2369-15-155.

[17] Ghasemi F, Abdi A, Salari N, Tohidi MR, Faraji A. Comparing the effects of intravenous and subcutaneous erythropoietin on blood indices in hemodialysis patients. Sci Rep. 2019; 9(1): 1-8. doi:10.1038/s41598-018-38193-z.

[18] Kim S, Kim E, Ryu E. Illness perceptions, self-care management , and clinical outcomes according to age-group in Korean hemodialysis patients. Int J Environ Res Public Health. 2019; 16(22): 1-12. doi:10.3390/ijerph16224459.

[19] Gunarathne TGNS, Tang LY, Lim SK, Nanayakkara N, Damayanthi HDWT, Abdullah KL. Factors associated with aymptom burden in adults with chronic kidney disease undergoing hemodialysis: a prospective study. Int J Environ Res Public Health. 2022; 19(9): 1-13. doi:10.3390/ijerph19095540.

[20] Xu Y, Evans M, Barany P, James G, Sjölander A, Carrero JJ. Factors affecting pre-end-stage kidney disease haemoglobin control and outcomes following dialysis initiation: a nationwide study. Clin Kidney J. 2021; 14(7): 1780-8. doi:10.1093/ckj/sfaa213.

[21] Yin S, Du Y, Guo Y, Guo G, Sun D, Yao L. Multifactorial analysis of renal anemia-associated substandard hemoglobin levels and prevalence of anemia in patients on maintenance hemodialysis in Liaoning Province: a cross-sectional study. Ann Palliat Med. 2022; 11(12): 3743-54. doi:10.21037/apm-22-1348.

[22] Sutranto AL, Liwang F, Nugroho P, Harimurti K. Pos-620 factors influencing hemoglobin variability in chronic hemodialysis patients: A prospective cohort study in tertiary hospital, Jakarta, Indonesia. Kidney Int Reports. 2021; 6(Suppl 4): S270. doi:https://doi.org/10.1016/j.ekir.2021.03.649.

[23] Qian G, Zhu Y, Tao S, et al. Increased hemoglobin concentration and related factors in maintenance hemodialysis patients in Anhui, China. Med (United States). 2022; 101(46): E31397. doi:10.1097/MD.0000000000031397.

[24] Cuschieri S. The STROBE guidelines. Saudi J Anaesth. 2019;13(Suppl 1): S31-S34. doi:10.4103/sja.SJA_543_18.

[25] Indonesian Nephrology Association. Dialysis consensus. 1st Edition; 2003. https://www.pernefri.org/konsensus/Konsensus Dialisis.pdf.

[26] Par'i HM, Wiyono S, Harjatmo TP. Nutritional status assessment. Center for Health Human Resources Education, Indonesian Ministry of Health; 2017.

[27] Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. J Psychosom Res. 2006; 60(6): 631-7. doi:10.1016/j.jpsychores.2005.10.020.

[28] Hadisa N, Susanti R, Robiyanto. Validity and

reliability test of Indonesian Version B-IPQ to CRF Outpatients in RSUD Soedarso Pontianak. J Manaj dan Pelayanan Farm. 2017; 7(4): 175-81. doi:10.22146/jmpf.33340.

[29] Indonesian Nephrology Association. Consensus on anemia management in chronic kidney disease. 2nd Edition; 2011. https://www.pernefri.org/konsensus/Konsensus Anemia - Isi.pdf.

[30] Brar SK, Perveen S, Chaudhry MR, AlBabtain S, Amreen S, Khan S. Erythropoietin-induced hypertension: A review of pathogenesis, treatment, and role of blood v iscosity. Cureus. 2021; 13(1): e12804. doi:10.7759/CUREUS.12804.

[31] Puspita AA, Setianingrum ELS, Lidia K. The effect of hemodialysis frequency on the differences hemoglobin levels and erythrocyte indices in chronic kidney failure patients pre and post hemodialysis at Prof. Dr. W. Z. Johannes Hospital. Cendana Med J. 2019; 7(1): 102-111. doi:10.35508/cmj.v7i1.1462.

[32] Kim HJ, Kim TE, Han M, et al. Effects of blood urea nitrogen independent of the estimated glomerular filtration rate on the development of anemia in non-dialysis chronic kidney disease: the results of the KNOW-CKD Study. PLoS One. 2021; 16(9): e0257305. doi:10.1371/journal.pone.0257305.

[33] Tayebi A, Biniaz V, Savari S, et al. Effect of vitamin B12 supplementation on serum homocysteine in patients undergoing hemodialysis: a randomized controlled trial. Saudi J Kidney Dis Transpl. 2016; 27(2): 256-62. doi:10.4103/1319-2442.178255.

[34] Nahas ARFM, Al Zarzour RH, Abu Laila ASM, et al. Effect of B12 supplementation on renal anemia among hemodialysis patients at El-Najar hospital, Gaza strip. J Ren Inj Prev. 2022; x(x): 1-5. doi:10.34172/jrip.2022.32009.

[35] Vittoria V. The relationship between food intake, nutritional status and hemodialysis adequacy with anemia in terminal renal failure patients undergoing routine hemodialysis at RSUP Dr. Sardjito Hospital, Yogyakarta. Published online 2016. https://etd.repository.ugm.ac.id/penelitian/detail/103973.

[36] Gvili AG, Schechter A, Zvi BR. Iron deficiency anemia in chronic kidney disease. Acta Haematol. 2019; 142(1): 44-50. doi:10.1159/000496492.

[37] Loutradis C, Skodra A, Georgianos P, et al. Diabetes mellitus increases the prevalence of anemia in patients with chronic kidney disease: a nested case-control study. World J Nephrol. 2016; 5(4): 358. doi:10.5527/wjn.v5.i4.358.

[38] Cappuccilli M, Bergamini C, Giacomelli FA, et al. Vitamin B supplementation and nutritional intake of methyl donors in patients with chronic kidney disease : a critical review of the impact on epigenetic machinery. MDPI Nutr. 2020; 12(5): 1234-53. doi:10.3390/nu12051234.

[39] Raza ST, Singh N, Kumar D, Singh S, Chatterji T, Zaidi Z. A study to compare hemoglobin levels and body mass index in normal and diagnosed diabetic stages of CKD patients visiting a tertiary care hospital. Indian J Clin Anat Physiol. 2021; 8(3): 230-4. doi:10.18231/J.IJCAP.2021.051.

[40] Brzózka AG, Franczyk B, Olszewski R, Rysz J. The influence of inflammation on anemia in CKD patients. Int J Mol Sci. 2020; 21(3): 725-49. doi:10.3390/IJMS21030725.

[41] Ginting L, Aziz A, Wulandari, Lasono DB, Koerniawan D. The relationship between hemodialysis adequacy achievement and the incidence of anemia in chronic kidney disease (CKD) patients on hemodialysis at the Hemodialysis Unit, RSUP Fatmawati. Diklit RS Fatmawati Jakarta. Published online 2018; 2(2): 3-12. http://diklitrsfatmawati.com/jurnal/single.php?id=27.

[42] Somji SS, Ruggajo P, Moledina S. Adequacy of hemodialysis and its associated factors among patients undergoing chronic hemodialysis in Dar es Salaam, Tanzania. Int J Nephrol. 2020; 2020: 1-6. doi:10.1155/2020/9863065.

[43] Richardson SR, O'Malley GF. Glucose 6 phosphate dehydrogenase deficiency. StatPearls Publishing; 2022. Accessed July 30, 2022. https://www.ncbi.nlm.nih.gov/books/NBK470315/.

[44] Broadbent E, Wilkes C, Koschwanez H, Weinman J, Norton S, Petrie KJ. A systematic review and meta-analysis of the Brief Illness Perception Questionnaire. Psychol Heal. 2015; 30(11): 1361-85. doi:10.1080/08870446.2015.1070851.

[45] Sheeran P, Harris PR, Epton T. Does heightening risk appraisals change people's intentions and behavior? A meta-analysis of experimental studies. Psychol Bull. 2014; 140(2): 511-43. doi:10.1037/a0033065.

[46] Mollaoglu M, Candan F, Mollaoglu M. Illness perception and hopelessness in hemodialysis. Arch Clin Nephrol. 2016; 2(1): 044-8. doi:10.17352/acn.000014.

Journal of Associated Medical Sciences 2026; 59(1): 184-190

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

Manufacturing and effectiveness evaluation of effective breast suit for breast radiotherapy

Yi-Seong Lee[1] and Jeong-Koo Kim[2*]

[1]Dept. of Radiation Oncology, Catholic Kwandong University International St. Mary's Hospital, Incheon, Republic of Korea.
[2]Dept. of Radiological Science, Hanseo University, Seosan-si, Chungnam, Republic of Korea.

ARTICLE INFO

Article history:
Received 22 May 2025
Accepted as revised 9 November 2025
Available online 27 November 2025

Keywords:
Breast cancer, IMRT, thermoplastic, setup error, commercial bra.

ABSTRACT

Background: Various techniques and devices are used to immobilize the breast satisfactorily during radiation therapy for breast cancer, and they are adopted and utilized in treatment centers according to their preferences.

Objectives: In this study, a breast suit was manufactured using the advantages of thermoplastics used in breast cancer and thoracic radiotherapy, and the breast suit was validated by analyzing the reproducibility of setup errors.

Materials and methods: A computed tomography simulator and phantom were used to acquire reference images of a breast suit and a commercial bra, respectively. After acquiring the images in the CBCT system, the setup error reproducibility was analyzed using a Volume View™ 3D image acquisition tool. The image analysis method used automated mode and manual mode, and the validity was evaluated by analyzing the change of setup error reproducibility in the X, Y, and Z axis directions and pitch, roll, and yaw rotation directions of each matched image.

Results: The reproducibility of setup errors between the breast suit and the commercial bra ranged from 0.4 mm to 5.6 mm on the X-axis and from 0.3 mm to 5.8 mm on the Y-axis. In contrast, the Z-axis exhibited larger variations, ranging from 0.8 mm to 16.2 mm. Rotational setup errors were all less than 1°, indicating no significant differences. While the breast suit showed comparable setup errors to the commercial bra in the X and Y directions, it demonstrated improved reproducibility along the Z-axis ($p = 0.09$) and significantly lower rotational errors ($p < 0.05$).

Conclusions: The validity of the reproducibility change of setup errors was confirmed in the X, Y, and Z directions for the breast suit, and valid results were obtained for more stable reproducibility change in the Z-axis direction. Therefore, it is inferred that the breast suit can be used as a fixation device to reduce setup error reproducibility by limiting breast movement.

Introduction

According to national cancer screenings from 2004 to 2018, stomach, liver, colon, breast, and cervical cancers were the top five cancers among Koreans. The incidence of breast cancer has increased among women, accounting for 20.6% of all female cancers according to the National Cancer Registry Project of the Ministry of Health and Welfare.[1] Breast cancer is a disease in which malignant cells form in the breast organ. The most common type of breast cancer is ductal carcinoma *in situ* (DCIS), which starts from cells

* **Corresponding contributor.**
Author's Address: *Dept. of Radiological Science, Hanseo University, Seosan-si, Chungnam, Republic of Korea.*
E-mail address: jkkim@hanseo.ac.kr
doi: 10.12982/JAMS.2026.023
E-ISSN: 2539-6056

in the ducts. Cancer that starts in the lobes or lobules is called lobular carcinoma and occurs more often in both breasts than other types of breast cancer.[2]

The standard treatment for breast cancer is mastectomy and partial mastectomy for DCIS. In addition, there is a method to improve the local control rate (up to 95%) with radiation therapy, which is expected to be as effective as mastectomy.[3] In general, radiotherapy for breast cancer is treated using the tangential technique of three-dimensional conformal radiation therapy (3D-CRT). However, in complex radiotherapy involving axillary nodes and internal mammary nodes (IMNs), high doses are delivered to surrounding organs at risk (OARs) and the heart to distribute the maximum dose to the planning target volume (PTV). Previous studies have shown that volume modulated arc therapy (VMAT) and helical tomotherapy techniques with Intensity-modulated radiation therapy (IMRT) can improve dose distribution to the PTV and OARs, and show more uniformity dose distribution compared to 3D-CRT.[4] As more sophisticated treatment delivery techniques than 3D CRT are used to treat breast cancer, additional means of immobilization are needed to properly align the patient during treatment. Thermoplastic immobilization systems can reduce chest wall motion in lung cancer patients under extreme breathing conditions, and thermoplastic immobilization that extends from the top of the breast to the upper torso can effectively reduce intra-fraction and inter-fraction motion under normal breathing conditions in patients undergoing breast or chest wall radiation therapy.[5]

As radiation therapy for breast cancer transitions from 3D-CRT to IMRT, VMAT, and helical tomotherapy, the Monitor Units (MUs) increase by 2 to 3 times, leading to increased scatter and leakage radiation doses, as well as increased exposure to the surrounding tissue, which can elevate the risk of secondary malignancies. Overall, IMRT may increase the risk of secondary malignancy occurrence by nearly doubling, from approximately 1% to 1.75%, compared to conventional radiation therapy over a 10-year survival period.[6,7] IMRT also requires high accuracy in treatment setup reproducibility due to the presence of steep intervals where the radiation dose changes rapidly. When the accuracy of positional reproducibility is not high, high doses of radiation can be delivered to surrounding normal organs. Therefore, to deliver uniform doses and high doses of radiation to the target volume, reliable image guidance devices and accurate patient positioning devices within the treatment room are necessary.[8]

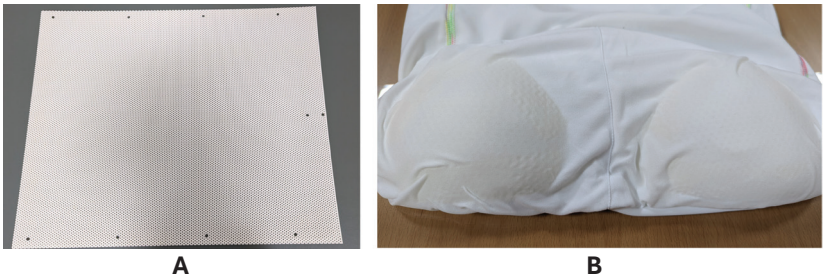
Various techniques and devices are used to immobilize the breast satisfactorily during radiation therapy for breast cancer, and they are adopted and utilized in treatment centers according to their preferences. These immobilization devices include; 1) prone breast board, 2) supine breast board, 3) thermoplastic shells, 4) wireless bra, 5) breast ring, 6) breast cup, and 7) vacuum bags, etc. Several previous studies, such as Strydhorst JH *et al.*, and Kovac *et al.*, have published data on the accuracy and reproducibility of patient setup using different thermoplastic materials or validated setup errors across different breast immobilization techniques.[9-12]

In this study, we fabricated a breast suit using the advantages of thermoplastic materials, a fixation technique used in breast cancer radiotherapy. The reproducibility and validity of setup errors in radiation therapy were verified by comparing the fabricated breast suit with a commercial bra. This study was based on the premise that the reproducibility of setup errors, which is an advantage of thermoplastic materials, is excellent, but it may cause skin side effects due to thermoplastic materials.

Materials and methods

Materials

Radiation therapy for breast cancer involves the use of several different fixation devices. Different fixation devices are used depending on the preference of the radiotherapy center. In this study, we used thermoplastic materials that are effective for fixation of the head & neck, and chest among various fixation devices. Use of a standard-sized thermoplastic sheet (approx. 550×450 mm) heated at 70°C. Application of elastic, breathable polyethylene fabric to enhance comfort and fit. Molding procedure directly over the breast pad. a breast suit was made by combining light, breathable, and elastic polyethylene materials that are quick to absorb and dry (Figure 1).



A B

Figure 1. Breast suit made using thermoplastic in breast radiotherapy. A: thermoplastic plate, B: breast suit.

To evaluate the reproducibility of setup errors in the breast suit, a thorax phantom (Radiology Support Devices; RSD, USA) and a water bag were fabricated and attached between the 2nd and 6th rib to create a custom-made breast phantom (Figure 2).

A B

Figure 2. Breast phantom made using a thorax phantom and a water bag. A: thorax phantom, B: brest phantom.

Methods

A breast board commonly used in breast cancer radiotherapy was used, and a breast suit and a commercial bra (Chabner XRT® Radiation Bra, USA) were worn on the breast phantom, respectively. Four Region of Interest (ROI) points were marked on the breast of the breast phantom to acquire baseline images, which were acquired using a 16-channel CT simulator (Somatom Definition AS CT, Siemens, Germany) at 120 kV, 87 Eff. mAs, 5 mm thickness, and 16.69 scan time. Images of each breast suit and commercial bra were acquired using the Volume View™ 3D image acquisition tool on a cone-beam computed tomography (CBCT) system (XVI, Elekta Infinity™, UK) with a linear accelerator (Elekta Infinity™, UK). The Volume View™ 3D image acquisition was performed using a 360° counterclockwise (CCW) rotation with a tube voltage of 120 kV and a tube current of 40 mA, and a volumetric image was acquired using 1040 mAs, 740 frames, field size medium(M20), and F1(filter), (Figure 3).

A B

Figure 3. Acquisition of breast suit and commercial bra. A: acquisition setup error evaluation, B: acquisition setup error evaluation for commercial bra for breast suit.

To evaluate the validity of the reproducibility of the setup error of the breast suit, the reference images of the breast suit and commercial bra acquired from the computed tomography simulator were registered with the CBCT system, an image guidance device. Using a Volume View™ 3D image acquisition tool, the reference images of the breast suit and commercial bra were compared and analyzed with the CBCT acquired images. The image analysis method used a manual method and an automated method of bone registration and gray value registration in the Volume View™ 3D image acquisition tool. Each matched image was validated by analyzing the change in reproducibility of six-dimensional setup errors in the X, Y, and Z axis directions and pitch, roll, and yaw rotation directions. Data were collected in Microsoft Excel sheets and statistically analyzed using SPSS version 20.0 for Windows (SPSS Inc., Chicago). Continuous variables were expressed as mean ± standard deviation (SD). A *P* value of < .05 was considered to reject the null hypothesis.

Results

Changes in reproducibility of setup errors in X, Y, and Z coordinate systems

The images acquired using a breast suit and a commercial bra in the 3D image acquisition mode of the CBCT system were compared and evaluated. The acquired images were analyzed by setting the region of interest using the 3D imaging tool and checking the change in reproducibility of setup errors in the X, Y, and Z axis coordinate system using manual and automated methods. In the manual method for the breast suit, the change was 0.4 mm in the X-axis, 0.3 mm in the Y-axis, and 0.35 mm in the Z-axis. For the Bone automation method, the change was 2.3 mm in the X-axis, 2.1 mm in the Y-axis, and 5.1 mm in the Z-axis. For the Gray automation method, the change was 3.7 mm in the X-axis, 0.2 mm in the Y-axis, and 3.9 mm in the Z-axis.

In the manual method of Commercial bra, the variation was 5.6 mm in the X-axis, 1.5 mm in the Y-axis, and 0.8 mm in the Z-axis. In the automated method of bone, the variation was 1.8 mm in the X-axis, 3.3 mm

in the Y-axis, and 16.2 mm in the Z-axis. For the Gray automation method, the change was 1.0 mm in the X-axis, 5.8 mm in the Y-axis, and 13.6 mm in the Z-axis.

The paired t-test analysis of translational setup errors showed no statistically significant differences between the breast suit and the commercial bra along the X-axis and Y-axis. However, along the Z-axis, the breast suit exhibited a lower mean error than the commercial bra, with a p=0.0881. Although this result did not meet the threshold for statistical significance, it indicates a clinically relevant trend, suggesting that the breast suit may provide improved reproducibility in the superior-inferior direction (Table 1 and 2, Figure 4).

Table 1. Setup error reproducibility results in the coordinate system (mm).

Axis	Breast suit			Commercial breast bra		
	Manual	Auto		Manual	Auto	
		Bone	Gray		Bone	Gray
X	0.4	2.3	3.7	5.6	1.8	1.0
Y	0.3	2.1	0.2	1.5	3.3	5.8
Z	0.3	5.1	3.9	0.8	16.2	13.6

Table 2. Setup error of mean absolute shift between breast suit and commercial bra (mm).

Axis	Breast suit	Commercial bra	*p* value
	Mean±SD (mm)	Mean±SD (mm)	
X	1.81±1.32	2.90±1.68	0.98
Y	1.11±0.78	3.10±1.62	0.29
Z	2.74±1.87	7.21±6.48	0.09

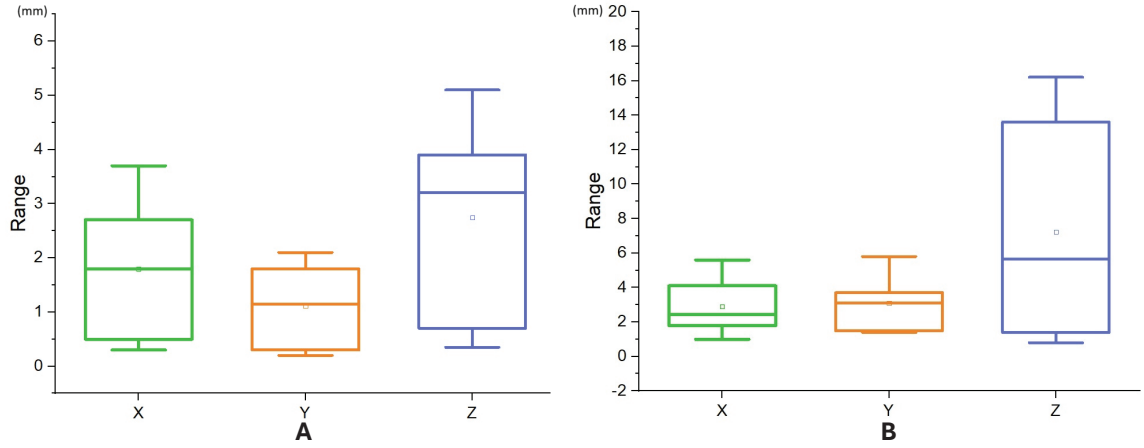


Figure 4. Setup error evaluation of breast suit and commercial bra. A: breast suit, B: commercial bra.

Changes in the reproducibility of setup errors in pitch, roll, and yaw rotation directions

The change in rotation setup error reproducibility of the breast suit and commercial bra was analyzed using the 3D image acquisition mode of the CBCT System, setting the same region of interest as the previous analysis method, and checking the change in pitch, roll, and yaw using the manual and automated methods. For the Breast suit, the change in the manual method was 0.28° for Pitch, 0.07° for Roll, and 0° for Yaw, while the change in the bone automation method was 0.19° for Pitch, 0° for Roll, and 0° for Yaw. For the Gray automation, the change was 0.19° for Pitch, 0.16° for Roll, and 0° for Yaw.

For the manual method of Commercial bra, the variation was 0° for pitch, 0.01° for roll, and 0.24° for yaw. For the automated method of bone, the variation was 0.74° for pitch, 0.1° for roll, and 0.1° for yaw. For the gray automation method, the variation was 0.64° for Pitch, 0.1° for Roll, and 0.1° for Yaw.

Paired t-test analysis of rotational setup errors revealed statistically significant differences between the breast suit and the commercial bra in all rotational directions. In the pitch direction, the breast suit showed an average error of 0.23±0.07°, compared to 0.29±0.39° for the commercial bra (p<0.05). In the roll direction, the breast suit exhibited a significantly higher stability with an error of 1.10±0.07°, while the

commercial bra showed 0.14±0.12° (p<0.05). For yaw, the breast suit and commercial bra recorded 0.03±0.04° and 0.16±0.09°, respectively (p<0.05). These results indicate that the breast suit provides superior reproducibility in rotational setup accuracy across all axes (Table 3 and 4, Figure 5).

Table 3. Setup error reproducibility results in rotation direction (°).

Axis	Breast suit			Commercial breast bra		
	Manual	Auto		Manual	Auto	
		Bone	Gray		Bone	Gray
Pitch	0.28	0.19	0.19	0	0.74	0.64
Roll	0.07	0	0.16	0.01	0.10	0.10
Yaw	0	0	0.24	0.24	0.10	0.10

Table 4. Rotation setup error of mean absolute shift between breast suit and commercial bra.

Axis	Breast suit	Commercial bra	*p* value
	Mean±SD (mm)	Mean±SD (mm)	
Pitch	0.23±0.07	0.29±0.39	0
Roll	1.10±0.07	3.10±1.62	0
Yaw	0.03±0.04	7.21±6.48	0

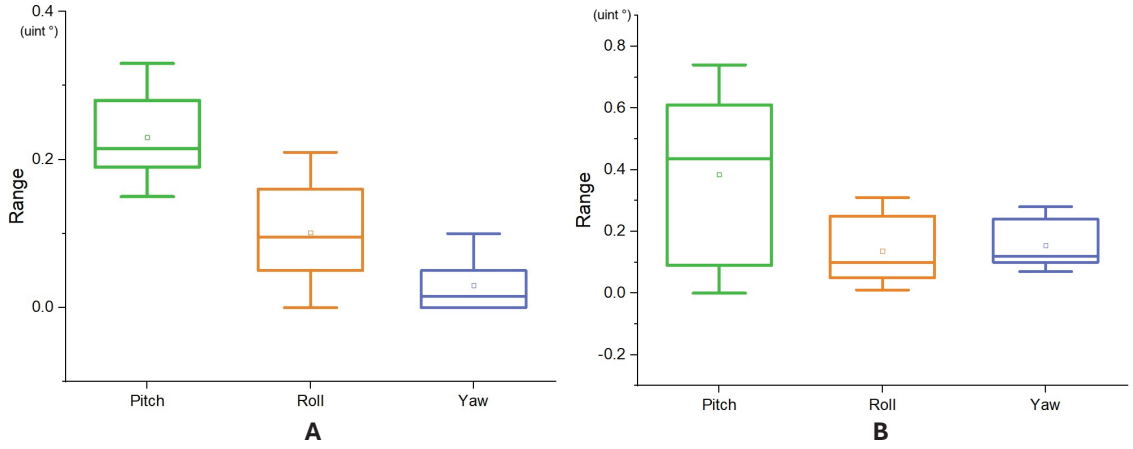


Figure 5. Rotation setup error of breast suit and commercial bra. A: breast suit, B: commercial bra.

Discussion

With the use of sophisticated radiotherapy techniques, proper alignment of patient posture is an important factor. Thermoplastic immobilization devices can effectively reduce intra-fraction and inter-fraction motion during normal breathing in patients undergoing radiotherapy to the breast or chest wall and chest.[13]

In this study, a breast suit was fabricated using thermoplastic fixtures to effectively reduce movement, and the reproducibility of setup errors in breast cancer radiotherapy was verified.

The setup error reproducibility of the breast suit showed variation of minimum 0.4 mm to maximum 5.6 mm in all directions in the X, Y, and Z axes. The setup error reproducibility of the commercial bra did not show a significant difference when comparing the changes in the breast suit with the changes in the X and Y axes, ranging from minimum 0.8 mm to maximum 5.8 mm. However, the Z axis showed a significant result with a minimum of 0.8 mm and a maximum of 16 mm.

Like the studies of Kubo *et al.*, and Kataria T *et al.*, that it is necessary to limit the setup error with significant vertical movement in chest radiotherapy, the breast suit showed better setup error reproducibility changes in the z-direction than the commercial bra.[14,15] These results demonstrate the validity of setup error reproducibility in the X, Y, and Z axes, and are expected to help improve setup error reproducibility by restricting breast movement in radiation therapy for breast cancer.

Breast suit rotation setup error reproducibility was no more than 1° at most in all directions of pitch, roll, and yaw for the breast suit and commercial bra, and did not show significant differences. Matthias

Guckenberger *et al*. showed in a study on the clinical relevance of size and rotation setup errors that the influence of weight and size on rotation errors is small.[16] However, it is believed that changes in the X, Y, and Z axes and corresponding changes in rotation can reduce setup error reproducibility as much as possible in radiation therapy for breast cancer.

Statistical analysis of the setup error reproducibility between the breast suit and the commercial bra showed no significant differences along the X- and Y-axes. However, in the Z-axis, the breast suit demonstrated a lower mean error, indicating a tendency to more effectively suppress gravitational sagging (p=0.09).

In terms of rotational directions, the breast suit demonstrated significantly lower setup errors in pitch, roll, and yaw, all of which reached statistical significance. These findings suggest that the breast suit provides superior stability in both translational and rotational positioning, supporting its clinical applicability in advanced radiotherapy techniques such as IMRT and VMAT. As this study was conducted using a phantom, the results may differ from those observed in actual clinical settings. Furthermore, no patient data were used in this study.

Conclusion

In this study, we analyzed the reproducibility of setup errors in breast cancer radiotherapy, which is a critical factor influencing treatment outcomes. A breast suit was fabricated in-house and evaluated by examining changes in setup error reproducibility. The breast suit showed no significant changes along the X and Y axes; however, in the Z-axis, it demonstrated smaller variations (2-5 mm) compared to the commercial bra (8-16 mm). No significant differences were observed in rotational setup error reproducibility. Overall, the breast suit exhibited greater stability, particularly in the Z-axis direction, and may serve as an effective fixation device in breast cancer radiotherapy by minimizing breast movement and gravitational sagging.

Ethical approval

Our study didn't involve humans or animals and didn't require ethics approval

Funding

This study received no research funding.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Yi-Seong Lee: conceptualization, investigation, resources, validation, and writing: original draft; **Jeong-Koo Kim**: conceptualization, methodology development, supervision, statistical analysis, and writing: review and editing.

Acknowledgements

The authors would like to thank Jiyeon medical for their assistance with this research.

References

[1] Hong S, Lee YY, Lee J, Kim Y, Choi KS, Jae K, Jun IK *et al*. Trends in cancer screening rates among Korean men and women. Results of the Korean National Cancer Screening Survey 2004-2018. 2021;53(2): 330-38. doi: 10.4143/crt.2020.263.

[2] Fonseca R, Hartmann LC, Petersen IA, Donohue JH, Crotty TB, Gisvold JJ. Ductal carcinoma in situ of the breast. Ann Intern Med. 1997; 127(11): 1013-22. doi: 10.7326/0003-4819-127-11-199712010-00013.

[3] McDnonald ES, ClarkAS, Tchou J, Zhang P, Freedman GM. Clinical diagnosis and management of breast cancer. J Nucl Med. 2016; 57: 9S-16S. doi: 10.2967/jnumed.115.157834.

[4] Arsene-Henry, Fourquet A, Kirova YM. Evolution of radiation techniques in the treatment of breast cancer (BC) patients: from 3D conformal radiotherapy (3D CRT) to intensity-modulated RT(IMRT) using Helical Tomotherapy(HT). Radiother Oncol. 2017; 124(2): 333-4. doi: 10.1016/j.radonc.2017.07.002.

[5] Strydhorst JH, Caudrelier JM, Clark BG, Montgomery LA, Fox G, MacPherson MS. Evaluation of thermoplastic immobilization system for breast and chest wall radiation therapy. Med. Dosim. 2011; 36: 81-4. doi: 10.1016/j.meddos.2010.01.001.

[6] Sakthivel V, Kadirampatti Mani G, Sunil M, Boopathy R, Selvaraj J. Estimating second malignancy risk in intensity-modulated radiotherapy and volumetric-modulated arc therapy using mechanistic radiobiological model in radiotherapy for carcinoma of left breast. J Med Phys 2017; 42(4): 234-240. doi: 10.4103/jmp.JMP_89_17.

[7] Hall EJ. Intensity-modulated radiation therapy, protons, and the risk of second cancers. Int J Radiat Oncol Biol Phys. 2006; 65: 1-7. doi: 10.1016/j.ijrobp.2006.01.027.

[8] Caudrelier JM, Morgan SC, Montgomery L., Lacelle M, Nyiri B, Macpherson M. Helical tomotherapy for locoregional irradiation including the internal mammary chain in left-sided breast cancer. Dosimetric evaluation. Radiother Oncol. 2009; 90: 99-105. doi: 10.1016/j.radonc.2008.09.028.

[9] Jassal K, Bisht S, Kataria T, Sachdev K, Choughle A, Supe S. Comparison of geometrical uncertainties in breast radiation therapy with different immobilization methods. J Nucl Med Radiat Ther. 2013; 4: 140. doi: 10.4172/2155-9619.1000140.

[10] Barrett-Lennard MJ, Thurstan SM. Comparing immobilization methods for the tangential treatment of large pendulous breasts. Radiographer. 2008; 55: 7-13. doi: 10.1016/j.meddos.2004.10.006.

[11] Kovacs A, Hadjiev J, Lakosi F, Vallyon M, Cselik Z, Bogner P, *et al*. Thermoplastic patient fixation: Influence on chest wall and target motion during radiotherapy of lung cancer. Strahlenther Onkol. 2007; 183: 271-8. doi: 10.1007/s00066-007-1681-6.
[12] Xiugen FU, Xiong H, Zheng Z, Shen FU, Chao T, Liu K, et al. Analysis of setup errors during radiotherapy for breast cancer patients immobilized with neck and breast thermoplastic mask. Cancer Res Clinic. 2018; 30: 374-8.
[13] Strydhorst JH, Caudrelier JM, Clark BG, Montgomery LA., Fox G, MacPherson MS. Evaluation of a thermoplastic immobilization system for breast and chest wall radiation Therapy. Med Dosim. 2011; 36(1): 81-4. doi: 10.1016/j.meddos.2010.01.001.
[14] Kubo HD, Hill BC. Respiration gated radiotherapy treatment: a technical study. Phys Med Biol 1996; 41(1): 83-91. doi: 10.1088/0031-9155/41/1/007
[15] Kataria T, Abhishek A, Chadha P, Nandigam J. Set-up uncertainties: online correction with X-ray volume imaging. J Cancer Res Ther. 2011; 7(1): 40-6. doi: 10.4103/0973-1482.80457.
[16] Guckenberger M, Meyer J, Vordermark D, Baier K, Wilbert J, Flentje M. Magnitude and clinical relevance of translational and rotational patient setup errors: a cone-beam CT study. Int J Radiat Oncol Biol Phys. 2006; 65(3): 934-42. doi: 10.1016/j.ijrobp.2006.02.019.

Journal of Associated Medical Sciences 2026; 59(1): 191-200

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

Review and cost-effectiveness analysis of rapid tuberculosis screening: Implementing the sputum pooling method in Indonesia

Febri Sembiring[1*], Hafizah Ilmi Sufa[2], Yuli Yantika[3], Wardati Humaira[1]

[1]Politeknik Kesehatan Kementerian Kesehatan Medan, North Sumatera, Indonesia.
[2]Politeknik Kesehatan Kementerian Kesehatan Bandung, West Java, Indonesia.
[3]The Integrated Laboratory, Politeknik Kesehatan Kementerian Kesehatan Medan, North Sumatera, Indonesia.

ARTICLE INFO

Article history:
Received 12 May 2025
Accepted as revised 28 November 2025
Available online 11 December 2025

Keywords:
Pooling sputum, rapid molecular tests, clinical bacteriology, tuberculosis, cost-effectiveness.

ABSTRACT

Background: Indonesia continues to face significant challenges in meeting its tuberculosis (TB) case detection targets, partly due to the underutilization of diagnostic resources like Rapid Molecular Tests (RMTs). Innovative, cost-effective strategies are urgently needed to scale up screening efforts.

Objectives: This study aims to review the diagnostic accuracy, cost-effectiveness, and implementation considerations of the pooling sputum strategy to assess its potential for enhancing TB screening in Indonesia.

Materials and methods: This study was conducted using publicly available data from the Indonesian Ministry of Health and findings from existing literature. The projected impact was analyzed by implementing a 4-sample pooling sputum method for RMT (Xpert MTB/RIF Ultra) compared to the current individual testing strategy.

Results: The analysis projects that the pooling sputum strategy could increase RMT screening capacity fourfold. The national average monthly tests per RMT unit could rise from 120 to 480 individuals screened. Consequently, the projected time required to screen 100,000 individuals per province could decrease from an average of 5.5 months to 1.4 months. The strategy offers potential cost savings averaging 36.71% (±14.97%). A review of diagnostic data from previous studies shows high accuracy (sensitivity 97.35%; specificity 99.25%).

Conclusion: The pooling sputum method presents a viable and powerful strategy to significantly enhance the efficiency of TB screening in Indonesia. This approach can substantially increase RMT utilization, accelerate case detection, and reduce costs, thereby helping Indonesia progress towards its 2030 TB elimination targets. A pilot study is recommended to validate these findings within the Indonesian healthcare context.

Introduction

The Indonesian Ministry of Health has set ambitious targets to reduce tuberculosis (TB) incidence and mortality, aiming for a decrease to 65 new cases and 6 deaths per 100,000 people, respectively.[1] To meet these goals, the Ministry has prioritized TB screening as a key strategy. Since 2021, following the Director General of Disease Prevention and Control's Circular Letter (No. HK.02.02/III.I/936/2021), the Rapid Molecular Test (RMT) has become the primary diagnostic tool for TB in Indonesia. Under this directive, individuals suspected of having TB are first examined using the RMT.[2]

* **Corresponding contributor.**
Author's Address: *Politeknik Kesehatan Kementerian Kesehatan Medan, North Sumatera, Indonesia.*
E-mail address: febrisembiring.kemenkes@gmail.com
doi: 10.12982/JAMS.2026.024
E-ISSN: 2539-6056

By 2022, Indonesia had deployed 2,167 RMT devices, surpassing the National Strategy for TB Control's target of 2,133 units for 2020-2024. Each RMT unit, when functioning optimally, can conduct up to 200 tests per month. However, despite the increase in the number of devices, the utilization rate of RMTs has not seen a corresponding rise. In 2022, the average utilization rate of these devices was only 60%, equating to just 120 tests per month per unit. Furthermore, in 23 provinces, the RMT utilization rate was below 60%, with some areas reporting as low as 13%. This is significantly below the national target of 80% utilization.[1,3]

The underutilization of RMTs has serious implications. In 2022, 23% of TB suspects, approximately 832,351 individuals, did not undergo laboratory testing. This gap is concerning, especially as the proportion of TB cases diagnosed through RMTs is increasing, reaching at least 12% in 2022.[1]

Transmission of TB remains a critical public health issue. Individuals with active TB can infect 10-15 others annually.[4] Research by McCreesh and White (2018) published in Nature highlights the significant role of household, repeated, and non-repeated contacts in TB transmission.[5] The Indonesian Ministry of Health aims to achieve a 90% TB case detection rate by 2030, aligning with global targets for 2030.[1,2] Given TB's high transmissibility, optimizing RMT usage is crucial for identifying new cases and curbing the disease's spread.

Globally, the World Health Organization (WHO) estimated 10.6 million TB cases in 2021, with 39.7% remaining undiagnosed. In Indonesia, about one-third of TB cases are missed or not reported, often due to delayed, missed, or unattempted diagnoses. The government's current strategy primarily targets symptomatic individuals for testing. However, the 2022 Joint External Monitoring Mission (JEMM) recommended expanding screening to high-risk populations to enhance TB detection.[6]

Financial constraints pose another challenge. Indonesia's national strategy for TB control requires IDR 47.3 trillion for 2020-2024, but only IDR 15.7 trillion is available. Annually, the WHO estimates that Indonesia needs IDR 8.1 trillion for TB prevention, diagnosis, and care, yet it receives only IDR 1.7 trillion in funding. Since 2009, Indonesia has consistently met just 41% of its annual TB program funding needs, hindering efforts to detect new cases.[7,8]

To address this challenge, this study provides a review of pooling sputum and a cost-effectiveness analysis of its potential implementation in Indonesia to provide evidence-based support for health policy decisions.

Materials and methods

Study design and data sources

This study employed a strategic literature review framework. A comprehensive search was conducted across multiple electronic databases, including PubMed, Scopus, and Google Scholar, to identify relevant articles published up to 2024. The search strategy utilized a combination of keywords: ("pooling sputum" OR "pooled testing") AND ("tuberculosis" OR "TB") AND ("Xpert MTB/RIF" OR "Xpert Ultra") AND ("cost-effectiveness" OR "diagnostic accuracy"). Additionally, official reports and technical guidelines from the Indonesian Ministry of Health (2022-2023)[1,3] and the World Health Organization (WHO)[6] were manually searched to gather data on the Indonesian context, RMT utilization, and programmatic costs.

Inclusion and exclusion criteria

Studies were included if they met the following criteria: 1) evaluated the pooling of sputum samples for TB diagnosis, 2) utilized Xpert MTB/RIF or Xpert Ultra assays, 3) reported quantitative data on diagnostic accuracy (sensitivity, specificity), or 4) provided data on cost-effectiveness or cost savings. Only full-text articles published in English were considered. Studies were excluded if they were editorials, case reports without systematic data, or qualitative studies not focused on diagnostic performance or cost.

Data extraction and synthesis

Data from the included studies were extracted into a standardized template. Key information extracted included: author, year, country of study, study design, number of samples, pooling ratio (e.g., 4-to-1), and reported outcomes (sensitivity, specificity, and percentage of cost savings). The data extracted on diagnostic accuracy (Table 1) and cost savings (Table 2, Table 3) were descriptively synthesized and summarized. A qualitative summary of implementation challenges, such as the dilution effect and logistical hurdles, was also compiled from the literature.[9]

Table 1. Diagnostic accuracy data from previous studies on pooling sputum.

References	All Negative		Positive*	
	Detected (%)	Not detected (%)	Detected (%)	Not detected (%)
13	0	77 (100)	32 (100)	0
12	0	16 (100)	19 (95)	1
14	0	6 (100)	95 (96)	4
15	NR	NR	87 (100)	0
10	1	34 (97.1)	61 (95)	3
Total	1	133	294	8
Sensitivity	97.35% (95% CI: 94.85-98.85%)			
Specificity	99.25% (95% CI: 95.91-99.98%)			

Note: *positive: classified by ct<38, NR: not reported

Table 2. Theoretical cost savings of a pooling sputum method for 4 samples in 1 Xpert cartridge using Xpert MTB/RIF Ultra.

Reference	Country	Total of the sample	Individual suspected TBC (%)	Percentage of cost savings	Test sensitivity (%)	Test specificity (%)
18	Nigeria	738	115 (16)	31	NR	NR
10	Brazil	396	95 (24)	12.4	95	97.1
13	Lao	436	199 (45)	46	100%	100%
19	Cameroon	4156	274 (6.6)	48	99.4%	NR
20	Cambodia	584	91 (15,6)	27	NR	NR
17	Lao	3076	NR	35.6	97.6	97
12	Brazil	1280	320 (25)	57	94	100
Cost saving average based on references (%±SD)					36.71±14.97	
Cost saving maximum (%)					57	
Cost saving minimum (%)					12.4	

Note: *NR: not reported.

Table 3. Theoretical cost savings of a pooled testing strategy using 4 pooled samples.[14]

Study setting (disease prevalence)	No. of pooled tests + No. of individual tests required[a]	Cost of pooled testing strategy ($)	Cost savings with pooled testing strategy ($)	Percentage of cost savings (%)
Community (11%)	185+276	4,600.78	2,764.46	59.6
District hospital (26%)	185+520	7,035.90	329.34 (4)	4.7
Total population (16%)	185+372	5,558.86	1,806.38 (25)	32.5

Results

The pooling sputum design method: A technical overview

In the original scheme, TB screening was reserved for individuals suspected of TB by collecting personal sputum for further examination using the RMT instrument. In this innovation, the pooling sputum method is a strategy where sputum samples from multiple individuals are combined into a single tube for a single diagnostic test.[17] This approach is designed to increase testing throughput and reduce costs, particularly for screening high-risk populations.

Figures 1 and 2 represent a schematic of the innovation strategy to accelerate and improve new TB case finding using RMT, specifically for TB risk groups within households. Identity data of known active TB patients are traced to addresses and family members within a household. The most common approach, and the focus of this review, involves pooling samples from four individuals for testing with a single Xpert MTB/RIF Ultra cartridge.[17,10-12] In this scheme, 0.5 mL of sputum from each of four individuals is combined to create a 2 mL pooled sample. This pooled sample is then tested using only one new Xpert cartridge with Xpert MTB/RIF Ultra and processed in the village GeneXpert platform (Figure 1).

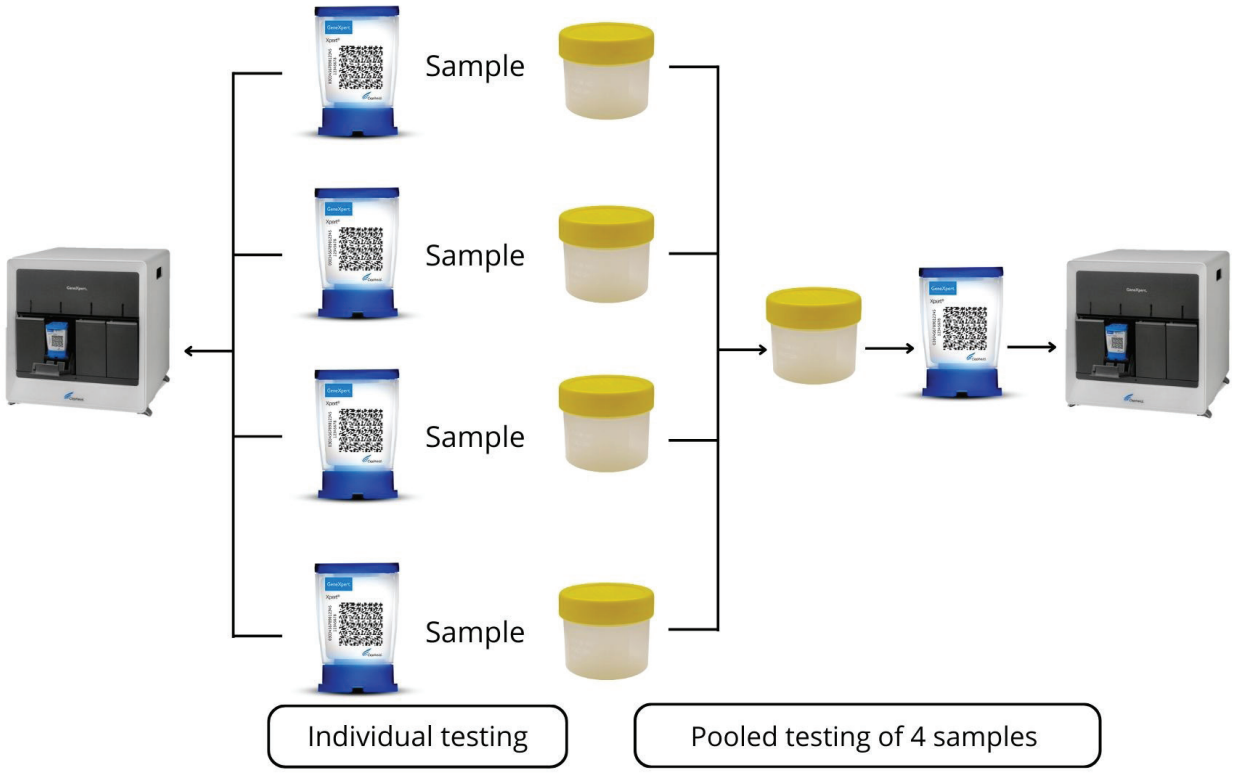


Figure 1. Innovative method of accelerating new TB cases-finding using the pooling sputum method (4 samples for 1 cartridge, a total of 16 samples for 1 test using Xpert MTB/RIF Ultra). *This figure illustrates the implementation of the pooling sputum method, wherein four individual sputum samples are combined into a single pooled sample and tested using one Xpert MTB/RIF Ultra cartridge. The diagram demonstrates how 16 individual sputum samples can be grouped into four pooled tests, thus increasing screening throughput while conserving testing resources. This strategy enhances case-finding efficiency, particularly in high-risk populations, and optimizes utilization of available molecular testing devices.*

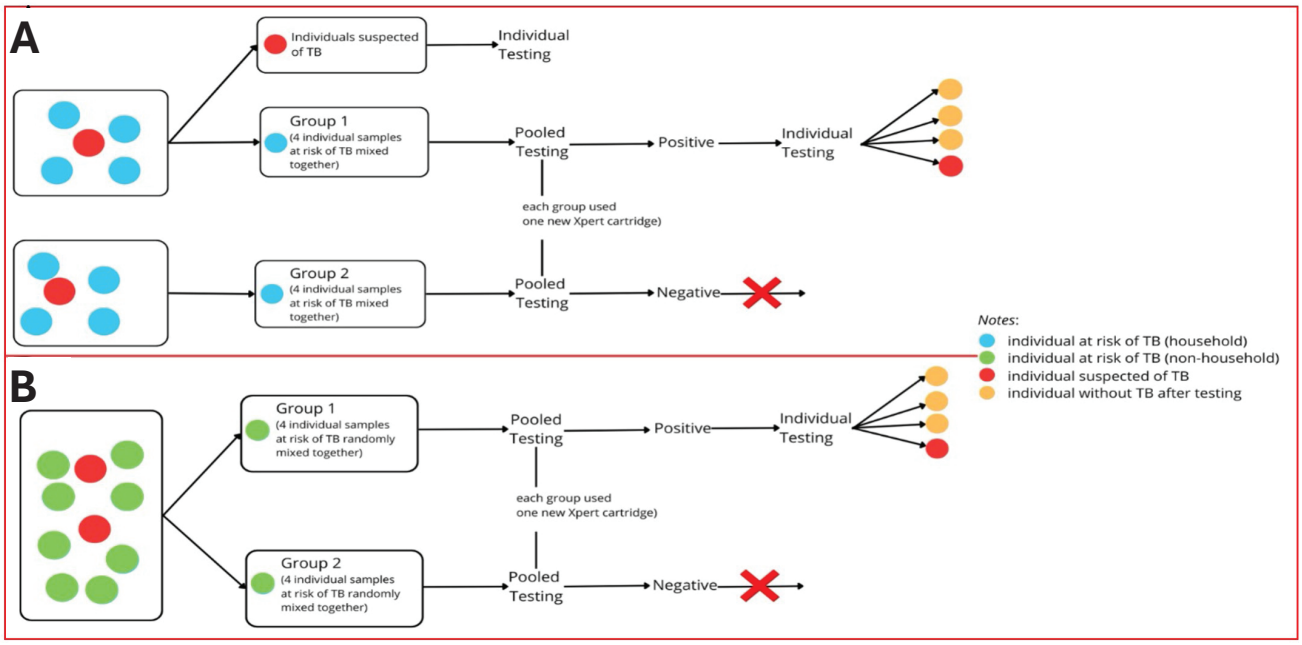


Figure 2. TB elimination strategy using screening of population groups at risk of TB. *The figure depicts the stratification of at-risk populations for tuberculosis screening using a pooling approach. A: household-based screening involves collecting sputum samples from individuals residing in households with a known TB-positive case, B: non-household-based screening targets community members or individuals with frequent contact outside the home, organized into risk groups. In both scenarios, sputum samples from four individuals are pooled for a single Xpert MTB/RIF Ultra test. If a pooled test yields a positive result, follow-up individual testing is conducted to identify the specific TB-positive case, thereby enabling efficient and targeted case detection.*

The workflow typically follows a two-step process (Figure 2). In Part A, two groups of households with TB-positive family members (Red Round) are grouped into two TB-risk population groups (Blue Round). If one of the groups tested TB-negative, it was not continued to individual testing. However, if the group is positive, then proceed to test each individual in the TB-positive group until finding an individual with TB. In Part B, the population at risk of TB (non-household) was randomly grouped into two groups. This approach ensures that resources for individual testing are only used on high-probability groups, which were defined according to WHO requirements. Next, the screening strategy was carried out in the same way as Part A to obtain suspected TB individuals.

Diagnostic accuracy

A primary concern for any screening strategy is its diagnostic accuracy. The pooling sputum method must be able to reliably detect TB without generating an unacceptable number of false negatives or false positives. Based on the literature review, the pooling method, when paired with the high-sensitivity Xpert MTB/RIF Ultra assay, maintains excellent diagnostic performance.

A summary of findings from key studies that met the inclusion criteria is presented in Table 1. Through these studies, which collectively analyzed hundreds of pooled samples, the method demonstrated a high probability of correctly identifying positive cases. The pooled sensitivity was calculated to be 97.35% (95% CI: 94.85-98.85%), with a specificity of 99.25% (95% CI: 95.91-99.98%). This means it correctly detected over 97% of TB-positive individuals within the pools. The specificity was even higher at 99.25% (95% CI: 95.91-99.98%), indicating a very low rate of false-positive results (Table 1).[10,12-15]

Evidence for cost-effectiveness

Beyond accuracy, the primary driver for adopting pooling sputum is its potential for significant cost savings. This is particularly relevant for countries like Indonesia facing funding gaps for their TB programs.[7,8] Multiple studies conducted in diverse settings have quantified these economic benefits.

Table 2 summarizes cost-saving analyses from studies in Nigeria, Brazil, Lao, Cameroon, and Cambodia. The reported percentage of cost savings varied, ranging from a minimum of 12.4% in a Brazilian study to a maximum of 57%, also in Brazil.[10,12] On average, the implementation of a 4-sample pooling strategy resulted in a cost saving of 36.71% (±14.97%).[17,10,12,13,19,20]

The degree of cost savings is heavily influenced by the underlying prevalence of TB in the screened population, as illustrated in Table 3. In a low-prevalence community setting (11%), the strategy saved nearly 60% of testing costs.[18] However, in a higher-prevalence hospital setting (26%), where more pools test positive and require individual follow-up, the savings were much

lower at 4.7%. This indicates that the strategy is most economically advantageous for mass screening in general or high-risk communities rather than in settings where a high positivity rate is already expected.

Projected impact on screening capacity and time in Indonesia

Pooled sputum testing at a suitable ratio using Xpert MTB/RIF Ultra provides a rapid, efficient, and cost-effective method for active TB case finding among high-risk groups in low-incidence and high-incidence areas.[21,10] Implementation of the innovation strategy will affect the total sample size of each province and automatically improve screening capability with a four times larger sample size. Data on the improvement of each province's screening capability is presented in Figure 3. With the previous strategy, Riau Province had a screening capability of 11,040 tests per month and increased to 44,160 tests per month after implementing the innovative strategy.[1,3] Thus, the province with the highest TB prevalence in Indonesia can prevent transmission by increasing the utility of one RMT instrument to the maximum every month.

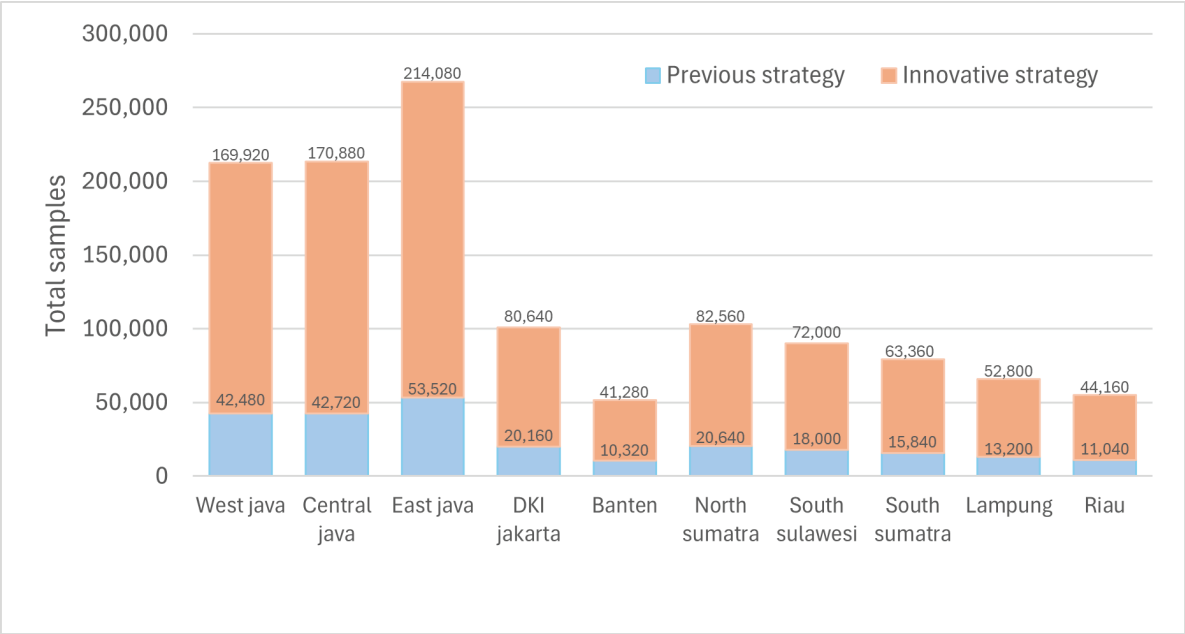


Figure 3. Comparison of the Number of samples potentially screened per one MRT per month between the previous strategy and the innovative strategy. *This figure compares the estimated number of individuals that can be screened per month using one Rapid Molecular Test (RMT) device under two strategies: the previous individual testing strategy versus the innovative pooling sputum method. The pooling approach demonstrates a fourfold increase in screening capacity per RMT unit, highlighting its potential to significantly expand diagnostic coverage in resource-constrained settings with high TB burdens.*

According to the Indonesian Ministry of Health (2022), an RMT instrument can only be used for 120 tests per month, using the previous strategy of screening individuals with suspected TB. With the implementation of the innovative strategy, the screening capacity will increase 4 times to 480 individuals per month.[1,2] Thus, 1 RMT device in good condition can screen 100,000 individuals within 8 months. When broken down further, each province can accelerate testing by 4 times to reach a total number of 100,000 individuals. In Riau Province, the implementation of the previous strategy took 9.1 months, then the innovative strategy can shorten the examination time to 2.3 months for 100,000 samples with 46 RMT devices. [1,3] In fact, the province with the highest TB prevalence in Indonesia was able to complete 100,000 tests in 0.6 months with the implementation of the innovative strategy (Figure 4).

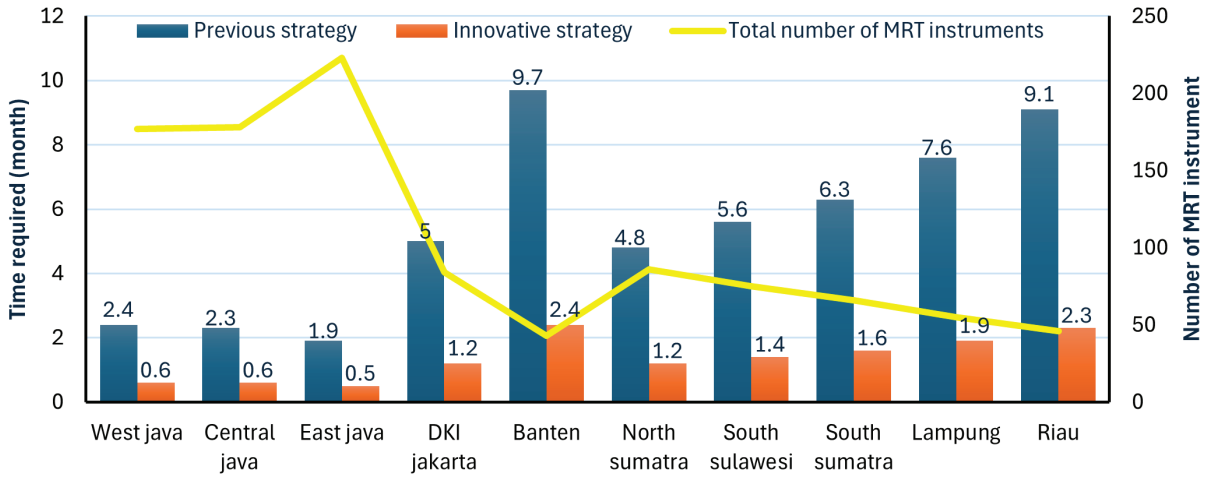


Figure 4. Comparison of the time required for each province based on the number of MRT instruments between the previous strategy and the innovative strategy. *This figure presents a comparative analysis of the time needed to screen 100,000 individuals across various provinces, based on existing RMT device availability. The innovative pooling sputum strategy dramatically reduces the time required for large-scale screening, particularly in high-burden provinces, thus enabling more rapid identification and treatment of active TB cases.*

Projected cost-effectiveness analysis for Indonesia

Based on the Technical Manual for Tuberculosis Testing Using GeneXpert Molecular Rapid Tests, the cost components of RMT testing are calculated based on the number of patients, including the cost of sending test samples to the laboratory and the cost of testing, with a total cost of IDR 50,000. [1,3] With the current screening strategy, the cost of conducting 100,000 tests is assumed to be IDR 5 billion. However, with the implementation of the proposed screening innovation strategy, the cost is IDR 1.25 billion per 100,000 screening tests. In addition, based on the theoretical study (Table 2), Indonesia can save costs with an average estimate of 36.71% by implementing the innovation strategy. If assumed based on the cost-saving average, the cost incurred is 1.84 billion per 100,000 tests.

Discussion

These performance metrics meet the minimum requirements set by the WHO for TB diagnostic tests, which call for at least 80% sensitivity and 97% specificity.[16] Even RT-PCR, the minimum proposed criteria have a sensitivity of 98.5% (95% CI: 97.5-99.5%) and specificity of 70% (95% CI: 65.8-74.2%). In fact, Sorsa and Kaso reported the sensitivity and specificity of GeneXpert to be 72-77% and 99% in smear-negative adults and 98-99% and 99-100% in smear-positive adults, respectively.[22] When compared to the previous 3 references, the pooling sputum method using Xpert MTB/RIF Ultra meets the specificity and sensitivity requirements to be implemented in the community. In fact, the specificity and sensitivity of the pooling sputum method are classified as very strong compared to Ag-RDT and RT-PCR in the COVID-19 case detection strategy.

Potential impact on TB case detection

The findings of this strategic review and analysis strongly suggest that pooling sputum is a transformative strategy for Indonesia's TB control program. By increasing screening capacity fourfold, this approach directly addresses the critical issue of RMT underutilization. This acceleration in case finding is crucial for breaking chains of transmission, as each active TB case can infect 10-15 others annually.

Under-reporting of TB is a major problem, as only 5.8 million individuals of the estimated 10 million individuals who developed TB were reported in 2020, and over 40% were missed by health services.[23] In Indonesia, the number of people suspected of having TB who did not undergo laboratory testing was 23% or 832,351 individuals.[1,2]

As explained in the feasibility and practicality section, the implementation of the innovative strategy can complete 100,000 tests with an average assumed time of 1.4 months per province, with the number of RMT instruments available. Compared to the previous strategy, which took 5.5 months per 100,000 tests per province. Thus, the implementation of the screening strategy using the pooling sputum method can save time by about 74.5%.

With the implementation of innovative strategies, individuals suspected of TB can be quickly identified in diverse and at-risk populations. Individuals found to be positive for TB in this screening method will be quickly treated, so that TB transmission can be suppressed in groups at risk of TB, such as family members, HIV sufferers, and others. By identifying suspected TB individuals, preventive measures can be taken for both suspected TB individuals and their populations.[24] The speed of screening in innovative strategies will increase the percentage of new case finding in Indonesia, which

is targeted by 2030. Improvements in RMT screening capabilities and new case finding will negatively correlate with the increase in TB cases in Indonesia.

Implications for health inclusivity and equity

In most low-income countries, direct sputum smear microscopy is the mainstay of TB diagnostics, as this test is inexpensive and highly specific, but it has low to moderate sensitivity. Conversely, sputum culture, in particular, automated liquid culture, is the most sensitive and specific diagnostic tool available for TB and facilitates drug susceptibility testing. However, culture requires a laboratory infrastructure, including biosafety equipment, not widely available in low-resource settings, and results typically take 2 to 6 weeks and, therefore, are rarely helpful for initial treatment decisions.[9] If sputum smear microscopy and culture are maintained to achieve 90% case finding in low-income countries, the speed of TB transmission will defeat this ambition. Therefore, the proposed innovation strategy that is fast and cheap can be a solution for low-income countries.

As this innovation strategy emphasizes screening using a pooling sputum method more quickly and economically, all levels of society have equal opportunity and a greater chance of detection. The innovation strategy can still be applied to both high and low prevalence population groups. There was no significant difference in the results of the two groups (Table 3).

In populations with the least number of devices, for example, North Kalimantan Province, which only has 14 devices, and Bengkulu, which has the lowest RMT utility, this innovation strategy can improve device performance 4 times faster and the number of tests 4 times more. A simulation of the improvement in time and number of tests in implementing the innovative strategy can be seen in Figures 3 and 4. Thus, equitable health services at every level of society can be achieved.

Limitations

The authors acknowledge several limitations. First, as this analysis relies on projections from secondary data, the actual operational performance may vary. Second, the cost-saving and diagnostic accuracy data were derived from studies conducted in other countries and require validation through a pilot study within the Indonesian healthcare system. Third, the analysis does not quantify the potential for reduced sensitivity due to sample dilution (the "dilution effect")[9], especially in cases with low bacterial loads. Finally, the logistical complexity of implementation, including staff training and sample tracking, presents an operational hurdle that requires careful planning.

Conclusion

This strategic review and cost-effectiveness analysis demonstrates that the implementation of a pooling sputum strategy holds substantial potential to overcome critical barriers in Indonesia's TB control program. Based on strong evidence from international literature, the method is highly accurate and cost-effective, particularly in lower-prevalence settings. The projected fourfold increase in screening capacity, significant time reduction for mass screening, and considerable cost savings present a compelling case for its adoption. For Indonesia, a country struggling with RMT underutilization and a significant TB burden, these findings are highly relevant. While logistical challenges must be addressed, the evidence strongly supports the consideration and piloting of pooling sputum as a core component of Indonesia's national TB control strategy. To translate these promising projections into practice, we strongly recommend conducting a pilot study in diverse Indonesian settings to validate the findings and assess operational feasibility for a national scale-up. Adoption of this innovative approach could be a pivotal step in accelerating Indonesia's progress toward its 2030 TB elimination goals.

Ethical approval

This article is an analysis based on publicly available data and did not involve the collection of primary data from human participants or animals. Therefore, ethical approval was not required.

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 conflicts of interest related to the content of this manuscript.

CRediT authorship contribution statement

Febri Sembiring: conceptualization, methodology, formal analysis, writing: original draft; **Hafizah Ilmi Sufa**: data curation, visualization; **Yuli Yantika Syahputri** and **Wardati Humaira**: validation, writing: original draft.

Acknowledgements

Special thanks to all professionals and researchers whose published work has supported the evidence-based foundation of this article.

References

[1] Indonesian Government, Ministry of Health. Tuberculosis control program report. 2022 Available from: https://tbindonesia.or.id/wp-content/uploads/2023/09/Laporan-Tahunan-Program-TBC-2022.pdf.

[2] Indonesian Government, Ministry of Health. Through the 4th INA - TIME 2022 Activity, Minister of Health Budi requests 90% of TB patients can be detected in 2030. 2023 Available from: http://p2p.kemkes.go.id/melalui-ina-time2022-ke-4-menkes-budi-minta-90-penderitatbc-dapat-terdeteksi-di-

tahun-2024.

[3] Indonesian Government, Ministry of Health. Technical manual for tuberculosis testing using GeneXpert molecular rapid test. 2023 Available from:https://tbindonesia.or.id/wp-content/uploads/2024/02/Buku-Petunjuk-Teknis-Pemeriksaan- TBC-Menggunakan-Alat-RMT-GeneXpert_2023.pdf.

[4] Ahmed N, Hasnain SE. Molecular epidemiology of tuberculosis in India: Moving forward with a systems biology approach. Tuberculosis. 2011; 91(5): 407-13. doi: 10.1016/j.tube.2011.03.006.

[5] McCreesh N, White RG. An explanation for the low proportion of tuberculosis that results from transmission between household and known social contacts. Sci Rep. 2018; 8: 5382. doi: 10.1038/s41598-018-23797-2.

[6] World Health Organization. Tb reaches the 3 million find treat cure TB. 2014. Available from: http://www.stoptb.org/assets/documents/resources/publications/acsm/WORLD_TB_DAY_BROCHURE_14March.pdf.

[7] Indonesian Government, The House of Representatives. DPR members appreciate increased budget for TB control. 2021. Available from: https://www.dpr.go.id/berita/detail/id/29969/t/Anggota+DPR+Appreciates+Increased+Budget+for+TB+Control.

[8] The Conversation. Indonesia needs to triple funding to control tuberculosis - Here's a start. 2023. Available from: https://theconversation.com/indonesia-perlu-tingkatkan-pendanaan-tiga-kali-lipat-untuk-kendalikan-tuberkulosis-inilah-awal-yang-harus-dilakukan-216266.

[9] Cuevas LE, Santos VS, Almeida LSVM, et al. Systematic review of pooling sputum as an efficient method for Xpert MTB/RIF tuberculosis testing during the COVID-19 pandemic. Emerg Infect Dis. 2021; 27(3): 1-9. doi: 10.3201/eid2703.204090.

[10] Santos VS, Allgayer MF, Kontogianni K, Rocha JE, et al. Pooling of sputum samples to increase tuberculosis diagnostic capacity in Brazil during the COVID-19 pandemic. Int J Infect Dis. 2023; 129: 10-4. doi: 10.1016/j.ijid.2023.01.009.

[11] Zishiri V, Chihota V, McCarthy K, Charalambous S, Churchyard GJ, Hoffmann CJ. Pooling sputum from multiple individuals for Xpert® MTB/ RIF testing: A strategy for screening high-risk populations. Int J Tuberc Lung Dis. 2015; 19(1): 87-90. doi: 10.5588/ijtld.14.0372.

[12] Dos Santos PC P, da Silva Santos A, de Oliveira RD, et al. Pooling sputum samples for efficient mass tuberculosis screening in prisons. Clin Infect Dis. 2022; 74(12): 2115-21. doi: 10.1093/cid/ciab847.

[13] Iem V, Chittamany P, Suthepmany S , et al. Pooled testing of sputum with Xpert MTB/RIF and Xpert Ultra during tuberculosis active case finding campaigns in Lao People's Democratic Republic. BMJ Global Health. 2022; 7(2): e007592. doi: 10.1136/bmjgh-2021-007592.

[14] Phuong NTB, Anh NT, Van SN, et al. Effect of two alternative methods of pooling sputum prior to testing for tuberculosis with Genexpert MTB/RIF. BMC Infect Dis. 2019; 19: 347. doi: 10.1186/s12879-019-3778-9.

[15] Zeng J, Huang H, Liu X. et al. Pooling sputum samples for the Xpert MTB/RIF assay: A practical screening strategy for highly infectious tuberculosis cases. BMC Infect Dis. 2024; 24: 122. doi: 10.1186/s12879-024-09020-w.

[16] World Health Organization. Antigen-detection in the diagnosis of SARS-CoV-2 infection: Emergencies preparedness, WHO Headquarters (HQ). 2021. Available from: https://www.who.int/publications/i/item/antigen-detection-in-the-diagnosis-of-sars-cov-2infection-using-rapid-immunoassays.

[17] Iem V, Bimba JS, Santos VS. et al. Pooling sputum testing to diagnose tuberculosis using Xpert MTB/RIF and Xpert Ultra: A cost-effectiveness analysis. BMC Infect Dis. 2023; 23: 341. doi: 10.1186/s12879-023-08330-9.

[18] Abdurrahman ST, Mbanaso O, Lawson L, et al. Testing pooled sputum with Xpert MTB/RIF for diagnosis of pulmonary tuberculosis to increase affordability in low-income countries. J Clin Microbiol. 2015; 53(8): 2502-8. doi: 10.1128/JCM.00864-15.

[19] Vuchas C, Teyim P, Dang BF, et al. Implementation of large-scale pooled testing to increase rapid molecular diagnostic test coverage for tuberculosis: a retrospective evaluation. Sci Rep. 2023; 13(1): 15358. doi: 10.1038/s41598-023-41904-w.

[20] Chry M, Smelyanskaya M, Ky M, et al. Can the high sensitivity of Xpert MTB/RIF Ultra be harnessed to save cartridge costs? Results from a pooled sputum evaluation in Cambodia. Trop Med Infect. 2020; 5 : 27. doi: 10.3390/tropicalmed5010027.

[21] Ma Z, Zhang L, Li S, et al. Pooled sputum testing by Xpert® MTB/RIF Ultra for active tuberculosis case finding among high-risk groups in a low-incidence area: A prospective study. Infect Dis. 2024; 56 (6): 434-40. doi: 10.1080/23744235.2024.2320324.

[22] Sorsa A, Kaso M. Diagnostic performance of GeneXpert in tuberculosis-HIV Co-infected patients at Asella Teaching and Referral Hospital, Southeastern Ethiopia: A cross sectional study. PLoS ONE. 2021; 16(1): e0242205. doi: 10.1371/journal.pone.0242205.

[23] World Health Organization. Global tuberculosis report 2021. 2022. Available from: https://www.who.int/publications/i/item/9789240037021.

[24] World Health Organization. Tuberculosis key facts. 2022. Available from: https://www.who.int/indonesia/news/campaign/tb-day-2022/fact-sheets.

[25] Abdurrahman ST, Mbanaso O, Lawson L, et al. Testing Pooled Sputum with Xpert MTB/RIF for diagnosis of pulmonary tuberculosis to increase affordability in low-income countries. J Clin

Microbiol. 2015; 53(8): 2502-8. doi: 10.1128/JCM.00864-15.

[26] Vuchas C, Teyim P, Dang BF, et al. Implementation of large-scale pooled testing to increase rapid molecular diagnostic test coverage for tuberculosis: a retrospective evaluation. Sci Rep. 2023; 13(1): 15358. doi: 10.1038/s41598-023-41904-w.

[27] Chry M, Smelyanskaya M, Ky M, et al. Can the High Sensitivity of Xpert MTB/RIF Ultra Be Harnessed to save cartridge costs? Results from a pooled sputum evaluation in Cambodia. Trop Med Infect. 2020; 5: 27. doi: 10.3390/tropicalmed5010027

[28] Ma Z, Zhang L, Li S, , et al. Pooled sputum testing by Xpert® MTB/RIF Ultra for active tuberculosis case finding among high-risk groups in a low-incidence area: A prospective study. Infect Dis. 2024; 56 (6): 434-440. doi: 10.1080/23744235.2024.2320324.

[29] Sorsa A, Kaso M. Diagnostic performance of GeneXpert in tuberculosis-HIV co-infected patients at Asella Teaching and Referral Hospital, Southeastern Ethiopia: A cross sectional study. PLoS ONE. 2021; 16(1): e0242205. doi: 10.1371/journal.pone.0242205.

[30] World Health Organization. Global tuberculosis report 2021. 2022. Available from: https://www.who.int/publications/i/item/9789240037021.

[31] World Health Organization. Tuberculosis key facts. 2022. Available from: https://www.who.int/indonesia/news/campaign/tb-day-2022/fact-sheets.

Journal of Associated Medical Sciences 2026; 59(1): 201-209

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 automated flow cytometer single-platform for absolute CD4⁺ T-lymphocytes enumeration in HIV patients, Trat Province, Thailand

Sirichai Pookkapund

Division of Virology, Medical Technology Department, Trat Hospital, Trat Province, Thailand.

ARTICLE INFO

Article history:
Received 11 February 2025
Accepted as revised 25 November 2025
Available online 11 December 2025

Keywords:
CD4⁺ T-lymphocyte, single-platform, dual-platform, flow cytometer.

ABSTRACT

Background: The routine analysis of CD4⁺ T-lymphocyte percentages and absolute counts in HIV patients commonly employs the dual-platform method (blood cell analyzer and flow cytometer). However, variability related to equipment, methodology, or operator performance may affect accuracy. This study introduces the single-platform method, which relies solely on a flow cytometer to reduce variability, streamline workflow, and shorten turnaround times, in alignment with the Rational Laboratory Use guidelines of the Department of Medical Sciences, Ministry of Public Health.

Objectives: To evaluate and compare the single-platform and dual-platform methods for determining CD4⁺ T-lymphocyte percentages and absolute counts.

Materials and methods: The study was conducted from January 24 to November 29, 2024, encompassing the entire research process — from conceptual development, problem analysis, and study design to data collection, statistical analysis, interpretation, and application of findings. Routine diagnostic data were collected from HIV-infected patients in Trat Province, Thailand, between April 1 and July 31, 2024, as part of the data collection phase. Samples were analyzed using a semi-automated blood cell analyzer and a flow cytometer, and results were compared with those obtained using the single-platform method, which employed only the flow cytometer. Statistical analyses were performed to assess the accuracy, precision, and reliability of the single-platform method compared with the conventional approach. Descriptive statistics, correlation, linear regression, and Bland–Altman analysis were applied to evaluate agreement and systematic bias. All statistical tests were conducted at a 95% confidence interval, with *p*-values <0.05 considered statistically significant.

Results: Both methods produced data that followed a normal distribution (Kolmogorov–Smirnov test; $p>0.05$). Correlation coefficients demonstrated excellent agreement for CD4⁺ T-lymphocyte percentages (r=0.9914) and absolute counts (r=0.9697). Linear regression analysis showed a strong association, with r^2=0.9403 for absolute counts. Bland–Altman analysis indicated a mean difference of 61.06 cells/µL (95% CI: -73.91 to 196.04), with most values falling within the confidence limits. Among patients with CD4⁺ T-lymphocyte percentages ≤ 20%, the mean difference was 28.56 cells/µL (95% CI: -61.98 to 119.09).

Conclusion: The single-platform method is comparable to the dual-platform method for analyzing CD4⁺ T-lymphocytes in HIV patients. Both methods demonstrated normal data distribution, confirming statistical robustness. High correlations for percentages and absolute counts ensured consistent and reliable results, while regression analysis indicated strong predictive capability of the single-platform method for dual-platform results. Bland–Altman analysis further confirmed the equivalence of the two methods,

***Corresponding contributor.**
Author's Address: *Division of Virology, Medical Technology Department, Trat Hospital, Trat Province, Thailand.*
E-mail address: Sirichai_p@trathospital.go.th
doi: 10.12982/JAMS.2026.025
E-ISSN: 2539-6056

supporting the reliability of the single-platform approach. Overall, the single-platform method offers a reliable and efficient alternative, reducing variability, workload, and turnaround time in laboratory settings while maintaining analytical accuracy.

Introduction

The Joint United Nations Programme on HIV/AIDS (UNAIDS) has urged nations to intensify efforts to end the HIV/AIDS epidemic, following concerns that HIV services had been neglected during the COVID-19 crisis. If this trend continues, it could result in an estimated 7.7 million deaths from HIV/AIDS over the next decade. In Thailand, during 2017–2018 and amidst the COVID-19 outbreak, an average of 16 new HIV infections per day was reported, largely due to reduced testing and a decline in the number of patients receiving antiretroviral therapy. This rate remains below the UNAIDS target of eliminating AIDS by 2030. Despite this decline, UNAIDS remains concerned about achieving the goal.[1] In 2023, Thailand reported approximately 436,170 people living with HIV, 2,503 AIDS-related deaths, and 4,148 new HIV cases. Key affected populations include men who have sex with men, transgender individuals, people who inject drugs, and individuals engaging in unprotected sexual activities. In Health Region 6, about 50,090 people were living with HIV, with 655 new cases and 1,458 AIDS-related deaths. Chonburi Province reported the highest number of new cases (~360) and the largest number of people living with HIV (~18,226). In comparison, Trat Province reported 18 new HIV cases and 2,302 people living with HIV.[2] These statistics underscore the need for ongoing diagnostic testing and treatment.

$CD4^+$ T-lymphocyte testing measures both the absolute count and the percentage of $CD4^+$ T-lymphocytes in blood. Flow cytometry has been developed as a diagnostic tool for determining $CD4^+$ T-lymphocyte counts.[3] This system functions by detecting cells in suspension as they pass through a laser beam, where light scattering and fluorescence signals emitted by labeled cells are detected by sensors and analyzed by computer systems.[4] Absolute $CD4^+$ T-lymphocyte counts are critical for disease prognosis, clinical decision-making, and monitoring antiretroviral therapy (ART) outcomes.[5] Current measurements typically require data from both flow cytometers and blood cell analyzers. The latter provides complete blood count (CBC) parameters, including white blood cell (WBC) count and lymphocyte percentage (% lymphocyte), which are incorporated into $CD4^+$ percentages calculations.[6] However, reliance on two instruments often leads to delays, particularly during urgent testing or when CBC workloads are high. Accuracy may also be compromised by differences in analyzer models or testing methodologies. Testing costs are approximately 23 THB per CBC (government charge: 130 THB) and 125 THB for $CD4^+$ testing (government charge: 900 THB).[7,8] Physicians requiring both absolute and percentage $CD4^+$ T-lymphocyte results must therefore request additional CBC tests, unnecessarily increasing costs.

International studies have explored alternative methodologies for calculating absolute $CD4^+$ T-lymphocyte counts without reliance on CBC results. The single-platform method, which relies solely on flow cytometry, has been introduced as an alternative to the conventional dual-platform approach. This method offers advantages such as simplified workflows, reduced costs, and minimized errors associated with multi-instrument dependency.[9]

In Thailand, research institutions and major hospitals equipped with flow cytometers have begun evaluating single-platform and dual-platform methods.[10] These assessments focus on accuracy, precision, processing efficiency, and cost-effectiveness. The single-platform method employs polymer beads of known size and concentration, which are mixed with patient samples to determine $CD4^+$ percentages. Through light scattering and fluorescence detection, flow cytometers directly measure $CD4^+$ T-lymphocyte counts without reliance on WBC counts and % lymphocytes, thereby streamlining the analytical process.[11] Successful implementation of single-platform testing requires skilled personnel, high-precision equipment, and meticulous sample preparation to ensure accurate and reliable results.[12] Although both international and domestic studies indicate promising outcomes, single-platform methods are not yet widely accepted due to limited research validating their reliability. Given these potential benefits, the present study aims to compare single-platform and dual-platform methods for measuring both absolute counts and percentages of $CD4^+$ T-lymphocytes.

Materials and methods

Sample collection and preparation

A total of 865 whole blood samples (3 mL, EDTA-anticoagulated) were collected from HIV-infected patients in Trat Province between April 1, 2024, and July 31, 2024, for the analysis of $CD4^+$ T-lymphocyte percentages and absolute counts. All sample collection and analyses were conducted at the Division of Virology, Department of Medical Technology, Trat Hospital. White blood cell counts and lymphocyte percentages were obtained using a semi-automated hematology analyzer UniCel DxH 900 (Beckman Coulter, USA) while $CD4^+$ T-lymphocyte counts and percentages were measured with a flow cytometer Cytomics FC500 (Beckman Coulter, USA). All analytical procedures were validated through rigorous quality control measures.

For immune-phenotyping, each EDTA-treated whole blood sample was stained with the monoclonal antibody CYTO-STAT® triCHROME™ CD45-FITC/CD4-RD1/CD3-PC5 (Beckman Coulter, USA). The immunophenotyping procedure involved mixing 10µL of antibody with 100µL of whole blood in a 12×75 mm plastic tube,

followed by thorough mixing using a vortex mixer. The stained samples were then processed using the TQ-Prep™ Workstation and the IMMUNOPREP Reagent System (Beckman Coulter, USA) for automated sample preparation. Subsequently, the prepared samples were analyzed by flow cytometry using two distinct analytical methods.

Method 1: Dual-platform technique

A total of 100µL of EDTA-treated whole blood was stained with 10µL of CYTO-STAT® triCHROME™ reagent (Beckman Coulter, USA). The samples were incubated for 10 minutes at room temperature (20–25°C) in the dark to allow adequate antibody binding, ensuring reproducibility and standardization. Red blood cells were lysed using the lyse–no-wash protocol with the TQ-Prep™ Workstation. The prepared samples were then analyzed on a flow cytometer Cytomics FC500 (Beckman Coulter, USA). The gating strategy for identifying lymphocyte populations and $CD4^+$ T-lymphocyte subsets is shown in Figure 1.

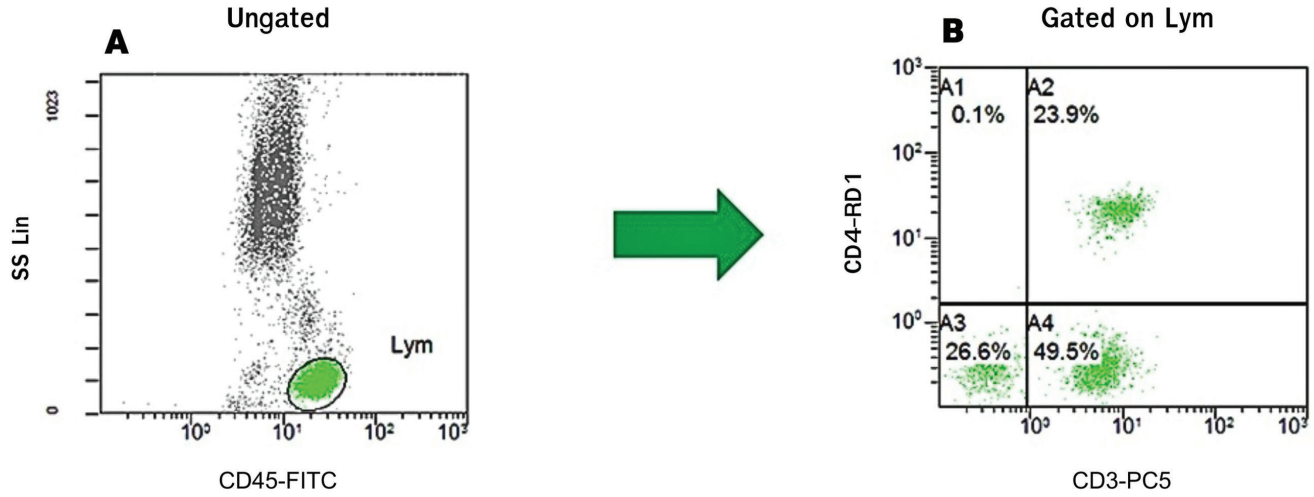


Figure 1. *CD45/SSC and CD3/CD4-based gating strategy for dual-platform $CD4^+$ T-lymphocyte enumeration. A: lymphocyte populations were gated by plot of CD45 against SSC. Lymphocytes are low side scatter and have high CD45, B: $CD4^+$ T-lymphocyte (%) in lymphocyte populations were analyzed by plotting CD3 against CD4. $CD4^+$ T-lymphocyte are positive with CD3 and CD4 (A2 quadrant). Absolute $CD4^+$ T-lymphocyte count derived from % $CD4^+$ T-lymphocyte x %lymphocyte x WBC count/10,000 (%$CD4^+$ T-lymphocyte from flow cytometer, %lymphocyte and WBC count from hematology analyzer).*

Method 2: Single-platform technique

CD4⁺ T-lymphocytes were stained using the same protocol as in the dual-platform technique. After red blood cell lysis, 100 µL of Flow-Set™ Fluorospheres (Beckman Coulter, USA) was added to an equal volume (100µL) of blood sample. The prepared samples were then analyzed on a flow cytometer Cytomics FC500 (Beckman Coulter, USA). The bead-based gating strategy used for absolute CD4⁺ T-lymphocyte enumeration is presented in Figure 2.

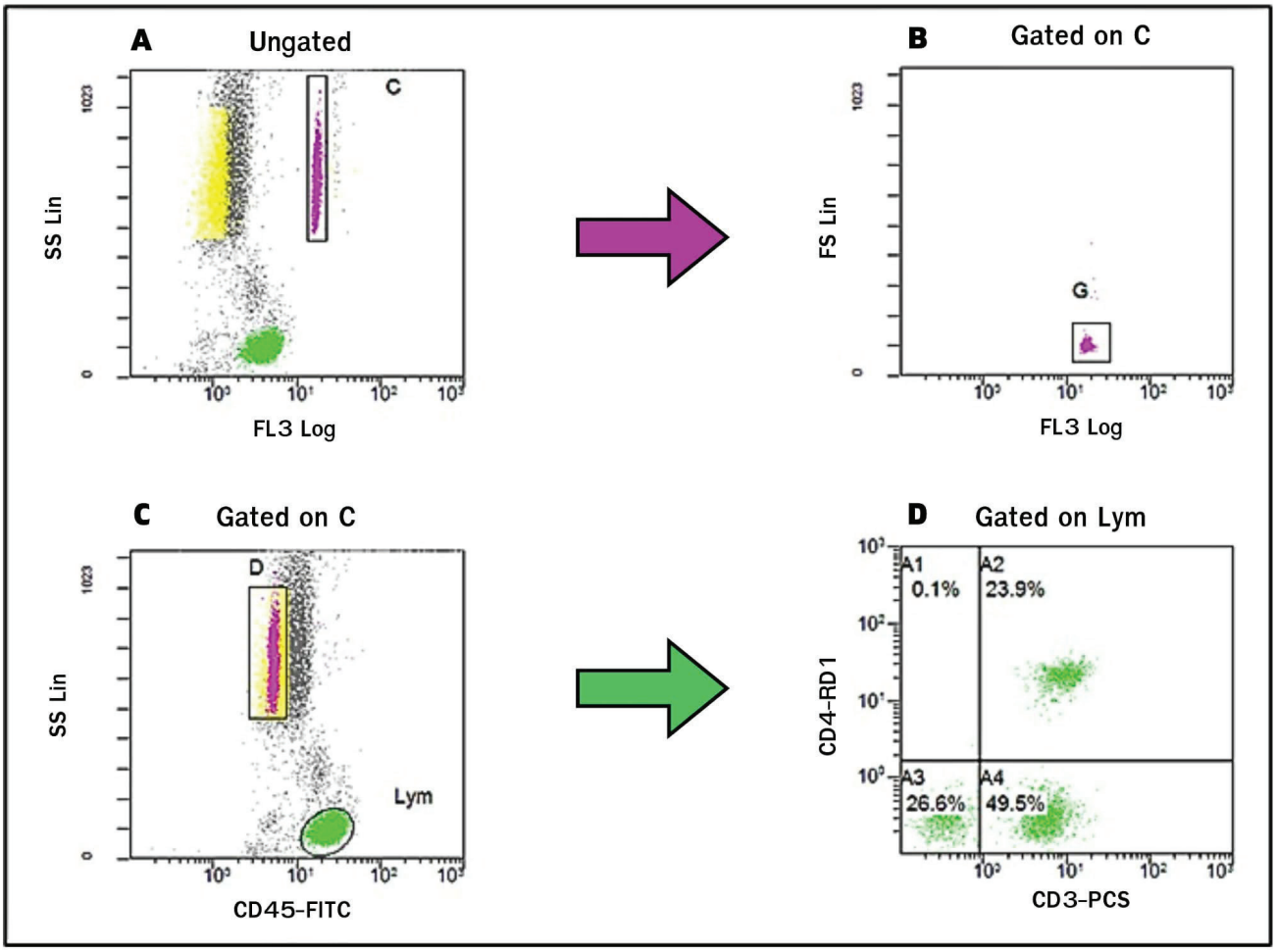


Figure 2. Bead-based gating procedure for absolute CD4⁺ T-lymphocyte enumeration using Flow-Set™ Fluorospheres. *The beads contain fluorescent polystyrene particles that emit at 525-700 nm when excited at 488 nm. The fluorescence signal was detected through the FL3 (ECD) channel using a 610-nm long-pass filter. A: Flow-Set™ Fluorospheres were primary gated by plot of SSC against FL3 (gate C). The fluorospheres from cell events could be separatedroughly, B: only singlet beads were gated from gate C events by plot of FSC against FL3 (gate G), C: lymphocyte populations were gated by plotting CD45 against SSC. In gate D, the pink events are Flow-Set™ Fluorospheres. (D) %CD4⁺ T-lymphocyte cell events in A2 quadrant were interested. Absolute CD4⁺ T-lymphocyte count derived from (CD4⁺ T-lymphocyte events x bead concentration) / bead events.*

Data analysis

Statistical analyses were conducted using Microsoft Excel 2013. For each variable, the mean and SD were calculated. Normality of data distribution was assessed using the Kolmogorov-Smirnov test prior to performing correlation and linear regression analyses. The correlation between absolute CD4⁺ T-lymphocyte counts and percentages obtained from the dual-platform and single-platform methods was evaluated using Pearson's correlation coefficient (r). Linear regression analysis was employed to examine the linear relationship between the absolute CD4⁺ T-lymphocyte counts obtained from both methods. Additionally, Bland-Altman plots were used to assess the agreement between the two methods. The results were presented as the mean bias and limits of agreement (LOA) to determine whether the methods could be used interchangeably.

Results

The percentage and absolute counts of CD4⁺ T-lymphocytes were measured in 865 blood samples from patients using the single-platform volumetric flow cytometry method with the FC-500 flow cytometer. The results were then compared with those obtained using

the standard dual-platform flow cytometry method in conjunction with the semi-automated blood cell analyzer UniCel DxH 900. The findings of this study are presented as follows.

Data distribution testing

The results of the data distribution analysis are presented in Table 1.

As summarized in Table 1, all parameters followed a normal distribution (Kolmogorov-Smirnov test, $p>0.05$). A normality distribution analysis was conducted on the percentage and absolute values of $CD4^+$ T-lymphocytes obtained from HIV patients in Trat Province who submitted samples to the Medical Technology Department at Trat Hospital. The Kolmogorov-Smirnov (K-S) test was applied to assess data normality. The K-S test values for single-platform $\%CD4^+$, dual-platform $\%CD4^+$, single-platform absolute $CD4^+$ T-lymphocyte count, and dual-platform absolute $CD4^+$ T-lymphocyte Count were 0.153, 0.109, 0.308, and 0.207, respectively. All values exceeded the critical threshold at the 0.05 significance level, confirming that the data for all four parameters followed a normal distribution.

Table 1. Statistical analysis and Kolmogorov–Smirnov test of samples.

Measurement parameter	N	Mean	SD	Kolmogorov–Smirnov test (K-S Test)	*p* value (N>40)
Single-platform $\%CD4^+$	865	25.0229	8.9762	0.153	0.05
Dual-platform $\%CD4^+$	865	25.0090	9.0796	0.109	0.05
Single-platform absolute $CD4^+$ count	865	490.00	241.46	0.308	0.05
Dual-platform absolute $CD4^+$ count	865	551.18	269.69	0.207	0.05

The correlation analysis

Correlation analysis was conducted to evaluate the relationship between single- platform $\%CD4^+$, dual-platform $\%CD4^+$, single-platform absolute $CD4^+$ T-lymphocyte count, and dual-platform absolute $CD4^+$ T-lymphocyte Count. The correlation coefficients (r) obtained from the statistical tests are presented in Table 2.

Table 2. Pearson correlation coefficient analysis of single and dual-platform methods.

Variable	Single-platform $\%CD4^+$	Single-platform absolute $CD4^+$
Single-platform $\%CD4^+$	1.0000	
Single-platform absolute $CD4^+$	0.7315	1.0000
Dual-platform $\%CD4^+$	0.9914	0.7233
Dual-platform absolute $CD4^+$	0.7430	0.9697

The statistical analysis demonstrated a strong correlation between the two methods for measuring the percentage of $CD4^+$ T-lymphocytes, with a correlation coefficient (r) of 0.9914. Similarly, the analysis of absolute $CD4^+$ T-lymphocyte counts using both methods exhibited a very high correlation, yielding a correlation coefficient (r) of 0.9697. The correlation coefficients are presented in Table 2 and illustrated in Figure 3**,** showing excellent agreement between single- and dual-platform methods.

The linear regression

Statistical analysis using linear regression demonstrated a strong correlation between the single-platform absolute $CD4^+$ T-lymphocyte count and the dual-platform absolute $CD4^+$ T-lymphocyte Count, with an R^2 value of 0.9403. The regression analysis demonstrated a strong relationship between the two methods, as illustrated in Figure 3B.

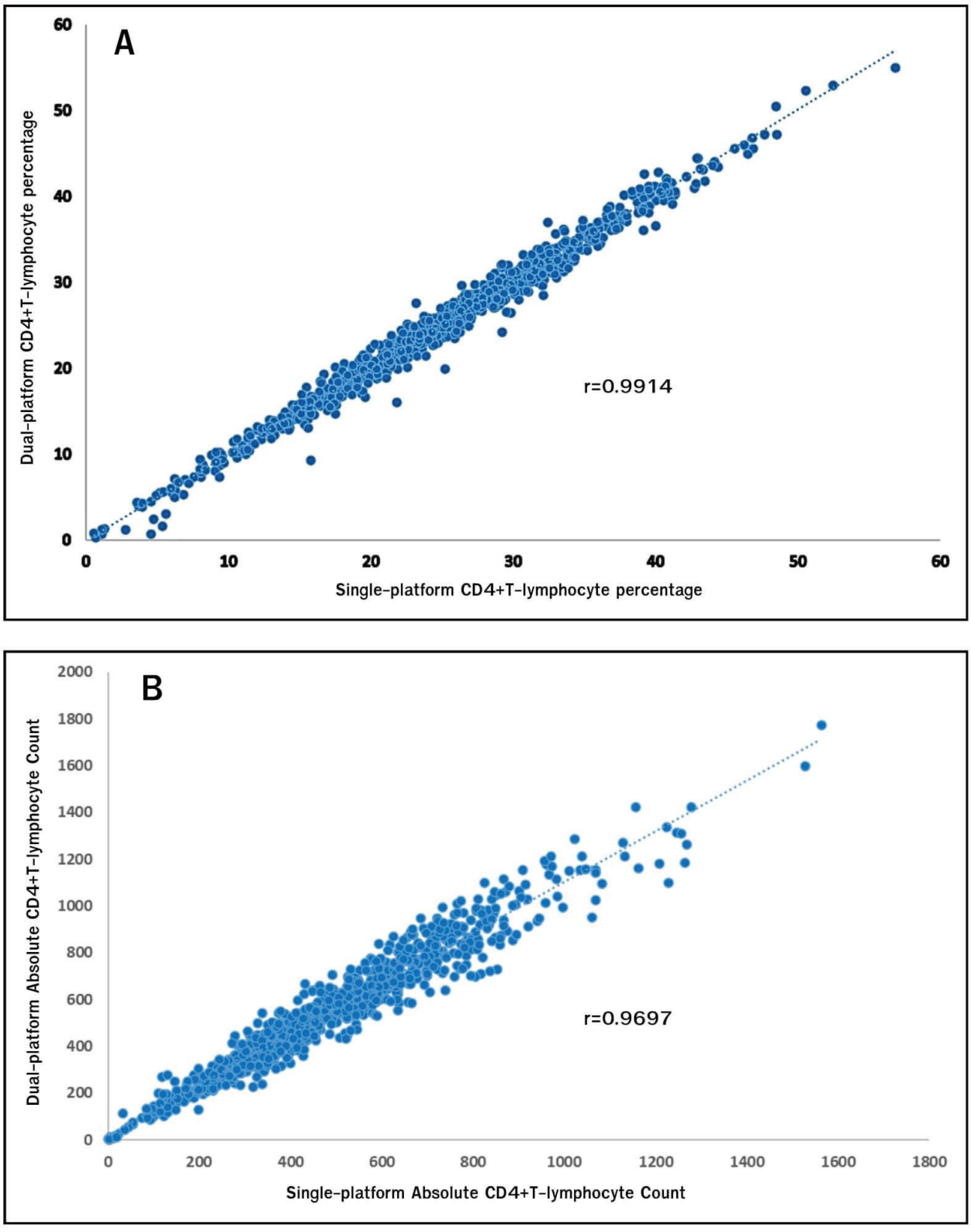


Figure 3. ***Correlation analysis of CD4⁺ T-lymphocyte values obtained using the single-platform and dual-platform techniques.*** *A: scatter plot illustrating the correlation between CD4⁺ T-lymphocyte percentages measured by both methods, B: scatter plot depicting the relationship between absolute CD4⁺ T-lymphocyte counts from the two analytical approaches.*

Bland-Altman plot analysis

Analysis of the single-platform and dual-platform absolute CD4⁺ T-lymphocyte counts using the Bland-Altman method revealed that the mean differences for both the percentage and absolute values of CD4⁺ T-lymphocytes were predominantly within the 95% confidence interval (95% CI). For the absolute CD4⁺ T-lymphocyte count, the mean difference was 61.06 cells/μL, with a 95% CI ranging from -73.91 to 196.04 cells/μL. The Bland-Altman plot in Figure 4 confirms that most values lie within the 95% confidence interval, with a mean bias of 61.06 cells/μL.

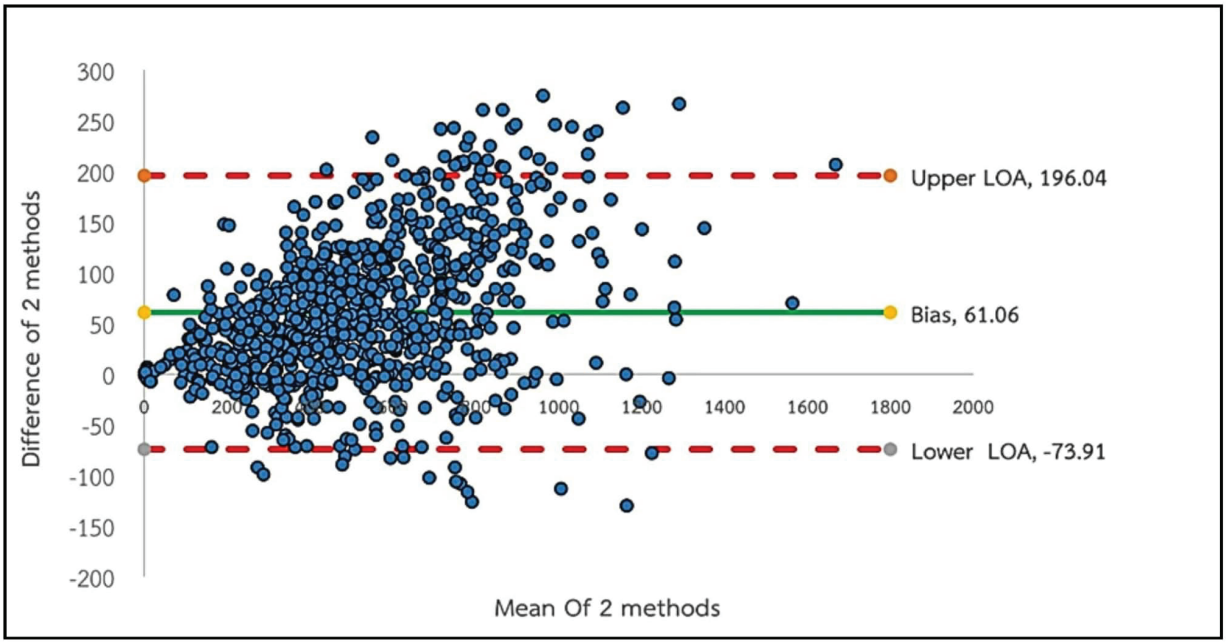


Figure 4. *Bland-Altman plot diagram analysis of absolute CD4⁺ T-lymphocyte count from single and dual-platform methods (N=865).*

When considering the analysis of absolute CD4⁺ T-lymphocyte count from a sample of 246 participants with a CD4⁺ percentage ≤20%, who represent a target group requiring close monitoring and timely treatment, using a Bland-Altman plot, the mean difference was found to be 28.56 cells/μL (95% CI: -61.98 to 119.09) For patients with CD4⁺ T-lymphocyte percentages ≤20%, the Bland-Altman analysis (see Figure 5) revealed a smaller mean difference of 28.56 cells/μL, with most results remaining within acceptable limits when compared to the normal CD4⁺ T-lymphocyte range (500-1500 cells/μL).[13]

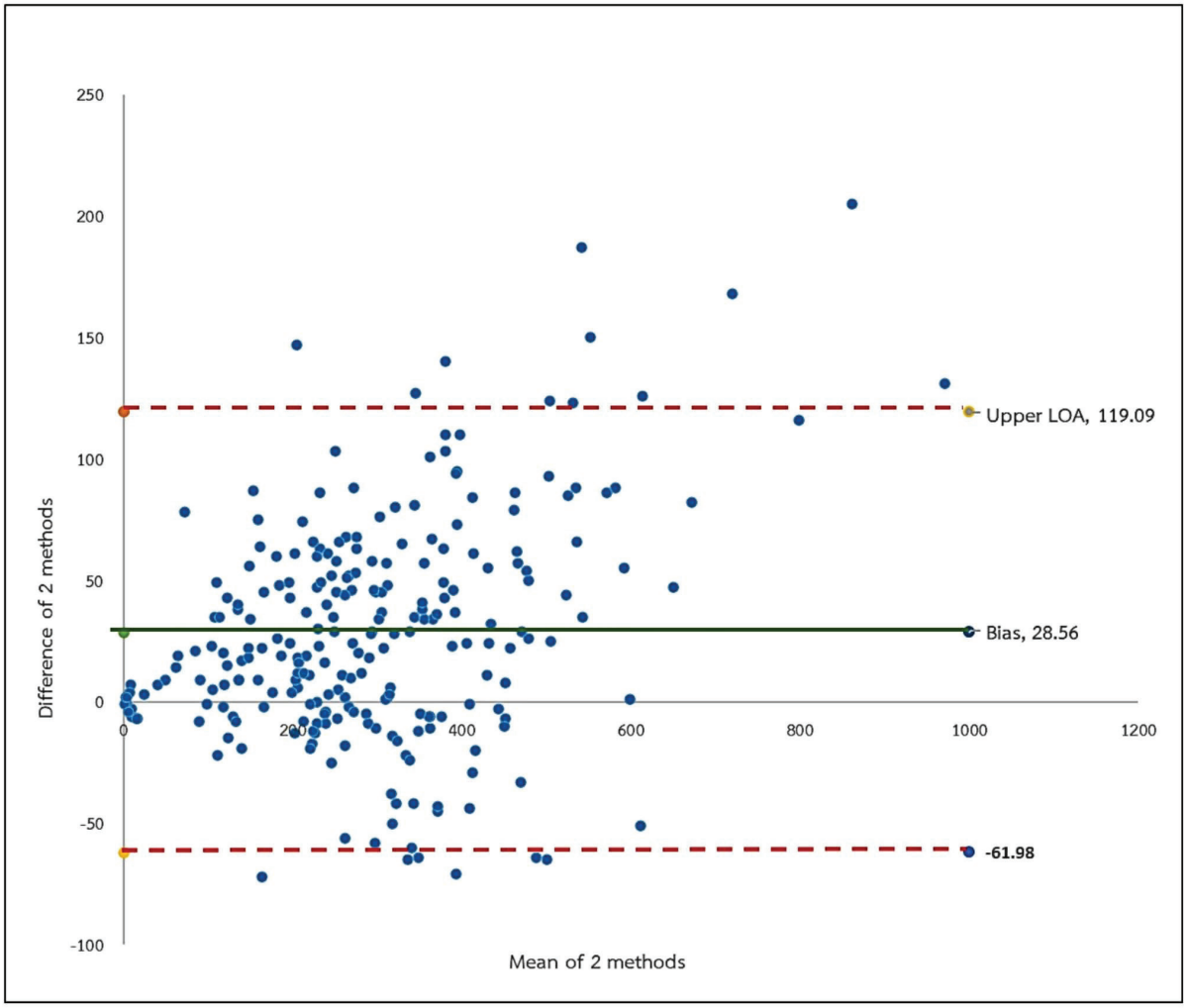


Figure 5. *Bland-Altman plot diagram of absolute CD4⁺ T-lymphocyte count analysis from single and dual-platform methods when %CD4⁺ ≤20% (N=246).*

Discussion

In Thailand, CD4⁺ T-lymphocyte enumeration remains a cornerstone of HIV patient management, guiding decisions on initiating antiretroviral therapy and preventing opportunistic infections. The conventional dual-platform technique, which integrates results from hematology analyzers and flow cytometers, has long been regarded as the standard approach. However, its reliance on multiple instruments, high costs, the requirement for skilled personnel, and stringent quality control procedures limit its accessibility in smaller hospitals and resource-limited laboratories. Moreover, variability arising from differences in reagents, instrument models, calibration methods, or operator performance can further undermine result reliability.

The single-platform method was evaluated in this study as a practical and efficient alternative. By relying solely on flow cytometry, this approach eliminates the need for complete blood count testing, thereby streamlining workflow, reducing analytical variability, shortening turnaround times, and lowering laboratory workload while maintaining high accuracy. International experience with volumetric flow cytometry supports its reliability, and in this study, polymer beads were used as standardized calibration materials. Their defined size and concentration served as a consistent internal reference, enhancing precision and reproducibility in CD4⁺ T-lymphocyte enumeration.

From an economic perspective, the single-platform method offers distinct advantages. The estimated cost per test was approximately 125 THB, compared to 148 THB for the dual-platform approach, representing a cost reduction of approximately 15.5%. This cost-saving is particularly significant in resource-limited settings, where efficient use of diagnostic resources is essential for sustaining HIV care programs. The robustness of this approach is further supported by the strong correlation and agreement observed between the two methods, as illustrated in Table 2 and Figures 3-5, which together provide consistent evidence confirming the reliability of the single-platform technique.

Taken together, these findings suggest that the single-platform method yields results that are clinically comparable to those obtained with the dual-platform approach, while also providing operational and economic

benefits. Nonetheless, challenges remain, including the relatively high cost and limited availability of reagents, as well as the need for adequately trained personnel to ensure consistent implementation. Addressing these limitations will be essential for broader adoption.

Conclusion

In conclusion, the single-platform method represents a reliable, accurate, and cost-effective alternative for $CD4^+$ T-lymphocyte monitoring in HIV-infected patients. Its implementation has the potential to enhance laboratory efficiency, expand diagnostic accessibility, and support national strategies for strengthening HIV/AIDS management in Thailand.

Conflict of Interest Statement

The authors declare no conflicts of interest.

Ethical approval

The study was approved by the Trat Provincial Human Research Ethics Committee (Approval No. **013/2567**, dated March **29**, **2024).**

Acknowledgements

The author would like to express sincere gratitude to Asst. Prof. Sakchai Dechatrairat for providing the conceptual framework for this research, Ms. Saowanit Chairatanapiwong for the research data supporting, and Ms. Tharathip Mukdaphetcharat, Head of the Medical Technology Department, Trat Hospital, for her invaluable support and encouragement. Special thanks are extended to all collaborators who contributed their time, effort, and motivation to this study. The author also expresses deep appreciation to Dr. Suchart Tantiniramai, Director of Trat Hospital, for his support in facilitating access to research equipment and facilities.

References

[1] World AIDS Day: UN warns nations focused on conquering COVID but neglect HIV/AIDS. BBC News Thai [Internet]. Available from: https://www.bbc.com/thai/59478416 (In Thai).

[2] AIDS surveillance system [Internet]. Moph.go.th. 2022 [cited 2024 Nov 26]. Available from: http://aidsboe.moph.go.th/aids_system/index.php?link=estimated

[3] Clift IC. Diagnostic flow cytometry and the AIDS pandemic. Lab Med. 2015; 46(3): e59-64. Available from: https://academic.oup.com/labmed/article/46/3/e59/2657953

[4] Lao-araya M. Flow cytometry in medical research and clinical applications. Biomed Sci Clin Med.; 53(2): 99-109. Available from: https://he01.tci-thaijo.org/index.php/CMMJ-MedCMJ/article/view/87433/0

[5] Thailand national guidelines on HIV/AIDS diagnosis, Treatment and prevention [Internet]. Available from:https://www.prepthai.net/Paper/HIVAIDS_Guidelines.pdf

[6] Li R, Gbadamosi-Akindele MF. CD4 count [Internet]. Nih.gov. Stat Pearls Publishing; 2019. Available from: https://www.ncbi.nlm.nih.gov/books/NBK470231/

[7] CD4 [Internet]. 196.119. 2024 [cited 2024 Nov 26]. Available from: http://122.155.196.119/laboratorynrh/?page_id=2095.

[8] Ministry of Public Health. Service fees for public health facilities under the Ministry of Public Health for Thai citizens 2019. Sunee Chawachalasai, Malinee Phattarajiraprasert, Nattawadee Wangsaeng, Korapat Thongsanit, Ratthana Songlaktong, Sirinan Sawangsuk, Editors. Nonthaburi: Health Administration Division, Ministry of Public Health; 2019. (In Thai)

[9] Storie I, Sawle A, Goodfellow K, Whitby L, Granger V, Ward RY, et al. Perfect count: A novel approach for the single-platform enumeration of absolute CD4+ T-lymphocytes. Cytometry. 2003; 57B (1): 47-52.

[10] Lolekha R, Phoolsawat T, Wannachai S, Sutthent R, Chaovavanich A, Nitayaphan S, et al. Multisite evaluation of PARTEC single-platform volumetric CyFlow counter system for determining percentage and absolute numbers of CD4 T lymphocytes in Thailand. J Virol Methods. 2004; 120(1): 49-53. doi: 10.1016/j.jviromet.2004.04.004.

[11] Kao-rian udom S, Dettrairat S. Multisite evaluation of Partec single-platform volumetric CyFlow® counter system for determining percentage and absolute numbers of CD4 T lymphocytes in Thailand. Thai AIDS J. 2015; 27(3):113-27.

[12] Montes M, Jaensson EA, Orozco AF, Lewis DE, Corry DB. A general method for bead-enhanced quantitation by flow cytometry. Journal of Immunological Methods [Internet]. 2006 Dec [cited 2019 Nov 4];317(1-2):45–55. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2680352/

[13] Battistini Garcia SA, Guzman N. Acquired immune deficiency syndrome $CD4^+$ count [Internet]. PubMed. Treasure Island (FL): StatPearls Publishing; 2020. Available from: https://www.ncbi.nlm.nih.gov/books/NBK513289

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 impact of parents' subjective happiness on children's executive functions: A longitudinal study

Seri Oh[1], Geonwoo Kim[2], Hyewon Jeong[2], Narae Lee[3], Jong-Sik Jang[1*]

[1]Department of Occupational Therapy, College of Health Science, Kangwon National University, Gangwon-do 25949, Republic of Korea.
[2]Department of Occupational Therapy, Kangwon National University Graduate School, Chuncheon, Gangwon-do, Republic of Korea.
[3]Department of Occupational Therapy, U1 University, Chungcheongbuk-do, Republic of Korea.

ARTICLE INFO

Article history:
Received 24 April 2025
Accepted as revised 29 September 2025
Available online 12 December 2025

Keywords:
Executive function, parental happiness, longitudinal study, child development.

ABSTRACT

Background: Executive function (EF) undergoes rapid development during childhood and is shaped by parenting context. Parental subjective happiness has been linked to children's socioemotional outcomes, but its long-term influence on EF remains unclear.

Objectives: To determine whether mothers' and fathers' subjective happiness at the child's age of 7 predict the child's EF at age 12, and to identify which facets of parental happiness (overall, relative, general, independent) are associated with later EF.

Materials and methods: A longitudinal cohort of 1 598 Korean families was drawn from the Education Panel Study on Korean Children. Parental happiness at baseline (child age 7) was assessed with the four-item Subjective Happiness Scale and classified into overall, relative, general, and independent domains. Children's EF was measured five years later (age 12) using the 40-item Child-dolescent Self-Reported Executive Function Difficulty Screening Questionnaire (higher scores = better EF). One-way ANOVA tested EF differences across low, normal, and high parental happiness groups for each domain.

Results: Maternal overall, relative, and independent happiness at age 7 were each positively associated with children's EF at age 12 (all $p<0.05$). In contrast, the corresponding paternal happiness domains showed no significant associations ($p>0.20$). For the general happiness domain-reflecting mood-related well-being both higher maternal and higher paternal scores predicted better child EF five years later (mother: F=5.34, $p=0.005$; father: F=3.04, $p=0.048$). Effect sizes were small to moderate ($\eta^2=0.01$-0.04).

Conclusion: In this Korean cohort, maternal happiness exerted a broader and stronger influence on children's later EF than paternal happiness, while both parents' general (mood-related) happiness independently predicted EF. Interventions that enhance parental-particularly maternal-subjective happiness may yield downstream cognitive benefits for school-aged children.

Introduction

Executive function includes neurophysiological, cognitive, and developmental elements, and is an essential behavioral function in daily life.[1] The ability to predict, prepare, plan, and organize situations in advance, the ability to concentrate attention and maintain working memory to focus on specific activities according to situational demands, the ability

* **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, College of Health Science, Kangwon National University, Gangwon-do 25949, Republic of Korea.*
E-mail address: JS_OT@kangwon.ac.kr
doi: 10.12982/JAMS.2026.026
E-ISSN: 2539-6056

to control behavior, and emotions have been noted as key subfactors of executive function.[2] These executive functions are activated and influenced by mental content, such as knowledge, beliefs, norms, values, and preferences, and understood as skills to achieve specific goals.[3] As such, it is a higher cognitive and executive function that is not only related to cognitive processes important for the active regulation of thinking, behavior, and emotions but also to social-emotional and behavioral areas. It can help one function adaptively in their environment.[1,4]

From a whole-life developmental perspective, children's executive functions are the developmental area with the greatest change during childhood, which a period in which self-regulation abilities expand and independence increases.[5,6] A child's executive function plays an important role in experiencing and adapting to new challenges during the development process. The child's ability to control some of their reactions, pay attention to specific stimuli, and follow rules is the standard for executive function.[7,8] The role of executive function is particularly important for school-age children because it is a time when they must adapt to a new social space called school, experience control and rules, and perform adaptation.[9] Children's executive function is related to school adaptation, affects academic level and relationships with peers, and can reduce problem behaviors by reducing impulsive behavior.10 The higher the level of executive function, the higher the peer competence; further, the actual performance of executive function leads to smooth school adaptation through the child's self-efficacy.[11]

Neurophysiologically, executive function is based on neural development in early life. The attachment created between parents and children, and parents' nurturing behavior are important factors affecting their children's neural development.[12,13] In particular, the emotional bond between parents and children, and consistent parenting behavior are key factors.[10] During childhood, parents gradually change from providing physical to psychological help to their children while performing their parental roles. They become encouragers and helpers for their children, and provide them with emotional stability.[14]

To provide appropriate psychological encouragement and support to children during this period, and become positive models, parents' psychological factors must first be healthy.[15] Parental psychological factors are related to the subjective happiness perceived by parents. These factors can determine parents' parenting attitudes, and have a lasting impact on the health of the family and psychological stability of the children (14). One such important parental psychological factor is parents' subjective happiness. It is very important because it affects parenting attitude and health; these are important for ensuring smooth interaction with their children while having the emotional stability of a caregiver. Crucially, parents' subjective happiness can affect the psychological stability of their child.[16]

Furthermore, parents' subjective happiness is related to the child's peer competence and stable attachment with the child, and can improve family functioning. It reduces issues such as anxiety, depression, defiance, and aggression in children, and positively affects their executive function.[14] In addition, when parents' subjective happiness improves, they engage in positive and responsive parenting behavior, and provide an environment that stimulates and supports their child's cognitive development. This helps the child in improving their executive function.[10] In this way, parents' subjective happiness reduces parenting stress and improves children's executive function. Indeed, Hong and Moon (2019) show that the happiness felt by parents affects children's executive function.[14]

This brief review of prior research suggests that the subjective happiness of fathers and mothers can affect their children's executive function. However, to the best of our knowledge, most studies focus on mothers. Moreover, few scholars independently analyze mothers and fathers. Next, recognizing changes in the development of children over time, who are constantly affected by the environment, is almost impossible because few studies individually analyze the type of subjective happiness of mothers and fathers. Finally, most research on the relationship between variables is cross-sectional in nature.

This study longitudinally analyzes the effect of mothers' and fathers' happiness when their children were in the first year of elementary school (7 years old) on the children's executive function in the sixth year of elementary school (12 years old). The aim is to determine the causal relationship between parents' subjective happiness and children's executive function.

Materials and methods

Study subjects

This study was conducted by the Korean Institute of Child Care to analyze the effects of mothers' and fathers' happiness when their children were in the first grade of elementary school (7 years old) on the children's executive function in the sixth grade of elementary school (12 years old). Data from the Education Panel Study on Korean Children's 8th (7 years old) and 13th years (12 years old) were used.

Both parents and children were selected as the research subjects. The sociodemographic characteristics in Table 1 revealed that the gender ratio of children was tilted towards males. The most common age group of both mothers and fathers was the 30s. Finally, the highest proportion of both mothers and fathers were college graduates.

Table 1. Sociodemographic background of study subjects.

Specification			N	%
Child	Gender ratio	Male	817	51.1
		Female	781	48.9
Mother	Age	Over 50	6	0.4
		40-49	491	30.9
		30-39	1075	67.6
		29 or less	18	1.1
	Education	Graduate school	93	5.8
		University (4 years or higher)	599	37.7
		College	435	27.4
		High school	456	28.4
		Middle school	7	0.5
Father	Age	Over 50	19	1.2
		40-49	549	35.0
		30-39	988	63.0
		29 or less	10	0.6
	Education	Graduate school	175	11.2
		University (4 years or higher)	658	42.0
		College	316	20.2
		High school	410	26.2
		Middle school	7	0.4

Research tools

Parents' subjective happiness

Parents' subjective happiness was measured by the mother and father questionnaires using the instrument developed by Lyubornirsky and Lepper (1999), and translated by the Korean Children's Panel researchers.[17] The items for subjective happiness are "Overall, I am...," "Compared to other people, I am...," "Some people are very happy overall. No matter what happens, I tend to have fun without worrying about it. I am...," and "Some people are very unhappy overall. They don't seem happy, even though they are not particularly depressed." In this study, 1 point, 2 points, and 3 points were made up of 1 point, 4 points, 5 points were made up of 2 points, 6 points, and 7 points were made up of 3 points. The higher the total score, the higher the happiness. The Cronbach's α values for the item reliabilities of mother's and father's happiness were 0.87 and 0.85, respectively (Table 2).

Table 2. Sociodemographic background of study subjects.

Variable		Item	Number of questions
Parental happiness	Overall happiness	Overall, I am...	1
	Relative happiness	Compared to other people, I am...	1
	General happiness	Some people are very happy overall. No matter what happens, I tend to have fun without worrying about it. I am...	1
	Independent happiness	Some people are very unhappy overall. They don't seem happy even though they are not particularly depressed.	1
		Entire	**4**

Difference between parents' general well-being on child's executive function

Children's executive function was evaluated using the Child-Adolescent Self-reported Executive Function Difficulty Screening Questionnaire developed by Song.[18] This scale comprises 40 questions in four areas: 11, 11, 8, and 10 questions on planning-organizing, behavioral control, emotional control, and attention-concentration difficulties, respectively. Each question is rated on a Likert scale ranging from one to three points. The higher the score by reverse-coding partial questions, the higher the executive function of the child. The Cronbach's α was 0.94.

Statistical analysis

Data from the Korean Children's Panel were analyzed using SPSS 25.0 after obtaining IRB exemption approval from the Gangwon University Bioethics Committee (KWNURB-2023-06-009). First, to analyze the demographic background of the study subjects, a frequency analysis was conducted on their gender, age, age group, educational background, and age of the child. Next, the Cronbach's α was used to analyze the reliability of items on the happiness of mothers and fathers, and children's executive functions. Subsequently, ANOVA was conducted to analyze the effect of mothers' and fathers' happiness when the children were in the first grade (7 years old) of elementary school on their executive function in the sixth grade (12 years old).

Results

Differences in the overall happiness of parents on children's executive functions

The difference between the mother's overall happiness on children's executive function five years later is from "very unhappy" (2.34±0.39) to "normal" (2.48±0.33) and "very happy" (2.53±0.32) in the first grade of elementary school. The higher the overall happiness of the mother in the first grade of elementary school, the more significant the executive function of child in the sixth grade of elementary school five years later (Table 3).

The difference between the father's overall happiness on children's executive function five years later is "very unhappy" (2.55±0.27), followed by "normal" (2.48±0.33) and "very happy" (2.51±0.33). Notably, the father's overall happiness was insignificant for children's executive function in the sixth grade of elementary school (Table 3).

Table 3. The differences between parents' overall happiness on children's executive function after five years.

Variable		N	Mean±SD	F	p	Scheffe
Mother	Very unhappy	47	2.34±0.39	8.542	0.000*	a<b
	Normal	607	2.48±0.33			
	Very happy	625	2.53±0.32			
Father	Very unhappy	266	2.55±0.27	1.575	0.208	-
	Normal	556	2.48±0.33			
	Very happy	628	2.51±0.33			

*Note: *p<0.05*

Differences in the relative happiness of parents on children's executive functions

The difference between the mother's relative happiness on children's executive function after five years is from "very unhappy" (2.38±0.37) to "normal" (2.49±0.33) and "very happy" (2.53±0.32) in the first grade of elementary school. The higher the mother's relative happiness in the first grade of elementary school, the more significant the executive function of sixth-grade elementary school children after five years (Table 4).

The difference between the father's relative happiness on the child's executive function after five years was "very unhappy" (2.54±0.31), "normal" (2.48±0.33), and "very happy" (2.51±0.32). Thus, the father's relative happiness was not significantly related with the child's executive function after five years (Table 4).

Table 4. The differences between the relative happiness of mothers and fathers on the executive function of their children after five years

Variable		N	Mean±SD	F	p	Scheffe
Mother	Very unhappy	67	2.38±0.37	7.240	0.001*	a<b
	Normal	626	2.49±0.33			
	Very happy	586	2.53±0.32			
Father	Very unhappy	31	2.54±0.31	1.399	0.247	-
	Normal	596	2.48±0.33			
	Very happy	588	2.51±0.32			

*Note: *p<0.05*

Differences between parents' general well-being on children's executive functions

The difference between the mother's general happiness on the executive function of children five years later is from "very unhappy person" (2.42±0.33) to "normal person" (2.53±0.32) and "very happy person" (2.53±0.32). The higher the mother's third happiness in the first grade of elementary school, the better the executive function of the sixth-grade elementary school children five years later (Table 5, Figure 3). The difference between father's general happiness on children's executive function after five years was "very unhappy" (2.43±0.33), "normal" (2.49±0.33), and "very happy" (2.52±0.32). Thus, the higher the father's general happiness, the better the executive function of sixth-grade elementary school children after five years (Table 5).

Table 5. The differences in the general happiness of the mother and father on the executive function of their children after five years.

Variable		N	Mean±SD	F	p	Scheffe
Mother	Very unhappy	100	2.42±0.33	8.542	0.000*	a<b
	Normal	700	2.49±0.33			
	Very happy	479	2.53±0.32			
Father	Very unhappy	70	2.43±0.33	1.575	0.208	-
	Normal	691	2.49±0.33			
	Very happy	454	2.52±0.32			

*Note: *p<0.05*

Differences in parents' independent happiness on children's executive function

The difference between mother's independent happiness on children's executive function five years later is from "very unhappy" (2.47±0.35) to "very happy" (2.47±0.34) and "very happy" (2.53±0.31) in the first grade of elementary school. The higher the mother's independent happiness in the first grade of elementary school, the more significant the executive function of sixth-grade elementary school children five years later (Table 6, Figure 4). The difference between the father's independent happiness on the child's executive function five years later was “very unhappy” (2.44±0.35), “normal” (2.49±0.33), and “very happy” (2.52±0.32). Thus, the father's independent happiness was not significantly related with the child's executive function five years later (Table 6).

Table 6. The differences between the mother's and father's independent happiness on the executive function of their children five years later

Variable		N	Mean±SD	F	p	Scheffe
Mother	Very unhappy	103	2.47±0.35	4.247	0.015*	a<b
	Normal	537	2.47±0.34			
	Very happy	639	2.53±0.31			
Father	Very unhappy	106	2.44±0.35	2.997	0.050	-
	Normal	535	2.49±0.33			
	Very happy	574	2.52±0.32			

Discussion

Using data from the 8th and 13th years of the Education Panel Study on Korean Children, this study longitudinally analyzed the effect of parents' happiness when their child was in the first grade of elementary school (7 years old) on the child's executive function in the sixth grade (12 years old).

First, mothers' overall and relative happiness showed a significant association with children's executive function, whereas fathers showed no association. Although fathers' participation in parenting has increased, they tend to be satisfied with supportive parenting. Hence, mothers are still more involved in the parenting process than fathers in Korean society.[19] Therefore, the degree of the mother's happiness may have a greater impact on the child's executive function (20). These results are consistent with those of studies showing that mothers' overall and relative happiness affect their children's executive function.[14,21,22]

Second, both mothers' and fathers' general happiness showed significant correlations with children's executive function. This is a feeling of happiness related to depression. Depression, expressed as anxiety, sadness, and lethargy, does not appear only at a specific time but continues to appear throughout an individual's life.[23] Therefore, parents with depression may negatively affect family functioning. These results are consistent with studies showing the negative influence of parental depression on children's executive function.[24-26]

Third, mothers' independent happiness was significantly correlated with children's executive function five years later, but fathers' independent happiness was not. This is because independent happiness is a feeling of happiness related to self-esteem; hence, if one spends more time with their mother than their father, they will naturally imitate the mother's attitude and behavior, and will be greatly influenced by their mother in various ways.[27]

This study has several limitations. Besides parental psychological factors, factors related to the roles of family members, workplace, and social environment may also be important. However, we did not include these in our analysis. Additionally, due to the characteristics of panel research, questions about happiness were simplified; therefore, we could not consider various subdomains of happiness. Finally, this study focused only on childhood; further research is needed to explore the effects of parental happiness on children's executive function at different developmental stages.

Policy suggestions and implications

Using child panel survey data, this study attempted to provide clinicians with clinical evidence for intervention by analyzing the effects of differences in parental happiness on child executive function. The results revealed that the mother's happiness is more important to the child's executive function than the father's happiness. Among the various feelings of happiness, overall and relative happiness emerged as the most important. National policies for mothers' feelings will be important, especially activities to increase the mother's overall and relative happiness. In addition, further studies should be conducted based on other parameters and ages of children not considered here.

Conclusion

This study is valuable in that it longitudinally analyzed the effects of mothers' and fathers' happiness when their children were in the first grade (7 years old) on their children's executive function in the sixth grade (12 years old). Since executive function develops continuously from infancy to adolescence, a longitudinal approach is useful for exploring mutual causal relationships over time.

This study holds academic significance as it identifies the longitudinal pathway through which parents' subjective happiness influences their children's executive function. Furthermore, it reveals the heterogeneity in the relative influence of different happiness factors between mothers and fathers, emphasizing the need to consider both parental roles when designing interventions aimed at improving children's executive function.

Ethical approval

Data from the Korean Children's Panel were analyzed

using SPSS 25.0 after obtaining IRB exemption approval from the Gangwon University Bioethics Committee (Approval No: KWNURB-2023-06-009)

Funding

None

Conflict of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CRediT authorship contribution statement

Seri Oh: conceptualization, methodology, writing; original draft, statistical analysis, data collection; **Geonwoo Kim:** manuscript editing, data collection, review and editing; **Hyewon Jeong:** manuscript editing, Graphic abstract, Narae Lee: manuscript editing, review; Jongsik Jang: supervision, validation, review and editing

Acknowledgements

The author expresses gratitude to the Korea Institute of Child Care and Education for providing the panel data.

References

[1] Baggetta P, Alexander PA. Conceptualization and operationalization of executive function. Mind Brain Educ. 2016;10(1):10-33. doi:10.1111/mbe.12100.

[2] Kim HO, Park JH, Kim KH. Comparison of income polarity groups of community factors control effects on mothers' parenting stress on children's executive functions. Korea J Child Care Educ. 2020;121:91-124.doi:10.37918/kce.2020.03.121.91.

[3] Doebel S. Rethinking executive function and its development. Perspect Psychol Sci. 2020;15(4): 942–56. doi:10.1177/1745691620904771.

[4] Gilbert SJ, Burgess PW. Executive function. Curr Biol. 2008; 18(3): R110-4. doi:10.1016/j.cub.2007.12.014.

[5] Hofmann W, Schmeichel BJ, Baddeley AD. Executive functions and self-regulation. Trends Cogn Sci. 2012; 16(3): 174-80. doi:10.1016/j.tics.2012.01.006.

[6] Fiske A, Holmboe K. Neural substrates of early executive function development. Dev Rev. 2019; 52: 42-62. doi:10.1016/j.dr.2019.100866.

[7] Lund JI, Toombs E, Radford A, Boles K, Mushquash C. Adverse childhood experiences and executive function difficulties in children: A systematic review. Child Abuse Negl. 2020; 106: 104485. doi:10.1016/j.chiabu.2020.104485.

[8] Luna B, Sweeney JA. The emergence of collaborative brain function: fMRI studies of the development of response inhibition. Ann N Y Acad Sci. 2004; 1021: 296-309. doi:10.1196/annals.1308.035.

[9] Kim CI. A study of factors influencing school adjustment of the first grade elementary school student: Focusing on characteristics of children, parents, and teachers. Korean J Child Educ. 2019; 28(4): 171-87. doi:10.17643/KJCE.2019.28.4.10.

[10] Lee HM, Han YJ. The sequential mediation effect of children's happiness and media device addiction on the relationship between mothers' happiness and children's executive function difficulty. Korean J Child Stud. 2020; 41(5): 85-98. doi:10.5723/kjcs.2020.41.5.85.

[11] Jeong JE, Shin N. Executive function and school adjustment in upper elementary school students: Mediating effect of self-efficacy. Korean J Child Stud. 2019; 40(4): 123-35. doi:10.5723/kjcs.2019.40.4.123.

[12] Zelazo PD. Executive function and psychopathology: A neurodevelopmental perspective. Annu Rev Clin Psychol. 2020; 16: 431-54. doi:10.1146/annurev-clinpsy-072319-024242.

[13] Caro-Morente C. The impact of child-parent relationship quality on neurodevelopmental problems in adoptees [dissertation]. Southampton (UK): University of Southampton; 2021.

[14] Hong G, Moon SR. The mediating effects of parenting stress on the relationship between parents' psychological variables, children's problem behavior, and executive function. J Korean Council Child Rights. 2019; 23(3): 497-519. doi:10.21459/kccr.2019.23.3.497.

[15] Molitor SJ, Dvorsky MR. Ethical considerations for assessing parent mental health during child assessment services. Ethics Behav. 2019; 29(2): 87-100. doi:10.1080/10508422.2018.1482746.

[16] Kim JI. Study of the relationship between the child's feeling of happiness, mother's feeling of happiness, parenting behaviors, and the elements of familial environment. Korea J Child Care Educ. 2018; 109: 53-77. doi:10.37918/kce.2018.03.109.53.

[17] Lyubomirsky S, Lepper HS. A measure of subjective happiness: Preliminary reliability and construct validation. Soc Indic Res. 1999; 46(2): 137-55. doi:10.1023/A:1006824100041.

[18] Song H. Validity of child-adolescent self-reported executive function difficulty screening questionnaire. Korean J Clin Psychol. 2014; 33(1): 121-37. doi:10.15842/kjcp.2014.33.1.008.

[19] Cho HS, Yoon EM. The relationship among parental subjective happiness, parent-child interaction, and self-esteem of children: Application of the actor-partner interdependence model (APIM). J Korea Acad Ind Coop Soc. 2022; 23(3): 183-93. doi:10.5762/KAIS.2022.23.3.183.

[20] An SY, Park SM. Longitudinal study of mothers' and fathers' parenting stress with early childhood children. J Humanit Soc Sci. 2022; 13(4): 295-310. doi:10.22143/HSS21.13.4.21.

[21] Murdock KW, Pittman LD, Fagundes CP.

Maternal and paternal predictors of child depressive symptoms: An actor-partner interdependence framework. J Child Fam Stud. 2018; 27(2): 559-68. doi:10.1007/s10826-017-0910-6.

[22] Newland LA. Family well-being, parenting, and child well-being: Pathways to healthy adjustment. Clin Psychol. 2015; 19(1):3-14. doi:10.1111/cp.12059.

[23] Suh GW. The influence of fathers' and mothers' depression and drinking behavior on children's development: The mediated role of family functioning and the moderated role of child sex. J Fam Relat. 2018; 23(2): 3-28. doi:10.21321/jfr.23.2.3.

[24] Yang JH, Kim YC. Actor and partner effect of young children-parents' family interaction, child care activities, and parenting stress. Asian J Educ. 2016; 17(3): 81-106. doi:10.15753/aje.2016.09.17.3.81.

[25] Wang J, Mansfield AK, Zhao X, Keitner G. Family functioning in depressed and non-clinical control families. Int J Soc Psychiatry. 2013; 59(6): 561-9. doi:10.1177/0020764012445260.

[26] Meyer S, Raikes HA, Virmani EA, Waters S, Thompson RA. Parent emotion representations and the socialization of emotion regulation in the family. Int J Behav Dev. 2014; 38(2): 164-73. doi:10.1177/0165025413519014.

[27] Han JW, Lee H. The effect of parental self-esteem on children's emotional responsiveness and attention: Through the child's self-esteem. J Korea Acad Ind Coop Soc. 2017; 18(11): 628-36. doi:10.5762/KAIS.2017.18.11.628.

Journal of Associated Medical Sciences 2026; 59(1): 218-226

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

Seroprevalence and potential risk factors associated with severe scrub typhus infection: an experience from a tertiary care set up in Odisha, India

Rout D[1], Panda NR[2], Otta S[3*]

[1]Department of Microbiology, IMS and SUM Hospital, Kalinga Nagar, Bhubaneswar, Philippines.
[2]Department of Medical Research, IMS and SUM Hospital, Kalinga Nagar, Bhubaneswar, Philippines.
[3]IMS and SUM Hospital, Kalinga Nagar, Bhubaneswar, Philippines.

ARTICLE INFO

Article history:
Received 12 December 2024
Accepted as revised 14 October 2025
Available online 12 December 2025

Keywords:
Scrub typhus, IgM ELISA, severe scrub typhus, scrub meningitis.

ABSTRACT

Background: Scrub typhus a mite-borne disease is being reported across globe in increasing numbers in known endemic as well as non-endemic areas. Non-specific clinical manifestations make it one of the most underreported diseases.

Objectives: Owing to the lack of adequate prevalence data from our region, the present study aims to denote the prevalence of scrub typhus in Odisha and factors associated with development of various complications of this disease.

Material and Methods: All patients who were clinically suspected as scrub typhus during the study period of one year coming to our hospital were tested for IgM antibodies against *Orientia tsutsugamushi* by commercially available enzyme-linked immunosorbent assay (ELISA) kit. Other common endemic diseases were ruled out by appropriate tests. Clinical data for admitted patients were obtained from case sheets. The patient demographics, treatment and laboratory data of the complicated and uncomplicated scrub typhus cases were noted. Data was entered in excel sheet and tabulated by SPSS software.

Results: In the study among 2449 suspected cases, 398(16.3%) were positive for scrub typhus by IgM ELISA with peak incidence from August to November months. There was male preponderance of the illness with peak age of affection being 51-60 years. Eschar was noted only in 6 cases, and 49 cases went on to have various complications, commonest being meningoencephalitis. Mortality was very low in our study. Longer duration of fever before seeking treatment was significantly associated with development of complications.

Conclusion: The prevalence of scrub typhus was 16.3 % among the study population. Taking scrub typhus diagnostics to the peripheral level is the need of the hour to ensure early diagnosis and prevent severe disease.

Introduction

Scrub typhus is one of the important causes of acute febrile infectious illness globally. It is caused by bacteria named *Orientia tsutsugamushi,* other closely related species *Candidatus Orientia chuto*[1]and *Candidatus Orientia chiloensis*.[2] It is transmitted by mite bite particularly from that of chigger of *Leptotrombidium deliense* in India. It is one of the world's most underdiagnosed diseases and the burden of this disease is vastly underreported due to unavailability of adequate testing facilities, scarce

* **Corresponding contributor.**
Author's Address: *Department of Microbiology, IMS and SUM Hospital, Kalinga Nagar, Bhubaneswar, Philippines.*
E-mail address: saritaotta@gmail.com
doi: 10.12982/JAMS.2026.027
E-ISSN: 2539-6056

research and non-distinctive clinical manifestations.

Every year about 1 million infected cases of scrub typhus are reported globally with mortality of around 30% when not treated.[3] The cases of scrub typhus in India is increasing steadily since past years and it is being detected outside the previous catchment areas. This is probably attributed to climate change and incessant travelling.

The present study aims to update the knowledge of prevalence of scrub typhus in our area and denote the significant risk factors associated with severe scrub typhus.

Materials and methods

Study setting

Present study was designed as a retrospective study from September 2023 to August 2024 in our tertiary care teaching hospital in Bhubaneswar, Odisha, India. This is a 1400-bedded hospital which caters mostly to the middle and lower socioeconomic status patients of the state. Odisha is a coastal state located in the eastern part of India. The weather is humid for most of the months of the year and receives rainfall from mid-June to September and is pleasant to humid in winter months between December and March.

Sample collection

All patients who were suspected during the study period of having scrub typhus and were sent for IgM ELISA (Scrub Typhus Detect™ IgM ELISA system, InBios International Inc.) for *Orientia tsustugamushi* to the central laboratory of our hospital were considered for the study. Central laboratory of the hospital receives samples from all the outpatient departments, wards and ICUs of the hospital and test are performed in house. The IgM ELISA is performed on the serum sample of the patients as per manufacturers' instructions and quality control measures are taken care of for each run of the ELISA test.

Suspected case was defined as a patient having fever ≥5 days without any other etiology with/without eschar. Possible case was defined as a suspected case with OD>0.5 for IgM by ELISA. Confirmed case is a possible case which showed effervescence on treatment with doxycycline/ azithromycin.

Patient data was collected from the lab register of the central laboratory and was used for finding the prevalence of the disease. Clinical data of the patients were extracted from the case sheets obtained from the medical record department for the possible scrub typhus patients admitted into wards or ICU and data was noted in a predesigned Performa. All these patients were evaluated for the demographics, clinical features, course in the hospital, laboratory parameters, treatment and various complications were carefully noted. Data for basic laboratory tests like complete blood count, renal function tests (blood urea and serum creatinine), blood glucose, and liver function tests serum bilirubin (direct and indirect), aspartate aminotransferase (AST), alanine aminotransferase (ALT) and alkaline phosphatase (ALP) was noted.

Exclusion criteria

Patients where tests for malaria, typhoid fever, dengue, leptospirosis, and pneumonia were positive and those where the cause of fever was known as any chronic illnesses, HIV or malignancy were excluded from the study.

Data thus obtained was entered in MS excel and analysis was done for various parameters. Qualitative variables were presented as frequency while quantitative variables were denoted with mean±SD. Associations was determined by using Chi-square test or Fisher's exact test. Univariate and multivariate regression analysis was performed for the independent predictors of severity. The $p<0.005$ was considered as significant. All the statistics were done using SPSS software version 22.0 (SPSS Inc., Chicago, IL).

The study was approved by the Ethics Committee of the hospital via Ref.no/IEC/IMS.SH/SOA/2O23/523. Dt- 2.5.23. The participant information was kept confidential. Because of the retrospective nature of the study, the diagnostic and treatment of the patients was not interfered in any way.

Results

Prevalence of scrub typhus

During the study period total 2449, clinically suspected cases were received in the lab from which 398 (16.3%) samples were positive by IgM ELISA taking a cut off OD of 0.5**.** Among this 339 (85.2%) were adults and 59 (14.8%) belonged to pediatric population of less than14 years. Mean and SD of the cut off OD of the cases was 2.13±1.64 (Figure 1).

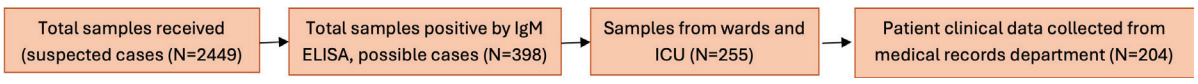


Figure 1. *Summary of the samples the study.*

Spatio-temporal distribution of cases

Scrub typhus cases are spread across the year except for winter months of December, January and February in our area. A rise in the number of cases from August was noted with peak incidence in November (Figure 2).

Similarly, the samples were mostly from Eastern part and the northern part of the state distributed across various districts as denoted in Figure 3.

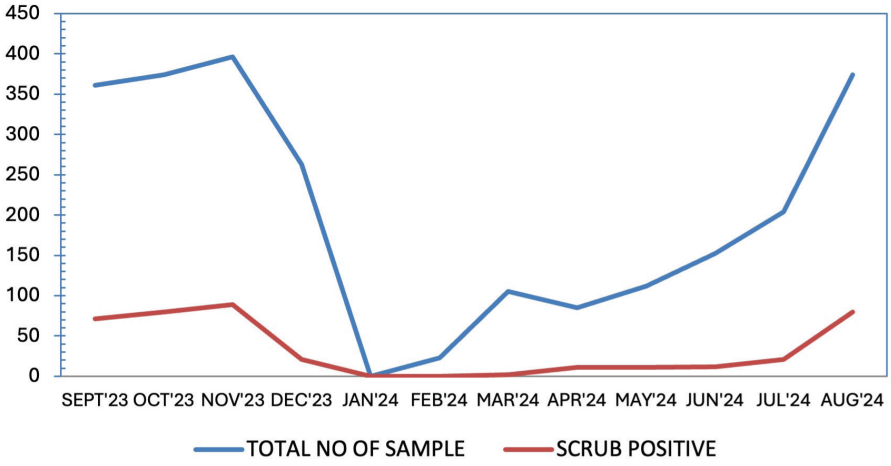


Figure 2. *Temporal distribution of scrub typhus cases in the study.*

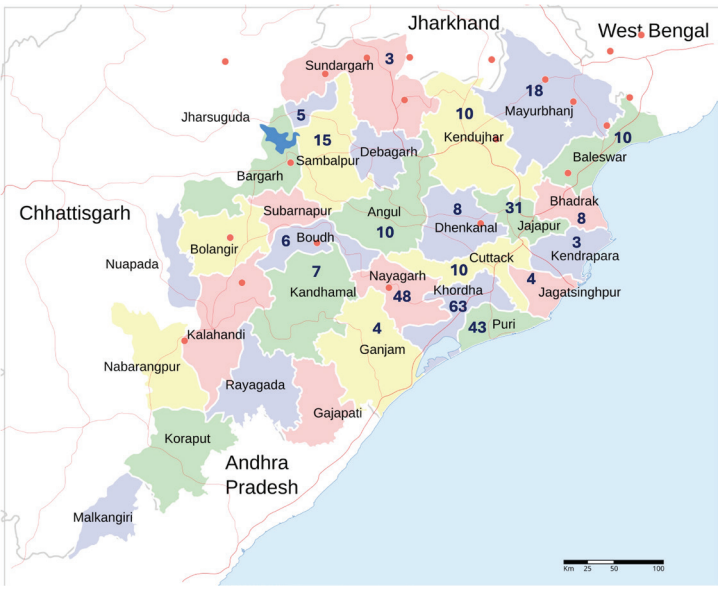


Figure 3. *Distribution of cases in the state.*

The samples were from Khordha (15.8%), Puri (10.8%), Nayagarh (12.1%) and Jajpur (7.8%) districts. The Southern and western districts of the state had a very little representation in our patient data.

Demographics of the patients

Among the total positives included (204) 64.4% were males. Mean age of patients was 35.5±23 years and median age being 35 years. The commonest age group of afflictions of the patients was 51-60 (19.8 %) years followed by 31-40 (15.5%), 41-50 (15.8 %) and 71-80 (18.5%) years.

Presenting symptoms

Fever was the most common presenting illness seen in 94.6% of patients. 65.8% of these patients presented to the hospital after 5 days fever. Mean duration of fever before presentation to the hospital was 8.4±6.8 days. Chill rigor was associated in 37.8% of cases. Other nonspecific presenting features were headache (27.9 %) and weakness (28.9%). Pain of the abdomen was noticed in about 38 % of patients while respiratory symptoms with breathlessness and cough were seen in 22.5 % and 35.8% cases respectively. Sensorial alteration was the presenting complaint in 25.5 % cases (Table1). Mean duration of the hospital stay in the scrub typhus cases in our set up was

7.29±5.5 days median being 6 days. ICU stay was noted in 49 (24.01%) patients and mean duration of ICU stay was 1.01±2.1 days. Mortality was seen in 2(0.9 %) of cases. Among the total cases (204), 49 (24.01%) progressed to complicated or severe scrub typhus. Some patients developed more than one complication or organ system affection. The various complications are summarized in Table 2.

Table1. Different presenting symptom of the cases (N=204).

Symptom	Number (% total)	Complicated scrub typhus patients (% total)	Uncomplicated scrub typhus patients (% total)
Fever	193 (94.6)	98	97
Fever <5 days	66 (34.2)	30.6	33
Fever >5 days	127 (65.8)	69.4	64
Chill and rigor	73 (37.8)	69.4	45
Headache	57 (27.9)	20.4	28
Weakness	59 (28.9)	14.2	29
Joint pain	24 (11.8)	4.08	12
Breathlessness	46 (22.5)	26.5	15
Cough	73 (35.8)	36.7	37
Pain abdomen	77 (37.8)	2.9	38.1
Diarrhea	45 (22.1)	28.6	19
Vomiting	81 (39.7)	1.5	37
Altered sensorium	52 (25.5)	23	18.4
Eschar	8 (3.92)	6.1	6
Hepatomegaly	16 (7.8)	16.3	9
Splenomegaly	23 (11.3)	22.4	12

Table 2. Various complications encountered in severe scrub typhus cases in the study.

Complication	Number (%)
Pneumonia	4 (8.1)
Acute kidney Injury	7 (14.2)
Anemia	1 (2.0)
Thrombocytopenia	1 (2.0)
Meningoencephalitis	14 (28.5)
Shock	12 (24.5)
Acute respiratory distress syndrome	4 (8.2)
Dyselectrolytemia	8 (16.3)
Myocarditis	2 (4.1)
GI bleeding	1 (2.0)
Hepatitis	6 (12.2)
MODS	3 (6.1)

Fever was the commonest symptom in the scrub typhus patients. While weakness, myalgia or flu like illness was the predominant picture in uncomplicated cases, severe cases presented with chill rigor. Other features like shortness of breath; diarrhea was seen more often in complicated cases. The uncomplicated cases had cough with pain abdomen and vomiting (Table1).

The demographic data of severe and uncomplicated scrub typhus cases are compared in Table 3 and 4. There is no significant difference in the mean age and sex distribution of the severe and uncomplicated scrub patients. The duration of illness before seeking health care was significantly higher in severe scrub cases than their counterpart. The severe cases had higher association with diabetes mellitus and hypertension as co-morbidities. ICU stays and the duration was significantly higher in severe cases than uncomplicated ones.

Among laboratory parameters, liver enzymes SGOT and SGPT was raised in significant number of uncomplicated and severe scrub typhus patients (89, 79.6 for OT; 78, 73.4% for PT) respectively. Total bilirubin was raised in 23%, 97.9 % of uncomplicated and severe cases respectively while Alkaline phosphatase was raised to more than 140 IU in 30% and 34.7% respectively. Platelets less than 1 lakh/ were seen in 31% and 30.6 % of uncomplicated and severe scrub cases.

Table 3. Demographic, clinical characteristics and laboratory findings on admission of the patients.

	Severe scrub typhus	Uncomplicated scrub typhus
Demographic of the patients		
Age>60 (Mean±SD)	36.5±24.04	34.8±23.1
Male:Female	1.45	2.13
Duration of illness before admission	8.5±5.6	8.4±7.2
Length of hospitalization (Mean±SD)	7.7±3.2	6.7±5.8
Mean duration from disease onset to effective antibiotic therapy (days)	9.73±5.8	10.23±3.7
Mortality (% total)	2 (4.08%)	0 (0)
Underlying diseases		
Diabetes mellitus (%)	20.4	8
Hypertension (%)	10.2	7
Liver ailments (%)	0	0
Lung disease (%)	0	0
Alcoholism (%)	8.2	1
Chronic kidney diseases (%)	2.04	1
Heart disease (%)	0	1
Sickle cell disease (%)	0	2
ICU stay (%)	44.9	12
Duration of ICU stay in days (Mean±SD)	2.08±2.9	0.49±1.35
Mortality (%)	2	0
Eschar (%)	4.08	6
Laboratory data		
Haemoglobin (gm/dL)	10.2±1.8	10.2±1.9
Total leukocyte count (x10^3/ mm^3)		8.8±5.8
Total platelet count (x10^5/mm^3)	2.3±6.8	1.9±1.9
Total bilirubin (mg/dL)	4.2±15.	1.4±1.86
Direct bilirubin (mg/dL)	2.23±14.9	1.75±10.2
SGOT(IU)	95.2±53.7	125.6±290.5

Table 3. Demographic, clinical characteristics and laboratory findings on admission of the patients. (Continue)

	Severe scrub typhus	Uncomplicated scrub typhus
SGPT (IU)	96.7±97	107.6±205.4
ALP(IU/L)	147.7±105.2	126.5±118.9
Urea (mg/dL)	44.9±44.7	26.2±24.5
Creatinine (mg/dL)	2.33±6.8	4.7±15.2
Mean OD	2.02±0.9	1.92±1.25

Table 4. Unadjusted relative risk and p value for selected factors on severe scrub typhus.

	Severe scrub typhus	Uncomplicated scrub typhus	Odds ratio	*p* value
Demographic of the patients				
Age (Mean±SD)>60	17	32	1.03	0.681
Male:Female	29:20	107:48		0.235
Duration of illness before admission	8.5±5.6	8.4±7.2	1.31	<0.001
Length of hospitalization (Mean±SD)	7.7±3.2	6.7±5.8	1.037	0.274
Mean duration from disease onset to effective antibiotic therapy (days)	9.73±5.8	10.23±3.7	1.02	0.32
Underlying diseases (diabetes, liver cirrhosis, COPD)	38.8%	18%	1.365	0.036
ICU stay (%)	44.9	12	1.569	<0.001
Duration of ICU stay in days (Mean±SD)	2.08±2.9	0.49±1.35	1.457	<0.001
Eschar (%)	4.08	6	0.236	0.569
Laboratory data				
Hb (<10gm/dL)	7	47	0.995	0.955
TLC (>10.0 x10^3/mm^3)	6	30	1.001	0.162
TPC (<1x10^5/mm^3)	4	0	1.024	0.540
Total bilirubin (>1.0 mg/dL)	5	47	1.092	0.083
Direct bilirubin (>0.2 mg/dL)	8	78	1.010	0.482
SGOT>40 IU	13	90	0.997	0.179
SGPT>40 IU	12	78	0.999	0.507
ALP (>200 U/L)	4	19	1.002	0.291
Urea (>24 mg/dL)	11	34	2.018	0.001
Creatinine (>1.4 mg/dL)	4	22	0.940	0.121

Discussion

A recent sero-epidemiological data showed that global pooled prevalence rate of scrub typhus in acute febrile illness was 24.93% [4] and that in Asia is 22.2%.[5] There has been a significant increase in sero-prevalence of scrub typhus since 2010 and it is estimated to be about 30% in community set up in India.[6] The epidemiology of scrub typhus varies as per change in climatic conditions and arthropod vector involved in transmission.[7,8] In India southern (55.5%)

and northern (31.5%) parts bear the major brunt of cases but, Eastern Indian literature is nominal.[6] In our study the prevalence of scrub typhus in clinically suspected cases was 16.3% similar to few other studies from Eastern India.[9] Increased travel have expanded the reach of scrub typhus into previously unreported areas for scrub typhus.[10] July to December months which correspond with monsoon and post-monsoon seasons correspond to high occurrences of scrub typhus cases in most of the states of India as in our case.[9,11-13]

Eschar is uncommon (9.5%-45%) in patients from Southeast Asia and Indian subcontinent,[14,15] probably due to presence in areas of tight fitted clothing as well as dark complexion in Indian set up making their elucidation difficult. Absence of eschar is an independent predictive factor for fatal outcomes.[16] But, in the present study eschar was detected in 3.9% cases with no significant difference of its presence among severe and uncomplicated cases. Variation in prevalence of eschar is because of different infecting *O. tsutsugamushi* strains[17,18] and vary across different ethnic populations.[19,20]

Most cases of scrub typhus in India are confined to young adults probably due to occupational or recreational exposure to mite infested vegetation and the median age of affliction is 28.1 years.[6]Predominant age group affected in our study was higher of >30 years age. Studies from countries like Japan also report 62% of scrub positive cases to be under 51-75 age group.[24] In the study by Liu et al. the highest number of cases (21%) were in the age group of 41-50 years.[22] Males constituted 64% of cases which agreed with previous studies by Liu *et al.* (56%) and Bal *et al.* (61.7%). [22,23] However, increased inclination of scrub typhus towards females was reported by studies carried out in few countries like South Korea.[24,25]

In our study hepatosplenomegaly was seen less often than other studies.[26] Splenomegaly is often not seen in dengue, thus helping to distinguish scrub typhus from dengue.[26] In our set up 24% cases developed various complications the most common being meningoencephalitis like other studies.[27,28] Meningoencephalitis is due to vasculitis caused by the organism. In scrub typhus CSF shows lymphocyte predominance cellular picture with increased protein levels. Other common complication was shock. We also found 2 cases of rare complication that is myocarditis in our study. The most common complication in scrub typhus in India in a review was hepatitis seen in 40.5% of scrub typhus cases[6] unlike in the present study. Raised liver enzymes (<3 fold) was nevertheless found in many patients in our study which never gave rise to severe scrub typhus.

Wang *et al*.reported the mortality of 6% in their review of various studies from India.[29] However, in our study, only 2 patients died of scrub typhus despite our patients being >30 years with about 20% belonging to age group of 51-60. Similar low mortality was also noted in another study.[30] Increased duration of illness before health care seeking was significantly different and was associated with higher odds of presenting with severe form of illness in our study. As expected, the severe cases had higher ICU admission and increased duration of stay in intensive care as well. In the present study initial hemoglobin, total platelet count or total leukocyte count and presence of underlying illnesses were not significantly associated with severity of infection.

Limitation

Present study has several limitations like genotypes or serotypes of *O. tsutsugamushi* have not been addressed and the samples were from a hospitalized setup in a single geographic region limiting its generalization.

Conclusion

The study highlights the growing prevalence of scrub typhus in endemic regions and its non-specific clinical manifestations. Although mortality can be averted by timely management, but complications can increase morbidity. The disease has heterogeneity in diagnostic methods, interpretation cutoff points, variable clinical presentation which thus points at further need for research.

Ethical approval

The study was approved by the ethics committee of the institute via letter no: Ref.no/IEC/IMS.SH/SOA/2O23/523. Dt- 2.5.23

Conflict of interest

None

CRedIT authorship contribution statement

Rout D: design, definition of intellectual contents, literature search, manuscript preparation, review, concepts and data acquisition; **PandaNR**: data analysis, statistics and manuscript editing; **Otta S:** design, definition of intellectual contents, literature search and manuscript preparation, manuscript editing, and manuscript review. All the authors have read and approved the final manuscript.

Acknowledgements

We acknowledge the support of all staff of IMS and SUM Hospital for their help in carrying out this work. We are also thankful to SOA University for your financial assistance with this work.

References

[1] Rout D, Praharaj I, Dalai SP, Mishra S, Otta S. Rising menace of scrub typhus – Current status and challenges. J Pure Appl Microbiol. 2023; 17(4): 2060-73. doi:10.22207/JPAM.17.4.39.

[2] Abarca K, Martínez-Valdebenito C, Angulo J, Jiang J,

Farris CM, Richards AL, et al. Molecular description of a novel Orientia species causing scrub typhus in Chile. Emerg Infect Dis. 2020; 26(9): 2148-56. doi:10.3201/eid2609.200918.

[3] Kala D, Gupta S, Nagraik R, VermaV, Thakur A, Kausal A. Diagnosis of scrub typhus: recent advancements and challenges. Biotech. 2020; 10: 396. doi:10.1007/s13205-020-02389-w.

[4] Dasgupta S, Asish PR, Rachel G, Bagepally BS, Chethrapilly Purushothaman GK. Global seroprevalence of scrub typhus: a systematic review and meta-analysis. Sci Rep. 2024;14(1): 10895. doi:10.1038/s41598-024-61555-9.

[5] Bonell A, Lubell Y, Newton PN, Crump JA, Paris DH. Estimating the burden of scrub typhus: A systematic review. PLoS Negl Trop Dis. 2017; 11(9) :e0005838. doi:10.1371/journal.pntd.0005838.

[6] 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.

[7] Rajapakse S, Rodrigo C, Fernando D. Scrub typhus: pathophysiology, clinical manifestations and prognosis. Asian Pac J Trop Med. 2012; 5(4): 261-4. doi:10.1016/S1995-7645(12)60045-4.

[8] Varghese GM, Janardhanan J, Mahajan SK, Tariang D, Trowbridge P, Prakash JA, et al. Molecular epidemiology and genetic diversity of Orientia tsutsugamushi from patients with scrub typhus in 3 regions of India. Emerg Infect Dis. 2015; 21(1): 64-71. doi:10.3201/eid2101.140727.

[9] Behera B, Biswal M, Das RR, Dey A, Jena J, Dhal S, et al. Clinico-epidemiological analysis of scrub typhus in hospitalised patients presenting with acute undifferentiated febrile illness: a hospital-based study from Eastern India. Indian J Med Microbiol. 2019; 37(2): 278-80. doi:10.4103/ijmm.IJMM_18_217.

[10] Xu G, Walker DH, Jupiter D, Melby PC, Arcari CM. A review of the global epidemiology of scrub typhus. PLoS Negl Trop Dis. 2017; 11(11): e0006062. doi:10.1371/journal.pntd.0006062.

[11] Das P, Singh D, Das M, Nayak RK, Mohakud NK. Epidemiological and clinical features of scrub typhus in Odisha, Eastern India. Med J Dr DY Patil Univ. 2019; 12(5): 419-23. doi:10.4103/mjdrdypu.mjdrdypu_76_19.

[12] Lakshmanan S, Sagayaraj B, Sujatha B, Vasudevan LD. Clinical and laboratory profile of pediatric scrub typhus in a tertiary care teaching hospital in Southern India. Int J Contemp Pediatr. 2018; 5(5): 2092-6. doi:10.18203/2349-3291.ijcp20183500.

[13] 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.i02.023.

[14] Rodkvamtook W, Ruang-Areerate T, Gaywee J, Richards AL, Jeamwattanalert P, Bodhidatta D, *et al*. Isolation and characterization of Orientia tsutsugamushi from rodents captured following a scrub typhus outbreak at a military training base, Bothong district, Chonburi province, central Thailand. Am J Trop Med Hyg. 2011; 84(4): 599-606. doi:10.4269/ajtmh.2011.10-0466.

[15] Thapa S, Hamal P, Chaudhary NK,Sapkota LB, Singh JP.Burden of scrub typhus among patients with acute febrile illness attending tertiary care hospital in Chitwan, Nepal. BMJ Open. 2020; 10: e034727. doi:10.1136/bmjopen-2019-034727.

[16] Lee HB, Kwon KS. Risk factors leading to fatal outcome in scrub typhus patients. Am J Trop Med Hyg. 2009; 81(3): 484-8. doi:10.4269/ajtmh.2009.81.484.

[17] Kim DM, Yun NR, Neupane GP, Shin SH, Ryu SY, Yoon HJ, et al. Differences in clinical features according to Boryoung and Karp genotypes of *Orientia tsutsugamushi*. PLoS One. 2011; 6(8): e22731. doi:10.1371/journal.pone.0022731.

[18] Park J, Woo SH, Lee CS. Evolution of eschar in scrub typhus. Am J Trop Med Hyg. 2016; 95(6): 1223-5. doi:10.4269/ajtmh.16-0644.

[19] Yoo JS, Kim D, Choi HY, Yoo S, Hwang JH, Hwang JH, et al. Prevalence rate and distribution of eschar in patients with scrub typhus. Am J Trop Med Hyg. 2022; 106(5): 1358-66. doi:10.4269/ajtmh.22-0040.

[20] Paris DH, Phetsouvanh R, Tanganuchitcharnchai A, Jones M, Jenjaroen K, Vongsouvath M, et al. *Orientia tsutsugamushi* in human scrub typhus eschars shows tropism for dendritic cells and monocytes rather than endothelium. PLoS Negl Trop Dis. 2012; 6(1): e1466. doi:10.1371/journal.pntd.0001466.

[21] Lalrinkima H, Lalremruata R, Lalchhandama C, Khiangte L, Siamthara FH, Lalnunpuia C, et al. Scrub typhus in Mizoram, India. J Vector Borne Dis. 2017; 54(4): 369-71. doi:10.4103/0972-9062.219888.

[22] Liu YX, Feng D, Suo JJ, Xing YB, Liu G, Liu LH, et al. Clinical characteristics of the autumn-winter type scrub typhus cases in south of Shandong province, northern China. BMC Infect Dis. 2009; 9: 1. doi:10.1186/1471-2334-9-1.

[23] Bal M, Mohanta MP, Sahu S, Dwibedi B, Pati S, Ranjit M. Profile of pediatric scrub typhus in Odisha, India. Indian Pediatr. 2019; 56: 304-6. doi:10.1007/s13312-019-1448-7.

[24] Xu G, Walker DH, Jupiter D, Melby PC, Arcari CM. A review of the global epidemiology of scrub typhus. PLoS Negl Trop Dis. 2017; 11(11): e0006062. doi:10.1371/journal.pntd.0006062.

[25] Lee HW, Cho PY, Moon SU, Na BK, Kang YJ, Sohn Y, et al. Current situation of scrub typhus in South Korea from 2001-2013. Parasites & Vectors. 2015;

8: 1-4. doi:10.1186/s13071-015-1154-9.
[26] Gupta N, Chaudhry R, Kabra SK, Lodha R, Mirdha BR, Das BK, et al. In search of scrub typhus: a prospective analysis of clinical and epidemiological profile of patients from a tertiary care hospital in New Delhi. Adv Infect Dis. 2015; 5(4): 140-7. doi:10.4236/aid.2015.54015.
[27] Vivekanandan M, Mani A, Priya YS, Singh AP, Jayakumar S, Purty S. Outbreak of scrub typhus in Pondicherry. J Assoc Physicians India. 2010; 58(1): 24-8. doi:10.4103/0972-6748.77644.
[28] Varghese GM, Trowbridge P, Janardhanan J, Thomas K, Peter JV, Mathews P, et al. Clinical profile and improving mortality trend of scrub typhus in South India. Int J Infect Dis. 2014; 23: 39-43. doi:10.1016/j.ijid.2013.11.016.
[29] Wang Q, Ma T, Ding F, Lim A, Takaya S, Saraswati K, et al. Global and regional seroprevalence, incidence, mortality of, and risk factors for scrub typhus: a systematic review and meta-analysis. Int J Infect Dis. 2024; 146: 107151. doi:10.1016/j.ijid.2024.107151.
[30] Kim DM, Kim SW, Choi SH, Yun NR. Clinical and laboratory findings associated with severe scrub typhus. BMC Infect Dis. 2010; 10: 1-7. doi:10.1186/1471-2334-10-7.

Journal of Associated Medical Sciences 2026; 59(1): 227-234

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Impact of density gradient technique and *Glycyrrhiza glabra* medium mixed with autologous platelet rich plasma on certain sperm function parameters and sperm DNA fragmentation index of asthenoteratozoospermic men: A comparative cross-sectional study

Noor Kadhim Khudhur*[1], Saad S. Al-Dujaily[2], Laith Amer Al-Anbari[3]

[1]*Applied Embryology, Medical Laboratory Techniques Department, Al-Farahidi University, Baghdad, Iraq.*
[2]*Clinical Reproductive Physiology and ART, Higher Institute for Infertility Diagnosis and Assisted Reproductive Technology, Baghdad, Iraq.*
[3]*Surgery (Urology Surgery), College of Medicine, Al-Nahrain University, Baghdad, Iraq.*

ARTICLE INFO

Article history:
Received 14 July 2025
Accepted as revised 28 November 2025
Available online 12 December 2025

Keywords:
DNA fragmentation, Glycyrrhiza glabra, platelet rich plasma, sperm preparation.

ABSTRACT

Background: Although there is evidence of a global decline in sperm motility and an increasing frequency of male infertility, there is minimal documentation regarding prospective treatments and techniques to activate and enhance certain sperm parameters to overwhelm male infertility.

Objectives: This study used Glycyrrhiza glabra (Gg) and Platelet Rich Plasma (PRP) media to establish a culture medium that aimed to improve sperm function parameters by preparing in vitro men's semen complaining of asthenoteratozoospermic factor.

Materials and methods: Semen samples from 80 participants were investigated in a comparative cross-sectional study. After semen fluid analysis 34 samples were revealed normozoospermia and 46 samples were with asthenoteratozoospermia. Certain sperm parameters and DNA fragmentation index were assessed. PRP and Gg were used as a sperm culture media. Certain sperm parameters and DNA fragmentation index were assessed using World Health Organization 2021 and 1999 guidelines.

Results: After activation using the density gradient method and Gg plus PRP, sperm function parameters improved significantly (p<0.001), including increased motility, morphologically normal sperm percentage, and decreased DNA fragmentation index.

Conclusion: It is concluded that the density gradient technique with any medium used in the present study revealed a significant improvement in certain sperm parameters and a reduction in DNA fragmentation index.

Introduction

Assisted reproductive technology treatments have improved in recent decades. However, intrauterine insemination, in vitro fertilization, and intracytoplasmic sperm injection have low pregnancy rates.[1] Gene polymorphism and insufficient semen treatment in vitro can affect the latter. For instance, oxidative stress, an imbalance between reactive oxygen species (ROS) and antioxidant defense, affects sperm. The imbalance can impede spermatozoa's ability to produce an

* **Corresponding contributor.**
Author's Address: *Applied Embryology, Medical Laboratory Techniques Department, Al-Farahidi University, Baghdad, Iraq.*
E-mail address: Noor.kadhim@uoalfarahidi.edu.iq
doi: 10.12982/JAMS.2026.028
E-ISSN: 2539-6056

embryo.[2] The sperm culture medium, incubation time, centrifugation, and sperm preparation method can all raise ROS levels. Density gradient centrifugation is an effective way to pick healthy spermatozoa and decrease ROS.[3]

The density gradient procedure is the most recommended for selecting a higher number of mobile sperm cells in cases of severe low sperm count, abnormal sperm morphology, or reduced sperm motility.[4-5]

On the other hand, Glycyrrhiza glabra (Licorice, Gg) is a tiny perennial widely used medicinal plant from the Fabaceae family, that is regularly used as feed and food.[6] Among its many beneficial components are glycyrrhizin, glycyrrhetinic acid, and others.[7-8] Since 2003, licorice has been introduced in vitro to gamete preparation and embryonic development culture mediums.[9-10] Some licorice elements may modulate steroidogenesis and influence testosterone and cortisol levels, according to research. These hormones are necessary for sperm production and motility.[11]

Platelet-Rich Plasma (PRP) is a form of plasma that is derived from the blood and contains a higher concentration of platelets compared to regular plasma.[12] Platelets release seven essential protein growth factors that activate all processes involved in tissue repair and medical disciplines e.g., dermatology, and reproductive medicine.[13] The supplementation of sperm culture media with PRP-derived compounds triggered a cascade of reactions including cell growth and protection against freezing, and improving motility.[14]

Thus, the current study aims to create a culture medium composed of Gg and PRP to improve certain sperm function parameters and reduce the DNA fragmentation index through in vitro activation of men's semen suffering from asthenoteratozoospermia factor.

Materials and methods

Participants

This comparative cross-sectional study (STROBE) examined 80 samples of semen including asthenoteratozoospermic (N=46) and normozoospermic (N=34) individuals. Prior to gathering each sample, all participants provided their consent and were informed of the purpose of the questionnaire as well as the goal of the research. The study was performed in the High Institute for Infertility Diagnosis and Assisted Reproductive Technologies, Al-Nahrain University and Al-Harithiya International Lab in Baghdad Iraq and extended from October 2023 to January 2024.

Inclusion criteria represented by fertile normozoospermic men and infertile asthenozoospermic men. On the other side exclusion criteria include azoospermic men, oligozoospermia/teratozoospermia, and men with history of undescended testes, hyperprolactinemia, genetic anomalies, and diabetes mellitus.

Sample size

The sample size was determined considering the given time frame for completing the research and the challenges associated with collecting samples during this timeframe, in line with the research's selection criteria. Hence, we formulated the study to include a sample size ranging from 80-100 samples. However, due to insufficient results and unforeseen errors, 13 samples were eventually removed.

Semen collection

Semen was obtained through masturbation and collected in a sterile wide-mouth container. The specimen was then placed in an incubator at a temperature of 37 °C and gently mixed. After the semen had liquefied, macroscopic and microscopic measurement was done as recommended by WHO.[15] Additionally, the evaluation of morphologically normal sperm was performed based on the instructions provided in the WHO manual because computer aided semen analysis (CASA) was not offered in the laboratories when the research done. [16]

Platelet rich plasma preparation

Blood from participants was processed to extract PRP.[17] Eight mL of venous blood yielded 2-3 mL of PRP [15] blood-drawing recommendations were followed for PRP procedures. Protocol specificity, equipment, and platelet concentration determined PRP centrifugation speed and time. This study used conventional criteria of 3200 RPM for 15 mins. Separated PRP were frozen at -8 °C.[18]

Preparation of Gg

To generate the working solution, 10 mg of Gg aqueous extract was added to 10 mL Phosphate Buffered Saline (PBS) in plastic plain tubes with broad- spectrum antibiotic (Ampicillin 0.004g) to prevent bacterial growth. The solution was filtered with a Millipore filter (0.22M) and adjusted to pH 7.2-7.4. Gg medium was then moved to a suitable storage container and stored in a cold, dark place away from sunlight and heat. [19]

Control media preparation

The ready-to-use medium FertiCult™ (FertiPro Company, Belgium) were also available. The medium was stored appropriately according to manufacturer instructions, typically in a refrigerator at the recommended temperature (4-8°C). For each sample, 4.5 mL of flushing medium was warmed in the incubator (37 °C) and used in the semen preparation and activation in vitro.

Density gradient centrifugation protocol

Sil-Select PlusTM gradients consist of silane-coated colloidal silica particles suspended in HEPES-buffered EBSS (Earle's balanced salt solution) supplemented

with 0.4-2.2 gm/L human serum albumin (medicinal substance derived from human blood plasma). Some product codes are supplemented with phenol red or 10mg/L gentamicin (medicinal substance). Our technique was done by adding 1 mL of 80% lower layer medium in a conical tube, followed by 1 mL of 40% upper layer medium (Sil-Select Plus™, FertiPro Company, Belgium) as a second layer, and then 1 mL of liquefied human semen. Four replicates of this initial step have been done with different superficial layers above the semen, as follows:

1. First tube: Density gradient media + liquefied semen + control medium (0.5 mL) Second tube: Density gradient media + liquefied semen + PRP (0.5 mL)
2. Third tube: Density gradient media + liquefied semen + Gg (0.5 mL)
3. Fourth tube: Density gradient media + liquefied semen + PRP and Gg (0.5+0.5 mL)

As soon as loading media and semen was finished, all tubes were moved to a centrifuge at 4000 rpm for 15 mins. Then the samples were prepared as described by Mohammed and his colleagues.[9] Immediately after the centrifugation process was completed, the supernatant layers of all tubes were removed and only the pellets were kept by using a plastic Pasteur pipette.

The next step in the protocol was the addition of flushing media (0.5 mL) to each tube and repeat centrifugation at 2000 rpm for 5 min. After the second centrifugation was accomplished, all supernatant layers from the tubes have also been discarded and replaced by 1 mL of Sil-Select Plus™ washing media (FertiPro Company, Belgium) above the final pellets. In the last step, tubes were placed in the incubator for 20 min, and then sperm parameters were assessed after activation by taking 10 µL droplets from each tube and analyzing them with a light microscope.

Normozoospermic men were enrolled in this study as a control group for the purpose of evaluating the efficiency of PRP, Gg, the new combination medium and the control medium. Semen samples of this group were activated and handled totally just as asthenoteratozoospermic individuals.

DNA integrity evaluation with Acridine Orange assay

According to Dutta and his colleagues.[20] Briefly, a 10 µL semen sample was smeared on the slide and air-dried for 20 mins. The slide was then treated with Carnoy's solution for 2-24 hrs. A brief drying period followed before staining. All stain preparations were done in the dark at room temperature with a pH of 2.5. 3-5 milliliters of dye were applied to the slide. The slide was carefully washed with distilled water and allowed to dry.[21] The slide was seen using a fluorescent microscope with a 40X objective lens the same day it was stained. Each participant's 5 slides of spermatozoa DNA were tested before and after activation, and the results were reported as green, red, or yellow.[10-15] Normal spermatozoa were green-fluorescent, whereas abnormal sperm heads were yellow red as appear in Figure 1.

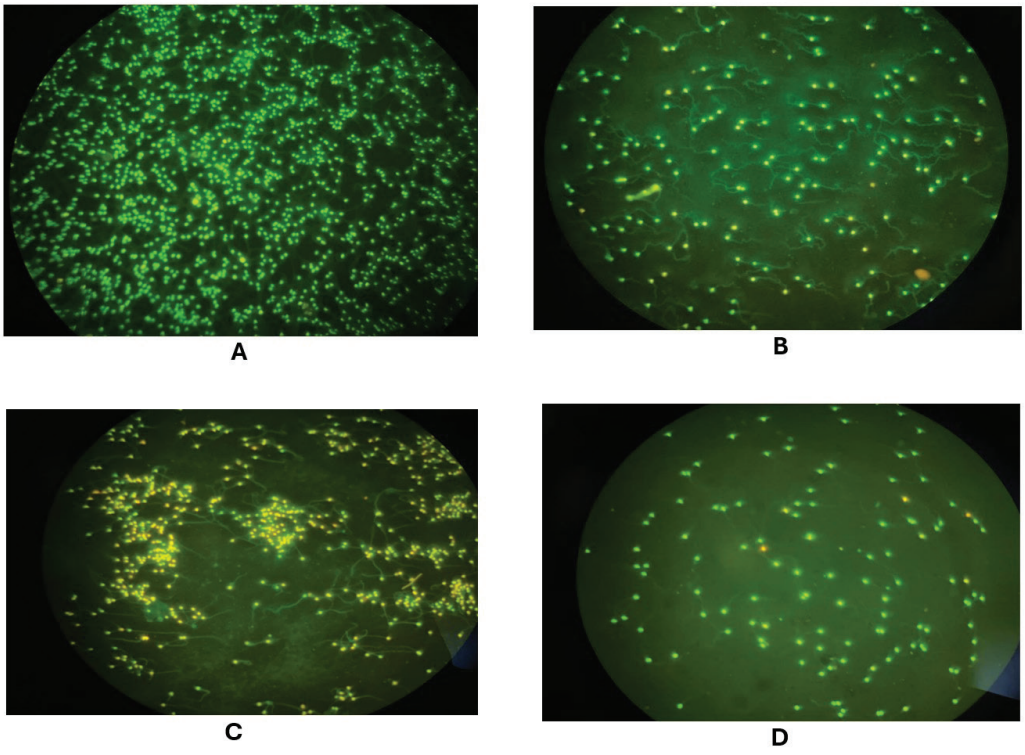


Figure 1. *DNA fragmentation assessment.*
A and B: DNA fragmentation assessment using Acridine Orange stain in normozoospermic men,
C and D: DNA fragmentation assessment using Acridine Orange stain in asthenoteratozoospermic men.

Statistical analysis

The data were analyzed using Statistical Package for Social Sciences (SPSS) version 23.0 and Microsoft Office 2010. The descriptive statistics including frequency, range, mean and standard error were measured to describe the data. The groups were compared by applying analysis of variance (ANOVA). The least significant test (LSD) was used to compare the results. $p \leq 0.05$ were considered statistically significant.

Ethical consideration

The current study was authorized by the ethical committee of the High Institute of Infertility Diagnosis and Assisted Reproductive Technologies/Al-Nahrain University, Baghdad, Iraq, with the approval code (0702-MF-2024N27). Informed consent was obtained from all participants involved in the study. Before gathering each sample, all participants provided their consent and were informed of the purpose of the questionnaire as well as the goal of the research.

Results

Basal (pre-activation) seminal fluids analysis parameters

The baseline seminal fluids analysis parameters were presented in Table 1., accordingly, the mean sperm concentration was 39.70±2.73 (10^6/mL), progressive motile sperms percent was 25.56±1.78, non-progressive motile sperms percent was 26.41±1.00, immotile sperm percent was 48.07±2.18, and morphologically normal sperm (MNS) percent was 22.70±1.32.

Comparison of certain sperm function parameters after activation between density gradient medium alone, Gg with and without PRP

Comparison of certain sperm function parameters after activation with density gradient medium was demonstrated in tables 2&3. In group 1 patients (Normozoospermia), the comparison between density gradient medium alone, density gradient mixed with Gg, density gradient mixed with PRP, and density gradient mixed with both Gg& PRP showed no significant ($p>0.05$) differences regarding sperm concentration, progressively motile sperms percent, non-progressively motile sperms percent, immotile sperms and DFI percentage respectively. However, there was significantly ($p=0.011$) higher MNS percent by using density gradient medium mixed with both Gg & PRP compared to using DG medium alone as shown in Table 2.

In Table 3., men with asthenoteratozoospermia had significantly ($p \leq 0.05$) higher sperm concentration (35.43±3.22), higher progressively motile sperms percent (73.24±1.77) and significantly ($p=0.031$) lower non- progressively motile sperms percent (15.80±1.15) after preparation by density gradient medium mixed with Gg; on the other hand, there was significantly ($p=0.032$) higher MNS(60.57±2.12) in density gradient media with PRP; however there were no significant ($p>0.05$) differences between all activation media regarding immotile sperms percent and DFI percent as presented in Table 3.

Table 1. Basal (pre-activation) seminal fluids analysis parameters and DFI (%).

Seminal fluids parameters	Range	Mean±SE
Volume (ml)	1.0-9.0	3.52±0.15
Sperm concentration (10^6/ml)	7.0 -116.0	39.70±2.73
Progressive motile sperms (%)	0.0-75.0	25.56±1.78
Non progressive motile sperms (%)	10.0-62.0	26.41±1.00
Immotile sperm (%)	5.0-85.0	48.07±2.18
Morphologically normal sperm (%)	5.0-75.0	22.70±1.32
DNA fragmentation index, DFI (%)	6.0-62.0	22.31±1.39

Table 2. Comparison of certain sperm function parameters after sperms activation by density gradient media alone and media contain *Gg,* PRP and mixed *Gg* with PRP in normozoospermic men.

Certain Sperm Parameters	DG alone	DG mixed with Gg	DG mixed with PRP	DG mixed with Gg and PRP	p value
Sperm concentration(m/mL)	33.57±4.90	24.14±3.21	38.93±7.24	22.21±5.46	0.106*
Progressive motile sperms (%)	74.21±2.72	81.07±3.03	77.21±3.30	78.57±3.29	0.470*
Non-progressivemotile sperms (%)	16.21±2.15	11.93±1.74	14.00±2.41	13.79±2.15	0.568*
Immotilesperm (%)	9.43±1.25	6.57±1.29	8.50±1.49	7.00±1.20	0.392*

Certain Sperm Parameters	DG alone	DG mixed with Gg	DG mixed with PRP	DG mixed with Gg and PRP	p value
Morphologically normal sperm (%)	62.36±2.59	70.93±2.11	71.43±2.86	73.29±1.99	0.011*
DNA fragmentation Index, DFI (%)	15.43±2.90	12.93±2.39	10.43±1.57	11.14±1.96	0.414*

Note: Data were analyzed with ANOVA and presented as mean±SE, *significant, DG: density gradient, Gg: glycyrrhiza glabra, PRP: platelet-rich plasma.

Table 3. Comparison of certain sperm function parameters after in vitro sperms activation by using density gradient media alone and media contain *Gg* with and without PRP in asthenoteratozoospermic men.

Certain Sperm Parameters	DG alone	DG mixed with Gg	DG mixed with PRP	DG mixed with Gg and PRP	*p* value
Sperm concentration (m/mL)	33.89±2.95	35.43±3.22	26.95±2.79	22.17±1.78	0.002*
Progressive motile sperms (%)	64.57±2.03	73.24±1.77	70.43±1.81	69.91±1.82	0.011*
Non-progressivemotile sperms (%)	21.41±1.55	15.80±1.15	18.74±1.35	19.22±1.22	0.031*
Immotilesperm (%)	13.48±0.87	10.61±0.86	11.26±0.95	10.85±0.95	0.103*
Morphologically normal sperm (%)	52.39±2.08	57.96±2.12	60.57±2.12	59.41±2.09	0.032*
DNA fragmentation Index, DFI (%)	15.09±1.24	13.89±1.09	12.48±1.22	13.07±1.28	0.458

Note: Data were analyzed with ANOVA and presented as mean±SE, *significant, DG: density gradient, Gg: glycyrrhiza glabra, PRP: platelet-rich plasma.

Discussion

The findings of present work demonstrated a significant enhancement in certain sperm parameters besides a highly significant reduction in DNA fragmentation index after *in vitro* stimulation using a single technique (Density gradient) that was applied to three distinct media types, plus as well as using it alone as a control method.

The comparative between basal semen parameters before activation as appear in Table 1. and semen parameters after activation in both studied groups as shown in Table 2. and Table 3. and Figure 2. and Figure 3., demonstrated that density gradient alone reduced sperm concentration, non-progressive motility, and immotility and increased progressive motility and morphologically normal sperm following activation.

In 2023, similar results were obtained by other study,[21] when used density gradients to isolate a considerable amount of structurally intact and actively moving human sperm from semen fluid. Waste, non-functional sperm, and non-sperm cells have been eliminated from these separated sperm.[22]

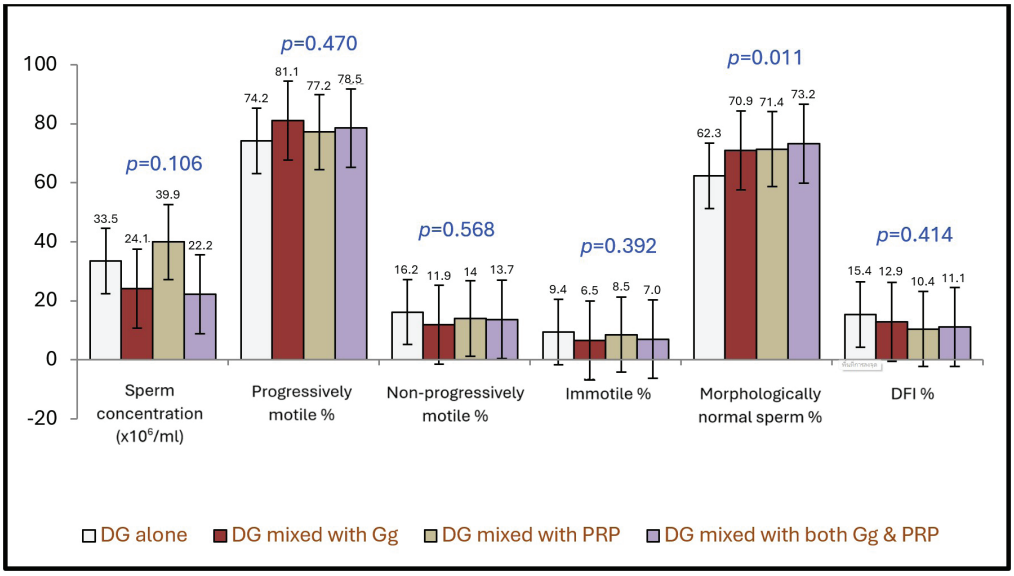


Figure 2. *Comparison of parameters in normozoospermic semen sample following the activation under various materials.* ***Open square:*** *density gradient media alone,* ***red square:*** *media contain glcyrrhiza glabra (Gg),* ***green square:*** *mixed with PRP,* ***purple square:*** *mixed with glcyrrhiza glabra and PRP.*

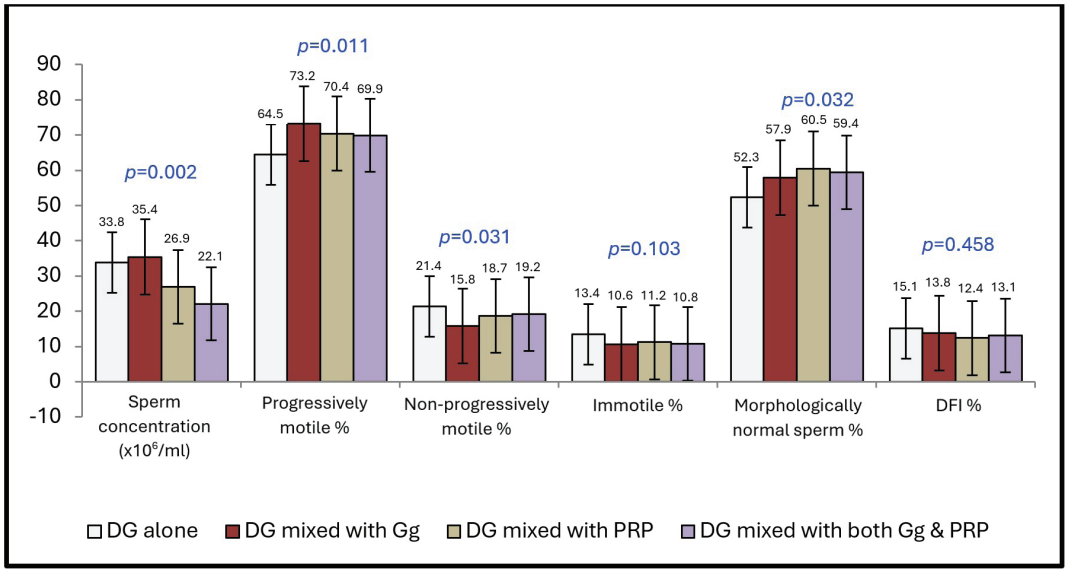


Figure 3. *Comparison of parameters in asthenoteratozoospermic semen sample following the activation under various materials.* ***Open square:*** *density gradient media alone,* ***red square:*** *media contain glcyrrhiza glabra (Gg),* ***green square:*** *mixed with PRP,* ***purple square:*** *mixed with glcyrrhiza glabra and PRP.*

Sperm activated with density gradient media and Gg have significantly improved progressive motile sperm and MNS percent, mainly because glabridin, a phytoestrogenic compound that binds to human estrogen receptors, has estrogenic activity. Estrogens increase sperm motility, grade activity, and hyperactivity.[19] Furthermore, Gg also contains calcium ions (Ca^{++}), potassium, glucose, fructose, vitamin E, vitamin C, zinc ions (Zn^{++}), sucrose, and amino acids. These nutrients boost sperm motility and forward migration *in vitro*.[23] Gg increases Ca^{++} in sperm medium to inhibit the phosphate diestrase enzyme, which limits cAMP breakdown, which increases sperm hyperactivation and motility.[9]

Sperms activated with density gradient media and PRP showed remarkable improvements in motility, morphology, and DFI. The various bioactive components of PRP are responsible for its benefits.[24] It has been noticed that PRP positively effects on goat semen parameters and found similar statistical findings for sperm motility and morphology improvement but a different finding regarding sperm concentration, which was significantly increased[25] whereas decreased in our findings. In the current study the reduction of sperm concentration is explained by the fact using the density gradient methodology which is known to lowers sperm concentration.[26] Like the other approaches, Gg coupled with PRP improves sperm parameters, however there is a slight difference in counts. This suggests that the newly developed medium did not enrich sperm parameters, particularly motility, as expected. In contrast, using Gg alone produces the maximum hyperactivation rating, as does PRP alone. Multiple researches suggested that PRP and Gg mode of intervention caused ineffective results. It has been postulated that platelet activation releases key granule components like P-selectin, ADP/ATP, GFs, and Ca^{++}.[27] In Gg, Glabridin may considerably inhibit activated platelet ATP release. Intracellular Ca^{++} increases platelet aggregation, while glabridin dramatically reduces it.

The first group i.e., normozoospermic was represented by thirty-four men with overall normal sperm parameters, including normal active motility and normal morphological sperm. Among all the certain sperm parameters that were analyzed after activation in the fourth culture media, there was no significant variation in the performance between these media across all parameters, with only one exception related to a significantly higher enhancement in sperm morphology when processed and cultured with both Gg and PRP. In contrast to the earlier studies, which demonstrated that Gg blocks platelet activity.[27] A study,[28] implies that the combination of PRP and Gg therapies in vivo has shown potential regenerative and healing effects, opening new therapy options.

The second group (asthenoteratozoospermic) expressed the largest population, with forty-six men. In addition, this group also exhibited the most diversity of variance in the decrease and increase of certain sperm function parameters throughout all activation media. Thus, after activation, this diversity involved significantly ($p \leq 0.05$) higher sperms concentration, higher progressively motile sperms percent and significantly ($p=0.031$) lower non- progressively motile sperms percent (15.80±1.15) at density gradient media with Gg, which confirmed the remarkable beneficial effects of Gg and the marvelous activation performance in the worst criteria group. However, according to Table 3, there were no significant differences ($p>0.05$) in terms of immotile sperms percent or DFI percent among all activation media. In the density gradient medium with PRP, there were significantly more morphologically

normal sperms than others.

The DFI percent were nearly equivalent in all media; however, for each specific method, the percent were statistically significantly decreased when comparing post-activation data to pre-activation data. These values indicate the high positive effectiveness of all methods that were used in eliminating the damaged cells and therefore reducing sperm DNA fragmentation.

Limitations

One of the study's limitations is its relatively small sample size, which would restrict how broadly the findings can be applied. Furthermore, the analysis was predicated on measurements and methodologies, which could cause biases in the measurements.

Conclusion

It is concluded that the application of Gg combined with PRP medium, along with the DG approach, demonstrates a significant enhancement in certain sperm function parameters as well as lowering DFI, in comparison to before activation in both groups of participants involved in this study. However, this combination is not as effective as using PRP medium alone or Gg medium alone.

Ethical approval

The study was approved by the Ethical Approval Committee with the approval code (0702-MF-2024N27).

Funding

There is no financial disclosure

Conflict of Interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Saad S. Al-Dujaily: designed, monitored, evaluated, and analyzed the result of the study; **Noor K. Khudhur**: performed the study; **Laith Amer Al-Anbary**: reviewed the article. All authors approved the final manuscript and take responsibility for the integrity of the data.

Acknowledgements

We would like to acknowledge the High Institute of Infertility Diagnosis and Assisted Reproductive Technologies/Al-Nahrain University, and Al-Harithiya International Lab Baghdad, Iraq.

References

[1] Raad G, Bakos HW, Bazzi M, Mourad Y, Fakih F, Shayya S, *et al.* Differential impact of four sperm preparation techniques on sperm motility, morphology, DNA fragmentation, acrosome status, oxidative stress, and mitochondrial activity: A prospective study. Andrology. 2021; 9(5): 1549-59. doi.org/10.1111/andr.13038.

[2] Agarwal A, Majzoub A, Baskaran S, Panner Selvam MK, Cho CL, Henkel R, *et al.* Sperm DNA fragmentation: A new guideline for clinicians. World J Men's Health. 2020; 38(4): 412-71. doi.org/10.5534/wjmh.200128.

[3] Pinto S, Carrageta DF, Alves MG, Rocha A, Agarwal A, Barros A, *et al.* Sperm selection strategies and their impact on assisted reproductive technology outcomes. Andrologia. 2020; *53*(2): e13725. https://doi.org/10.1111/and.13725

[4] Oseguera-López I, Ruiz-Díaz S, Ramos-Ibeas P, Pérez-Cerezales S. Novel techniques of sperm selection for improving IVF and ICSI outcomes. Front Cell Dev Biol. 2019; 7: 298 doi.org/10.3389/fcell.2019.00298.

[5] Natali I. Artificial insemination in farm animals. Google Books. 2025 [cited 2025 Jun 16]. Available from: https://books.google.iq/books?hl=en&lr=&id=ceGdDwAAQBAJ&oi=fnd&pg=PA115&dq=(5)%09Natali+I.+Sperm+preparation+techniques+for+artificial+insemination-comparison+of+sperm+washing

[6] El-Saber Batiha G, Magdy Beshbishy A, El-Mleeh A, M. Abdel-Daim M, Prasad Devkota H. Traditional Uses, Bioactive chemical constituents, and pharmacological and toxicological activities of Glycyrrhiza glabra L. (Fabaceae). Biomolecules. 2020; 10(3): 352. doi.org/10.3390/biom10030352.

[7] Ahmadi B, Bashir K, Rauf A, Hussain A, Herrera-Calderon O, Naseer Mabkhot Y. Cytotoxic effect of Picea smithiana (wall) Boiss on human tumor cell lines [Internet]. Googleusercontent.com. 2015 [cited 2025 Jun 16]. Available from: https://scholar.googleusercontent.com/scholar?q=cache:hWPGOI2TKbcJ:scholar.google.com/+(7)%09Ahmadi+B

[8] Pastorino G, Cornara L, Soares S, Rodrigues F, Oliveira MBPP. Liquorice (Glycyrrhiza glabra): A phytochemical and pharmacological review. PTR. 2018; 32(12): 2323-39. doi.org/10.1002/ptr.6178.

[9] Ethman Hashem M, Abu-El-Naga a, Aldujaily SS. Effect of Glycyrrhiza Glabra as motility stimulant on sperm function for patients with varicocele and diabetes before and *after in vitro* preparation for ICSI. Int J Pharm Res. 2020; 12(4). doi.org/10.31838/ijpr/2020.12.04.436.

[10] Al-Dujaily S, Selman M, Al-Hadithy R, Abdul-Ghafoor H. Impact of *in vitro* preparation with Citrullus Colocynthis on DNA normality of human Asthenozoospermic semen. IJEIR. 2015; 5: 36-45. Available from: https://web.archive.org/web/20220304070355id_/http://www.ijeir.nahrainuniv.edu.iq/images/attach/attach5af9574cc106d.pdf

[11] Wahab S, Annadurai S, Abullais SS, Das G, Ahmad W, Ahmad MF, *et al.* Glycyrrhiza glabra (Licorice): A comprehensive review on Its phytochemistry, biological activities, clinical evidence and toxicology. Plants. 2021; 10(12): 2751. doi.org/10.3390/

plants10122751.

[12] Franco D, Franco T, Schettino AM, Filho JMT, Vendramin FS. Protocol for obtaining platelet-rich plasma (PRP), platelet-poor plasma (PPP), and thrombin for autologous use. Aesthet Plast Surg. 2012; 36(5): 1254-9. doi.org/10.1007/s00266-012-9957-3.

[13] Hamdan DA, Rahim AI, Ula. *In Vitro* studying the effect of adding autologous platelet rich plasma (PRP) to the human semen on the sperm DNA integrity. Iraqi J Embryos Infertil Res. 2021; 11(2): 90-100. doi.org/10.28969/IJEIR.v11.i2.r7.

[14] Bader R, Ibrahim JN, Moussa M, Mourad A, Azoury J, Azoury J, *et al*. *In vitro* effect of autologous platelet-rich plasma on H_2O_2-induced oxidative stress in human spermatozoa. Andrology. 2019; 8(1): 191-200. doi.org/10.1111/andr.12648.

[15] World Health Organization. WHO laboratory manual for the examination and processing of human semen. (No Title) [Internet]. 2021 [cited 2025 Jun 16]. Available from: https://cir.nii.ac.jp/crid/1130282268700921472

[16] WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction [Internet]. Google Books. 2025 [cited 2025 Jun 16]. Available from: https://books.google.iq/books?hl=en&lr=&id=dEfWhZZcC0AC&oi=fnd&pg=PR8&ots=elfK1fhdYT&sig=gQSv8TCbXCDiFa7tZ430LHVDKBg&redir_esc=y#v=onepage&q&f=false

[17] World. WHO guidelines on drawing blood: best practices in phlebotomy. http://wwwwhoint/infection-prevention/publications/drawing_blood_best/en/ [Internet]. 2025 [cited 2025 Jun 16]. Available from: https://cir.nii.ac.jp/crid/1370286996906335104

[18] Croisé B, Paré A, Joly A, Louisy A, Laure B, Goga D. Optimized centrifugation preparation of the platelet rich plasma: Literature review. Journal of Stomatology, Oral and Maxillofacial Surgery. 2020 Apr;121(2):150–4. https://doi.org/10.1016/j.jormas.2019.07.001

[19] Al-Dujaily S, Mohammed S. DNA normality after *In Vitro* preparation with pentoxifylline and glycyrrhiza glabra of cryopreserved human sperms. World J Pharm Res. 2015; 4(4): 7105-277. Available from: https://wjpr.s3.ap-south-1.amazonaws.com/article_issue/1428406344.pdf

[20] Dutta S, Henkel R, Agarwal A. Comparative analysis of tests used to assess sperm chromatin integrity and DNA fragmentation. Andrologia. 2020; 53(2): e13718. doi.org/10.1111/and.13718.

[21] de N, Grasiele C, Gean Pier Panizzon, Motta P, Vânia Cibele Minguetti Câmara, Martins R, *et al*. Comparative sperm recovery rate after density gradient centrifugation with two media for *in vitro* fertilization. JBRA. 2022; 27(1): 25-28. https://doi.org/10.5935/1518-0557.20220008

[22] Dai C, Zhang Z, Shan G, Chu LT, Huang Z, Moskovstev S, *et al*. Advances in sperm analysis: techniques, discoveries and applications. Nat Rev Urol. 2021: 18(8): 447-67. https://doi.org/10.1038/s41585-021-00472-2

[23] Hasan MdK, Ara I, Mondal MSA, Kabir Y. Phytochemistry, pharmacological activity, and potential health benefits of Glycyrrhiza glabra. Heliyon. 2021; 7(6): e07240. doi.org/10.1016/j.heliyon.2021.e07240.

[24] Magalon J, Bausset O, Serratrice N, Giraudo L, Aboudou H, Veran J, *et al*. Characterization and comparison of 5 platelet-rich plasma preparations in a single-donor model. Arthroscopy. 2014; 30(5): 629-38. doi.org/10.1016/j.arthro.2014.02.020.

[25] Kashkool MH, AL-Delemi DHJ. Testicular injection of autologous platelet rich plasma (PRP) to improve of semen parameters of goats [Internet]. Googleusercontent.com. 2023 [cited 2025 Jun 16]. Available from: https://scholar.googleusercontent.com/scholar?q=cache:3HoGrvIzJZkJ:scholar.google.com/+(25)%09Kashkool+MH

[26] Al-Dujaily S, Al-Jubori W. Gender selection by intra-uterine insemination for infertile couples. World J Pharm Res. 2015; 4(8): 309-16. Available from: https://wjpr.s3.ap-south-1.amazonaws.com/article_issue/1438320555.pdf

[27] Seung Na Ko, Ji Won Son, Gyu Ri Kim, Min Seon Kim, Yea Jin Lee, Seung Ju Kim, *et al*. The inhibitory effects of Glycyrrhiza uralensis on human platelet aggregation and thrombus formation. Biomed Sci Lett. 2023; 29(4): 242-8. doi.org/10.15616/BSL.2023.29.4.242.

[28] Shafik NM, Rasha Osama El-Esawy, Mohamed DA, Deghidy EA, Omnia Safwat El-Deeb. Regenerative effects of glycyrrhizin and/or platelet rich plasma on type-II collagen induced arthritis: Targeting autophay machinery markers, inflammation and oxidative stress. Arch Biochem Bophys. 2019; 675: 108095–5. doi.org/10.1016/j.abb.2019.108095.

Journal of Associated Medical Sciences 2026; 59(1): 235-246

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

Assessment of ACR phantom image quality in mammography using multiple Deep learning models

Thunyarat Chusin[1,2], Supitcha Saengkaew[1], Suradet Aunnoi[1], Amarin Thuikham[1], Pratchayakan Hompeng[1,3], Siriprapa Kaewjaeng[4,5], Titipong Kaewlek[1,2*]

[1]*Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*
[2]*Interdisciplinary Health and Data Sciences Research Unit, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*
[3]*Regional Medical Sciences Center 7 Khonkaen, Khonkaen Province, Thailand.*
[4]*Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
[5]*Center of Radiation Research and Medical Imaging, Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*

ARTICLE INFO

Article history:
Received 16 August 2025
Accepted as revised 14 December 2025
Available online 17 December 2025

Keywords:
Mammography, ACR phantom, image quality assessment, deep learning, convolutional neural networks, artificial intelligence.

ABSTRACT

Background: Evaluating image quality in mammography—particularly using American College of Radiology (ACR) phantom images—is essential for maintaining diagnostic accuracy. Conventional evaluation relies on human visual inspection, which is prone to variability due to individual perception differences.

Objectives: This study examined the capability of multiple convolutional neural network (CNN)-based artificial intelligence (AI) models to assess the quality of ACR phantom images and address the limitations of human-based evaluation.

Materials and methods: Five CNN-based models—LeNet5, AlexNet, VGG19, GoogLeNet, and ResNet50—were used to classify 231 ACR phantom images acquired under different exposure settings. Dataset augmentation was performed by adding and removing artificial noise, increasing the dataset to 1,617 images. The dataset was then divided into training (70%), validation (10%), and testing (20%) subsets. Model performance was compared based on phantom image scoring.

Results: GoogLeNet showed the highest performance in evaluating fiber and mass groups, whereas ResNet50 achieved the best results for the speck group. The classification accuracy for scoring 16 object positions in ACR phantom images ranged from 31% to 100%, depending on object size.

Conclusion: To reliably classify acceptable mammographic image quality, model performance in detecting borderline cases (score 0.5) must improve. For clinical applicability, accuracy should exceed 80%.

*** Corresponding contributor.**
Author's Address: *Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*
E-mail address: titipongk@nu.ac.th
doi: 10.12982/JAMS.2026.029
E-ISSN: 2539-6056

Introduction

Mammography is a widely used imaging modality that employs low-dose X-rays for breast cancer screening and diagnosis. Women aged 40 and older are typically advised to undergo mammographic screening every 1-2 years.[1] Early detection through mammography reduces disease severity and mortality rates.[2] This highlights the importance of maintaining high image quality for effective breast cancer screening.

Consequently, dependable imaging systems and standardized image quality evaluation are essential components of quality control (QC) in mammography.[3]

Consistently high image quality is critical for accurate diagnosis. Routine quality control procedures are performed to ensure optimal system performance and to reduce diagnostic errors. A standard approach involves phantom testing using American College of Radiology (ACR) phantoms. These tests evaluate key image quality parameters, including density, contrast, uniformity, and artifact presence. However, these evaluations are generally performed through human visual inspection, which is time-consuming and inherently subjective. Variations in visual perception and professional experience can introduce inconsistencies in the results.

Artificial intelligence (AI) has increasingly transformed medical imaging. Deep learning methods, particularly convolutional neural networks (CNNs), have gained prominence with architectures such as LeNet, AlexNet, VGGNet, GoogLeNet, and Residual Network (ResNet). LeNet was originally designed for handwritten digit classification using the Modified National Institute of Standards and Technology database of 28×28 grayscale images.[4] Its relatively simple architecture is makes it suitable for low-complexity datasets and lightweight applications, including blood cell classification,[5] pneumonia detection,[6] and breast cancer histopathology analysis.[7] AlexNet demonstrated the power of deep CNNs in large-scale image classification.[8] Its innovations—such as deeper convolutional layers, ReLU activation, and dropout regularization—improved learning efficiency and reduced overfitting. In medical imaging, AlexNet has been applied to tasks such as chest disease classification (e.g., tumor vs. normal in chest images)[9] and brain tumor differentiation using magnetic resonance imaging (MRI).[10] VGGNet uses deep architectures composed of small 3×3 convolutional filters.[11] This structure allows effective hierarchical feature learning by stacking multiple layers. VGGNet has been widely adopted in transfer learning for medical imaging tasks, including breast cancer histopathology classification,[12] mammography interpretation,[13] and ultrasound image analysis,[14] and image quality assessment (IQA),[15] such as binary or multi-class scoring of ACR phantom images. GoogLeNet introduced Inception modules that integrate multiple convolutional filter sizes (1×1, 3×3, and 5×5) within a single module.[16] This design enables efficient multi-scale feature extraction with fewer parameters. As a result, GoogLeNet is effective in detecting microcalcifications and small lesions in mammograms[17] and is suitable for real-time classification due to its computational efficiency. ResNet introduced residual connections to address the vanishing gradient problem, allowing the training of very deep networks.[18] It has shown strong performance in complex medical image analysis,[11] including MRI-based studies,[19] cancer subtype classification from biopsy images,[20] and IQA tasks such as resolution assessment and phantom image scoring.[21]

Beyond abnormality detection in radiographic imaging,[22,23] CNNs have also demonstrated strong performance in radiological image quality assessment, including general radiography and mammography.[21,24-26] In one study, eight CNN architectures with 3–10 convolutional layers were trained to detect structures in ACR phantom images, achieving up to 95% classification accuracy using a six-layer model.[21] Another study modified the VGG16 architecture for ACR image scoring, reporting an F1-score of 0.69 for multi-class classification and an F1-score of 0.93 with an area under the receiver operating characteristic curve (AUC) of 0.97 for binary classification.[24]

This study aimed to evaluate and compare the performance of different CNN architectures in assessing ACR phantom image quality. Understanding their capabilities will support the development of AI-assisted tools that help reduce evaluator variability and improve consistency in mammography QC.

Materials and methods

Data collection

Phantom images were acquired using the Planmed Sophie Classic mammography system (Planmed Oy, Finland) equipped with a digital image receptor and the ACR accreditation phantom (Mammo 156 phantom, Gammex, USA). The phantom contains three test groups—fiber, specks, and masses—comprising 16 test objects, as illustrated in Figure 1. These objects differ in shape, diameter, and thickness and are designed to simulate common breast lesions observed in clinical mammograms. Specifically, the phantom includes 6 fibers (F1-F6) with diameters of 1.56, 1.12, 0.89, 0.75, 0.54, and 0.40 mm; 5 speck groups (S1-S5), each containing 6 specks with diameters of 0.54, 0.40, 0.32, 0.24, and 0.16 mm; and 5 mass objects (M1-M5) with decreasing diameters and thicknesses of 2.00, 1.00, 0.75, 0.50, and 0.25 mm.[27]

The phantom images were acquired using different tube voltage and tube current–time product settings controlled by the automatic exposure control system. Tube voltages ranged from 25 to 30 kVp (in 1 kVp increments), while tube current–time products ranged from 10 to 303 mAs, corresponding to seven optical density levels, as shown in Table 1. Each imaging condition was repeated three times. In total, 231 phantom images were collected, covering a wide range of image quality levels, from optimal (27 kVp, 71 mAs) to both lower and higher extremes.

Table 1. Number of ACR phantom images captured from the Planmed Sophie Classic mammography system with various exposure techniques.

Target/filter combination	Exposure techniques		Number of images
	kVp	mAs	
Mo/Mo	25	18-303	36
	26	10-225	39
	27	10-161	39
	28	10-114	39
	29	10-85	39
	30	10-67	39
Total			231

Note: Mo: molybdenum

Image augmentation

Artificial Gaussian noise was added to, and subsequently removed from, the original images using a median filter in ImageJ version 1.54g. This process was used to expand the dataset. Three empirically selected noise levels-10%, 30%, and 100%—were applied to simulate low, medium, and high noise conditions, respectively. In this context, the percentages represent the relative standard deviation parameter of the Gaussian noise within the software and do not be correspond to calibrated detector noise levels or specific clinical exposure settings.

For each of the 231 original images, three noise levels and two processing conditions (noisy and median-filtered) were generated, producing 1,386 augmented images (231×3×2). When combined with the original images (N=231), the total dataset consisted of 1,617 images. The dataset was then randomly divided into training (N=1,134; ~70%), validation (N=161; ~10%), and test (N=322; ~20%) subsets, as summarized in Table 2.

Table 2. Number of images used in the training, validation, and testing subsets.

	Number of images			
	ACR phantom	6 Fibers	5 Specks	5 Masses
Training	1,134 (~70%)	6,804	5,670	5,670
Validation	161 (~10%)	966	805	805
Testing	322 (~20%)	1,932	1,610	1,610
Total	1,617 (100%)	9,702	8,085	8,085
			25,872	

Image quality scoring by evaluators

This study was approved by the Institutional Review Board (IRB No. P1-0091/2567). All images in the training and validation subsets were scored by consensus between two researchers, while the testing subset was scored by consensus between two medical physicists. Each of the 16 phantom objects was assigned a score of 0, 0.5, or 1 according to the ACR digital mammography phantom scoring criteria.[28] All images were displayed on a 2-megapixel monitor, and the window width and level were adjusted for optimal viewing conditions based on evaluator judgment to ensure consistent and accurate scoring.

Image quality scoring by CNN-based AI models

All scored images in the training, validation, and testing subsets were cropped at the 16 predefined object locations, as shown in Figure 1. From each phantom image, 16 object-level images were extracted. This resulted in 18,144 training images (1,134×16), 2,576 validation images (161×16), and 5,152 testing images (322×16), as summarized in Table 2. Each object-level image was classified into one of three score categories (0, 0.5, 1) based on the corresponding image quality score. Five convolutional neural network (CNN)-based models—LeNet5,[4] AlexNet,[8] VGG19,[11] GoogLeNet,[16] and ResNet50[18]—were trained and validated for each of the 16 test objects. The LeNet5 model was trained from scratch using random weight initialization. In contrast, AlexNet, VGG19, GoogLeNet, and ResNet50 were initialized using pre-trained ImageNet weights and then fine-tuned on the phantom image dataset. This transfer learning strategy was employed to improve convergence of the deeper networks considering the

relatively small dataset size. All CNN models were trained on Google Colab with a Tesla T4 GPU. A learning rate of 0.000001 and a batch size of 32 were applied consistently across all models. The number of training epochs was set to 20 for VGG19, 20 for GoogLeNet, 30 for AlexNet, 50 for ResNet50, and 60 for LeNet5. Training and validation loss and accuracy curves were generated for each object and model. These curves were used to assess network convergence and to confirm the optimization of training parameters, including batch size and number of epochs. During inference, the trained CNN models predicted image quality scores of 0, 0.5, or 1 for each test object.

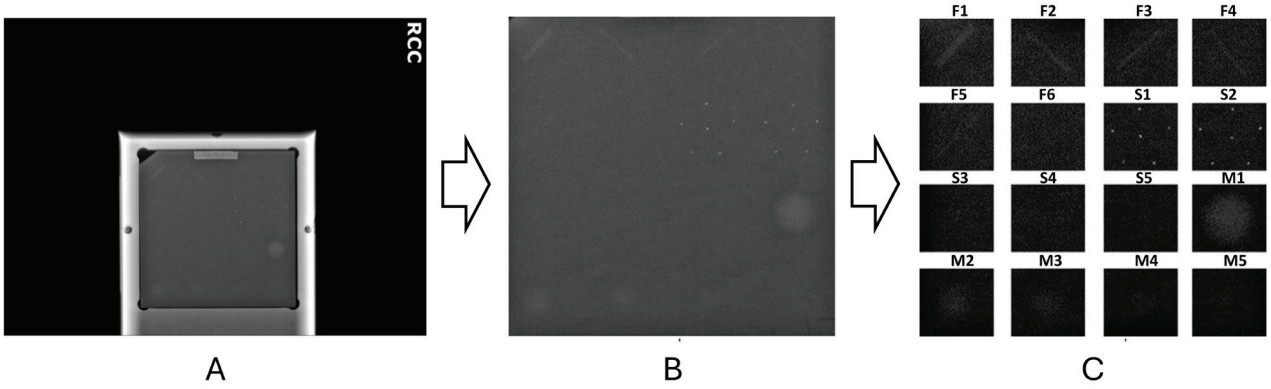


Figure 1. *A: original ACR phantom image, B: a cropped image of 1074x1044 pixels, C: A 268x261 pixel cropped image for the 16 test objects.*

Performance comparisons for CNN-based AI models

This study compared the performance of five CNN-based AI models: LeNet5, AlexNet, VGG19, GoogLeNet, and ResNet50. Model performance was evaluated using a multi-class confusion matrix, in which predicted scores were compared against three predefined image labels (0, 0.5, and 1) assigned by two medical physicists, as presented in Table 3.

Table 3. Analysis of the CNN-based AI models' performance based on multi-confusion matrix scoring.

Image scoring by CNN-based AI	Image scoring by medical physicists		
	1	0.5	0
1	TP_1, $TN_{0.5}$, TN_0	FP_1, $FN_{0.5}$, TN_0	FP_1, $TN_{0.5}$, FP_0
0.5	FN_1, $FP_{0.5}$, TN_0	TN_1, $TP_{0.5}$, TN_0	TN_1, $FP_{0.5}$, FP_0
0	FN_1, $TN_{0.5}$, FN_0	TN_1, $FN_{0.5}$, FN_0	TN_1, $TN_{0.5}$, TP_0

Where:

True Positive (TP): AI model correctly scored an image as 1 or 0.5, as marked by the medical physicists.
False Positive (FP): AI model incorrectly scored an image, assigning a higher score than the medical physicists.
False Negative (FN): AI model incorrectly scored an image, assigning a lower score than the medical physicists.
True Negative (TN): AI model correctly scored an image as 0, as marked by the medical physicists.

Performance of the five CNN-based AI models was then evaluated using accuracy, precision, sensitivity, specificity (recall), F1-score, and false positive rate (FPR), calculated using Equations (1) to (6) as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP+FP+FN+TN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (4)$$

$$\text{F1 - score} = \frac{2 \text{ x precision x sensitivity}}{\text{precision + sensitivity}} \quad (5)$$

$$\text{False Positive Rate} = 1 - \text{Specificity} \quad (6)$$

Equation (7) was used to calculate the percentage accuracy of each model. Unlike confusion matrix-based metrics, accuracy is calculated by comparing the number of correctly predicted images to the total number of images, without distinguishing between the FP and FN.

$$\%\text{Accuracy} = \frac{\text{Number of corrected prediction image}}{\text{Total number of images}} \quad (7)$$

Results

The performance of five CNN-based AI models—LeNet5, AlexNet, VGG19, GoogLeNet, and ResNet50—was evaluated using key statistical metrics: accuracy, precision, sensitivity, specificity (recall), F1-score, and FPR. These results are presented in Figures 2-5. The analysis focused on classifying fibers, specks, and masses in ACR phantom images.

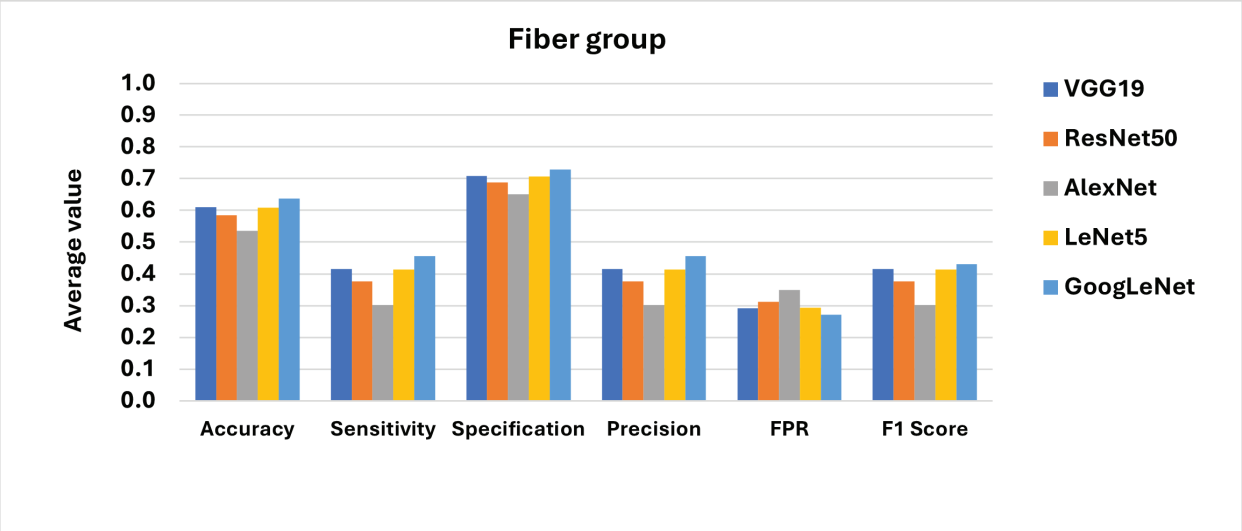


Figure 2. *Statistical metrics compared to all CNN-based AI models for classifying fiber objects.*

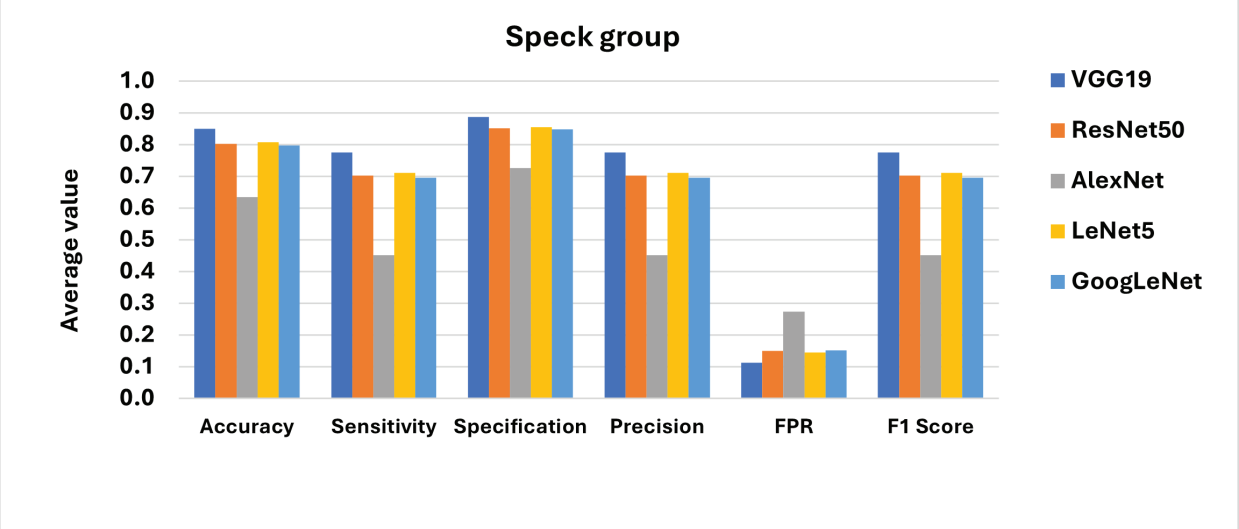


Figure 3. *Statistical metrics compared to all CNN-based AI models for classifying speck objects.*

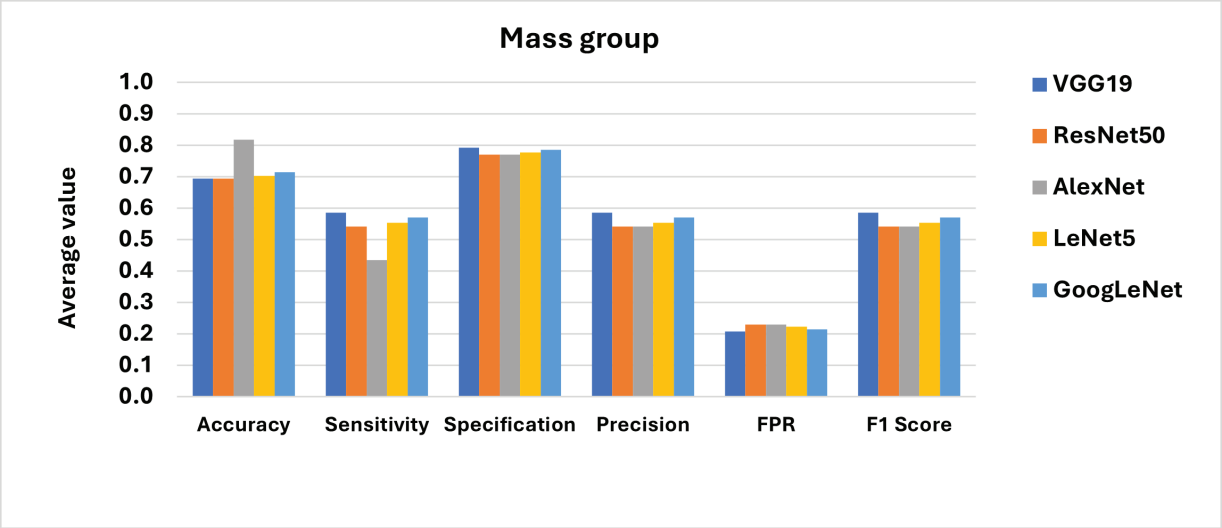
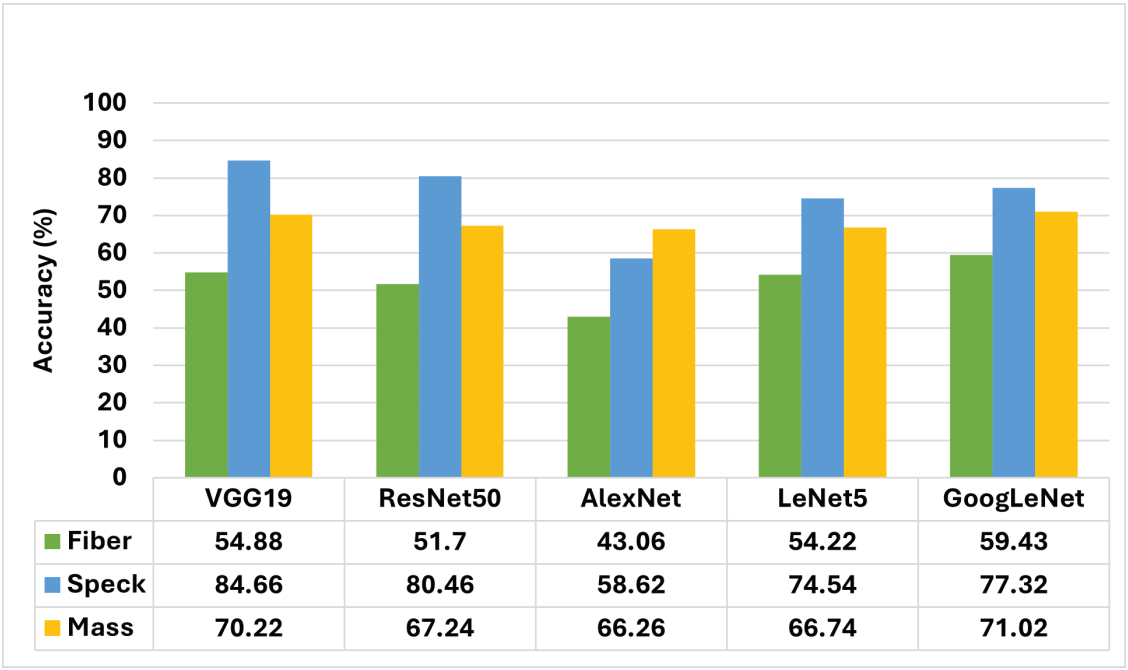


Figure 4. *Statistical metrics compared to all CNN-based AI models for classifying mass objects.*



	VGG19	ResNet50	AlexNet	LeNet5	GoogLeNet
Fiber	54.88	51.7	43.06	54.22	59.43
Speck	84.66	80.46	58.62	74.54	77.32
Mass	70.22	67.24	66.26	66.74	71.02

Figure 5. *Percentage accuracy compared to all CNN-based AI models for classifying fiber, speck and mass objects.*

The results showed that VGG19 and LeNet5 were unable to effectively differentiate object features, including size and density. Both models consistently produced the same output score of 1 for all test images, indicating a failure to distinguish among object classes. As a result, these two models were excluded as viable classifiers for fiber, speck, and mass objects in the phantom dataset.

The optimal CNN-based AI model for each object group was selected based on comparative statistical performance.

For the fiber group, GoogLeNet achieved the highest classification accuracy and underperformed the other models across most statistical metrics, while also demonstrating the lowest FPR. Therefore, GoogLeNet was identified as the most effective model for fiber evaluation.

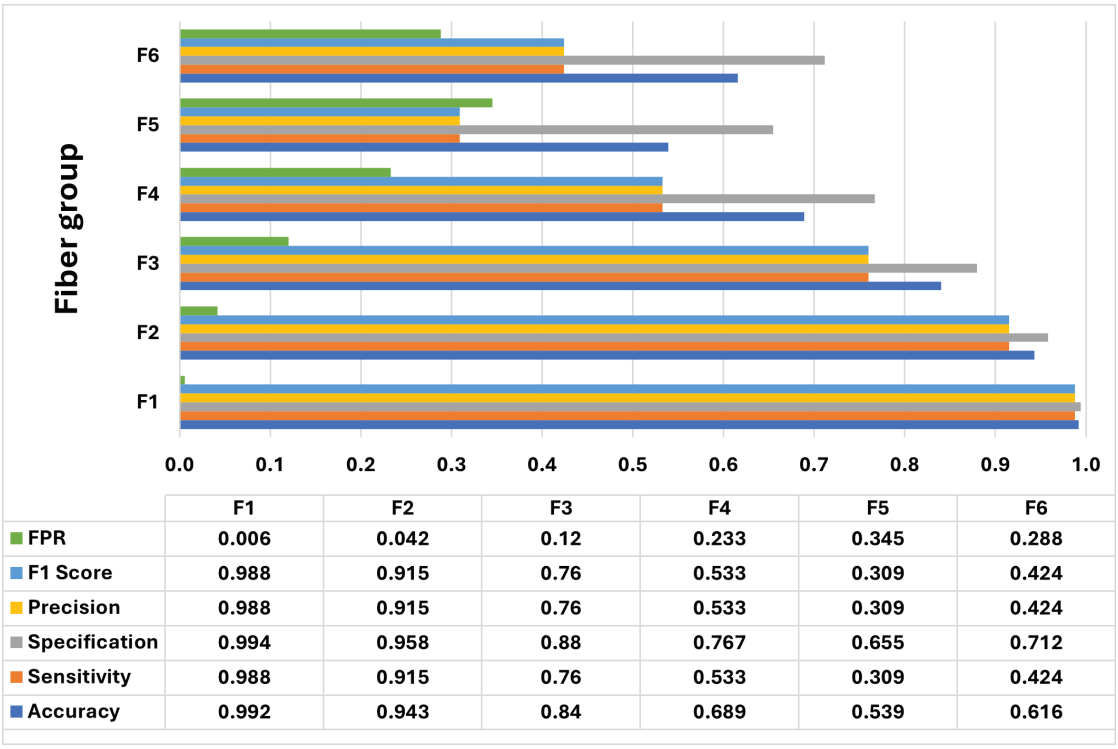
For the speck group, although VGG19 and LeNet5 produced high statistical values, their ability to differentiate object features (due to identical output predictions) limited their usability. In contrast, ResNet50 demonstrated the highest accuracy, superior performance across key metrics, and the lowest FPR. Thus, ResNet50 was selected as the best model for speck classification.

For the mass group, VGG19 and GoogLeNet exhibited similarly high statistical values. However, VGG19 was excluded due to its inability to distinguish object features, consistently outputting a score of 1. Although AlexNet achieved the highest accuracy, its performance across other statistical metrics was inferior to GoogLeNet. Therefore, GoogLeNet was determined to be the most suitable model for mass object evaluation because it combined the highest

accuracy with stronger overall metric performance and the lowest FPR.

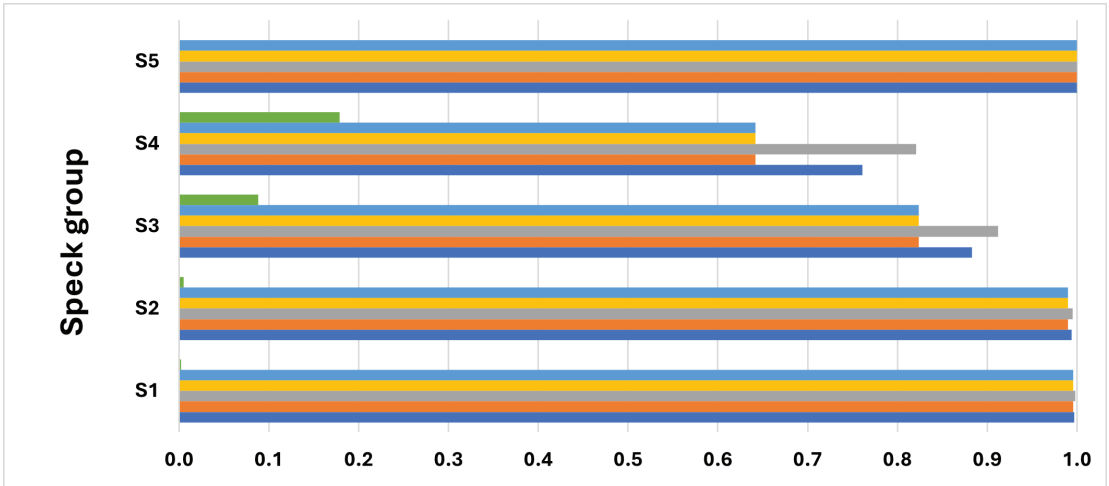
Figure 6 illustrates the performance of GoogLeNet in classifying fiber objects in ACR phantom images. In general, the FPR increased as fiber size and density decreased, except for F6, which had the smallest size and lowest density. In contrast, the other statistical metrics tended to decrease as fiber size and density decreased, except for F6. The GoogLeNet model demonstrated strong performance for F1, F2, and F3, with FPR values below 0.12 and other performance metrics exceeding 0.76.

Figure 7 shows the performance of the ResNet50 model in classifying speck objects. The model achieved FPR values below 0.09 and other statistical metrics greater than 0.82 for all speck groups, except S4.



	F1	F2	F3	F4	F5	F6
FPR	0.006	0.042	0.12	0.233	0.345	0.288
F1 Score	0.988	0.915	0.76	0.533	0.309	0.424
Precision	0.988	0.915	0.76	0.533	0.309	0.424
Specification	0.994	0.958	0.88	0.767	0.655	0.712
Sensitivity	0.988	0.915	0.76	0.533	0.309	0.424
Accuracy	0.992	0.943	0.84	0.689	0.539	0.616

Figure 6. *Performance of GoogLeNet model in classifying fiber objects in ACR phantom images.*

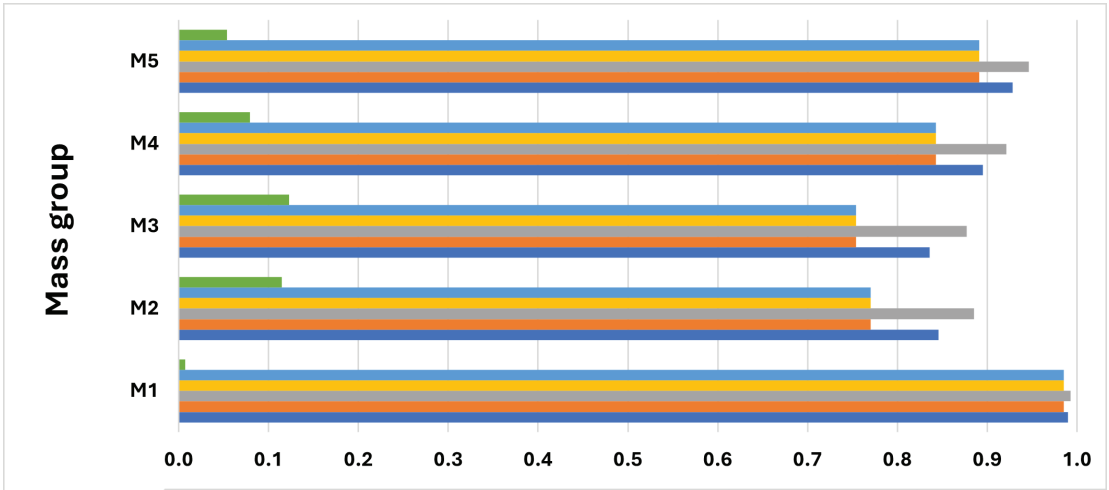


	S1	S2	S3	S4	S5
FPR	0.002	0.005	0.088	0.179	0.000
F1 Score	0.996	0.99	0.824	0.642	1.000
Precision	0.996	0.99	0.824	0.642	1.000
Specification	0.998	0.995	0.912	0.821	1.000
Sensitivity	0.996	0.99	0.824	0.642	1.000
Accuracy	0.997	0.994	0.883	0.761	1.000

Figure 7. *Performance of ResNet50 model in classifying speck objects in ACR phantom images.*

Figure 8 presents the performance of the GoogLeNet model in classifying mass objects. The FPR values were below 0.12, while the remaining statistical metrics exceeded 0.75 across all mass groups.

Figure 9 illustrates the classification accuracy of the GoogLeNet model for fiber and mass objects and the ResNet50 model for speck objects. The accuracy percentages for all 16 object positions were as follows: fibers-F1 (98.76%), F2 (91.51%), F3 (75.96%), F4 (53.33%), F5 (30.90%), and F6 (42.37%); specks-S1 (99.61%), S2 (99.03%), S3 (82.38%), S4 (64.20%), and S5 (100%); masses-M1 (98.53%), M2 (76.96%), M3 (75.37%), M4 (84.30%), and M5 (89.13%).



	M1	M2	M3	M4	M5
FPR	0.007	0.115	0.123	0.079	0.054
F1 Score	0.985	0.77	0.754	0.843	0.891
Precision	0.985	0.77	0.754	0.843	0.891
Specification	0.993	0.885	0.877	0.921	0.946
Sensitivity	0.985	0.77	0.754	0.843	0.891
Accuracy	0.99	0.846	0.836	0.895	0.928

Figure 8. *Performance of GoogLeNet model in classifying mass objects in ACR phantom images.*

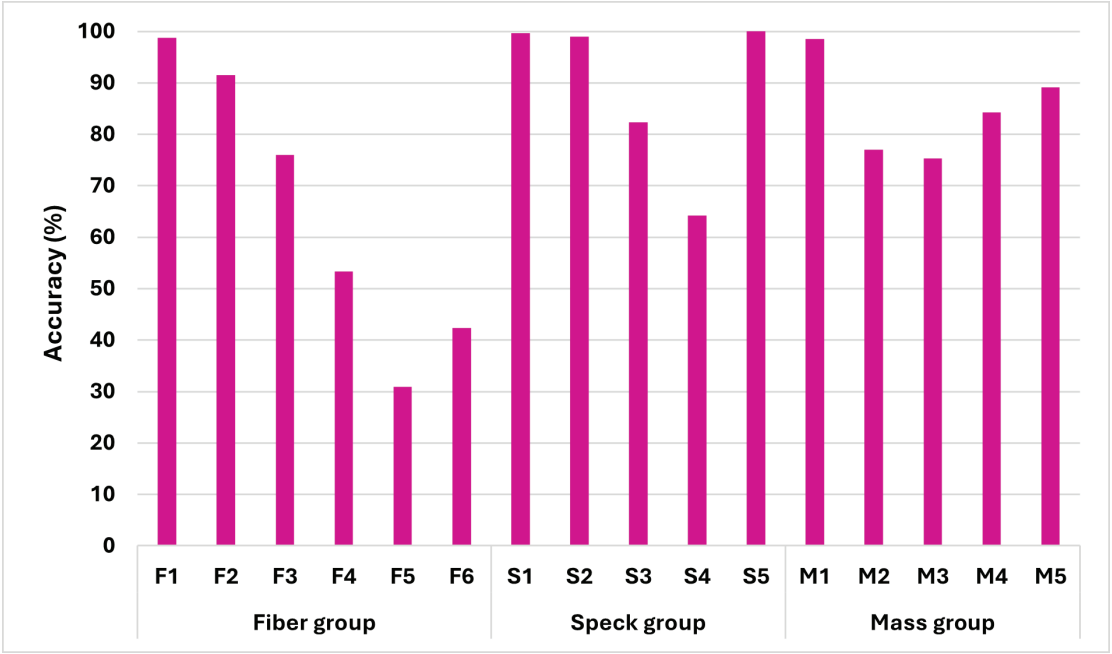


Figure 9. *Percentage accuracy of GoogLeNet in classifying fiber and mass objects, and the ResNet50 model in classifying speck objects.*

Discussion

This study evaluated the performance of five CNN-based AI models—LeNet5, AlexNet, VGG19, GoogLeNet, and ResNet50—in classifying fiber, speck, and mass objects in ACR phantom images. Model outputs were validated against object classifications performed by two medical physicists. The objects differed in shape, size, and density, which correspond to characteristics of typical lesions observed in clinical mammographic images.

VGG19 and LeNet5 failed to effectively differentiate object features because both models produced uniform outputs (score=1) across all test cases. This behavior indicates that the models did not learn meaningful decision boundaries. A possible cause is activation function saturation. Specifically, the original LeNet5 architecture uses the sigmoid activation function, which is prone to saturation. This leads to near-zero gradients, causing the vanishing gradient problem and limiting effective learning, particularly in deeper architectures. Furthermore, training VGG19 from scratch is known to be challenging. Issues such as improper weight initialization and unsuitable activation functions can prevent effective feature learning. These limitations may result in constant model outputs, indicating a failure to capture discriminative information from the input images.

According to ACR guidelines, mammographic image quality is considered acceptable when the visual image score reaches 4 or at least four fibers (F1-F4) are visible, a score of 3 or at least three specks (S1-S3) are identified, and a score of 3 or at least three masses (M1-M3) are detected, covering objects from large to small sizes. Based on the results of this study, the GoogLeNet model accurately classified F1, F2, and M1 with accuracy above 80%. The ResNet50 model effectively classified specks S1, S2, and S3, also exceeding 80% accuracy. These findings suggest that the prediction performance of GoogLeNet and ResNet50 must be improved, particularly for F3, F4, M2, and M3, to fully meet clinical image quality standards.

Although this study provides meaningful contributions, several limitations must be considered. First, all mammographic images were acquired using a single mammography system from one manufacturer with a Mo/Mo target/filter combination. Although data augmentation increased the dataset size, it did not add diversity in terms of vendors, system models, or beam qualities. Consequently, the reported performance of the CNN-based models may not generalize to phantom images acquired on other mammography systems or different target/filter combinations (e.g., Mo/Rh or Rh/Rh), which are commonly used in clinical quality control. To achieve robust and generalizable performance, the CNN models must be retrained or fine-tuned using phantom images acquired under clinically relevant exposure and target/filter configurations for each system.

A further limitation involves the imbalance in object score distribution. Images with an intermediate score of 0.5 were relatively scarce in both the training and test sets. This imbalance was most evident for smaller objects, which are less likely to be captured by the mammography system or reliably detected by human observers. Consequently, objects such as F4, F5, S4, and M4 often received intermediate scores of 0.5. In contrast, larger objects are more consistently captured and detected, leading to a score of 1 for F1, F2, M1, M2, S1, and S2. The smallest objects, although often not captured by the imaging system, were typically classified as non-visible and assigned a score of 0, as observed for F6, S5, and M5. Consequently, images with scores of 0 and 1 were more prevalent than those with a score of 0.5. In this study, scores of 1, 0.5, and 0 accounted for 56.5%, 19.0%, and 24.5% of all images, respectively, confirming that 0.5-score samples were significantly underrepresented.

All images, including those containing artificial Gaussian noise, were scored after data augmentation. The added noise can reduce the effective signal-to-noise ratio and, in some cases, shift clearly visible objects (score 1) into borderline cases (score 0.5), particularly for low-contrast fibers and masses. However, these borderline classifications were influenced not only by Gaussian noise but also by intrinsic object size and contrast, exposure conditions, and system noise, rather than by Gaussian noise alone. Another limitation is that data augmentation was limited to additive Gaussian noise followed by median filtering. Although this method directly targets quantum noise, it does not simulate other real-world variations, such as changes in positioning, geometric distortions, or subtle differences in object shape and contrast. Future studies should incorporate more advanced augmentation techniques, including small rotations and translations, elastic deformations, and controlled contrast or blur adjustments, to generate a more realistic and diverse training dataset.

As previously mentioned, CNN-based AI models have limited learning capacity when trained on under-represented image categories, particularly those assigned a score of 0.5. This class imbalance contributed to the reduced detection performance for F4, F5, S4, and M4. Increasing the number of images across all score categories, particularly for the 0.5 class, would likely improve model generalization and overall classification performance.

Sung Soo Park *et al.*[24] proposed a deep neural network-based phantom scoring method using the VGG16Net model for ACR phantom image quality evaluation. They applied two classification approaches: Multi-Class Classification (MCC), using scores of 0, 0.5, and 1, and Binary Classification (BCC), where scores of 0.5–1 were labeled as passing and 0 as failing. Their reported F1-score was 0.69 for MCC and 0.93 for BCC. This result indicates inferior performance for MCC,

primarily due to class imbalance, especially the limited number of images with a score of 0.5. In the present study, the average F1-score for MCC was 0.592, which supports the same conclusion. The scarcity of 0.5-score images restricted model performance in both studies, emphasizing the need for a more balanced dataset.

Veli-Matti Sundell *et al.*[21] performed ACR phantom image quality scoring using a CNN-based AI model composed of six convolutional layers and achieved an overall accuracy of 95%. Their model was trained on a large dataset of 90,288 images, including 4,752 images with visible objects (score of 1) and 14,256 images with non-visible objects (score of 0), collected from eight mammography systems manufactured by three manufacturers. The substantial dataset size and system diversity enhanced model learning and improved generalization. However, their model sensitivity decreased as object size decreased, which is consistent with the reduced sensitivity for smaller objects observed in this study.

To address these limitations, future studies should use a larger and more diverse dataset that includes images of varying quality from multiple imaging systems to improve both model training and generalizability of CNN-based AI models. Beyond increasing dataset size, specific strategies are needed to enrich borderline cases (score of 0.5), which are critical for quality control decisions. Potential approaches include: (1) acquiring additional phantom images at exposure settings that produce near-threshold visibility for specific objects; (2) applying controlled contrast reduction and localized mild blurring to high-quality images to simulate subtle loss of conspicuity; (3) generating synthetic phantom images or using generative models to simulate near-threshold object visibility while preserving realistic background and noise characteristics; and (4) employing training strategies such as class-balanced sampling, focal loss, and cost-sensitive learning to reduce the impact of residual class imbalance during model optimization.

Conclusion

This study evaluated five CNN-based AI models for classifying objects in ACR phantom images. GoogLeNet demonstrated the best performance for fiber and mass classification, whereas ResNet50 was most effective for speck classification. GoogLeNet accurately classified fibers F1-F3 with an FPR below 0.12 and other statistical metrics above 0.76. It also demonstrated strong performance for mass classification, achieving an FPR below 0.12 and other statistical values exceeding 0.75. ResNet50 performed best for speck classification, with an FPR below 0.09 and other statistical metrics above 0.82, except for S4. Using the best-performing models across all 16 object positions, classification accuracies were as follows: fibers-F1 (98.76%), F2 (91.51%), F3 (75.96%), F4 (53.33%), F5 (30.90%), and F6 (42.37%); specks-S1 (99.61%), S2 (99.03%), S3 (82.38%), S4 (64.20%), and S5 (100%); and masses-M1 (98.53%), M2 (76.96%), M3 (75.37%), M4 (84.30%), and M5 (89.13%). Overall, GoogLeNet achieved accuracies ranging from 30.9% to 98.8% for fibers and masses, whereas ResNet50 achieved 64.2% to 100% for specks. To improve clinical applicability and align with ACR guidelines, future work should prioritize increasing prediction accuracy for intermediate objects, particularly for F3, F4, M2, and M3.

Ethical approval

This research was approved by the ethics committee of Naresuan University, Thailand (IRB No. P1-0091/2567.

Funding

This project is funded by the National Research Council of Thailand (NRCT) under contract number N42A670908 and awarded to Thunyarat Chusin.

Conflict of interest

This work has no conflicts of interest.

CRediT authorship contribution statement

Supitcha Saengkaew, Suradet Aunnoi, Amarin Thuikham, Pratchayakan Hompeng, and **Thunyarat Chusin**: methodology, investigation, data curation, and formal analysis; **Thunyarat Chusin**: original draft of the manuscript; **Siriprapa Kaewjaeng** and **Titipong Kaewlek**: supervision. All authors contributed to writing: review and editing, conceptualization and study design. All authors read and approved the final manuscript.

Acknowledgements

The authors acknowledge the Faculty of Allied Health Sciences, Naresuan University, for providing the mammography system used in this study.

References

[1] Calvocoressi L, Sun A, Kasl S V., Claus EB, Jones BA. Mammography screening of women in their 40s: Impact of changes in screening guidelines. Cancer 2008; 112: 473-80. doi: 10.1002/cncr.23210.

[2] Myers ER, Moorman P, Gierisch JM, Havrilesky LJ, Grimm LJ, Ghate S, et al. Benefits and harms of breast cancer screening: A systematic review. JAMA 2015; 314: 1615-34. doi: 10.1001/jama.2015.13183.

[3] Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition - Summary document. Ann. Oncol. 2008; 19: 614-22. doi: 10.1093/annonc/mdm 481.

[4] Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc. IEEE 1998; 86: 2278-324. doi: 10.1109/5.726791.

[5] Kousalya K, Krishnakumar B, Mohana RS, Karthikeyan N. Comparative analysis of White Blood Cells Classification using Deep Learning

Architectures. 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), 2021; pp. 1220-5. doi: 10.1109/ICOSEC 51865. 2021.9591771.

[6] Jaganathan D, Balsubramaniam S, Sureshkumar V, Dhanasekaran S. Concatenated Modified LeNet Approach for Classifying Pneumonia Images. J Pers Med 2024;14: 328. doi: 10.3390/jpm14030328.

[7] Balasubramaniam S, Velmurugan Y, Jaganathan D, Dhanasekaran S. A modified LeNet CNN for breast cancer diagnosis in ultrasound images. Diagnostics 2023; 13: 2746 doi: 10.3390/diagnostics 13172746.

[8] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Adv Neural Inf Process Syst. 2012; 25: 1097-105. doi: 10.1145/3065386

[9] Guo C, Chen Y, Li J. Radiographic imaging and diagnosis of spinal bone tumors: AlexNet and ResNet for the classification of tumor malignancy. J Bone Oncol 2024; 48: 100629. doi: 10.1016/j.jbo. 2024.100629.

[10] Azhagiri M, Rajesh P. EAN: enhanced AlexNet deep learning model to detect brain tumor using magnetic resonance images. Multimed Tools Appl 2024; 83: 66925-41. doi: 10.1007/s11042-024-181 43-w.

[11] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2015 International Conference on Learning Representations (ICLR), 2015; pp. 1114. https://arxiv.org/abs/1409. 1556v6.

[12] Parvin F, Hasan MAM, Ahmed B, Mamun MA, Parvej SMK. Breast cancer histopathological image classification using an ensemble of deep convolutional neural networks. 2023 26th International Conference on Computer and Information Technology (ICCIT), 2023, pp. 1-5. doi. org/10.1109/ICCIT60459.2023.10441273.

[13] Harrison P, Hasan R, Park K. State-of-the-Art of Breast Cancer Diagnosis in Medical Images via Convolutional Neural Networks (CNNs). J Healthc Inform Res. 2023; 7: 387-432. doi: 10.1007/s416 66-023-00144-3.

[14] Chen Y, Shao X, Shi K, Rominger A, Caobelli F. AI in breast cancer imaging: An update and future trends. Semin Nucl Med. 2025; 55: 358-70. doi: 10.1053/j.semnuclmed. 2025. 01. 008.

[15] Amanova N, Martin J, Elster C. Explainability for deep learning in mammography image quality assessment. Mach Learn Sci Technol. 2022; 3: 025015. doi: 10.1088/2632-2153/ac7a03.

[16] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015; pp. 1-9. doi: 10.1109/ CVPR.2015.7298594.

[17] Teoh JR, Hasikin K, Lai KW, Wu X, Li C. Enhancing early breast cancer diagnosis through automated microcalcification detection using an optimized ensemble deep learning framework. PeerJ Comput Sci 2024;10:e2082. doi: 10.7717/peerj-cs. 2082.

[18] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, p. 770–8. doi: 10.1109/CVPR.2016.90.

[19] Zhang Y, Liu YL, Nie K, Zhou J, Chen Z, Chen JH, et al. Deep learning-based automatic diagnosis of breast cancer on MRI using mask R-CNN for detection followed by ResNet50 for classification. Acad Radiol. 2023; 30: S161-71. doi: 10.1016/j.acra. 2022.12.038.

[20] Tafavvoghi M, Sildnes A, Rakaee M, Shvetsov N, Bongo LA, Busund L-TR, et al. Deep learning-based classification of breast cancer molecular subtypes from H&E whole-slide images. J Pathol Inform 2025; 16: 100410. doi: 10.1016/j.jpi.2024. 100410.

[21] Sundell VM, Mäkelä T, Vitikainen AM, Kaasalainen T. Convolutional neural network -based phantom image scoring for mammography quality control. BMC Med Imaging 2022; 22: 216. doi: 10.1186/s12 880-022-00944-w.

[22] Al-antari MA, Al-masni MA, Choi MT, Han SM, Kim TS. A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. Int J Med Inform. 2018; 117: 44-54. doi: 10.1016/ j.ijmedinf.2018.06.003.

[23] Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra Acharya U. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med. 2020; 121: 103792. doi: 10.1016/j.compbiomed.2020.103792.

[24] Park SS, Ku YM, Seo KJ, Whang IY, Hwang YS, Kim MJ, et al. Devising a deep neural network-based mammography phantom image filtering algorithm using images obtained under mAs and kVp control. Sci Rep. 2023;13: 3545. doi: 10.1038/s41598-023-30780-z.

[25] Ho PS, Hwang YS, Tsai HY. Machine learning framework for automatic image quality evaluation involving a mammographic American College of Radiology phantom. Phys Med. 2022; 102:1-8. doi: 10.1016/j.ejmp.2022.08.004.

[26] Tanaka R, Nozaki S, Goshima F, Shiraishi J. Deep learning versus the human visual system for detecting motion blur in radiography. J Med Imaging. 2022; 9: 015501. doi: 10.1117/1.jmi.9.1.015501.

[27] Song SE, Seo BK, Yie A, Ku BK, Kim HY, Cho KR, et al. Which phantom is better for assessing the image quality in full-field digital mammography?: American College of Radiology accreditation phantom versus digital mammography accreditation phantom. Korean J Radiol. 2012; 13: 776-83. doi: 10.3348/kjr.2012.13.6.776.

[28] Berns EA, Pfeiffer DE, Adent C, Baker JA, Bassett LW, Hendrick FRE, et al. Quality control manual for 2D and digital breast tomosynthesis. Radiologist's Section, Radiologic Technologist's Section, Medical Physicist's Section. American College of Radiology Subcommittee on Quality Assurance in Mammography of the Committee on Mammography Accreditation, American College of Radiology; 2018.

Journal of Associated Medical Sciences 2026; 59(1): 247-253

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: Music therapy for enhancing executive functions in children with autism spectrum disorder

Natwipa Wanicharoen[1], Thanasak Kalaysak[1*], Supaporn Chinchai[1], Kanokporn Kitisupornphan[2], Vich Boonrod[3]

[1]Communication Sciences and Disorders Division, Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Nakornping Hospital, Chiang Mai Province, Thailand.
[3]Department of Music, Faculty of Humanities, Naresuan University, Phitsanulok Province, Thailand.

ARTICLE INFO

Article history:
Received 15 July 2025
Accepted as revised 7 December 2025
Available online 17 December 2025

Keywords:
Music therapy, executive function, autism spectrum disorder.

ABSTRACT

Executive functions (EFs) are higher-order cognitive processes essential for learning, self-regulation, and social adaptation. Children with autism spectrum disorder (ASD) frequently exhibit EF deficits, which contribute to challenges in communication, behavior, and daily functioning. Growing attention has been directed toward alternative therapeutic approaches for enhancing EF skills, with music therapy (MT) increasingly recognized as a promising intervention. Music engages neural circuits within the frontal lobe-an area central to EF-while simultaneously providing a multisensory, motivating, and socially interactive medium for intervention. This review synthesizes current research across three key areas: (1) the association between EFs, social communication, and behavior; (2) the relationships between music, the brain, and EFs; and (3) empirical studies on MT-based EF training in children with ASD. Collectively, findings indicate that MT holds considerable potential as a complementary approach for strengthening EF skills and, in turn, improving social communication and adaptive behavior in this population.

Introduction

Autism spectrum disorder is classified in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) as a neurodevelopmental disorder defined by persistent deficits across two core domains. The first domain involves impairments in social communication and interaction, which may include limited eye contact, reduced emotional reciprocity, difficulties in forming and sustaining relationships, restricted imaginative play, and atypical use of verbal or nonverbal language in social contexts. The second domain concerns restricted and repetitive patterns of behavior, encompassing repetitive motor movements or use of objects, strong adherence to routines, highly fixated interests, and atypical responses to sensory input-such as hypersensitivity, hyposensitivity, or unusual fascination with sensory stimuli. These defining characteristics shape the clinical presentation of ASD and present substantial challenges for children's learning, social integration, and adaptive functioning.[1] Importantly, the two core domains are closely linked with deficits in executive functions (EFs). Growing evidence suggests that many of the behavioral and communicative features of ASD are not only from social or language impairments but also from underlying EF.[2]

*** Corresponding contributor.**
Author's Address: *Communication Sciences andDisorders Division, Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: thanasak.k@cmu.ac.th
doi: 10.12982/JAMS.2026.030
E-ISSN: 2539-6056

This recognition has shifted research attention toward interventions that can strengthen EF in more engaging and ecologically valid ways.[3] Music has emerged as evidence that is capable of stimulating attention, emotional regulation, and cognitive flexibility through its multisensory and social nature.[4] Consequently, understanding EFs provides a necessary foundation for exploring why music therapy may be particularly suited to enhance these higher-order cognitive processes in children.[5]

Executive functions

Executive functions-sometimes referred to as executive control or cognitive control-are higher-order cognitive processes that guide goal-directed behavior, regulate actions, and enable adaptive responses to changing demands.[6] They are typically conceptualized as comprising three interrelated core components: cognitive flexibility,[7] inhibition,[8] and working memory.[9]

Cognitive flexibility refers to the capacity to shift thinking and adapt to new or unexpected situations across different contexts.[10] Inhibition is the ability to suppress impulsive responses, resist distractions, and sustain attention on relevant tasks.[8] Working memory involves the temporary storage, organization, and retrieval of information, supporting essential skills such as planning, reasoning, comprehension, and problem-solving.[9-11] These three core components of EFs undergo rapid development during early childhood, particularly between ages two and five, and continue to mature throughout adolescence into adulthood.[12]

Brain maturation plays a critical role in the development of EFs. Core EF skills previously mentioned, such as cognitive flexibility, inhibitory control, and working memory, are primarily supported by the prefrontal cortex (PFC). These processes also rely on interconnected neural circuits involving the cingulate cortex, parietal cortex, basal ganglia, amygdala, hippocampus, and other subcortical structures. Together, these networks enable individuals to solve complex problems, regulate behavior, and adapt to changing demands. The amygdala, situated within the limbic system, plays a crucial role in linking cognition with emotion and motivation. The rapid exchange of signals between the PFC and limbic system allows individuals to sustain interest, regulate attention, and prioritize meaningful information while filtering out less relevant stimuli.[13] This dynamic interaction illustrates how brain maturation underpins both the cognitive and emotional dimensions of EF development.

Accordingly, understanding how music engages EF-brain systems would provide a strong foundation for investigating its therapeutic potential for ASD. The following section will therefore address: (1) the association between EFs, social communication, and behavior; (2) the relationships between music, the brain, and EFs; and (3) empirical studies on MT-based EF training in children with ASD.

Executive functions, social communication, and behavior

Executive functions are closely associated with the social communication difficulties and restricted, repetitive behavior patterns observed in children with ASD. Effective social interaction relies on several EF skills, including the ability to inhibit inappropriate responses, plan conversational turns, and flexibly shift topics.

When these skills are impaired, children often struggle to engage appropriately with others, and their everyday functioning might be disrupted.

Howard *et al*. investigated the relationship between EFs, sustained attention, language skills, and social communication in 180 children with ASD aged 2-8 years.[14] The study found that children with stronger EF abilities—particularly in working memory and planning/organization—demonstrated higher levels of social communication comprehension and expression. Elevated EF skills were also linked to fewer pragmatic difficulties in communication. Furthermore, children who were able to sustain longer attention spans exhibited stronger expressive language abilities. Collectively, these findings highlight that EFs and attentional control play pivotal roles in supporting language development and social communication in children with ASD.

Executive functions also play a critical role in the language and speech development of children with ASD. In terms of receptive language, these children often demonstrate reduced vocabulary knowledge and usage compared with typically developing peers. This limitation is linked to difficulties in sustaining attention, regulating focus, and exercising self-control, all of which are necessary for consistently engaging with learning stimuli.[15]

Ellis *et al*. reported that school-aged children with ASD frequently display EF deficits in inhibition, cognitive flexibility, and working memory, which in turn are strongly associated with poorer language comprehension.[16] Consequently, children with ASD often show lower levels of receptive vocabulary. Expressive language skills are similarly influenced by EF abilities-particularly planning, organization, lexical retrieval, and sentence formulation. Consistent with these findings, studies by McEvoy *et al*. and Gilotty *et al*. demonstrated that EF deficits negatively affect both communication and social skills.[17,18] Specifically, children with ASD often struggle to maintain joint attention during conversations and may exhibit inappropriate initiation patterns, further complicating effective social interaction.

Restricted and repetitive behavioral patterns-such as repetitive movements, rigid adherence to routines, and highly focused interests-are strongly associated with deficits in EFs, particularly cognitive flexibility and inhibition. Iversen and Lewis conducted a meta-analysis examining the relationship between these behaviors, EF performance, and parental reports.[19]

Their findings indicated that higher levels of restricted and repetitive behaviors were consistently linked to greater EF deficits in flexibility, inhibition, and caregiver-reported EF skills.

Similarly, Berenguer *et al*. investigated the interplay between ASD, EFs, Theory of Mind (ToM), and behavioral problems.[20] ToM, defined as the ability to understand that others hold beliefs, thoughts, and perspectives different from one's own, was found to be impaired in children with both ASD and ADHD. The study showed that deficits in EF and ToM contribute jointly to difficulties in interpreting others' intentions, which in turn exacerbate behavioral problems.

While these studies show consistent links between executive function performance and social or behavioral outcomes, most of them are cross-sectional and correlational rather than experimental studies, and several rely on parent reports. These factors might make it difficult to determine whether executive function directly causes improvements in behavior or communication. Moreover, differences in age, intelligence, and co-occurring conditions such as ADHD may also influence how strongly executive function relates to social and behavioral skills.[2] Further longitudinal or intervention studies are needed to clarify these relationships.

In conclusion, children with ASD commonly experience deficits in EFs-particularly cognitive flexibility, inhibition, and working memory-that contribute to difficulties in social communication as well as the presence of restricted and repetitive behaviors, which lead to challenges in daily adaptability and behavioral regulation. Strengthening EF skills is therefore a critical therapeutic priority, as doing so can promote more effective communication, support social integration, and encourage adaptive behavioral patterns.

Music, brain, and executive functions

A variety of interventions are available to support development in children with ASD, including speech therapy, occupational therapy, and applied behavior analysis, each of which should be tailored to the individual child's needs and characteristics.[3] Among these, music therapy (MT) has received increasing attention as a complementary approach.[21] Emerging evidence suggests that MT is effective in fostering language, social, and behavioral development,[22] as well as engaging both cognitive and emotional systems, which may offer a promising pathway for enhancing EFs in children with ASD.[23-24]

Early childhood is often described as a golden period of intellectual, emotional, and behavioral growth. During this stage, the brain is highly flexible and particularly sensitive to environmental stimuli.[25] Rapid synaptic formation occurs as children engage with experiences and external inputs, shaping neural architecture and cognitive development. Among the most powerful of these stimuli is music, a complex multisensory experience that promotes integrated brain learning. Music engages both hemispheres of the brain simultaneously, activating networks involved in auditory processing, movement, language, and emotion.[4] When children are exposed to music-whether through listening, singing, or movement-the brain coordinates these functions in parallel, strengthening connections between neural pathways. This integrated stimulation supports the development of essential cognitive and behavioral skills, including attention, short-term memory, self-control, and social interaction.[26]

Beyond its immediate effects, music training and regular music exposure can lead to lasting structural and functional changes in the brain. These neuroplastic adaptations are particularly evident in regions associated with movement and auditory processing, where music training has been shown to increase cortical representation. Such training also enhances a range of cognitive skills, including verbal memory, phonological awareness, and spatial reasoning. Moreover, research indicates that attentive listening to music can induce changes in the auditory cortex, supporting the view that music exposure may positively influence executive functioning by strengthening neural networks involved in attention, memory, and cognitive control.[27-28]

Music therapy is a structured therapeutic process in which trained therapists use musical experiences-and the therapeutic relationship itself-to promote health and facilitate change according to individualized goals.[29] The American Music Therapy Association (AMTA) defines MT as an intervention designed to meet physical, emotional, cognitive, and social needs through musical activities such as composing, singing, moving, and listening.[30] Evidence from a systematic review by Rodriguez-Gomez and Talero-Gutiérrez[5] indicates that MT can stimulate activity in the frontal lobe, the brain region central to EFs. Such stimulation has been shown to significantly support the development of cognitive flexibility, inhibitory control, and working memory in children.

In conclusion, music therapy is more than an enjoyable activity; it is a powerful therapeutic tool for strengthening EFs, particularly in children who require support for learning and behavior. By integrating structured musical experiences into developmental programs, MT can promote cognitive flexibility, inhibitory control, and working memory while also fostering social and emotional growth. Consequently, incorporating music as a sustained component of child development initiatives should be actively encouraged and further expanded in both clinical and educational settings.

Music therapy effects on executive functions in children with ASD

A growing body of research, conducted both internationally and in Thailand, demonstrates that music therapy interventions targeting EFs can significantly enhance EF skills in children with ASD.[23-24] Recent publications reinforce MT as a promising intervention for cognitive development, providing empirical support for its effectiveness across diverse cultural and clinical contexts.

For example, Sharma and Mehta[23] conducted a study in India examining the effects of music therapy (MT) on executive functions in 100 children with ASD. Participants in the experimental group attended MT sessions twice weekly for 40 minutes over eight weeks, while the control group received no intervention. The sessions incorporated rhythm- and melody/swar-based activities, with both children and caregivers engaging in interactive exercises. Tasks included passing small cushions and large pillows in time with musical rhythms to encourage attention and coordination, singing familiar songs such as Chanda Mama Door Ke, Chiriyan Di Chaal, and Jingle Bells, and participating in movement-based games (e.g., "sing and stop," "clap and go," "clap and jump").

Results showed that children in the experimental group demonstrated significant improvements in EF performance, as indicated by reduced post-intervention dysfunction scores. However, the study had several limitations: it did not account for children's baseline interest in music, the short intervention period may have limited the effect size, and individual rather than group therapy was used due to variability in hyperactivity and sensory sensitivities among participants.[23]

Rupsuwan *et al*. conducted a single-case study in Thailand to examine the effects of a virtual music therapy (MT) model on executive functions in a five-year-old boy with ASD.[24] The child and his mother participated in eight sessions delivered twice weekly for 30-45 minutes for four weeks. Sessions were conducted via telehealth, with the therapist using a laptop equipped with a camera and condenser microphone, while the mother joined through a smartphone.

The intervention targeted EF skills by combining intellectual and emotional engagement through music listening, structured musical activities, and behavioral reinforcement strategies. Activities included instrument play, singing, movement, and musical games. The therapist provided live music by singing and playing the guitar, adjusting rhythm, tempo, and dynamics to match the child's progress. Favorite songs were incorporated to enhance motivation, and the child was encouraged to choose instruments and participate actively. Both live and recorded music supported movement-based activities, with lyrics, rhythm, and tempo providing cues for attention, initiation, and stop–go responses.[24]

Results showed improvements in EF performance, including working memory, inhibitory control, and cognitive flexibility, as well as reductions in behavioral problems. However, the results of this study were limited by its single-subject design without a control group, short intervention period, and the inherent limitations of virtual delivery (e.g., reduced nonverbal interaction, internet instability, and limited camera angles, which could restrict the generalizability of the results.[24]

Rupsuwan *et al*.[24] highlighted the multifaceted impact of virtual music therapy (MT) on executive functions (EFs).[24] First, working memory improved through repetitive melodies and lyrics that supported song recall, particularly when paired with role modeling and therapist guidance. Second, inhibitory control was strengthened via musical play and movement activities that incorporated role switching and rhythmic start–stop patterns, thereby enhancing sustained attention. Third, cognitive flexibility was fostered through varied musical elements-such as changes in rhythm, tempo, lyrics, and volume-that encouraged adaptation to auditory shifts and behavioral adjustment in social contexts. Fourth, emotional regulation benefited, as music and movement provided calming effects, facilitated emotional expression, and supported the management of negative emotions. Finally, planning and organization were promoted through composing and engaging in songs tied to daily routines and target behaviors, helping the child sequence tasks and structure actions effectively.

Despite these positive outcomes, the study faced several limitations. Its single-subject design without a control group restricts generalizability, while the short intervention period precludes long-term follow-up. Additionally, constraints of virtual delivery-including reduced nonverbal interaction, restricted camera angles, and internet instability-introduced time lags and limited engagement.[24]

Consistent with the findings of Srinivasan and Bhat,[31] rhythm and melody stimulate reward- and emotion-related brain circuits. Regular participation in rhythm-based activities helps children regulate arousal, reduce anxiety, and sustain attention-effects that are particularly valuable during therapeutic or educational sessions. Activities such as clapping to rhythm, playing percussion instruments, or moving in response to melodic cues also support gross and fine motor coordination, thereby strengthening timing, sequencing, and inhibitory control-core aspects of executive function frequently impaired in ASD.

Similarly, Sutikno *et al*. demonstrated that singing activities in preschool children enhance memory, increase learning motivation, and foster creativity.[32] Participation in singing activates neural processes that facilitate the absorption and retention of information. Moreover, the repetitive use of words in theme songs reinforces recall, enabling children to remember and integrate material more effectively.

Although existing studies indicate that MT can enhance several aspects of executive functioning in ASD, current evidence on music therapy targeting executive functions in autistic children remains limited, with few studies explicitly evaluating EF outcomes and mostly relying on small or single-case samples. The group-based interventions[23] demonstrated general feasibility, whereas individualized, virtual approaches[24] provided insight into mechanisms such as rhythm-based attention and emotional engagement. Active, participatory forms of music therapy, particularly incorporating rhythm, singing, and movement, appear most effective for improving inhibitory control and working memory, while effects on planning and cognitive flexibility are less consistent.

Taken together, these studies demonstrate that music therapy, whether delivered in person to large groups or virtually in single-case formats, consistently benefits the development of EFs in children with ASD. Repetitive, rhythm- and melody-based activities, along with active caregiver participation, appear to be key mechanisms underlying improvement. However, methodological differences emphasize the need for future research to balance the scalability of group interventions with the individualized, domain-specific insights afforded by case studies.

Practical implications

This study highlights several important implications for practice and research. Rhythm- and melody-based activities can be embedded into interventions to strengthen working memory, attention, inhibition, and cognitive flexibility. Simple strategies such as repetitive songs and musical games serve as effective tools for reinforcing EF skills both in therapy and in daily routines. Caregiver involvement enhances engagement and promotes generalization, while parent training extends benefits into the home environment.

The Indian study demonstrated the feasibility of structured, group-based MT, whereas the Thai case illustrated the potential of telehealth delivery, suggesting that MT can be adapted to diverse contexts, including resource-limited settings. Incorporating children's preferred music increases motivation, while rhythmic cues help regulate behavior and reduce repetitive patterns. Collectively, these findings position MT as a practical, child-friendly intervention for improving cognitive and behavioral outcomes. Nonetheless, further research is needed to evaluate longer intervention periods, develop standardized protocols, design scalable models, and determine how factors, such as age, baseline cognitive ability, and sensory sensitivities, e.g., auditory hypersensitivity commonly found in ASD, influence responsiveness and comfort for broader implementation of the intervention.

Conclusion

Children with ASD commonly experience deficits in EFs-particularly cognitive flexibility, inhibition, and working memory-that adversely affect social communication and behavior. Music therapy has emerged as a promising and engaging intervention that stimulates brain regions involved in EF, most notably the prefrontal cortex. Interestingly, it integrates cognitive, motor, and emotional engagement within a single activity, activating multisensory and brain systems that support EF. Evidence from studies conducted across different countries demonstrates that MT, whether delivered in person or virtually, can significantly enhance EF skills in children with ASD. Reported benefits include improvements in working memory, inhibitory control, emotional regulation, and behavioral planning. Additionally, structured and individualized MT interventions have been shown to reduce repetitive behaviors and foster broader social and communicative development. Collectively, these findings underscore the therapeutic value of integrating music-based EF training into intervention programs for children with ASD. By strengthening cognitive and behavioral capacities through enjoyable, child-centered activities, MT represents a practical and adaptable approach to supporting development across diverse contexts.

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 authors declare no conflict of interest.

CRediT authorship contribution statement

Natwipa Wanicharoen: conceptaulization, investigation, project administration, supervisor, writing original draft, review and edit; **Thanasak Kalaysak:** writing original draft, review and edit, correspondence; **Supaporn Chinchai, Kitisupornphan and Vich Boonrod:** writing review and edit.

References

[1] American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. 5th ed. Washington, DC: American Psychiatric Association; 2013.

[2] Demetriou EA, Lampit A, Quintana DS, Naismith SL, Song YJC, Pye JE, et al. Autism spectrum disorders: a meta-analysis of executive function. J Autism Dev Disord. 2018; 8)48): 73-2555. doi:10.1007/s2-3523-018-10803.

[3] Schreibman L, Dawson G, Stahmer AC, Landa R, Rogers SJ, McGee GG, et al. Naturalistic developmental behavioral interventions: Empirically

validated treatments for autism spectrum disorder spectrum disorder. J Autism Dev Disord. 2015; 45: 2411-28. doi:10.1007/s10803-015-2407-8.

[4] Patel AD. Music, language, and the brain. Oxford: Oxford University Press; 2010. doi:10.1093/acprof: oso/9780195123753.001.0001.

[5] Rodriguez-Gomez DA, Talero-Gutiérrez C. Effects of music training on executive function performance in children: a systematic review. Front Psychol. 2022; 13 :968144. doi:10.3389/ fpsyg.2022.968144.

[6] Diamond A. Executive functions. Annu Rev Psychol. 2013; 64: 135-68. doi: 10.1146/ annurev-psych-113011-143750.

[7] Davidson M,Amso D, Anderson L, Diamond A. Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. Neuropsychologia. 2006; 44(11): 2037-78. doi: 10.1016/j.neuropsychologia.2006.02.006.

[8] Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. Cogn Psychol. 2000; 41(1): 49-100. doi:10.1006/cogp.1999.0734.

[9] Baddeley A. Working memory. Curr Biol. 2010; 20(4): 136-40. doi:10.1016/j.cub.2009.12.014.

[10] Canas JJ, Fajardo I, Salmeron L. Cognitive flexibility. Int Encycl Ergonom Hum Factors. 2006; 1(3): 297-301.

[11] Cowan N. Working memory underpins cognitive development, learning, and education. Educ Psychol Rev. 2014; 26: 197-223. doi:10.1007/ s10648-013-9246-y.

[12] Zelazo PD, Müller U. Executive function in typical and atypical development. In: Goswami U, editor. Blackwell handbook of childhood cognitive development. Oxford: Blackwell Publishing; 2002. pp. 445-69. doi: 10.1002/9780470996652.ch20.

[13] Blair C. Educating executive function. Wiley Interdiscip Rev Cogn Sci. 2017; 8(1-2): e1403. doi: 10.1002/wcs.1403.

[14] Howard J, Herold B, Major S, Leahy C, Ramseur K, Franz L, et al. Associations between executive function and attention abilities and language and social communication skills in young children with autism spectrum disorder. Autism. 2023; 27(7): 2135-44. doi:10.1177/13623613231154310.

[15] Cascia J, Barr JJ. Associations among vocabulary, executive function skills and empathy in individuals with autism spectrum disorder spectrum disorder. J Appl Res Intellect Disabil. 2017; 30(4): 627-37. doi:10.1111/jar.12257.

[16] Ellis Weismer S, Kaushanskaya M, Larson C, Math√©e J, Bolt D. Executive function skills in school-age children with autism spectrum disorder spectrum disorder: Association with language abilities. J Speech Lang Hear Res. 2018; 61(11): 2641-58. doi: 10.1044/2018_jslhr-l-rsaut-18-0026.

[17] McEvoy RE, Rogers SJ, Pennington BF. Executive function and social communication deficits in young children with autism spectrum disorder. J Child Psychol Psychiatry. 1993; 34(4): 563-78. doi: 10.1111/j.1469-7610.1993.tb01036.x.

[18] Gilotty L, Kenworthy L, Sirian L, Black DO, Wagner AE. Adaptive skills and executive function in autism spectrum disorder spectrum disorders. Child Neuropsychol. 2002; 8(4): 241-58. doi: 10.1076/chin.8.4.241.13504.

[19] Iversen RK, Lewis C. Executive function skills are linked to restricted and repetitive behaviors: Three correlational meta-analyses. Autism Res. 2021; 14(6): 1163-85. doi:10.1002/aur.2468.

[20] Berenguer C, Rosell√≥ B, Colomer C, Baixauli I, Miranda A. Children with autism spectrum disorder and attention deficit hyperactivity disorder: Relationships between symptoms and executive function, theory of mind, and behavioral problems. Res Dev Disabil. 2018; 83: 260-9. doi:10.1016/j. ridd.2018.10.001.

[21] Brondino N, Fusar-Poli L, Rocchetti M, Provenzani U, Barale F, Politi P. Complementary and alternative therapies for autism spectrum disorder spectrum disorder. Evid Based Complement Alternat Med. 2015; 2015: 258589. doi:10.1155/2015/258589.

[22] Boonrod V, Wanicharoen N. A systematic review of the effect of music therapy compared to speech therapy on social communication skills in preschool children with autism spectrum disorder spectrum disorder. J Assoc Med Sci. 2025; 58(2): 130-8. doi:10.12982/JAMS.2025.046.

[23] Sharma M, Mehta A. Music therapy improves executive functioning in children with autism spectrum disorder spectrum disorder. Int J Res Spec Educ. 2023;4(2): 46-50.

[24] Rupsuwan T, Chiengchana N, Buranaprapuk A. The effect of virtual music therapy interventions on the executive function of a child with autism spectrum disorder spectrum disorder: A Single-case study. J Ratchasuda Inst Res Dev Persons Disabil. 2023; 19(1): 39-60.

[25] Shonkoff JP, Phillips DA. From neurons to neighborhoods: The science of early childhood development. Washington, DC: National Academy Press; 2000. doi:10.17226/9824.

[26] Gerry D, Unrau A, Trainor LJ. Active music classes in infancy enhance musical, communicative and social development. Dev Sci. 2012; 15(3): 398-407. doi:10.1111/j.1467-7687.2012.01142.x.

[27] Johansson BB. Music and brain plasticity. Eur Rev. 2006; 14(1): 49-64. doi:10.1017/ S1062798706000056.

[28] Reybrouck M. Music and the plastic brain. How sounds can generate neuroplastic adaptations. Biomed J Sci Tech Res. 2023; 52(3): 43689-93. doi:10.26717/BJSTR.2023.52.008246.

[29] Bruscia KE. Defining music therapy. 2nd ed. Gilsum, NH: Barcelona Publishers; 1989.
[30] American Music Therapy Association (AMTA). American Music Therapy Association (AMTA) [Internet]. 2015 [cited 2025 April 11]. Available from: https://musictherapy.org.
[31] Srinivasan SM, Bhat AN. A review of "music and movement" therapies for children with autism: embodied interventions for multisystem development. Front in integr Neurosci. 2013; 7: 22. doi:10.3389/fnint.2013.00022.
[32] Sutikno PY, Jazuli M, Utomo U. The effects of singing activities on children's memory, learning motivation, and creativity in an indonesian kindergarten for preschoolers. J Law Sustain Dev. 2024; 12(1): e3123. doi: 10.55908/sdgs.v12i1.3123.

Instructions for Authors

Instructions for Authors

Original article/thesis can be submitted through the on-line system via website https://www.tci-thaijo.org/index.php/bullet-

General Principles

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, Communication Disorders, and other aspects related to the health sciences are welcome. Before publication, the articles will go through a system of assessment and acceptance by at least three experts who are specialized in the relevant discipline. All manuscripts submitted to Journal of Associated Medical Sciences should not have been previously published or under consideration for publication elsewhere. All publications are protected by the Journal of Associated Medical

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.

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. However, AI or AI-assisted graphical abstracts are also

Manuscript categories

1. **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).
2. **Original articles** must not exceed 15 journal pages (not more than 3,500 words) including tables, figures, and 40 references (maximum 40, recent and relevant).
3. **Short communications** including technical reports, notes, and letter to editor must not exceed 5 journal pages (not more than 1,500 words), including 2 tables/figures, and references (maximum 10, recent and relevant).
4. **Register** base on institute of official email address only.

Manuscript files

To submit your manuscripts, you will need the following files:

1. A Title page file with the names of all authors and corresponding authors*
2. Main document file with abstract, keywords, main text and references
3. Figure files
4. Table files
5. Any extra files such as Supplemental files or Author Biographical notes
6. Register base on institute of official email address only.

Manuscript Format

1. **Language**: English, Caribri 10 for text and 7 for all symbols. PLEASE be informed that the Journal only accept the submission of English manuscripts.
2. **Format**: One-side printing, double spacing. Use standard program and fonts and, add page and line number for all pages.
3. **A Title page**: Include article title, names of all authors and co-authors, name of the corresponding author and acknowledgements. Prepare according to 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 clearly. Present the authors' affiliation addresses (where the actual work was done) below the names. Indicate all affiliations with superscript number immediately after author's name and in front of 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*: Clearly 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. Contact details must be kept up to date by the corresponding author.
   - *Acknowledgements*: Acknowledgements will be collated in a separate section at the end of the article before the references in the stage of copyediting. 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 proof reading 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 Reference. Prepare according to following contents;
   - *Abstract*: Not exceeding 400 words, abstract must be structured with below headings in separated paragraph:
     - Background,
     - Objectives,
     - Materials and methods,
     - Results,
     - Conclusion, and
     - Keywords (3-5 keywords should be included)
   - *Introduction*: State the objectives of work and provide an adequate background, avoiding a detailed literature survey or a summary of the results.
   - *Materials and Methods*: Provide sufficient detail to allow the work to be reproduced. Methods already published should be indicated by a reference, only relevant modifications should be described. Ensure that each table, graph, or figure is referred in the text. According to the policy of ethical approval, authors must state the ethical approval code and conduct informed consent for human subject research (If any) and for animal research, authors must include a statement or text describing the experimental procedures that affirms all appropriate measures (if any) in this section.
   - *Results*: Results should be clear and concise. Present the new results of the study such as tables and figures mentioned in the main body of the article and numbered in the order in which they appear in the text or discussion.
   - *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.
   - *Conclusion*: The main conclusions of the study may be presented in a short Conclusions section, which may stand alone or form a subsection of a "Discussion" or "Results and Discussion".
   - *Conflict of interes*t: All authors must declare any financial and personal relationship with other people or organization that could inappropriately influence (bias) their work. If there is no interest to declare, then please state this: "The authors declare no conflict of interest".
   - *Ethic approval*: Ethic clearance for research involving human and animal subjects.
   - *References*: Vancouver's style.
5. **Artwork Requirements**
   - 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 per each and submit separately as supplementary file(s).
   - Save the figures as high resolution JPEC or TIFF files.

Note: Permission to reprint table(s) and/or 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 authors of the document 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 was accepted, the name of authors and affiliation and the name of the corresponding author must be included into the document and re-submitted in the copyediting stage.

Proof correction

The Editor-in-Chief and production team are in charge of the reprint preparation for online publication. The corresponding author will shortly be informed by email the proof of final reprint paper to approve for publication.

Page charge

No page charge.

References Format

1. References 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 are to 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 first and last numbers that are inclusive. Use commas (with spaces) to separate non-inclusive numbers in a multiple citation e.g. (2-5, 7, 10). Do not use a hyphen if there are no citation numbers in between inclusive statement e.g. (1-2). Use instead (1, 2).
3. References list: number the references (numbers in square brackets) in the list must be in the order in which they are mentioned in the text. In case of references source from non-English language, translate the title to English and retain "in Thai" in the parentheses.
4. Please note that if references are not cited in order the manuscript may be returned for amendment before it is passed

Examples of References list

Multiple Authors: List up to the first 6 authors/editors, and use "et al." for any additional authors.

Journal Articles (print): In case of reference source contains Digital Object Identifier (DOI), retain doi: at the end of reference. Vancouver Style does not use the full journal name, only the commonly-used abbreviation: "Physical Therapy" is cited as "Phys Ther". As an option, if a journal carries continuous pagination throughout a volume (as many medical journals do) the month and/or issue number may be omitted. Allow one space after semi-colon and colon then end each reference with full stop after page number.

[1] 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.
[2] 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.
[3] 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): 877-884. doi: 10.1002/mp.13263.
[4] 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.
[5] 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

[1] 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

[1] 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 of books and chapter in edited book should not be include in the research article, but others manuscript categories.

[1] Grove SK, Cipher DJ. Statistics for nursing research: A workbook for evidence-based practice. 3rd ed. St. Louis, Missouri: Elsevier; 2019.
[2] Haznadar M, editor. Cancer metabolism: Methods and protocols. New York: Humana Press; 2019. doi: 10.1007/978-1-4939-9027-6.
[3] Perrin DH. The evaluation process in rehabilitation. In: Prentice WE, editor. Rehabilitation techniques in sports medicine. 2nd ed. St Louis, Mo: Mosby Year Book; 1994: 253–276.

E-book

[1] 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
[2] 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

[1] 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].
[2] 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

[1] Lake M, Isherwood J, Clansey. Determining initial knee joint loading during a single limb drop landing: reducing soft tissue errors. Proceedings of 34th International Conference of Biomechanics in Sport; 2016 Jul 18-22; Tsukuba, Japan, 2016. Available from: https://ojs.ub.uni-konstanz.de/cpa/article/view/7126
[2] 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 38th Annual 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

[1] 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.
[2] 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).
[3] Department of Local Administration, Ministry of Interior Affairs. Standard of Sports Promotion. Bangkok. 2015: 7–9. (in Thai).
[4] World Health Organization. WHO guidelines on physical activity and sedentary behaviour. Geneva: World Health Organization; 2020. Licence: CC BY-NC-SA 3.0 IGO.
[5] World Health Organization. The epidemiology and impact of dementia: current state and future trends. 2015 [cited 2021 Mar

Journal History

Established in 1968

- 1968-2016 As the Bulletin of Chiang Mai Associated Medical Sciences
  - Vol1, No1 - Vol.49, No3
- 2017, the Journal of Associated Medical Sciences
  - Vol.50, No1 and forward.

Journal Sponsorship Publisher

Faculty of Associated Medical Sciences, Chiang Mai University


OURNAL OF ASSOCIATED
MEDICAL SCIENCES
AMS