JOURNAL OF ASSOCIATED MEDICAL SCIENCES

JAMS

JOURNAL OF ASSOCIATED MEDICAL SCIENCES

THE OFFICIAL PEER-REVIEWED ONLINE JOURNAL

Volume 58 Number 3 September -December 2025 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

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.

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. 58 No. 3 : September - December 2025

Content

Content

Content

Content

Journal of Associated Medical Sciences 2025; 58(3): 1-9

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Development of one channel-football formation neural network (OC-FFNet) for classification the breast ultrasound images

Titipong Kaewlek [1,2,3*], Thanatpat Jansaengsri[1], Yosita Leeju[1], Satawee Bootsapawanich[1]

[1]*Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*
[2]*Medical Physics Program, Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*
[3]*Interdisciplinary Health and Data Sciences Research Unit, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*

ARTICLE INFO

Article history:
Received 7 March 2025
Accepted as revised 15 April 2025
Available online 18 April 2025

Keywords:
Breast cancer detection, ultrasound image, convolutional neural networks, football formation model, deep learning.

ABSTRACT

Background: Breast cancer is a leading global health issue, with increasing incidence and mortality rates. In Thailand, it is the most common cancer among women, highlighting the need for better diagnostic methods. Traditional imaging techniques like mammography and ultrasound have limitations that hinder early detection. Recent advances in artificial intelligence (AI), particularly convolutional neural networks (CNNs), offer promising solutions for enhancing breast cancer detection in ultrasound images.

Objective: This study aims to develop a one-channel AI model for detecting breast cancer in ultrasound images, inspired by football formations to structure the CNN layers.

Materials and methods: The dataset comprises 18,000 breast ultrasound images categorized into normal, malignant, and benign cases. Data preprocessing involved image resizing, enhancement, and augmentation to address class imbalances. The proposed AI model, the One-Channel Football Formation Neural Network (OC-FFNet), was designed based on four distinct football formations: 4-3-3, 4-2-3-1, 4-4-2, and 5-4-1. Each formation guided the structuring of CNN layers, incorporating DenseNet-based modified dense blocks and transition layers. Model training was conducted with batch sizes ranging from 64 to 256 and epochs between 50 and 150. Performance evaluation metrics included accuracy, precision, recall, specificity, F1-score, false positive rate, and area under the curve (AUC).

Results: The models based on the 4-3-3 and 4-4-2 formations exhibited the highest classification performance, achieving an accuracy of 0.999, precision of 0.999, recall of 1.000, specificity of 0.999, F1-score of 0.999, and AUC of 0.999. The 4-2-3-1 model attained an accuracy of 0.963, while the 5-4-1 model achieved an accuracy of 0.968. Prediction times were consistent across all models, indicating computational efficiency. The findings suggest that formations with balanced positional distributions, such as 4-3-3 and 4-4-2, required fewer training iterations and larger batch sizes to achieve optimal performance.

Conclusion: The integration of football formation strategies into CNN architecture represents a novel approach to AI model design. The results indicate that strategically structured CNNs can improve breast cancer detection in ultrasound images.

∗ **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.2025.067
E-ISSN: 2539-6056

Introduction

Breast cancer represents significant global health concern, with its incidence and mortality rates escalating in numerous countries. In 2022, the World Health Organization (WHO) reported approximately 2.3 million new cases of breast cancer worldwide, resulting in 670,000

deaths.[1] This malignancy has now surpassed lung cancer as the most diagnosed cancer globally. The rising trend underscores the imperative for enhanced awareness, early detection, and comprehensive management strategies to mitigate the disease's impact.

In Thailand, breast cancer has emerged as the most prevalent cancer among women. The National Cancer Institute reported over 140,000 new cancer diagnoses in 2022, with breast cancer accounting for a substantial proportion of these cases. The annual incidence rate stands at 37.8 per 100,000 population, equating to approximately 22,158 new cases in 2020, which constituted 22.8% of all female cancer cases. This upward trajectory in breast cancer incidence necessitates targeted public health interventions and resource allocation to address the burgeoning burden effectively.[2]

Diagnostic modalities for breast cancer primarily include mammography and breast ultrasound.[3] Mammography, while considered the gold standard, exhibits reduced sensitivity in women with dense breast tissue, potentially leading to false-negative results and delayed diagnoses. Such delays can exacerbate disease progression and complicate treatment outcomes. Moreover, mammography involves exposure to ionizing radiation and requires breast compression, which may cause discomfort and deter some women from participating in regular screening programs. These limitations highlight the need for alternative or adjunctive imaging techniques to improve diagnostic accuracy and patient compliance.[4]

Breast ultrasound serves as a complementary diagnostic tool, particularly advantageous in detecting malignancies obscured in dense breast tissue that mammography might miss. Its non-invasive nature and absence of radiation exposure make it a favorable option for many patients. However, ultrasound is operator-dependent, and its efficacy can vary based on the technician's expertise. Integrating ultrasound into routine screening, especially for high-risk populations, could enhance early detection rates and improve prognostic outcomes.[3,4]

In recent years, artificial intelligence has been increasingly developed and applied to support medical diagnosis, particularly using convolutional neural networks, which are known for their powerful image analysis capabilities. Convolutional neural networks consist of multiple layers designed to extract hierarchical features from input data, enabling high accuracy in complex pattern recognition tasks.[5-10] Notably, the research team observed that the layered structure of convolutional neural networks bears a conceptual similarity to football formations, where each layer corresponds to specific positional roles such as goalkeeper, defender, midfielder, and forward.[11] This analogy provides a novel perspective on designing artificial intelligence architectures by aligning neural network layers with strategic football positions, thereby enhancing interpretability and potential model optimization.

For example, the widely recognized 4-4-2 football formation,[11-14] which comprises four defenders, four midfielders, and two forwards, can be mapped to the layers of a convolutional neural network as follows: four layers in the defense block, four layers in the middle block, and two layers in the forward block (Figure 1). This mapping is not merely conceptual but is strategically employed to organize the neural network layers to emulate hierarchical decision-making processes akin to positional roles on a football field. Such an approach aids in the systematic construction of convolutional neural networks, where each block is tailored to extract features relevant to specific tasks, thus enhancing the model's learning and predictive performance.

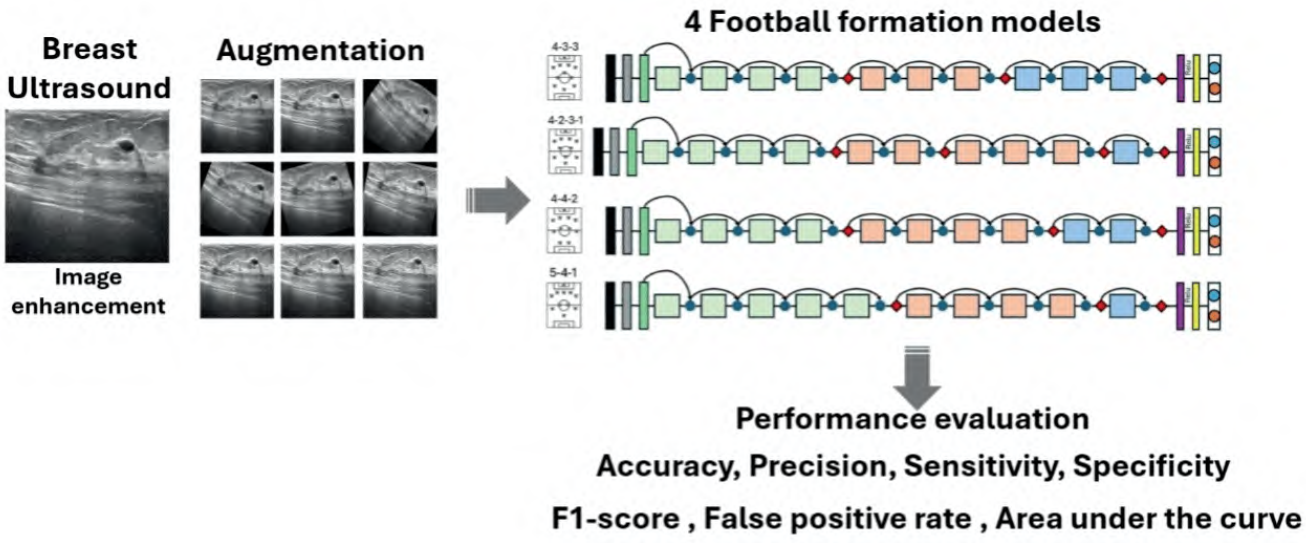


***Figure 1**. Workflows of this study.*

Additionally, the research team is investigating other football formations, encompassing both offensive and defensive strategies, to evaluate the efficacy of these architectural analogies in breast cancer detection. By experimenting with varied formations, such as 4-3-3 for more aggressive feature extraction or 5-4-1 for robust defensive classification layers, the research aims to identify optimal configurations that balance sensitivity and specificity in diagnostic outcomes.[12-14] This exploration of diverse configurations may reveal insights into the most effective architectures for accurate and efficient breast cancer detection using ultrasound images. The strategic alignment of football formations with neural network layers thus not only facilitates conceptualization but also serves as a foundation for optimizing artificial intelligence model structures for medical imaging applications.[15-17]

The primary objective of ongoing research is to develop a one-channel artificial intelligence model tailored for breast cancer detection in ultrasound imagery. This model aims to accurately differentiate between normal and abnormal breast tissues by emulating the strategic organization inherent in football formations. By leveraging artificial intelligence's analytical prowess, this approach aspires to enhance diagnostic precision, facilitate early intervention, and ultimately improve patient outcomes in breast cancer care.

Materials and methods

Data collection

The images utilized in this study were obtained from the online database Kaggle.com. All images were confirmed to be breast ultrasound scans and categorized into three groups: normal breast tissue, malignant tumors, and benign tumors. The dataset consisted of images collected from female patients aged 25 to 75 years. Each image had an average resolution of 500 × 500 pixels and was stored in PNG format. A total of 9,282 breast ultrasound images were sourced from two Kaggle databases,[18-19] with permission granted by the data owners for unrestricted use, eliminating the need for additional consent. The dataset comprised 266 images of normal breast tissue and 9,016 images of abnormal breast tissue, including both malignant and benign tumors.

Inclusion criteria

The study included single breast ultrasound images that were confirmed to depict either normal breast tissue or abnormal breast tissue, including malignant and benign tumors.

Exclusion criteria

Images diagnosed with pathological conditions unrelated to breast masses were excluded from the study to maintain dataset specificity.

To ensure uniformity, all selected images were resized to a consistent matrix size. Image enhancement and noise reduction techniques were subsequently applied to improve image quality and ensure suitability for breast tissue classification.

To address class imbalance and enhance the robustness of the artificial intelligence model, image augmentation techniques were applied to both normal and abnormal breast ultrasound images. Augmentation methods included rotation, translation, and zooming. This approach generated an equal number of images across classes, resulting in 9,000 images of normal breast tissue and 9,000 images of abnormal breast tissue, for a total of 18,000 images.

The dataset was divided into three subsets as follows:

- *Training set*: 70% of the images (12,600 images) were used to train the artificial intelligence model.
- *Validation set*: 20% of the images (3,600 images) were allocated for model validation to optimize hyperparameters and prevent overfitting.
- *Testing set*: 10% of the images (1,800 images) were reserved for evaluating the model's performance and generalization capability.

Development of the football formation-based artificial intelligence model

Concept and design

The proposed artificial intelligence model was developed using a one-channel convolutional neural network architecture, with its internal structure based on DenseNet.[20] The model design was inspired by football formations, categorized into offensive and defensive strategies. Four football formations were selected for implementation:

Attacking formations:

4-3-3 and 4-2-3-1: These formations are recognized as highly effective offensive strategies and are widely adopted by top-tier football clubs such as Manchester City, Real Madrid, Liverpool, and Bayern Munich.

Defensive formations:

5-4-1 and 4-4-2: These formations are well-established defensive strategies commonly employed by elite football clubs, including Manchester City, Real Madrid, and Liverpool.

Architecture of the one-channel football formation neural network (OC-FFNet)

The one-channel football formation neural network (OC-FFNet) was designed to align with football formations, where the number of blocks within the network corresponded to the number of players in each position. The architecture was primarily based on DenseNet, incorporating dense blocks and concatenation layers to enhance feature propagation and reuse.[20]

Key components of dense block

The fundamental components of dense block include:

1. Batch normalization
2. Activation function
3. Convolutional layer
4. Concatenation layer

Modified dense block structure in OC-FFNet

Each modified dense block in the OC-FFNet consists of six layers:

1. Batch normalization
2. ReLU activation function
3. Convolutional layer
4. Batch normalization
5. ReLU activation function
6. Convolutional layer

Transition layers

At the initial stage of the model, the architecture comprises an input layer, convolutional layers, and max pooling layers. Between each position block, transition layers are incorporated to facilitate information flow. Each transition layer consists of:

1. Batch normalization
2. Convolutional layer
3. Average pooling

Final processing layers

Before producing the final model output, the following layers are included:

1. Global average pooling layer
2. Flatten layer
3. Softmax function

Football formation-based network structure

The number of blocks within different sections of the network was determined based on each football formation (Further details are provided in Figure 2):

4-3-3 Formation: 4 Defensive blocks, 3 Midfield blocks, 3 Forward blocks

4-2-3-1 Formation: 4 Defensive blocks, 2 Defensive midfield blocks, 3 Attacking midfield blocks, 1 Forward block

5-4-1 Formation: 5 Defensive blocks, 4 Midfield blocks, 1 Forward block

4-4-2 Formation: 4 Defensive blocks, 4 Midfield blocks, 2 Forward blocks

Each model was trained using batch sizes ranging from 64 to 256, with the number of training epochs varied between 50 and 150.

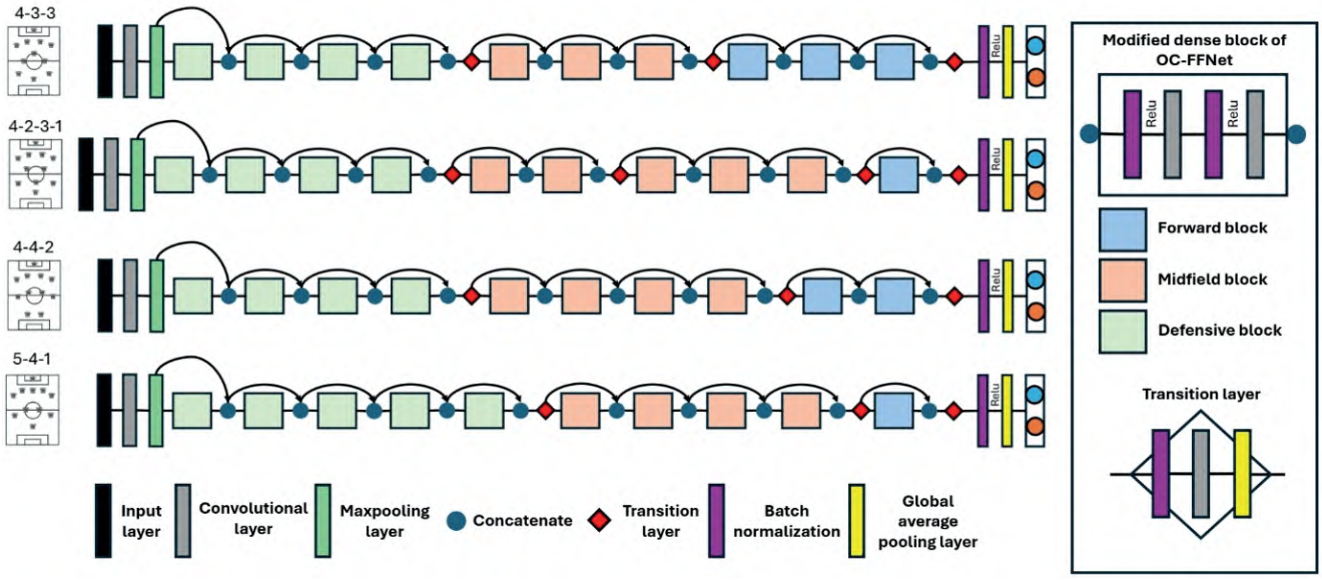


Figure 2*. illustrates the structure of four football formation networks, the structure of a modified dense block in OC-FFNet and the transition layer in OC-FFNet.*

Performance evaluation

The models' performance in detecting breast cancer in ultrasound images was evaluated using a confusion matrix. The classification outcomes were defined as follows:

True positives (TP): The model correctly identifies and classifies breast cancer in ultrasound images.

True negatives (TN): The model correctly identifies normal images without misclassifying them as cancerous.

False positives (FP): The model incorrectly detects breast cancer in a normal image or misclassifies a benign lesion.

False negatives (FN): The model fails to detect breast cancer in an image that contains a malignant tumor.

To assess model performance, the following metrics were calculated: accuracy, precision (positive predictive value), sensitivity (recall), specificity, F1-score, false positive rate (FPR), and area under the curve (AUC).

Results

Evaluation of the performance of four football formation models

Tables 1 to 2 present the performance evaluation of artificial intelligence models developed using four football formations: 4-3-3, 4-2-3-1, 4-4-2, and 5-4-1. These models were evaluated using various batch sizes ranging from 64 to 256 and training epochs ranging from 50 to 150. The performance metrics included accuracy, positive predictive value (precision), sensitivity (recall), specificity, F-1 score, false positive rate, and the area under the curve. Additionally, the prediction time for 1,800 test images was recorded.

The 4-3-3 formation model exhibited optimal performance with a batch size of 128 and 100 epochs. It obtained an accuracy of 0.999, a precision of 0.999, a sensitivity of 1.000, a specificity of 0.999, an F-1 score of 0.999, a false positive rate of 0.001, and an AUC of 0.999. The prediction time for this model was 6.588 seconds. The 4-2-3-1 formation model demonstrated optimal performance using a batch size of 64 and 150 epochs. It achieved an accuracy of 0.963, a precision of 0.927, a sensitivity of 1.000, a specificity of 0.932, an F-1 score of 0.962, a false positive rate of 0.068, and an AUC of 0.963. The prediction time for this model was 5.122 seconds, as shown in Table 1.

Table 1. The performance of attacking formations (4-3-3 and 4-2-3-1 football formation models)

	Batch size	Epoch	Accuracy	Precision	Sensitivity	Specificity	F1-score	False positive rate	AUC	time (sec)
4-3-3	64	50	0.500	1.000	0.500	0.000	0.667	0.000	0.500	5.044
		100	0.998	0.997	1.000	0.997	0.998	0.003	0.998	4.918
		150	0.967	1.000	0.939	1.000	0.968	0.000	0.967	4.956
	128	50	0.500	1.000	0.500	0.000	0.667	0.000	0.500	6.329
		100	**0.999**	**0.999**	**1.000**	**0.999**	**0.999**	**0.001**	**0.999**	**6.588**
		150	0.803	0.606	1.000	0.717	0.754	0.283	0.803	4.917
	256	50	0.500	1.000	0.500	0.000	0.667	0.000	0.500	6.360
		100	0.500	1.000	0.500	0.000	0.667	0.000	0.500	6.660
		150	0.988	0.977	1.000	0.977	0.988	0.228	0.988	4.701
4-2-3-1	64	50	0.547	0.853	0.529	0.621	0.653	0.379	0.547	5.102
		100	0.850	0.966	0.784	0.955	0.866	0.045	0.850	5.088
		150	**0.963**	**0.927**	**1.000**	**0.932**	**0.962**	**0.068**	**0.963**	**5.122**
	128	50	0.762	0.523	1.000	0.677	0.687	0.323	0.762	4.943
		100	0.545	1.000	0.524	1.000	0.678	0.000	0.545	4.871
		150	0.937	0.880	0.993	0.892	0.933	0.108	0.937	4.940
	256	50	0.788	0.577	1.000	0.703	0.732	0.297	0.788	4.935
		100	0.500	0.000	0.000	0.500	0.000	0.500	0.500	5.189
		150	0.928	0.857	1.000	0.875	0.923	0.125	0.928	4.962

The 5-4-1 formation model achieved its best performance using a batch size of 64 and 150 epochs. It recorded an accuracy of 0.968, a precision of 0.994, a sensitivity of 0.945, a specificity of 0.994, an F-1 score of 0.969, a false positive rate of 0.006, and an AUC of 0.968. The prediction time for the test images was 5.327 seconds. The 4-4-2 formation model demonstrated the highest performance using a batch size of 128 and 100 epochs. Specifically, it achieved an accuracy of 0.999, a precision of 0.999, a sensitivity of 1.000, a specificity of 0.999, an F-1 score of 0.999, a false positive rate of 0.001, and an AUC of 0.999. The prediction time for the test images was 6.647 seconds, as shown in Table 2.

The models based on the 4-4-2 and 4-3-3 formations exhibited the highest accuracy and AUC values, indicating superior predictive performance. The 4-2-3-1 model also demonstrated strong performance, although its specificity and false positive rate were comparatively lower. In contrast, the 5-4-1 formation model showed slightly lower accuracy and AUC metrics. Notably, the consistency in prediction times across most models suggests that computational efficiency was maintained regardless of the formation strategy employed. However, the faster prediction time of the 4-2-3-1 model warrants further investigation to understand the underlying factors contributing to this difference.

Table 2. The performance of defensive formations (5-4-1 and 4-4-2 football formation models)

	Batch size	Epoch	Accuracy	Precision	Sensitivity	Specificity	F1-score	False positive rate	AUC	Time (sec)
5-4-1	64	50	0.609	0.894	0.569	0.754	0.696	0.246	0.609	5.410
		100	0.529	1.000	0.515	1.000	0.680	0.000	0.529	5.400
		150	**0.968**	**0.994**	**0.945**	**0.994**	**0.969**	**0.006**	**0.968**	**5.327**
	128	50	0.500	1.000	0.500	0.000	0.667	0.000	0.500	9.030
		100	0.967	0.933	1.000	0.938	0.966	0.063	0.967	5.405
		150	0.912	0.823	1.000	0.850	0.903	0.150	0.912	5.353
	256	50	0.548	0.963	0.526	0.783	0.681	0.217	0.548	5.276
		100	0.726	0.979	0.650	0.957	0.781	0.043	0.726	5.189
		150	0.756	0.511	1.000	0.672	0.677	0.328	0.756	5.267
4-2-2	64	50	0.495	0.990	0.497	0.000	0.662	1.000	0.495	5.415
		100	0.994	0.999	1.000	0.999	0.999	0.001	0.999	5.928
		150	0.497	0.996	0.499	0.000	0.665	1.000	0.498	5.196
	128	50	0.751	0.917	0.688	0.875	0.786	0.125	0.751	6.706
		100	**0.999**	**0.999**	**1.000**	**0.999**	**0.999**	**0.001**	**0.999**	**6.647**
		150	0.980	0.999	0.963	0.999	0.980	0.001	0.980	6.635
	256	50	0.500	1.000	0.500	0.000	0.667	0.000	0.500	5.089
		100	0.500	1.000	0.500	0.000	0.667	0.000	0.500	5.033
		150	0.500	1.000	0.500	0.000	0.667	0.000	0.500	6.723

Discussion

This study introduces a novel approach to artificial intelligence model development, representing the first instance in which football formation strategies have been utilized as a structural basis for designing and constructing convolutional neural network architectures. By analyzing the spatial arrangement of football formations, the research team identified the interconnected positioning of players within various tactical strategies. This observation led to the hypothesis that the number of players in each position could be mapped onto the number of layers used for feature extraction in the hidden layers of the convolutional neural network model.

In football strategy planning, coaches select formations based on their emphasis on either offensive or defensive play. Tactical formations have evolved significantly over time. During the 20th century, the 5-3-2 formation was predominantly employed for defensive strategies, whereas the 4-2-4 formation gained prominence in the 1950s for its offensive effectiveness, achieving considerable success during that period[13]. In contemporary football, formations such as 4-3-3, 4-4-2, and 3-5-2 are widely implemented to optimize match performance.[13]

Empirical studies by Forcher have indicated that the 4-4-2 and 4-2-3-1 formations are among the most frequently utilized strategies in professional football.[23] Additionally, research published in 2019[24] and 2023[25] demonstrated that the 4-2-3-1 formation significantly influences the physical and tactical performance of wingers in amateur football. Zhang *et al*.[26] further examined the impact of formations on the physical and technical performance of players in the Chinese Super League, concluding that formation selection has a direct effect on match outcomes. Case studies have also highlighted this impact; for example, in the 2009 UEFA Champions League Final, FC Barcelona secured a 2-0 victory over Manchester United by employing their signature 4-3-3 formation.[27] Furthermore, defensive formations, such as the 4-5-1 strategy, have been strategically implemented to reinforce defensive stability.[28]

Each football formation offers distinct strategic advantages, rendering them suitable for different styles of play. In this study, football formations were integrated into the artificial intelligence model's architecture in a unidirectional manner, consistent with the convolutional neural network framework. The dense block structure was adapted from DenseNet and the number of modified dense blocks in the hidden layers was determined based on the number of players in each position.[20] Specifically, defenders, midfielders, and forwards were assigned distinct layers, while the goalkeeper was designated as the input layer.

In the development of the football formation model, each group of players sequentially transitions into the next layer: the defense block transitions into the midfield block, which subsequently transitions into the Forward Block. Following feature extraction, the processed data is propagated to the fully connected layers, which incorporate batch normalization with ReLU activation and global average pooling to mitigate data redundancy before classification using the Softmax function.

The artificial intelligence models based on the four football formations were trained to distinguish between normal and cancerous breast ultrasound images. The performance of these models was influenced by hyperparameter configurations, including batch size and the number of training epochs. The results indicate that the 4-3-3 and 4-4-2 formations achieved optimal classification performance when trained with a batch size of 128 and 100 epochs. In contrast, the 4-2-3-1 and 5-4-1 formations yielded the best results with a batch size of 64 and 150 epochs.

These findings suggest that formations with a higher number of defense and midfield blocks necessitate a smaller batch size and an increased number of training iterations to attain optimal performance. Conversely, formations such as 4-3-3 and 4-4-2, which exhibit a balanced number of blocks, perform effectively with a larger batch size and fewer training iterations. Furthermore, the models exhibited similar prediction times across all formations, indicating that computational efficiency was maintained regardless of the selected formation.

Table 3 provides a comparative analysis of the breast cancer detection performance in this study relative to related research. The proposed model demonstrates superior performance compared to the study conducted by Rakibul Islam, which utilized the EDCNN method to classify breast ultrasound images into three categories: normal, benign, and malignant. Their study evaluated classification performance across three distinct image datasets, with the highest performance observed on the BUSI dataset, achieving an accuracy of 87.82%, a sensitivity of 85.33%, and a precision of 87.33%[21].

Table 3. Comparison of our proposed models and previous work.

Authors	Study	Model	Method	Test dataset (Images)	Training and validation dataset (Images)	Performance (%)
Islam[21]	Normal, benign, and malignant classification	EDCNN	Deep learning	BUSI Dataset 156	624	ACC 87.82 SEN 85.33 PRE 87.33
				UDAIT Dataset 33	130	ACC 85.69 SEN 78.00 PRE 84.00
				UDAIT Dataset 33	130	ACC 85.69 SEN 78.00 PRE 84.00
Umapathi[22]	Normal, benign, and malignant classification	SRAD	Deep learning	60	100	ACC 99.99 AUC 99.00 SEN 94.10 SPEC 96.12
Our proposed	Normal and breast cancer classification	OC-FFNet (4-4-2)	Deep learning	1,800	16,200	ACC 99.99 AUC 0.99 PRE 99.99 SEN 100.00 SPEC 99.99 F-1 99.99
		OC-FFNet (4-3-3)	Deep learning	1,800	16,200	ACC 99.99 AUC 0.99 PRE 99.99 SEN 100.00 SPEC 99.99 F-1 99.99
		OC-FFNet (5-4-1)	Deep learning	1,800	16,200	ACC 96.80 AUC 0.968 PRE 99.40 SEN 94.50 SPEC 99.40 F-1 96.90
		OC-FFNet (4-2-3-1)	Deep learning	1,800	16,200	ACC 96.30 AUC 0.963 PRE 92.70 SEN 100.00 SPEC 93.20 F-1 96.20

***Note**: ACC: accuracy, AUC: area under the curve, PRE: precision, SEN: sensitivity, SPEC: pecificity, F-1: F-1 score.*

Additionally, the performance of the proposed model is comparable to that of Umapathi's study, which employed the SRAD method for breast ultrasound image classification. Umapathi's model achieved an accuracy of 99.99%, a sensitivity of 99.00%, a specificity of 84.00%, and an area under the curve (AUC) of 99.00%.[22]

However, while both studies utilized artificial intelligence models to classify breast ultrasound images into three categories—normal, benign, and malignant, our study combined benign and malignant images into a single category.[16,29] Future research will focus on evaluating the proposed artificial intelligence model's effectiveness in a three-class classification setting. Nevertheless, this study utilized a larger dataset than both studies, which likely contributed to reducing the risk of overfitting during model training.

Furthermore, certain studies specifically assess the ability to differentiate between benign and malignant breast ultrasound images to evaluate classification performance.

Limitations

The dataset utilized in this study exhibited a high degree of similarity between images, presenting challenges in distinguishing between normal and cancerous cases. To mitigate this issue, image augmentation techniques were implemented to expand the dataset and enhance the model's ability to recognize subtle differences. Enhancing the diversity of both normal and cancerous ultrasound images would further improve the artificial intelligence model's learning process. A more varied dataset could contribute to better generalization and higher classification accuracy in real-world applications.

Using the proposed method, the augmented images may not clearly represent the target population. Future testing should include a greater number of images with similar characteristics or apply minimal augmentation to ensure better representation in the study.

Conclusion

The development of artificial intelligence model architectures based on football formations, incorporating modifications to the dense block structure in accordance with the number of players in each formation, yielded notable performance results. Among the four formations examined, the 4-4-2 and 4-3-3 models exhibited the highest efficiency, as their balanced distribution of modified dense blocks enhanced feature extraction capabilities. This structural advantage resulted in superior performance compared to the 4-2-3-1 and 5-4-1 formations. Nonetheless, all four models effectively distinguished between normal and cancerous breast ultrasound images when optimal parameter configurations were applied.

Conflict of interest

None

Funding

This study received support from the Faculty of Allied Health Sciences, with a grant for general researchers (Grant No. AH-68-01-004)

Ethical Approval

This study was approved by the Ethics Committee of Naresuan University, Thailand (IRB No. P1-0114/2567).

References

[1] World Health Organization (WHO). Breast cancer [cited 2024 September 7]. Available from: https://www.who.int/news-room/fact-sheets/detail/breast-cancer.

[2] The Ministry of Public Health of Thailand, Mammography and ultrasound breast cancer screening to be accessible for Thai women with high-risk [cited 2024 October 14]. Available from: https://eng.nhso.go.th/view/1/DescriptionNews/Mammography-and-ultrasound-breast-cancer-screening-to-be-accessible-for-Thai-women-with-high-risk/592/EN-US.

[3] Brem RF, Lenihan MJ, Lieberman J, Torrente J. Screening Breast Ultrasound: Past, Present, and Future. AJR Am J Roentgenol. 2015; 204(2): 234-40. doi: 10.2214/AJR.13.12072.

[4] Chongthawonsatid S. Inequity of healthcare utilization on mammography examination and Pap smear screening in Thailand: Analysis of a population-based household survey. PLoS One. 2017; 12(3): e0173656. doi: 10.1371/journal.pone.0173656.

[5] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521(7553): 436-44. doi: 101038/nature14539

[6] Xin M, Wang Y. Research on image classification model based on deep convolution neural network. JIVP. 2019; 2019: 40. doi: 101186/s13640-019-0417-8.

[7] Sathiyapriya G, Shanthi SA. Image classification using convolutional neural network. First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichy, India, 2022, pp. 1-4. doi: 10.1109/ICEEICT53079.2022.9768622.

[8] Alotaibi M, Aljouie A, Alluhaidan N, Qureshi W, Almatar H, Alduhayan R, *et al*. Breast cancer classification based on convolutional neural network and image fusion approaches using ultrasound images. Heliyon. 2023; 9(11): e22406. doi: 101016/j.heliyon.2023.e22406.

[9] Qi X, Yi F, Zhang L, Chen Y, Pi Y, Chen Y, Guo J, *et al*. Computer-aided diagnosis of breast cancer in ultrasonography images by deep learning. Neurocomputing. 2022; 472: 152-65. doi: 101016/j.neucom.2021.11.047.

[10] Li Q, Lei J, Ren C, Peng Z, Hong J. Research on sports image classification method based on SE-RES-CNN model. Sci Rep. 2024; 14: 19087. doi: 101038/s41598-024-69965-5.

[11] Sotudeh H. The Principles of Tactical Formation Identification in Association Football (Soccer) - A Survey. Front. Sports Act. Living. 2025; 6 (2024). doi: 103389/fspor.2024.1512386

[12] White M. Formations in football: Every modern formation and system. [cited 2025 March 6]. Available

from: https://www.fourfourtwo.com/features/formations-in-football.

[13] Simpson J. Football formations: A Comprehensive guide to the most popular tactics in the game. [cited 2025 March 6]. Available from: https://www.team-stats.net/blog/football-formations-a-comprehensive-guide-to-the-most-popular-tactics-in-the-game.

[14] Izquierdo JM, Marqués-Jiménez D, Redondo JC. Running demands and tactical individual actions of wingers appear to depend on the playing formations within an amateur football team. Sci Rep. 2023; 13: 8927. doi: 101038/s41598-023-36157-6.

[15] He K, Zhang X, Ren S and Sun J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016; pp. 770-8. doi: 10.1109/CVPR.2016.90.

[16] Chen SH, Wu YL, Pan CY, Lian LY, Su QC. Breast ultrasound image classification and physiological assessment based on GoogLeNet. J Radiat Res Appl Sci. 2023; 16(3): 100628. doi: 101016/j.jrras.2023.100628.

[17] Lu SY, Wang SH, Zhang YD. BCDNet: An optimized deep network for ultrasound breast cancer detection. IRBM. 2023; 44(4): 100774. doi: 101016/j.irbm.2023.100774.

[18] Sairam VA. Ultrasound breast images for breast cancer 2022. [cited 2024 October 6]. Available from: www.kaggle.com/datasets/vuppalaadithyasairam/ultrasound-breast-images-for-breast-cancer.

[19] Shah A. Breast Ultrasound Images Dataset 2021. [cited 2024 October 6]. Available from: www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset.

[20] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 2261-9. doi: 10.1109/CVPR.2017.243.

[21] Islam MR, Rahman MM, Ali MS, Nafi AAN, Alam MS, Godder TK, *et al*. Enhancing breast cancer segmentation and classification: An ensemble deep convolutional neural network and U-net approach on ultrasound images. MLWA. 2024; 16: 100555. doi: 101016/j.mlwa.2024.100555.

[22] Umapathi K, Shobana S, Nayyar A, Justin J, Vanithamani R, Galindo MV, *et al*. Novel approach to breast tumor detection: Enhanced speckle reduction and hybrid classification in ultrasound imaging. CMC. 2024; 79(2): 1875-901. doi: 1032604/cmc.2024.047961.

[23] Forcher L, Forcher L, Wäsche H, Jekauc D, Woll A, Altmann S. The influence of tactical formation on physical and technical match performance in male soccer: A systematic review. Int J Sports Sci Coach. 2023; 18(5), 1820-49. doi: 101177/17479541221101363.

[24] Memmert D, Raabe D, Schwab S, Rein R. A tactical comparison of the 4-2-3-1 and 3-5-2 formation in soccer: A theory-oriented, experimental approach based on positional data in an 11 vs. 11 game Set-up. PLoS ONE. 2019; 14(1): e0210191. doi: 101371/journal.pone.0210191.

[25] Bauer P, Anzer G, Shaw L. Putting team formations in Association Football into Context. JSA. 2023; 9(1): 39-59. doi: 103233/JSA-220620.

[26] Zhang W, Gong B, Tao R, Zhou F, Ruano MA, Zhou C. The influence of tactical formation on physical and technical performance across playing positions in The Chinese Super League. Sci. Rep. 2024;14: 2538. doi: 101038/s41598-024-53113-0.

[27] Coaches' Voice. The 4-3-3: Football tactics explained. [cited 2025 February 24]. Available from: https://learning.coachesvoice.com/cv/4-3-3-football-tactics-explained-formation-liverpool-klopp-barcelona-guardiola/.

[28] Coaches' Voice. The 4-5-1 Formation: Football tactics explained. [cited 2025 February 24]. Available from: https://learning.coachesvoice.com/cv/4-5-1-formation-football-tactics/.

[29] Shen WJ, Zhou HX, He Y, Xing W. Predicting female breast cancer by artificial intelligence: combining clinical information and BI-RADS ultrasound descriptors. WFUMB. 2023; 1(2): 100013. doi: 101016/j.wfumbo.2023. 100013

Journal of Associated Medical Sciences 2025; 58(3): 10-24

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

Hepatotoxic effects of sildenafil-containing "Tiger King" herbal supplement in a rat model: An *in vitro* and *in vivo* study

Rana A. Al-Saadi*, Tuqa M. Abdul-Saheb, Ban Thabit, Nisreen Khazaal Flayyih, Ula Al-Kawaz

High Institute of Infertility Diagnosis and Assisted Reproductive Technologies, Al-Nahrain University, Kadhimiya, Baghdad, Iraq.

ARTICLE INFO

Article history:
Received 7 January 2025
Accepted as revised 10 April 2025
Available online 26 April 2025

Keywords:
Drug-induce, erectile dysfunction, herbal medicine, liver injury, sildenafil citrate.

ABSTRACT

Background: The global market for erectile dysfunction (ED) treatments has seen a rise in herbal supplements marketed as natural alternatives to prescription medications. However, many of these products contain undeclared pharmaceutical ingredients, posing significant health risks.

Objective: This study aimed to investigate the chemical composition and potential hepatotoxic effects of "Tiger King," a purported Chinese herbal supplement for sexual enhancement, using both *in vitro* and *in vivo* experiments.

Materials and methods: Chemical analysis of "Tiger King" tablets was performed using thin-layer chromatography and colorimetric tests. Antioxidant activities were evaluated using DPPH and FRAP assays. *In vivo* studies were conducted using male Wistar rats (N=20) divided into four groups: control, clomiphene, low-dose "Tiger King" (5 mg/kg) or TK1, and high-dose "Tiger King" (10 mg/kg) or TK2. Treatments were administered orally for 30 days. Serum testosterone levels, sperm parameters, oxidative stress markers, liver and kidney function tests, and histopathological changes were assessed.

Results: Chemical analysis revealed the presence of sildenafil in "Tiger King" tablets, with no detectable amounts of the claimed herbal ingredients. *In vivo* studies showed significant increases in sperm count and testosterone levels in treated groups. However, oxidative stress markers (MDA, GSH) were significantly altered, and liver function tests (ALT, AST, ALP, bilirubin) were elevated in treatment groups with ALT increased by 17.8% (from 45.0±1.1 to 53.0±1.1 U/L), AST by 12.5% (from 120.0±1.1 to 135.0±1.1 U/L) in the high-dose TK2 group. Histopathological examination revealed mild to moderate changes in liver, kidney, and reproductive organs of treated animals, including hepatic steatosis, renal glomerular congestion, and glandular atrophy in reproductive tissues.

Conclusion: This study provides evidence that "Tiger King" contains undeclared sildenafil and lacks the advertised herbal components. Its use is associated with improved reproductive parameters but also with significant biochemical and histopathological changes which draw attention to potential health risks of adulterated herbal supplements.

Introduction

The global market for erectile dysfunction (ED) treatments has seen a significant surge in the availability of herbal supplements marketed as "natural" alternatives to prescription medications.[1]especially with regard to the treatment of erectile dysfunction (ED These products, often promoted as safer options, have raised concerns among healthcare professionals and regulatory bodies due to the potential presence of undeclared pharmaceutical ingredients and their associated health risks. One such

* **Corresponding contributor.**
Author's Address: *High Institute of Infertility Diagnosis and Assisted Reproductive Technologies, Al-Nahrain University, Kadhimiya, Baghdad, Iraq.*
E-mail address: ranaalsaadi@st.nahrainuniv.edu.iq
doi: 10.12982/JAMS.2025.068
E-ISSN: 2539-6056

product, "Tiger King", a purported Chinese herbal remedy for sexual enhancement, has gained popularity in various countries, including Israel, Australia, Canada, and the United States.[2-4] Recent investigations have found a disturbing trend in the composition of these "herbal" supplements. Numerous reports have detected the presence of phosphodiesterase type 5 (PDE-5) inhibitors and their analogues in products labeled as "100% natural" or "herbal".[5,6] Sildenafil, the active ingredient in Viagra®, and its analogues are frequently found in these supplements, often in varying and potentially dangerous doses.[7,8] This adulteration poses significant health risks, particularly as consumers may be unaware of the actual contents of these products.

The case of "Tiger King" is particularly alarming. Chemical analyses conducted by health authorities in the United States, Israel, Australia, and Canada have consistently revealed the presence of sildenafil in these tablets, with doses ranging from therapeutic levels to potentially harmful amounts exceeding 200 mg per tablet.[9] More concerning is the complete absence of the herbal ingredients listed on the product label, further emphasizing the questionable nature of these supplements. However, and while the efficacy of sildenafil in treating ED is well-established, its safety profile, particularly in uncontrolled doses and in combination with undisclosed ingredients, remains a subject of concern. Recent case reports have suggested a possible link between sildenafil use and hepatotoxicity, a connection not widely recognized or reported in standard prescribing information.[2,10-13] These cases raise important questions about the safety of sildenafil, especially in patients with pre-existing liver conditions or when consumed in unregulated supplements. Similarly, recent studies have demonstrated the importance of rigorous testing methodologies when evaluating reproductive parameters and drug safety.[14]

The potential for hepatotoxicity associated with sildenafil is particularly worrisome given the prevalence of liver disease globally and the tendency of many patients to seek "natural" remedies without medical supervision. The mechanism by which sildenafil might induce liver injury remains unclear, with hypotheses ranging from idiosyncratic reactions to dose-dependent toxicity in susceptible individuals.[13] Despite these concerns, there is a small number of controlled studies investigating the hepatotoxic potential of sildenafil, particularly in the context of adulterated herbal supplements like "Tiger King". In this respect, we hypothesized that the undeclared presence of sildenafil in "Tiger King" supplements may cause health risks, particularly hepatotoxicity, despite potential benefits for sexual function. Therefore, this study aims to address this critical knowledge gap by investigating the potential hepatotoxic effects of "Tiger King" and its primary active ingredient, sildenafil, in a controlled animal model.

Materials and methods:

In vitro studies

Chemical analysis

Thin-layer chromatography was performed using silica gel 60 F254 plates (Merck, Germany). The mobile phase consisted of ethyl acetate:methanol:ammonium hydroxide (85:10:5 v/v/v).[15] Samples were prepared by dissolving one TK tablet in 10 mL of methanol, sonicating for 15 minutes, and filtering through a 0.45 µm membrane filter. Ten microliters of the sample solution and authenticated sildenafil standard (Sigma-Aldrich, USA) were applied to the TLC plate. After development, the plates were visualized under UV light at 254 nm and 366 nm.[16] Sildenafil identification was confirmed by co-chromatography with reference standard (Rf=0.68±0.02) and subsequent phosphomolybdic acid reagent spray followed by heating at 105 °C for 5 minutes. The presence of sildenafil was indicated by a blue spot.[17]

In addition to TLC analysis, colorimetric tests were also performed such as Dragendorff's test (for alkaloids),[18] vanillin-sulfuric acid test (for ginseng saponins),[19] and ninhydrin test (for amino acids and peptides),[20] as well as the visual inspection for synthetic dyes.[21] All tests were performed in triplicate to ensure reproducibility. Standard solutions of sildenafil, tadalafil, yohimbine, ginseng extract, and *Cordyceps* extract were used as positive controls.[22]

The TK tablets were ground into fine powder using a clean mortar and pestle (Figure 1). For *in vitro* experiments, the powder was dissolved in methanol (10 mg/mL), sonicated for 15 minutes, and filtered through a 0.45 µm membrane filter. For chemical analysis, 10 µL of this solution was used for TLC analysis and appropriate dilutions were made for other assays.

Figure 1. *Tiger king tablets from the local market.*

Antioxidant activities

The antioxidant activities of "Tiger King" were evaluated using multiple assays. The 2,2-diphenyl-1-picrylhydrazyl (DPPH) radical scavenging assay was performed to measure free radical scavenging capacity.[23] The ferric reducing antioxidant power (FRAP) assay was used to assess the supplement's reducing power.[24] To ensure measurement consistency, all assays were performed in duplicate, with ascorbic acid and Trolox as positive controls.

In vivo studies

Animal model and treatment

Male Wistar rats (8-10 weeks old, weighing 200-250 gm) were used for the *in vivo* studies. Animals were housed under standard laboratory conditions (12 hrs light/dark cycle, 22±2 °C, 55±5% humidity) with free access to food and water. Rats were randomly divided into four groups (N=5 per group): control, clomiphene, low-dose "Tiger King" (5 mg/kg), high-dose "Tiger King" (10 mg/kg). Treatments were administered orally once daily for 30 days. For oral administration, TK tablets were ground into fine powder and dissolved in sterile saline (10 mg/mL). The solution was sonicated for 15 minutes and filtered through a 0.45 µm membrane filter. Fresh solutions were prepared daily and administered via oral gavage according to the designated doses for each treatment group. All animal procedures were approved by the Institutional Ethics Committee and conducted in accordance with international guidelines for the care and use of laboratory animals.

Sperm parameters

On day 29, rats were euthanized, and epididymal sperm were collected. Sperm count was determined using a hemocytometer. Sperm motility was assessed using computer-assisted sperm analysis (CASA). Sperm morphology was evaluated by examining Eosin-Nigrosin stained smears under a light microscope, counting at least 200 sperm per sample and classifying them as normal or abnormal based on head, midpiece, and tail morphology.

Biochemical analysis

Blood samples were collected via cardiac puncture, and serum was separated by centrifugation. Liver function was assessed by measuring serum levels of alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), and total bilirubin using commercial kits. Oxidative stress markers, including malondialdehyde (MDA), reduced glutathione (GSH), and superoxide dismutase (SOD) activity, were measured in liver homogenates using spectrophotometric methods.

Histopathological changes

Liver and testicular tissues were fixed in 10% neutral buffered formalin, embedded in paraffin, sectioned at 5 µm thickness, and stained with hematoxylin and eosin (H&E). Slides were examined under a light microscope by a pathologist blinded to the treatment groups. Liver sections were assessed for signs of hepatotoxicity, including steatosis, inflammation, and necrosis. Testicular sections were evaluated for changes in seminiferous tubule morphology, spermatogenesis, and interstitial tissue.

Statistical analysis

All data were analyzed using SPSS version 27 (IBM, USA). Data normality was assessed using Shapiro-Wilk test and the results are presented as mean±SD. One-way analysis of variance (ANOVA) followed by Tukey's post-hoc test was used to compare differences between groups. For fertility rate, Chi-square test was employed. *The* $p<0.05$ was considered statistically significant. Graphs were generated using GraphPad Prism 9 (GraphPad Software, USA).

Results

Chemical analysis

Thin-layer chromatography (TLC) and colorimetric tests were performed to detect the presence of key compounds in the Tiger King (TK) supplement. The results are summarized in Table 1. The TLC analysis revealed a prominent spot (Rf=0.68) under UV light at 254 nm, characteristic of sildenafil. This spot also gave a positive reaction with phosphomolybdic acid spray reagent which additionally confirms the presence of sildenafil. No spots corresponding to tadalafil or other common PDE-5 inhibitors were observed under UV light at 366 nm. The Dragendorff's reagent, used to detect alkaloids such as yohimbine, did not produce any orange-brown spots, which indicates the absence of these compounds. Also, the vanillin-sulfuric acid reagent, typically used to detect ginseng saponins, did not produce any characteristic purple spots, which qualitatively means the absence of ginseng in the sample. Moreover, the ninhydrin reagent, used to detect amino acids and peptides found in *Cordyceps*, did not produce any purple spots. The visual inspection under normal and UV light (366 nm) revealed the presence of synthetic dyes, likely used for tablet coloration.

Table 1. Qualitative chemical analysis of Tiger King (TK).

Compound	Method of detection	Result
Sildenafil	TLC + UV (254 nm)	Present
Tadalafil	TLC + UV (366 nm)	Absent
Yohimbine	Dragendorff's reagent	Absent
Ginseng saponins	Vanillin-sulfuric acid reagent	Absent
Cordyceps markers	Ninhydrin reagent	Absent
Synthetic dyes	Visual + UV (366 nm)	Present

Antioxidant activities

The DPPH radical scavenging activity of TK extract exhibited a concentration-dependent response (Figure 2A). At the lowest concentration tested (10 µg/mL), TK showed 15.3±1.2% inhibition, increasing to 84.2±1.9% at the highest concentration (200 µg/mL). In comparison, ascorbic acid (AA), used as a positive control, demonstrated higher activity, 38.5±1.7% inhibition at 10 µg/mL and 98.9±0.5% at 200 µg/mL. The calculated IC_{50} value for TK was approximately 70 µg/mL, while for AA it was below 25 µg/mL.

The ferric reducing antioxidant power (FRAP) of TK extract also showed a concentration-dependent increase (Figure 2B). At 50 µg/mL, TK exhibited a FRAP value of 215.3±7.8 µmol Fe(II)/gm extract, which rose to 789.4±18.3 µmol Fe(II)/gm extract at 200 µg/mL. The standard antioxidant Trolox demonstrated higher reducing power, FRAP values of 456.7±11.3 and 1785.3±28.6 µmol Fe(II)/gm at 50 and 200 µg/mL, respectively.

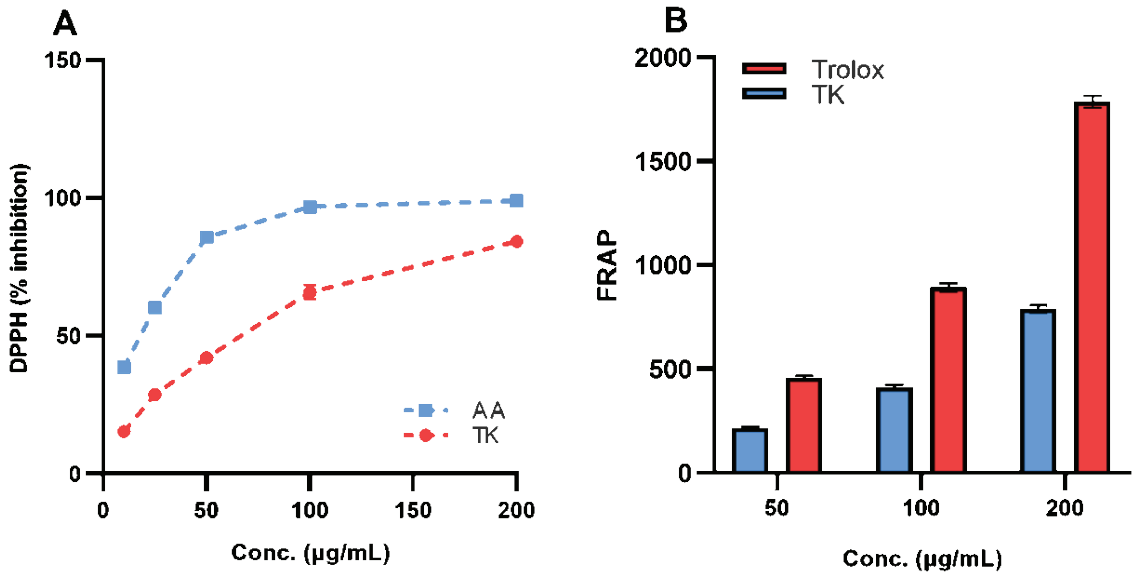


Figure 2. *Antioxidant activity of Tiger King (TK) extract compared to standard antioxidants. A: DPPH radical scavenging activity of TK extract and ascorbic acid (AA), B: FRAP of TK extract compared to Trolox.*

Weight change

Body weight changes in rats were monitored at three time points: day 1, day 14, and day 30, for all treatment groups (Figure 3). At the start of the experiment (day 1), the mean body weights were comparable among groups: control (300.0±1.6 gm), clomiphene (298.8±3.3 gm), TK1 (296.0±1.6 gm), and TK2 (301.0±1.6 gm). By day 14, all groups showed a slight increase in weight: control (305.0±1.6 gm), clomiphene (303.6±3.4 gm), TK1 (300.0± 1.6 gm), and TK2 (306.0±1.7 gm). The thing continued through day 30, with further weight gains observed in all groups: control (310.0±1.6 gm), clomiphene (309.8±3.4 gm), TK1 (306.0±1.7 gm), and TK2 (312.0±1.6 gm). Despite these changes, statistical analysis revealed no significant differences in body weight between the treatment groups and the control group at almost all time points (p>0.05 mostly). These results suggest that neither clomiphene nor the two doses of TK significantly affected body weight gain over the course of the experiment.

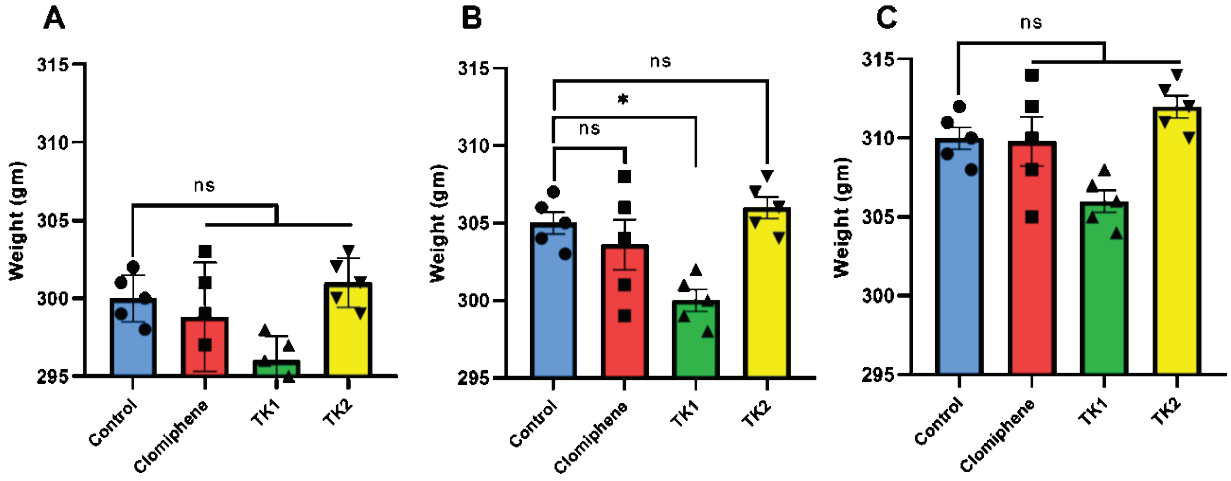


Figure 3. *Body weight changes in rats over the course of the experiment (N=5 per group). A: day 1 initial body weights, B: day 14 body weights, C: day 30 final body weights. Data are presented as mean±SE. No significant differences were observed among groups at almost all time points (p>0.05).*

Sperm parameters

Sperm parameters were affected by the treatments (Figure 4). The mean sperm count in the control group was 65.00±1.58 million/mL. Treatment with clomiphene resulted in the highest sperm count (75.80±1.92 million/mL, p<0.05), showing a significant increase of approximately 16.6% compared to the control. The TK2 group also showed a significant increase in sperm count (73.00±1.58 million/mL, p<0.05), while the TK1 group demonstrated a moderate but still significant elevation (70.00±1.58 million/mL, p<0.05) compared to the control. Interestingly, the clomiphene group showed significantly higher sperm counts compared to TK1 (p<0.05) but not compared to TK2 (p=0.08).

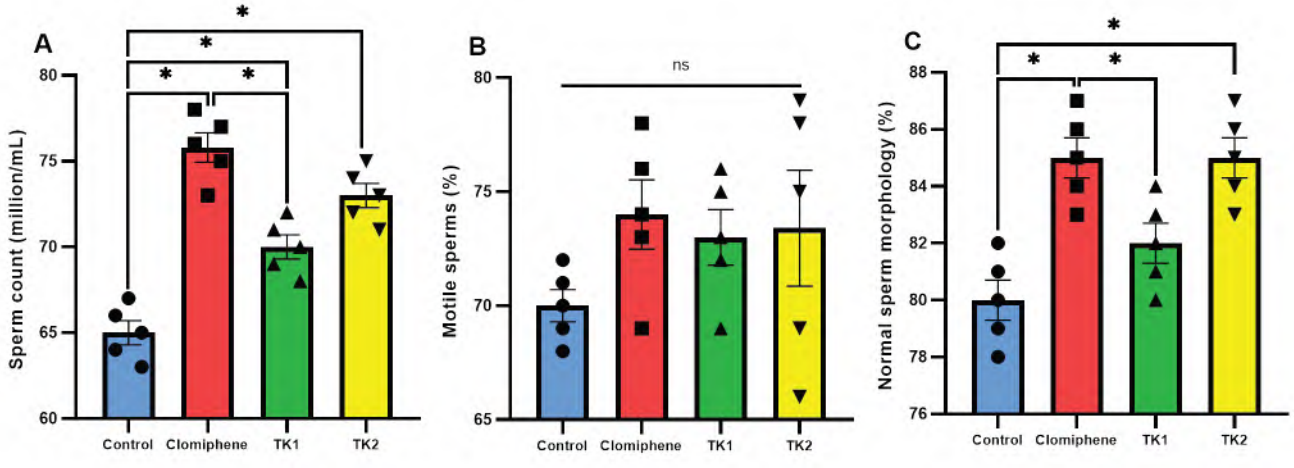


Figure 4. *Sperm parameters in rats after treatment (N=5 per group). A: sperm count (million/mL), B: percentage of motile sperm, C: percentage of sperm with normal morphology. Data are presented as mean±SE. *p<0.05.*

Testosterone levels

Testosterone levels were significantly altered by the treatments (Figure 5). The control group had a mean testosterone level of 450.0±15.81 ng/dL. Both the clomiphene and TK2 groups showed the highest increases in testosterone levels (520.0±15.81 ng/dL for both groups), representing a significant elevation of approximately 15.6% compared to the control. The TK1 group demonstrated a moderate but significant increase (490.0±15.81 ng/dL), showing an 8.9% elevation compared to control levels. The range of testosterone values across groups was consistent, with minimum values of 430.0 ng/dL in the control group and maximum values reaching 540.0 ng/dL in both the clomiphene and TK2 groups.

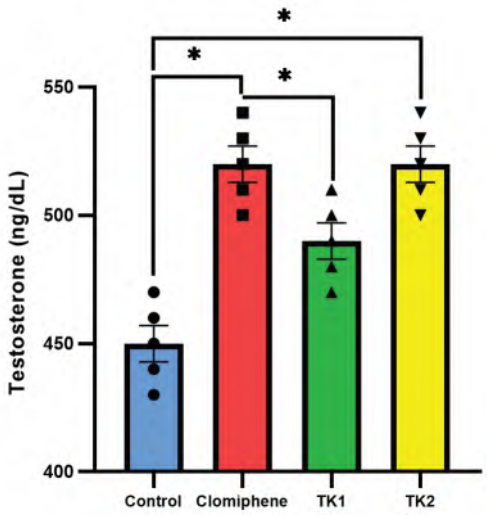


Figure 5. *Serum testosterone levels in rats after treatment (N=5 per group). Data are presented as mean SE. *p<0.05.*

Biochemical parameters

The oxidative stress markers, MDA and GSH, were significantly altered in both serum and liver tissue following treatment (Figure 6). In serum, MDA levels were markedly increased in all treatment groups compared to the control (0.300±0.011 µmol/L), with clomiphene (0.440±0.024 µmol/L, $p<0.05$), TK1 (0.390±0.011 µmol/L, $p<0.05$), and TK2 (0.490±0.011 µmol/L, $p<0.05$) showing progressively higher levels. Conversely, serum GSH levels were significantly decreased in TK1 (2.540±0.093 µmol/L, $p<0.05$) and TK2 (2.060±0.087 µmol/L, $p<0.05$) groups compared to the control (3.240±0.081 µmol/L), whereas the clomiphene group (3.100±0.045 µmol/L) did not differ significantly ($p>0.05$). Furthermore, liver tissue analysis revealed a similar finding. Hepatic MDA levels were significantly elevated in TK1 (1.150±0.092 nmol/mg protein, $p<0.05$) and TK2 (1.474±0.036 nmol/mg protein, $p<0.05$) groups compared to the control (0.796±0.083 nmol/mg protein), while the clomiphene group (0.904±0.035 nmol/mg protein) showed no significant difference ($p>0.05$). In addition, liver GSH levels were significantly reduced in both TK1 (5.380±0.080 µmol/gm tissue, $p<0.05$) and TK2 (4.440±0.108 µmol/gm tissue, $p<0.05$) groups compared to the control (6.780±0.086 µmol/gm tissue), whereas the clomiphene group (6.560±0.087 µmol/gm tissue) again showed no significant change ($p>0.05$).

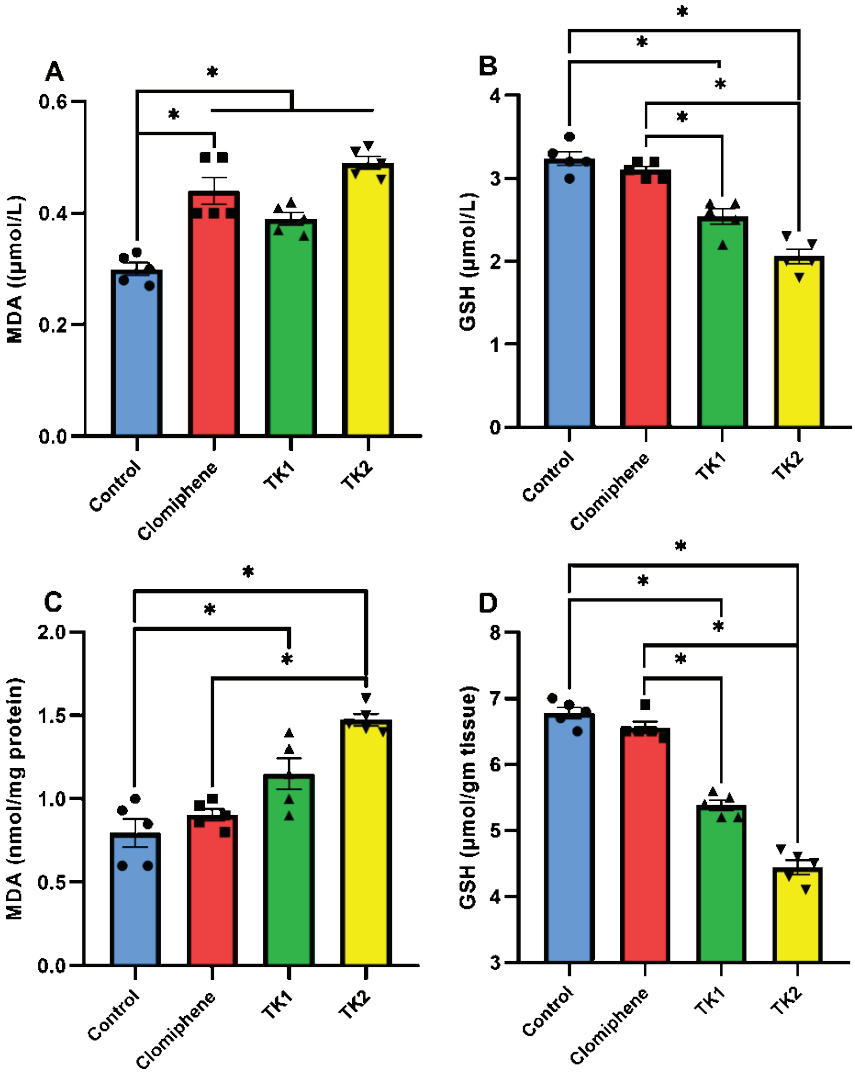


Figure 6. *Oxidative stress markers in serum and liver tissue (N=5 per group). A: serum malondialdehyde (MDA) levels, B: serum glutathione (GSH) levels, C: liver tissue MDA levels, D: liver tissue GSH levels. Data are presented as mean ± SE. *p<0.05.*

Liver function tests revealed significant alterations across treatment groups (Figure 7). Serum bilirubin levels (Figure 7A) were significantly elevated in all treatment groups compared to the control (0.300±0.011 mg/dL), with clomiphene (0.440±0.024 mg/dL, $p<0.05$), TK1 (0.390±0.011 mg/dL, $p<0.05$), and TK2 (0.490±0.011 mg/dL, $p<0.05$) showing progressively higher levels. Similarly, AST levels (Figure 7B) were significantly increased in all treatment groups compared to the control (120.0±1.1 U/L), with clomiphene (131.2±1.2 U/L, $p<0.05$), TK1 (125.0±1.1 U/L, $p<0.05$), and TK2 (135.0±1.1 U/L, $p<0.05$) demonstrating escalating values. ALT levels (Figure 7C) were also significantly elevated in clomiphene (50.0±0.7 U/L, $p<0.05$) and TK2 (53.0±1.1 U/L, $p<0.05$) groups compared to the control (45.0±1.1 U/L), whereas the TK1 group (47.0±1.1 U/L) showed no significant difference ($p>0.05$). Furthermore, ALP levels (Figure 7D) exhibited a similar trend, with significant increases observed in Clomiphene (75.0±0.7 U/L, $p<0.05$) and TK2 (78.0±1.1 U/L, $p<0.05$) groups compared to the control (70.0±1.1 U/L), while the TK1 group (72.0±1.1 U/L) showed no significant change ($p>0.05$).

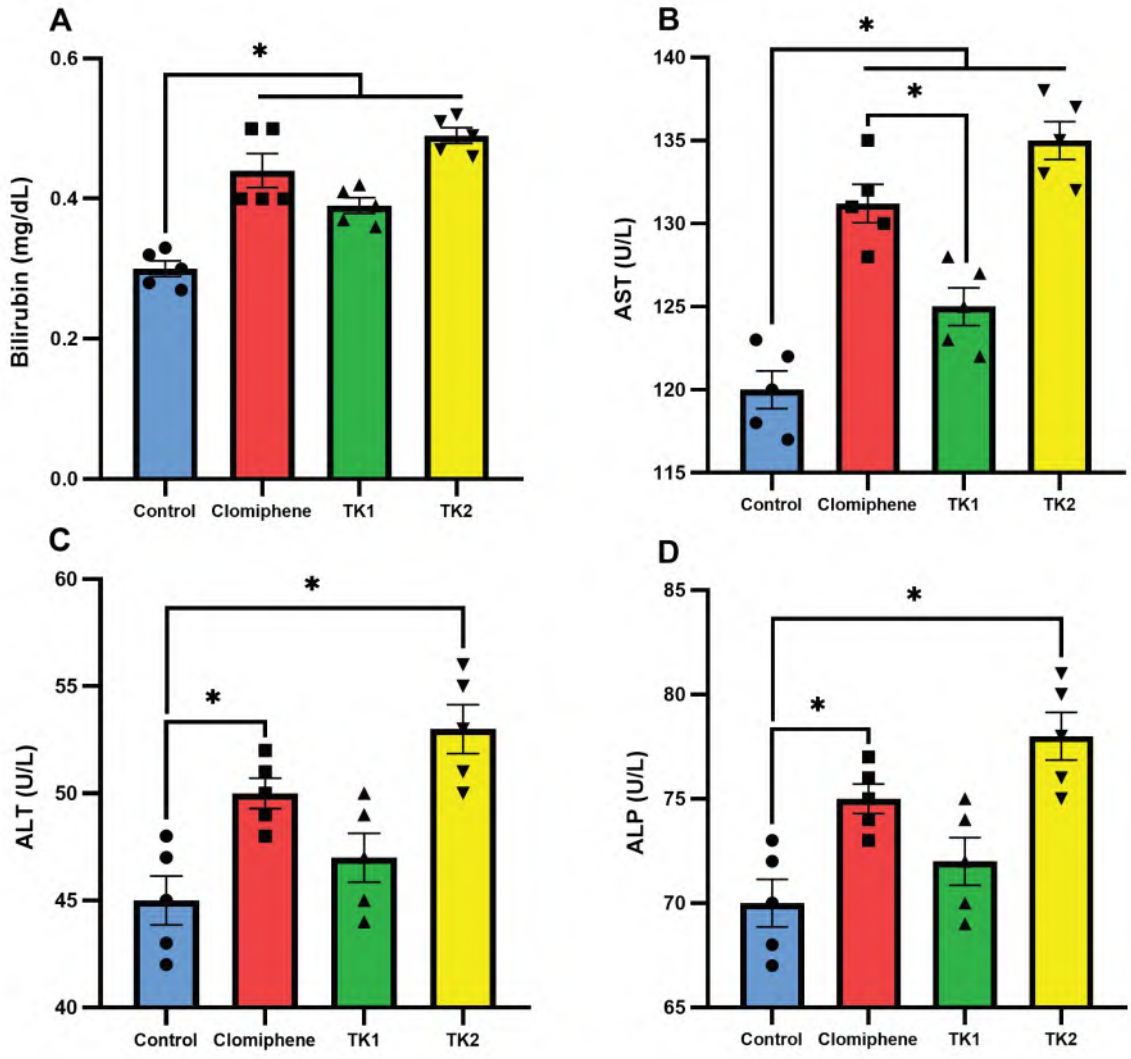


Figure 7. *Liver function tests (LFTs) in rats after treatment (N=5 per group). A: serum bilirubin levels, B: aspartate aminotransferase (AST) activity, C: alanine aminotransferase (ALT) activity, D: alkaline phosphatase (ALP) activity. Data are presented as mean±SD. *p<0.05.*

Renal function tests showed significant alterations as well in the treatment groups (Figure 8). Blood urea nitrogen (BUN) levels (Figure 8A) were significantly elevated in the clomiphene (20.0±0.3 mg/dL, $p<0.05$) and TK2 (21.0±0.4 mg/dL, $p<0.05$) groups compared to the control (18.0±0.4 mg/dL), while the TK1 group (19.0±0.4 mg/dL) showed no significant difference ($p>0.05$). Also, serum creatinine levels (Figure 8B) were significantly increased in the clomiphene (0.900±0.032 mg/dL, $p<0.05$) and TK2 (1.000±0.011 mg/dL, $p<0.05$) groups compared to the control (0.800±0.011 mg/dL), whereas the TK1 group (0.800±0.011 mg/dL) remained unchanged ($p>0.05$). Furthermore, uric acid levels (Figure 8C) exhibited a comparable result, with significant elevations observed in the clomiphene (5.300±0.071 mg/dL, $p<0.05$) and TK2 (5.360±0.068 mg/dL, $p<0.05$) groups compared to the control (5.000±0.071 mg/dL), while the TK1 group (5.100±0.071 mg/dL) showed no significant change ($p>0.05$).

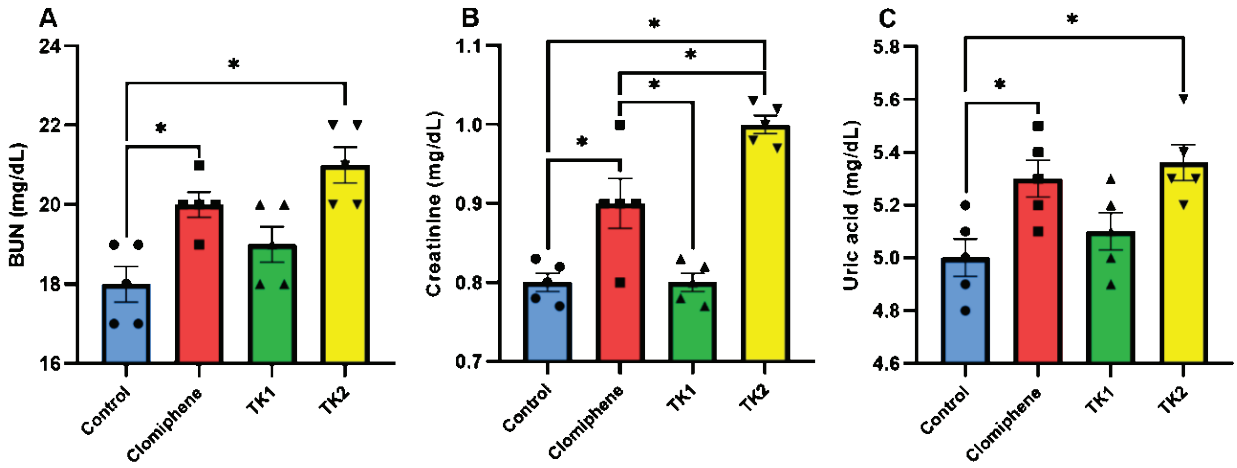


Figure 8. *Renal function tests (RFTs) in rats after treatment (N=5 per group). A: blood urea nitrogen (BUN) levels, B: serum creatinine levels, C: serum uric acid levels. Data are presented as mean±SD. *p<0.05.*

Histological changes

Histopathological examination revealed various changes in the reproductive organs of treatment groups (Figure 9). The testis showed normal size and arrangement of seminiferous tubules in all groups, comparable to the untreated control (Figure 9D). However, mild thickening of testicular interstitial connective tissue was observed in treated groups, although interstitial cells appeared normal. The germinal epithelium showed mild vacuolar degeneration of spermatogenic cells without necrosis in the clomiphene (Figure 9A), TK1 (Figure 9B), and TK2 (Figure 9C) groups. The seminal vesicles maintained normal histological features across all groups, displaying normal glandular fibro-muscular walls, mucosal folds, and glandular crypts with secretion, as exemplified in the normal (Figure 9E) and TK2 (Figure 9F) groups. In contrast, the bulbourethral gland exhibited notable changes in the TK2 group, characterized by moderate glandular atrophy associated with marked atrophy of alveolar cells and significant duct dilatation (Figure 9G, 9H). The control group's bulbourethral gland maintained normal histology (Figure 9I). Importantly, the other treatment groups showed no significant histological changes in the bulbourethral gland compared to the control.

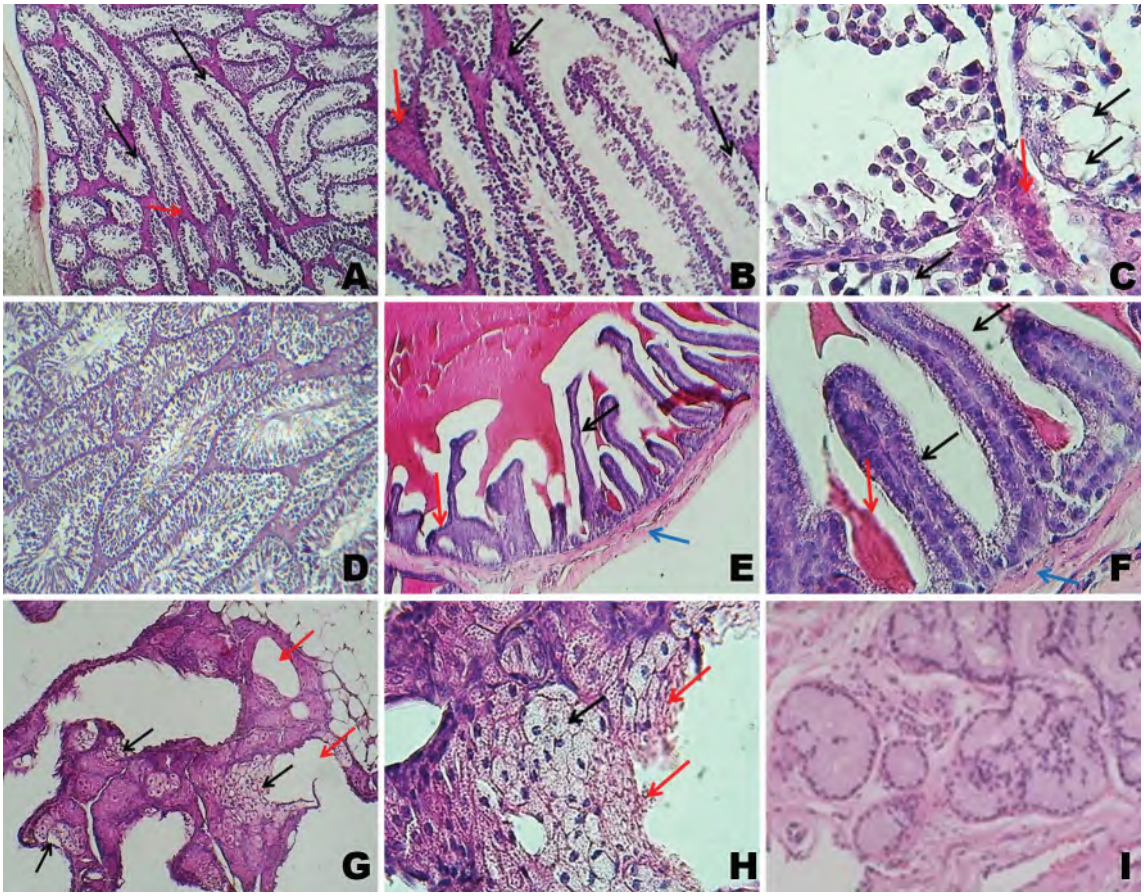


Figure 9. *Histopathological changes in reproductive organs following treatment. (A-D) Testis sections; A: clomiphene, showing normal diameter of seminiferous tubules with mild thickening of testicular interstitial connective tissue (red arrow) and mild vacuolar degeneration of germinal epithelium (black arrow), (H&E stain, 100X), B: TK1 treatment, similar findings as in A (H&E stain, 400X), C: TK2 treatment, showing mild thickening of testicular interstitial connective tissue (red arrow) and mild vacuolar degeneration of germinal epithelium (black arrow), (H&E stain, 400X), D: normal control, showing normal testicular interstitial connective tissue, (H&E stain, 400X). (E-F) Seminal vesicle sections; E: TK2 treatment, showing normal glandular fibro-muscular wall (blue arrow), mucosal folds (red arrow) and glandular crypts (black arrows), (H&E stain, 40X), F: TK2 treatment, showing normal glandular fibro-muscular wall (blue arrow), mucosal folds with normal lining epithelium (black arrow) and secretion (red arrow), (H&E stain, 400X). (G-I) Bulbourethral gland sections; G: TK2 treatment, showing moderate glandular atrophy associated with marked atrophy of alveolar cells (black arrows) and marked dilatation of ducts (red arrows), (H&E stain, 100X), H: TK2 treatment, showing moderate glandular atrophy associated with marked atrophy of alveolar cells (red arrows) and normal alveolar cells (black arrows), (H&E stain, 400X), I: normal control, showing normal bulbourethral gland structure, (H&E stain, 400X).*

The kidneys sections revealed both normal structures and some pathological changes among different regions (Figure 10). The renal medulla exhibited largely normal features, including a normal renal pelvis and intact renal tubules of the duct of Bellini among TK2 group (Figure 10A). Higher magnification of the renal medulla confirmed the presence of normal thick and thin segments of the loop of Henle and collecting tubules (Figure 10B), which appeared consistent with the control group (Figure 10C). In contrast, the renal cortex displayed some alterations. Severe subcapsular hemorrhage was observed, accompanied by congestion of the glomerular tuft (Figure 10D). Upon closer examination, the renal cortex exhibited mild intertubular vascular congestion and congestion of the glomerular tuft, although the renal tubules maintained their normal appearance (Figure 10E). These cortical changes were evident when compared to the control group (Figure 10F).

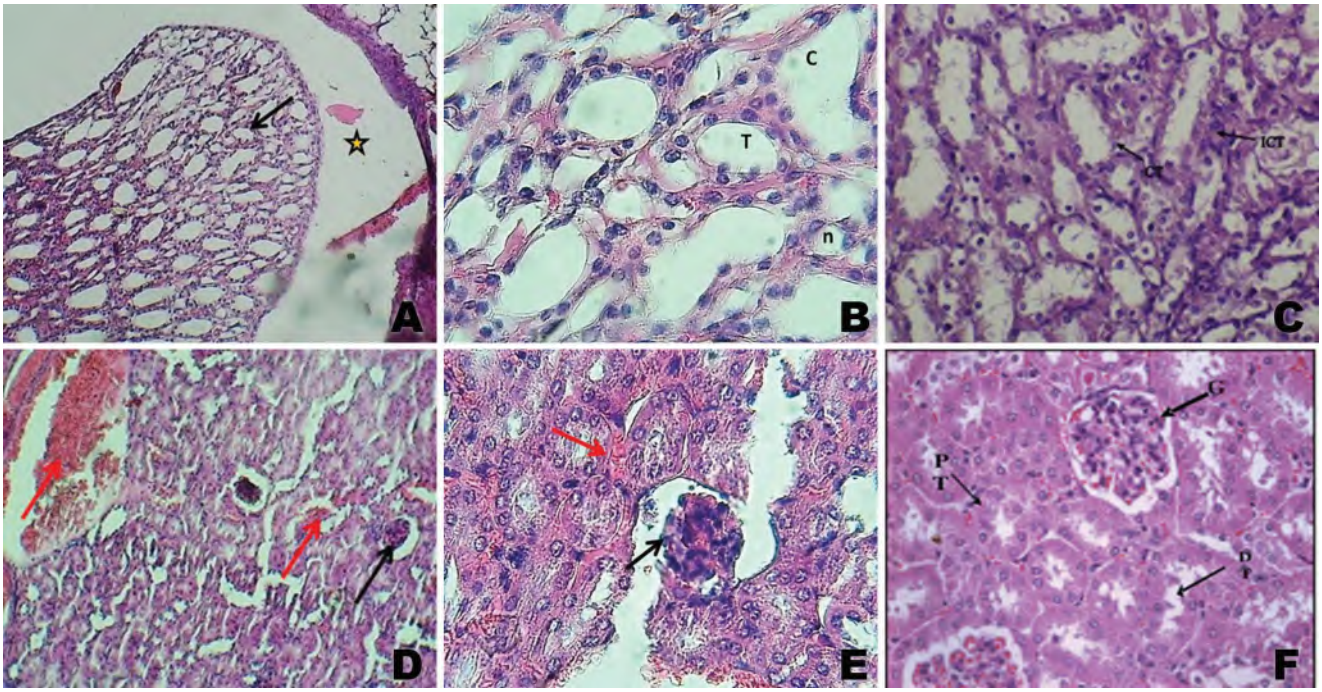


***Figure 10**. Histopathological changes in kidney following treatment. Renal medulla; A: normal renal pelvis (asterisk) and tubules of the duct of Bellini (arrow) in TK2 group, (H&E stain, 100X), B: higher magnification showing normal thick segment (T), thin segment (n) of the loop of Henle and collecting tubules (C) of TK2 group, (H&E stain, 400X), C: control group medulla. Renal cortex; D: severe subcapsular hemorrhage (red arrows) with glomerular tuft congestion (black arrow) of TK2 group, (H&E stain, 100X), E: higher magnification showing mild intertubular vascular congestion (red arrow) and glomerular tuft congestion (black arrow) with normal renal tubules of TK2 group, (H&E stain, 400X), F: control group cortex showing normal proximal tubules (P), distal tubules (T) and normal glomeruli (black arrows), (H&E stain, 100X).*

The liver and heart also revealed various pathological changes (Figure 11). The liver exhibited a spectrum of alterations, ranging from mild to moderate. Hepatic cords were generally normally arranged but showed mild congestion with dilation of the central vein and sinusoidal congestion after TK2 group (Figure 11A). Higher magnification revealed mild steatosis of hepatocytes, occasional necrotic hepatocytes, and sinusoidal dilation with congestion (Figure 11B). More severe changes were also observed, including moderate congestion with central vein dilation, focal necrosis accompanied by mononuclear leukocyte aggregation, marked dilation of portal triad blood vessels with perivascular cuffing, and cellular swelling of hepatocytes in TK2 group (Figures 11C and 11D). These pathological features contrasted sharply with the normal liver architecture observed in the control group (Figure 11E). The heart also displayed several alterations, most marked intrachamber thrombosis (Figure 11F). Also, the myocardium showed moderate vascular congestion, although the myofibers themselves appeared normal (Figures 11G, 11H, and 11I). For low dose TK, no histopathological changes were observed.

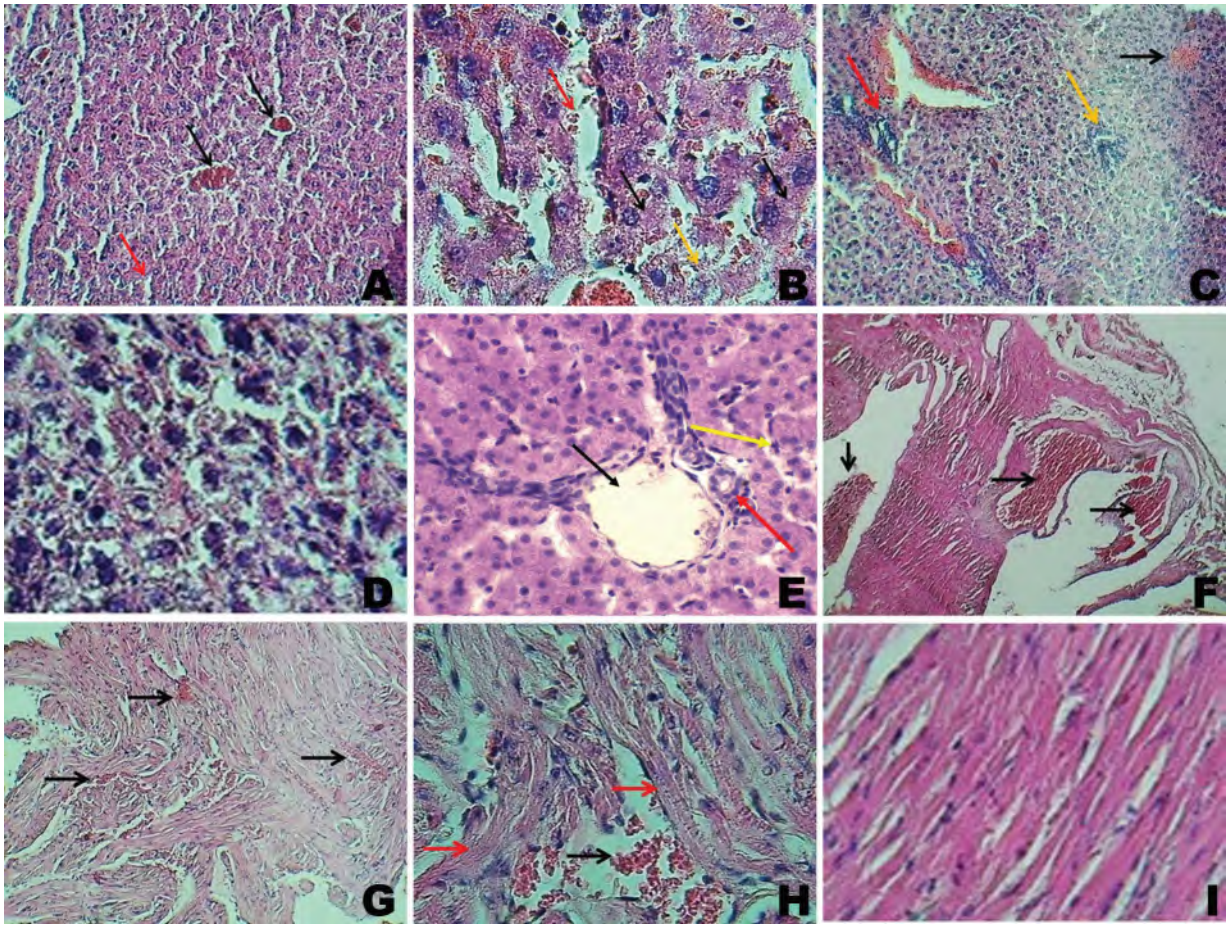


Figure 11: *Histopathological changes in liver and heart following treatment. Liver; A: mild congestion with central vein dilation (black arrows), sinusoidal dilation and congestion (red arrow) in TK2 group, (H&E, 100X), B: mild hepatocyte steatosis (black arrows), necrosis (yellow arrow), sinusoidal dilation and congestion (red arrow) in TK2 group, (H&E, 400X), C: central vein dilation (black arrows), focal necrosis with MNC aggregation (yellow arrow), portal triad vessel dilation with perivascular cuffing (red arrow) in TK2 group, (H&E, 400X), D: hepatocyte cellular swelling in TK2 group, (H&E, 400X), E: control group showing normal liver architecture: central vein (black arrow), reticular fiber (red arrow), hepatocyte (yellow arrow), portal vein, bile duct, hepatic artery, nucleus, Kupffer cells, sinusoids, binucleated hepatocytes, (H&E, 400X). Heart; F: marked cardiac thrombosis (black arrows) in TK2 group, (H&E, 400X), G: moderate vascular congestion (black arrows) in TK2 group, (H&E, 100X), H: vascular congestion (black arrows) with normal myofibers (red arrows) in TK2 group, (H&E, 400X), I: normal cardiac tissue structure, (H&E, 400X).*

Discussion

This study provides evidence for the potential hepatotoxic effects of the "Tiger King" herbal supplement, which was found to contain undeclared sildenafil. Our findings raise significant concerns about the safety of such adulterated "natural" products and draw attention to the need for stricter regulation and monitoring of dietary supplements marketed for sexual enhancement. The chemical analysis of the "Tiger King" tablets revealed the presence of sildenafil as the primary active ingredient, with no detectable amounts of the herbal components listed on the product label. This finding is consistent with previous reports of adulterated herbal supplements for ED.[5-7,9] The absence of the claimed natural ingredients not only constitutes fraudulent marketing but also poses serious health risks to consumers who may unknowingly ingest pharmaceutical-grade sildenafil without proper medical supervision.

Our *in vitro* studies demonstrated that the "Tiger King" extract influenced some antioxidant properties, as evidenced by the DPPH radical scavenging and FRAP assays. However, these antioxidant effects were significantly lower than those of standard antioxidants like ascorbic acid and Trolox. This suggests that any potential benefits from the antioxidant activity of the supplement are likely reduced by the risks associated with its undeclared pharmaceutical content.

On the other hand, the *in vivo* experiments in rat model revealed several concerning effects of "Tiger King" administration. While we observed improvements in some sperm parameters and increases in serum testosterone levels, these apparent benefits were accompanied by significant adverse effects on multiple organ systems. The histopathological changes observed in the liver, kidneys, and heart are particularly alarming and consistent with potential sildenafil-induced toxicity.

The liver histology showed evidence of steatosis, focal necrosis, and inflammatory cell infiltration, which are indicative of drug-induced liver injury. These findings align with previous case reports of sildenafil-associated hepatotoxicity in humans.[10-12] The mechanism of sildenafil-induced liver injury remains unclear, but our results support the hypothesis that it may involve oxidative stress, as evidenced by the increased levels of MDA and decreased GSH in liver tissue.

While our chemical analysis identified sildenafil as the primary active ingredient without detecting the claimed herbal components, we acknowledge that tablet excipients or trace compounds could potentially contribute to the observed hepatotoxicity. Future studies using more sensitive analytical methods and isolated sildenafil administration would help definitively establish causation.

Renal histopathology revealed glomerular congestion and tubular changes which suggests potential nephrotoxicity. While sildenafil is not typically associated with kidney injury, our findings indicate that the adulterated supplement may have renal effects, possibly due to the presence of undeclared ingredients or contaminants. This underlines the importance of comprehensive safety evaluations for such products. The cardiac histology showed evidence of intrachamber thrombosis and vascular congestion. While sildenafil is generally considered safe for cardiac patients when used as prescribed, our findings raise concerns about potential cardiovascular risks associated with uncontrolled use of adulterated supplements. This is particularly relevant given that many consumers of ED supplements may have underlying cardiovascular conditions.

The reproductive system effects observed in our study were mixed. While we noted improvements in sperm parameters and testosterone levels, consistent with the known effects of sildenafil, we also observed concerning histological changes in the testis and accessory glands. The moderate glandular atrophy and marked duct dilation in the bulbourethral gland are novel findings that require further investigation. These results, together, suggest that long-term use of such supplements may have complex and potentially detrimental effects on male reproductive health. Our findings of improved sperm parameters and increased testosterone levels need mechanistic discussion. While sildenafil is primarily known for PDE5 inhibition in erectile tissue, numerous studies have demonstrated broader reproductive effects. The observed increase in testosterone levels likely results from sildenafil's enhancement of testicular blood flow and nitric oxide signaling, which stimulates Leydig cell function.[25-27] Additionally, PDE-5 inhibition has been shown to improve spermatogenesis through increased cGMP levels in seminiferous tubules.[28] These mechanisms explain how a PDE5 inhibitor like sildenafil can affect both testosterone production and sperm parameters, beyond its classical vasodilatory role in erectile function.

The common availability of adulterated "herbal" supplements like "Tiger King" poses significant risks to consumers who may believe they are taking a natural and safe product. The undeclared presence of sildenafil, often in varying and potentially dangerous doses, can lead to adverse drug interactions, especially in patients with cardiovascular conditions or those taking nitrates.[2,9] The disagreement between the product's labeling and its actual contents draw attention to the urgent need for more stringent quality control and regulatory oversight in the dietary supplement industry. Current regulations in many countries, including the United States, do not require pre-market approval for dietary supplements, making it difficult to prevent adulterated products from reaching consumers.[1,3,4]

In addition to the current study's strengths of combining chemical analysis, *in vitro* assays, and *in vivo* experiments, it also has a limitation as it focused on relatively short-term effects, and longer-term studies are needed to assess the chronic impacts of such supplements.

Conclusion

In conclusion, our study provides evidence for the potential hepatotoxicity and broader systemic effects of the adulterated "Tiger King" herbal supplement in Iraqi market. These findings draw attention to the importance of consumer awareness, healthcare provider vigilance,

and regulatory action to address the risks associated with fraudulent and potentially dangerous sexual enhancement products.

Conflict of interest

The authors declare that they have no known competing interests.

Funding

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

Acknowledgements

The authors would like to thank the staff at the High Institute of Infertility Diagnosis and Assisted Reproductive Technologies, Al-Nahrain University, for their technical support throughout this study.

References

[1] Petre GC, Francini-Pesenti F, Vitagliano A, Grande G, Ferlin A, Garolla A. Dietary supplements for erectile dysfunction: Analysis of marketed products, systematic review, meta-analysis and rational use. Nutrients. 2023; 15(17): 3677. doi: 10.3390/nu15173677

[2] Nissan R, Poperno A, Y. Stein G, *et al*. A case of hepatotoxicity induced by adulterated "Tiger King", a Chinese herbal medicine containing sildenafil. Curr Drug Saf. 2016; 11(2): 184-8. doi: 10.2174/1574886311207040257

[3] FDA. Public Notification: Tiger King contains hidden drug ingredient.; 2014. https://www.fda.gov/drugs/medication-health-fraud/public-notification-tiger-king-contains-hidden-drug-ingredient

[4] Australian Government Department of Health. Tiger King Tablets.; 2015. http://www.tga.gov.au/alert/tiger-king-tablets

[5] Reeuwijk NM, Venhuis BJ, de Kaste D, Hoogenboom LAP, Rietjens IMCM, Martena MJ. Sildenafil and analogous phosphodiesterase type 5 (PDE-5) inhibitors in herbal food supplements sampled on the Dutch market. Food Addit Contam Part A. 2013; 30(12): 2027-34. doi: 10.1080/19440049.2013.848294

[6] Low MY, Zeng Y, Li L, *et al*. Safety and quality assessment of 175 illegal sexual enhancement products seized in red-light districts in Singapore. Drug Saf. 2009; 32(12): 1141-6. doi: 10.2165/11316690-000000000-00000

[7] Venhuis BJ, Zwaagstra ME, Keizers PHJ, de Kaste D. Dose-to-dose variations with single packages of counterfeit medicines and adulterated dietary supplements as a potential source of false negatives and inaccurate health risk assessments. J Pharm Biomed Anal. 2014; 89: 158-65. doi: 10.1016/j.jpba.2013.10.038

[8] Qahtan Mohammed B, Ali Hussaini H, Adnan abdulhameed W. The effect of aspirin and sildenafil on endometrial thickness, oocyte characteristic, embryo quality and pregnancy test in iraqi infertile women undergoing intracytoplasmic sperm injection. IraQi J Embryos Infertil Res. 2022; 12(2): 40-61. doi: 10.28969/IJEIR.v12.i2.r4.22

[9] State of Israel Ministry of Health. Counterfeit Medicines.; 2023. https://www.health.gov.il/English/Topics/PharmAndCosmetics/pharm_crime/Pages/default.aspx

[10] Wolfhagen FHJ, Vermeulen HG, de Man RA, Lesterhuis W. Initially obscure hepatotoxicity attributed to sildenafil. Eur J Gastroenterol Hepatol. 2008; 20(7): 710-2. doi: 10.1097/MEG.0b013e3282f2bbb5

[11] Daghfous R, El Aidli S, Zaiem A, Loueslati MH, Belkahia C. Sildenafil-associated hepatotoxicity. Am J Gastroenterol. 2005; 100(8): 1895-6. doi: 10.1111/j.1572-0241.2005.41983_6.x

[12] Enomoto M, Sakaguchi H, Ominami M, *et al*. Sildenafil-induced severe cholestatic hepatotoxicity. Am J Gastroenterol. 2009; 104(1): 254-5. doi: 10.1038/ajg.2008.18

[13] Graziano S, Montana A, Zaami S, *et al*. Sildenafil-associated hepatoxicity: a review of the literature. Eur Rev Med Pharmacol Sci. 2017; 21(Suppl1): 17-22. http://www.ncbi.nlm.nih.gov/pubmed/28379598

[14] Al-Maliki RS. COVID-19 vaccination doesn't influence sperm motility, concentration, and morphology Rehab. J Assoc Med Sci. 2025; 58(1): 185-91. doi: 10.12982/JAMS.2025.02

[15] Patel DN, Li L, Kee CL, Ge X, Low MY, Koh HL. Screening of synthetic PDE-5 inhibitors and their analogues as adulterants: Analytical techniques and challenges. J Pharm Biomed Anal. 2014; 87: 176-90. doi: 10.1016/j.jpba.2013.04.037

[16] Venhuis BJ, de Kaste D. Towards a decade of detecting new analogues of sildenafil, tadalafil and vardenafil in food supplements: A history, analytical aspects and health risks. J Pharm Biomed Anal. 2012; 69: 196-208. doi: 10.1016/j.jpba.2012.02.014

[17] Abourashed E, Abdel-Kader M, Habib AA. HPTLC determination of sildenafil in pharmaceutical products and aphrodisiac herbal preparations. J Planar Chromatogr – Mod TLC. 2005; 18(105): 372-6. doi: 10.1556/JPC.18.2005.5.7

[18] Wagner H, Bladt S. Plant Drug Analysis: A Thin Layer Chromatography Atlas. 2nd Ed. Springer-Verlag Berlin Heidelberg; 1996.

[19] V. Le A, E. Parks S, H. Nguyen M, D. Roach P. Improving the vanillin-sulphuric acid method for quantifying total saponins. Technologies. 2018; 6(3): 84. doi: 10.3390/technologies6030084

[20] Sherma J, Fried B, eds. Handbook of Thin-Layer Chromatography. 3rd Ed. Marcel Dekker; 1991.

[21] Rebane R, Leito I, Yurchenko S, Herodes K. A review of analytical techniques for determination of Sudan I–IV dyes in food matrixes. J Chromatogr A. 2010; 1217(17): 2747-57. doi: 10.1016/j.chroma.2010.02.038

[22] Savaliya AA, Shah RP, Prasad B, Singh S. Screening of Indian aphrodisiac ayurvedic/herbal healthcare products for adulteration with sildenafil, tadalafil

and/or vardenafil using LC/PDA and extracted ion LC–MS/TOF. J Pharm Biomed Anal. 2010; 52(3): 406-9. doi: 10.1016/j.jpba.2009.05.021

[23] Brand-Williams W, Cuvelier ME, Berset C. Use of a free radical method to evaluate antioxidant activity. LWT - Food Sci Technol. 1995; 28(1) :25-30. doi: 10.1016/S0023-6438(95)80008-5

[24] Benzie IFF, Strain JJ. The ferric reducing ability of plasma (FRAP) as a measure of "antioxidant power": the FRAP assay. Anal Biochem. 1996; 239(1): 70-6. doi: 10.1006/abio.1996.0292

[25] Janjic MM, Stojkov NJ, Bjelic MM, Mihajlovic AI, Andric SA, Kostic TS. Transient Rise of Serum Testosterone Level After Single Sildenafil Treatment of Adult Male Rats. J Sex Med. 2012; 9(10): 2534-43. doi: 10.1111/j.1743-6109.2012.02674.x

[26] Saraiva KLA, Silva AKSE, Wanderley MI, De Araújo AA, De Souza JRB, Peixoto CA. Chronic treatment with sildenafil stimulates Leydig cell and testosterone secretion. Int J Exp Pathol. 2009; 90(4): 454-62. doi: 10.1111/j.1365-2613.2009.00660.x

[27] Spitzer M, Bhasin S, Travison TG, Davda MN, Stroh H, Basaria S. Sildenafil increases serum testosterone levels by a direct action on the testes. Andrology. 2013; 1(6): 913-8. doi: 10.1111/j.2047-2927.2013.00131.x

[28] Andric SA, Janjic MM, Stojkov NJ, Kostic TS. Sildenafil treatment *in vivo* stimulates Leydig cell steroidogenesis via the cAMP/cGMP signaling pathway. Am J Physiol Metab. 2010; 299(4): E544-50. doi: 10.1152/ajpendo.00337.2010

Journal of Associated Medical Sciences 2025; 58(3): 25-32

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

Sensory processing patterns link alcohol consumption in undergraduate students

Natcha Chaikhamla[1], Nalinee Yingchankul[2], Tiam Srikhamjak[1], Pornpen Sirisatayawong[1*]

[1]Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 19 January 2025
Accepted as revised 2 May 2025
Available online 8 May 2025

Keywords:
Sensory processing patterns, Alcohol consumption, Undergraduate students.

ABSTRACT

Background: Alcohol consumption among undergraduate students is a public health concern. Sensory processing refers to how individuals receive, organize, and respond to sensory information from their environment, influencing their emotions and behavior. However, no research has yet explored the relationship between sensory processing patterns in specific modalities and alcohol consumption among undergraduate students.

Objective: The purpose of this study was to examine the relationship between sensory processing patterns and alcohol consumption among undergraduate students.

Materials and methods: Participants included 430 students. Participants completed a demographic questionnaire, the Functional-Belief-Based Alcohol Use Questionnaire and the Thai Sensory Patterns Assessment. Descriptive statistics, Spearman's correlation analysis, Mann-Whitney U test, Kruskal-Wallis H test, and chi-squared test were used to analyze the data.

Results: The findings indicated that alcohol consumption was positively correlated with sensory arousal in the auditory and proprioceptive modalities (r=0.176 and 0.186, respectively; $p<0.01$), as well as with sensory preferences across all modalities (rs ranged from 0.116 to 0.204; $p<0.01$; $p<0.05$).

Conclusion: This research found that sensory processing patterns may be associated with alcohol consumption. Further research and public health efforts could contribute to developing strategies to assist students in managing sensory cues and preventing alcohol abuse.

Introduction

Alcohol is the most widely used substance worldwide, including in Thailand. Globally, more than half (57%) of people consumed alcohol[1], while nearly 30% of Thais drinked alcohol.[2] In addition, it was a leading contributor to the burden of alcohol-attributable disability-adjusted life years (DALYs), with approximately 115.9 million DALYs in 2019.[1] In Thailand, alcohol consumption was the second highest risk factor for DALYs lost in 2019.[3] Alcohol use among undergraduate students was a significant concern, as they were at an age when many new drinkers emerge, and heavy consumption was common.[2] In Thailand, the prevalence of hazardous or harmful drinking among undergraduate students has increased from 10.8% to 15.6% between 2015 and 2019.[4,5]

Undergraduate students are in a transition period and are an important human resource in the future. Undergraduate students' experiences of alcohol use were linked with stressful times,[6] daily stressors,[7] influence of

∗ **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: pornpen.siri@cmu.ac.th
doi: 10.12982/JAMS.2025.069
E-ISSN: 2539-6056

friends,[8] and the lack of self-awareness of how to solve problems.[9] In addition, a study in Thailand illustrated that religion, attitudes, access to alcohol, age, academic success, and family assets were potential factors for alcohol consumption among undergraduate students.[10]

Sensation serves as a foundation for perception, learning, behavior, and emotion.[11] Sensory processing refers to an individual's ability to receive, organize, and respond to sensory information in daily activities.[12] Difficulties in these processes are associated with emotions and behaviors.[12] According to sensory restriction concepts, low or high levels of stimulation can cause difficulty processing sensory input. It increases physiological, perceptual, and affective impairments and has a maladaptive effect on behavior, including alcohol consumption.[12] Emotional and behavioral processing in the brain begins with the reception of sensory input. If the brain perceives this information as stressful or threatening, it triggers the release of glucocorticoids, leading to physiological and behavioral responses. These reactions may lead to alcohol use and addiction, both of which are associated with the brain's reward system.[13,14] Every person varies in their sensory preferences and in their style of processing sensory information, called sensory processing patterns (SPPs). Previous studies showed that the sensory processing patterns of people with substance use were significantly different from those of the general population[15], and young people with sensory sensitivity contributed to increasing stress, resulting in alcohol consumption to manage stress.[16] However, no research has examined the relationship between sensory processing patterns in specific modalities and alcohol consumption among undergraduate students.

Materials and methods

Participants

The current study included 430 students aged between 18-25 years old. The sample size was calculated using Yamane's formula.[17] The sampling method was purposive sampling followed by an inclusion criteria. The inclusion criteria were students at Chiang Mai University, Thailand, who had been exposed to alcohol consumption within the past year. The study protocol was approved by the Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University (AMSEC-67EX-007).

Measurement

Demographic questionnaire

The demographic questionnaire includes questions to gather information from the research participants, such as their gender, age, overall grade point average (GPAX), field of study, monthly income, and part-time job.

The Functional-Belief-Based Alcohol Use Questionnaire (FBAQ)

The FBAQ consists of three straightforward belief questions. The three items in the FBAQ address participants about self-regulation and focusing on consumption for stress relief or socialization. Participants would rate their agreement to the given statement on a 5-point Likert scale from 1 to 5 (totally disagreed to totally agreed). The total FBAQ score is then categorized into two groups, low-risk drinking and high-risk drinking. Three functional-belief items independently predict high-risk drinking (score ≥ 6). Previous research validated the FBAQ with Chiang Mai University undergraduates, which is the same group as the participants in this study. The FBAQ showed strong discriminative performance and high sensitivity in identifying high-risk drinkers. The pooled sensitivity and specificity at the FBAQ score cutoff of 6 were 92.8% (95% CI 88.0-95.7%) and 51.6% (95% CI 41.1-62.0%).[18]

The Thai Sensory Patterns Assessment (TSPA)

This self-reported assessment consists of two modules, including sensory preferences and sensory arousals. Each module has six categories based on the types of sensory modality (visual, auditory, smell and taste, tactile, proprioceptive, and vestibular senses). To interpret the scores, the items are rated on a 5-point scale from 1 to 5 (never to always), and the scores for each sensory modality are combined separately and are reported as a percentage. Each participant's types of sensory modalities mean scores are then categorized into the corresponding three classifications of score results: Low (<25%), Moderate (25-75%), and High (>75%). The TSPA had valid content (IOC=0.60-1.00 in modules I and II) and reliable tests (ICC=0.93 and 0.77, Cronbach's alpha=0.89 and 0.62 in modules I and II, respectively). Additionally, the Cronbach's alpha of the TSPA among undergraduate students was equal to 0.86 (Cronbach's alpha=0.85 and 0.74 in modules I and II, respectively).[19]

Data collection procedures

After the research received approval from the ethics committee, students were invited through advertisement via sheet posters, online posters, and face-to-face contact. All participants who agreed to participate in the study signed informed consent and completed the assessments in a paper-based format or through an online format.

Data analysis

Data analysis was conducted using SPSS (version 27). Descriptive statistics, including percentage, mean, and standard deviation, were used to describe sensory processing patterns, alcohol consumption scores, and demographic characteristics. The Mann-Whitney U test and Kruskal-Wallis H test analyzed differences in alcohol consumption across demographics, while the chi-squared test examined correlations between demographic characteristics and alcohol consumption scores. Spearman's correlation was used to explore the relationship between sensory processing patterns and alcohol consumption, as well as between demographic characteristics and alcohol consumption scores.

Results

Characteristics of participants

Table 1 showed that most participants were female (66.0%). High-risk drinkers had a slightly higher proportion of males (35.5%) compared to low-risk drinkers (26.1%). The average age was around 20.3 years. Humanities and Social Sciences had the highest representation (41.4%). High-risk drinkers slightly skewed toward lower GPAX ranges. High-risk drinkers had a slightly higher proportion in the lowest income category (<5,000 Bht). About 15.2% of students had part-time jobs, with a similar distribution occurring between high- and low-risk groups.

Table 1. Demographic characteristics among undergraduate students with alcohol consumption.

Variables	High risk (N=361)	Low risk (N=69)	Overalls (N=430)
	N (%)	N (%)	N (%)
Gender			
Male	128 (35.5)	18 (26.1)	146 (34.0)
Female	233 (64.5)	51 (73.9)	284 (66.0)
Age (mean±SD)	20.3±11.16	20.1±1.10	20.3±1.15
Field of study			
Science and technology group	122 (33.8)	25 (36.2)	147 (34.2)
Health science group	81 (22.4)	24 (34.8)	105 (24.4)
Humanities and social sciences group	158 (43.8)	20 (29.0)	178 (41.4)
Overall grade point average (GPAX)			
<2.00	10 (2.8)	3 (4.3)	13 (3.0)
2.00-2.49	65 (18.0)	6 (8.7)	71 (16.5)
2.50-2.99	98 (27.1)	20 (29.0)	118 (27.4)
3.00-3.49	115 (31.9)	21 (30.4)	136 (31.6)
3.50-4.00	73 (20.2)	19 (27.5)	92 (21.4)
Income per month (Bht)			
<5,000	83 (23.0)	18 (26.1)	101 (23.5)
5,001-10,000	207 (57.3)	32 (46.4)	239 (55.6)
10,001-15,000	50 (13.9)	14 (20.3)	64 (14.9)
15,001-20,000	19 (5.3)	3 (4.3)	22 (5.1)
>20,001	2 (0.6)	2 (2.9)	4 (0.9)
Part-time job			
Yes	55 (15.2)	10 (14.5)	65 (15.1)
No	306 (84.8)	59 (85.5)	365 (84.9)

Sensory processing patterns

Table 2 showed sensory preferences, and the auditory sense stood out: High-risk drinkers scored a high level (82.96±12.91), while low-risk drinkers scored a moderate level (79.83±15.22). The overall group had a high level (82.46±13.34). Both groups rated all other senses as moderate, with minor differences in mean scores. Sensory arousal for students across both high- and low-risk groups showed moderate levels in all senses. High-risk drinkers had slightly higher levels in the auditory (50.91±13.60 vs. 48.21±14.43) and proprioceptive (43.34±15.20 vs. 39.48±14.16) senses compared to low-risk drinkers.

Alcohol consumption

Table 3 revealed that most students (84.0%) were classified as high-risk drinkers.

Table 2. Sensory processing patterns in specific sensory modalities of undergraduate students with alcohol consumption.

Modules	Sensory processing patterns					
	High risk (N=361)		Low risk (N=69)		Overalls (N=430)	
	Mean±SD	Result	Mean±SD	Result	Mean±SD	Result
Sensory preferences						
Auditory	82.96±12.91	**High**	79.83±15.22	Moderate	82.45±13.34	**High**
Touch	72.71±14.35	Moderate	72.54±19.15	Moderate	72.68±15.19	Moderate
Smell and taste	72.89±12.40	Moderate	70.92±12.99	Moderate	72.57±12.51	Moderate
Visual	63.74±15.77	Moderate	61.30±15.33	Moderate	63.35±15.71	Moderate
Proprioceptive	59.99±14.86	Moderate	56.74±16.79	Moderate	59.47±15.21	Moderate
Vestibular	55.84±15.46	Moderate	57.13±18.65	Moderate	56.05±15.99	Moderate
Sensory arousals						
Visual	69.42±16.42	Moderate	66.47±16.65	Moderate	68.94±16.47	Moderate
Smell and taste	64.59±19.25	Moderate	62.90±20.05	Moderate	64.32±19.36	Moderate
Auditory	50.91±13.60	Moderate	48.21±14.44	Moderate	50.48±13.76	Moderate
Touch	46.04±16.21	Moderate	46.52±18.46	Moderate	46.12±16.56	Moderate
Vestibular	45.34±20.47	Moderate	43.08±19.21	Moderate	44.98±20.27	Moderate
Proprioceptive	43.34±15.20	Moderate	39.48±14.16	Moderate	42.72±15.09	Moderate

Table 3. FBAQ scores and levels of alcohol consumption among undergraduate students (N=430).

Variables	Minimum	Maximum	Mean	SD
Alcohol use behavior				
FBAQ scores	3	15	8.76	3.12
Alcohol use levels	**N**	**%**		
Low risk drinking	69	16.0		
High risk drinking	361	84.0		

Alcohol differences among certain demographic characteristics

Table 4 showed that there was no significant difference in FBAQ scores between males and females (p>0.05). As for the field of study, a significant difference was discovered among low-risk drinkers (p=0.005), but not in high-risk or general groups. Overall, there was a significant difference (p=0.049) in GPAX and alcohol use. In terms of monthly income, there were no significant differences overall. However, variances were observed among low-risk drinkers (p=0.033).

Correlation between demographic characteristics and alcohol consumption scores

Table 5 shows that the field of study has a significant correlation with alcohol consumption in both low-risk (p=0.019) and overall groups (p=0.020). Gender, age, GPAX, monthly income, and part-time job showed no significant correlation with alcohol consumption (p>0.05).

Table 4. Alcohol difference among certain demographic characteristics.

Variables	FBAQ score					
	High risk (N=361)		Low risk (N=69)		Overalls (N=430)	
	Mean±SD	*p* value	Mean±SD	*p* value	Mean±SD	*p* value
Gender						
Male	9.77±2.52	0.511[a]	3.94±0.94	0.883[a]	9.05±3.06	0.157[a]
Female	9.62±2.47		3.98±0.91		8.69±3.14	
Field of study						
STG	9.82±2.46		3.52±0.82		8.75±3.28	
HSG	9.73±2.71	0.601[b]	4.13±0.90	**0.005*[b]**	8.45±3.38	0.401[b]
HSSG	9.54±2.40		4.35±0.81		8.96±2.81	
Overall grade point average (GPAX)						
<2.00	10.90±2.42		4.33±1.15		9.38±3.59	
2.00-2.49	10.09±2.50		3.67±1.03		9.55±3.01	
2.50-2.99	9.79±2.55	0.092[b]	4.10±0.91	0.149[b]	8.82±3.18	**0.049*[b]**
3.00-3.49	9.60±2.59		3.62±0.86		8.68±3.24	
3.50-4.00	9.11±2.12		4.26±0.81		8.11±2.75	
Income per month (Bht)						
<5,000	9.17±2.36		3.50±0.86		8.16±3.07	
5,001-10,000	9.76±2.45		4.25±0.84		9.03±2.97	
10,001-15,000	9.80±2.70	0.165[b]	4.07±0.92	**0.033*[b]**	8.55±3.40	0.137[b]
15,001-20,000	10.42±2.65		4.00±1.00		9.55±3.35	
>20,001	11.50±2.12		3.00±0.00		7.25±5.06	
Part-time job						
Yes	9.80±2.59	0.706[a]	3.80±0.92	0.520[a]	8.88±3.25	0.724[a]
No	9.65±2.47		4.00±0.90		8.74±3.10	

***Note**: STG: science and technology Group, HSG: health science group, HSSG: humanities and social sciences group, a: Mann-Whitney U test, b: Kruskal-Wallis H test; *p<0.05.*

Table 5. Correlation between demographic characteristics and alcohol consumption score.

Variables	FBAQ score		
	High risk	Low risk	Overalls
Gender	0.665[a]	0.953[a]	0.684[a]
Age	0.303[b]	0.529[b]	0.851[b]
Field of study	0.206[a]	**0.019*[a]**	**0.020*[a]**
Overall grade point average	0.772[a]	0.196[a]	0.570[a]
Income per month	0.272[a]	0.078[a]	0.075[a]
Part-time job	0.209[a]	0.805[a]	0.398[a]

***Note**: a: Chi-squared test, b: Spearman correlation test, *p<0.05.*

The relationship between sensory processing patterns and alcohol consumption

Table 6 showed that for sensory preferences, significant positive correlations were found between alcohol consumption and all senses in the high-risk and overall groups (p<0.05 to p<0.001). No significant correlations existed in the low-risk group. As for sensory arousal, auditory and proprioceptive senses showed significant positive correlations with alcohol consumption in the high-risk and overall groups (p<0.01). Touch sense showed a weak correlation in the high-risk group (p=0.032). There were no significant correlations in the low-risk group.

Table 6. Correlation between sensory modalities and alcohol consumption score in undergraduate students.

Sensory processing patterns	FBAQ score		
	High risk rs (p-value)	Low risk rs (p-value)	Overalls rs (p-value)
Sensory preferences			
visual	0.235 (0.001)**	0.083 (0.499)	0.204 (0.000)**
auditory	0.145 (0.006)**	-0.147 (0.227)	0.146 (0.002)**
smell and taste	0.211 (0.001)**	0.157 (0.628)	0.186 (0.000)**
Touch	0.179 (0.001)**	0.059 (0.204)	0.117 (0.015)*
vestibular	0.176 (0.001)**	0.106 (0.387)	0.116 (0.016)*
proprioceptive	0.172 (0.001)**	0.159 (0.191)	0.167 (0.001)**
Sensory arousals			
visual	0.014 (0.791)	0.095 (0.436)	0.059 (0.222)
auditory	0.175 (0.001)**	0.112 (0.359)	0.176 (0.000)**
smell and taste	0.076 (0.150)	0.117 (0.338)	0.084 (0.081)
Touch	0.113 (0.032)*	-0.056 (0.649)	0.078 (0.105)
vestibular	0.030 (0.571)	0.139 (0.258)	0.051 (0.292)
proprioceptive	0.182 (0.001)**	-0.134 (0.272)	0.186 (0.000)**

***Note**: p: the Spearman correlation test; *p<0.05, **p<0.01 (significant).*

Discussion

Demographic characteristics and alcohol consumption

The findings indicated that certain demographic factors, such as the field of study, academic performance, and income, may influence alcohol consumption among undergraduate students. The field of study on alcohol consumption may be more significant among students. Previous research indicated that students majoring in social sciences and business consumed more alcohol than those majoring in arts,[20] and medical students drank alcohol more frequently. These findings emphasized the role of academic stress and the cultural context within various fields in shaping students' drinking habits.[21] Further research into how the field of study influences alcohol consumption among both low- and high-risk drinkers might give a deeper understanding of this relationship.

Moreover, GPAX may be more significant for alcohol use (p=0.049). Similarly, research conducted at Chiang Mai University in 2019 revealed that increased alcohol consumption among students was significantly related to lower GPAX (p<0.001). The study also noted associations between higher alcohol intake and behaviors such as skipping classes and withdrawing from courses.[22] Income may impact those in the low-risk drinking category (p=0.033). A previous study suggested that high socioeconomic status might increase the risk of problematic alcohol use among university students.[23] Overall, these findings emphasized the importance of considering academic factors, such as field of study and GPAX, when examining alcohol consumption patterns among students.

The relationship between sensory arousal and alcohol consumption

The findings showed a positive correlation between the high-risk and overall groups' alcohol consumption and sensory arousal in auditory and proprioceptive senses. Low and high levels of sensory stimulation can create challenges in processing sensory input, leading to maladaptive behaviors, including alcohol use as a coping mechanism. Individuals with high sensory arousal may experience sensory overload, resulting in stress, prompting them to use alcohol for relaxation and self-regulation. Conversely, those with lower sensory arousal may seek out intense sensory input, including alcohol's effects, to compensate for their reduced responsiveness to stimuli.[12,24] This result is consistent with prior research demonstrating a correlation between sensory arousal and alcohol use.[15,16,25] One research study found that high sensory arousal correlated with alcohol use (r=0.137; p<0.05).[25] However, this earlier research generally studied mentally healthy adults. Moreover, the prior studies did not explicitly determine

whether specific sensory modalities were associated with alcohol consumption. This study indicated that sensory arousals were correlated with attitudes and alcohol use behavior scores, especially in the proprioceptive sense. In Dunn's Model of Sensory Processing, individuals with high sensory arousal in the proprioceptive sense are more prone to sensory overload during physical activity and physical movement. This can contribute to a preference for a sedentary lifestyle, potentially leading to alcohol use as a coping mechanism.[26,27] Therefore, it is important to carefully regulate proprioceptive stimulation.

The relationship between sensory preferences and alcohol consumption

The results demonstrated a positive correlation between the high-risk and overall groups' alcohol consumption and sensory preferences in all senses. These results also indicated that individuals who have high sensory preference can drive increased alcohol consumption. This finding aligns with the mechanisms underlying the pleasurable effects of alcohol, as alcohol consumption activates reward circuits in the brain and increases dopamine levels, which are associated with feelings of pleasure and reward.[28,29] Preferred sensory stimuli promote pleasure in mammals.[30] While most previous research has not clearly established the relationship between specific sensory modalities and alcohol consumption scores, this study provided new insight into these associations. For instance, some individuals may prefer listening to music, whereas others may favor the taste of beverages to satisfy their sensory preferences. However, this study has limitations, as it does not delve deeply into why individuals drink, the environment in which alcohol consumption occurs, or the social contexts surrounding drinking behavior. As evidenced by current literature, external factors such as stress, peer influence, social environment, and cultural norms can play a significant role in explaining alcohol consumption patterns.[31] These factors, when combined with individual characteristics such as sensory processing patterns, can contribute to the development of drinking behaviors and influence how individuals use alcohol as a coping mechanism.

Limitation

Statistically significant findings were observed at low levels. Future research should increase sample size, minimize confounding variables, and use random sampling to improve generalizability. Longitudinal studies are needed to explore the temporal relationship between sensory processing and alcohol use. Methodological limitations, such as reliance on self-reports, should be addressed by incorporating peer reports and objective physiological measures. However, self-reports are useful for large-scale studies and capture the environmental and emotional context of alcohol use.

Conclusion

The findings indicated that sensory processing patterns, particularly sensory arousal in auditory and proprioceptive modalities, as well as sensory preference across all senses, were associated with alcohol consumption. This study identified only relationships. Future research incorporating perspectives of occupational therapists and assessment of public health initiatives could provide a deeper understanding of how to deal with drinking behaviors, further guiding the development of ways to assist students in regulating external and internal cues more successfully to prevent alcohol abuse.

Conflict of interest

There are no conflicts of interest to declare in this study.

Funding

This study received a thesis support scholarship for graduate students from the Faculty of Associated Medical Sciences, Chiang Mai University.

Acknowledgements

The authors were supported by the Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University.

References

[1] World Health Organization. Global status report on alcohol and health and treatment of substance use disorders. Geneva: World Health Organization; 2024. Available from: https://www.who.int/publications/i/item/9789240096745

[2] National Statistical office. Health behavior of population survey 2021. Bangkok: National Statistical Office; 2021. Available from: https://www.nso.go.th/nsoweb/nso/survey_detail/w6?set_lang=en

[3] Burden of Disease Thailand Study Group. Burden of disease attributable to risk factors in Thailand 2019. Nonthaburi: International Health Policy Planning; 2023. Available from: https://www.trc.or.th/en/burden-of-disease-attributable-to-risk-factors-in-thailand-2019/

[4] Pengpid S, Peltzer K. Prevalence of overweight and underweight and its associated factors among male and female university students in Thailand. Homo. 2015; 66(2): 176-86. doi: 10.1016/j.jchb.2014.11.002

[5] Booranasuksakul U, Singhato A, Rueangsri N, Prasertsri P. Association between alcohol consumption and body mass index in university students. Asian Pac Isl Nurs J. 2019; 4(1): 57. doi: 10.31372/20190401.1035

[6] Grab J, Green M, Norris J, Pilchik K, Fisher GS. Exploring the impact of OT student stress: Professor and student perspectives. J Occup Ther Educ. 2021; 5(1): Art. 3. doi: 10.26681/jote.2021.050103

[7] Thompson SM, Simmons AN, McMurray MS. The effects of multiple early life stressors on adolescent alcohol consumption. Behav Brain Res. 2020; 380: 112449. doi: 10.1016/j.bbr.2019.112449

[8] Promyod P. The correlation between self-confidence in refusing an invitation to drink alcohol and alcohol drinking behavior of undergraduate students of Rajabhat Maha Sarakham University for academic year 2020. MKHJ 2021; 18(2): 81-7. Available from: https://he02.tci-thaijo.org/index.php/MKHJ/article/download/251057/172216 [in Thai]

[9] Pourmohammadi B, Jalilvand MA. Prevalence of alcohol consumption and related factors among students of higher education centers in one of the northeastern cities of Iran. AIMS Public Health. 2019; 6(4): 523-33. doi: 10.3934/publichealth.2019.4.523

[10] Phoosuwan N. Prevalence and factors related to alcohol consumption among undergraduate students in a northeast Thailand province. PHJBUU. 2019; 14: 1-15. Available from: https://journal.lib.buu.ac.th/index.php/health/article/view/5785 [in Thai].

[11] Pagliano P. The multisensory handbook: a guide for children and adults with sensory learning disabilities. New York, NY: Taylor & Francis Group; 2012.

[12] Dunn W. Living sensationally: understanding your senses. London: Jessica Kingsley; 2008.

[13] Gutman S. Why addiction has a chronic, relapsing course. the neurobiology of addiction. Occup Ther Ment Health. 2006; 22(2): 1-29. doi: 10.1300/J004v22n02_01

[14] Becker HC. Influence of stress associated with chronic alcohol exposure on drinking. Neuropharmacology. 2017; 122: 115-26. doi: 10.1016/j.neuropharm.2017.04.028

[15] Kelly J, Meredith PJ, Taylor M, Morphett A, Wilson H. Substances and your senses: the sensory patterns of young people within an alcohol and drug treatment service. Subst Abus. 2021; 42(4): 998-1006. doi:10.1080/08897077.2021.1901177.

[16] Meredith P, Moyle R, Kerley L. Substance use: links with sensory sensitivity, attachment insecurity, and distress in young adults. Subst Use Misuse. 2020;55(11): 1817-24. doi: 10.1080/10826084.2020.1766502.

[17] Yamane T. Statistics: An introductory analysis. 3rd Ed. New York: Harper & Row; 1973.

[18] Phinyo P, Ungrungseesopon N, Namsongwong N, Visavakul O, Chaiya S, Wattananukorn T, *et al*. Pre-screening ability of the Functional-Belief-Based Alcohol Use Questionnaire (FBAQ) among Chiang Mai University undergraduates: an external validation study. Drug Alcohol Depend. 2021; 227: 109002. doi:10.1016/j.drugalcdep.2021.109002

[19] Pomngen I, Srikhamjak, T, Putthinoi, S. Development of the sensory patterns assessment [Thesis]. Faculty of Associated Medical Sciences, Chiang Mai University; 2020. [in Thai]

[20] Chen WL, Chen JH. College fields of study and substance use. BMC Public Health. 2020;20(1):1631. doi: 10.1186/s12889-020-09722-1

[21] Brito APA, Lima AS, Rocha ÁC, Gonçalves BM, Freitas DMCB, Oliveira GJ, *et al*. Alcohol consumption habits and their impact on academic performance: analysis of ethanol patterns among health students. A cross-sectional study. Sao Paulo Med. J. 2024; 142(6): e2023410. doi: 10.1590/1516-3180.2023.0410.R1.05062024

[22] Suparan K, Wongsuthipol J, Hiranwong T, Rojanasumapong A, Pinyopornpanish K, Angkurawaranon C, *et al*. A study of the association between alcohol consumption, road accidents, learning behavior and academic achievement in university students. BSCM 2019; 58(3): 133-43. Available from: https://he01.tci-thaijo.org/index.php/CMMJ-MedCMJ/article/view/121381 [in Thai].

[23] Mekonen T, Fekadu W, Chane T, Bitew S. Problematic alcohol use among university students. Front Psychiatry. 2017; 8: 86. doi: 10.3389/fpsyt.2017.00086

[23] Collier L, Staal J, Homel P. Multisensory environmental therapy (Snoezelen) for job stress reduction in mental health nurses: a randomized trial. Int j complement altern med. 2018; 11(2): 1-7. doi 10.15406/ijcam.2018.11.00346

[24] Borges JM, Castillo JAGd, Marzo JC, Castillo-López ÁGd. Relationship between sensory processing, resilience, attitudes and drug use in Portuguese adults. Paidéia (Ribeirão Preto). 2017; 27(68): 255-62. doi:10.1590/1982-43272768201702

[25] Niemelä O, Bloigu A, Bloigu R, Halkola AS, Niemelä M, Aalto M, *et al*. Impact of physical activity on the characteristics and metabolic consequences of alcohol consumption: A cross-sectional population-based study. Int J Environ Res Public Health. 2022; 19(22): 15048. doi: 10.3390/ijerph192215048

[26] Gligoroski A, Stojmanovska DS. Alcohol consumption habits among young adults, with sedentary and active lifestyle. RPESH. 2018; 7(2): 1857-8160. Retrieve from https://www.researchgate.net/publication/338829575

[27] Ashok K. Critical review of alcohol, alcoholism and the withdrawal symptoms I. mechanisms of addiction and withdrawal syndrome. Arch Addict Rehabil. 2017; 1(1): 11-30. doi:10.36959/843/421

[28] Supawat A, Phunikhom K, Sattayasai J. Alcohol addiction, alcohol withdrawal syndrome and treatment. Srinagarind Med Journal. 2015;30(6):618-29. Available from: https://www.thaiscience.info/Journals/Article/SRMJ/10981942.pdf [in Thai].

[29] Horii Y, Nagai K, Nakashima T. Order of exposure to pleasant and unpleasant odors affects autonomic nervous system response. Behav Brain Res. 2013; 243: 109-17. doi: 10.1016/j.bbr.2012.12.042

[30] Newbury-Birch D, Gilvarry E, McArdle P, Ramesh V, Stewart S, Walker J, *et al*. Impact of alcohol consumption on young people: a review of reviews. [internet] Nottingham: Department for Children, Schools and Families; 2009 [cited 2023 Mar 29]. Available from: https://www.drugsandalcohol.ie/12292/.

Journal of Associated Medical Sciences 2025; 58(3): 33-42

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Development and validation of a self-assessment instrument for sodium intake in the population aged 30-44 years in 4th Regional Health, Thailand

Krich Ruangchai* **and Artitaya Wangwonsin**

Faculty of Public Health, Naresuan University, Phitsanulok Province, Thailand.

ARTICLE INFO

Article history:
Received 3 February 2025
Accepted as revised 7 May 2025
Available online 13 May 2025

Keywords:
Validation, sodium intake, self-assessment, instrument development

ABSTRACT

Background: Laboratory-based urine tests are the current standard for assessing sodium intake but are complex and costly. This study aimed to develop a valid, self-assessment questionnaire aligned with laboratory findings using a six-step instrument development process.

Objective: To develop and validate a self-assessment instrument for estimating sodium intake among individuals aged 30–44 years.

Materials and methods: The six-step process involved (1) defining sodium intake by reviewing existing research; (2) creating a list of items based on different food types and how they are eaten; (3) deciding on the format and scale for assessment; (4) checking face validity with feedback from expert advisors; (5) evaluating content validity with five experts and calculating the Item-Objective Congruence (IOC); and (6) testing construct and concurrent validity using ROC analysis.

Results: We categorized a total of 157 food items for inclusion. The questionnaire used a checklist format with frequency ratings (1–7 times/day). Content validity yielded IOC values between 0.6 and 1.0, with 151 items retained. Construct validity showed strong alignment with sodium excretion levels, with a cut-off point of 3.46 and an ROC of 93.3%. Of 120 participants, 112 were correctly classified, yielding 92.8% sensitivity, 95.2% specificity, a 6.7% false positive rate, and 86.7% accuracy (95% CI = 0.880-0.987).

Conclusion: The developed self-assessment tool is practical for single-day dietary evaluation and demonstrates high specificity and accuracy. It offers a reliable, cost-effective alternative for assessing sodium intake at the population level.

Introduction

Sodium is an essential mineral for maintaining blood electrolyte balance. Consumed primarily through diet, excessive or insufficient sodium intake can disrupt blood pressure and contribute to health problems.[1] In developing countries, NCDs are causing a public health crisis, with rising rates of illness and premature death, especially among those aged between 30-60.[2,3] Globally, NCDs account for 41 million deaths annually, representing a 71% increase in recent years. The Easthand Southeast Asian regions are particularly hard hit, with approximately 10.4 million deaths attributable to NCDs each year.[4]

Hypertension (45.0%) and diabetes (28.7%) accounted for the highest mortality rates among individuals with non-communicable diseases (NCDs). Dietary habits high in sugar, fat, and salt were the primary risk factors.[5,6] Thailand has approximately 14 million NCD patients, with

* **Corresponding contributor.**
Author's Address: *Faculty of Public Health, Naresuan University, Phitsanulok Province, Thailand.*
E-mail address: krichrua@gmail.com
doi: 10.12982/JAMS.2025.070
E-ISSN: 2539-6056

over 300,000 deaths annually.[7] The working-age cohort aged 30-44 exhibits early hypertension risk indicators coupled with a diminished engagement in preventive health behaviors, encompassing self-assessment and routine examinations.[8,9] The average age of onset is 40.5 years, with hypertension (21.4%) and hyperlipidemia (19.4%) being the most common NCDs.[10] Lifestyle factors, particularly dietary habits deeply rooted in Thai culture, can be linked to the disease. The high and frequent consumption of sodium-rich foods, such as sukiyaki, shabu-shabu, and instant noodles, combined with the traditional use of salt in cooking, contributes significantly to the problem.[11]

To effectively assess sodium intake within a general population, it is necessary to employ assessment instruments with high specificity for dietary behavior. This approach facilitates the collection of comprehensive data, crucial for accurately evaluating population sodium intake and informing precise dietary behavior interventions and policies aimed at reducing consumption. Currently, two primary categories of instruments are available for this purpose. The first is 24-hour urinary sodium excretion, which is considered the gold standard for sodium intake assessment.[12] However, this method is expensive and not easily accessible to the general population. The urine collection process is also complex, requiring specialized training for sample collectors, making it impractical for large-scale population studies. For the second, four primary methods employing questionnaires or interviews are utilized in Thailand for the assessment of dietary behavior: the Food Record, the 24-hour Dietary Recall, the Food Frequency Questionnaire (FFQ), and the Diet History. Each of these methodologies presents distinct advantages and limitations, and the selection of a particular method is contingent upon the specific objectives of the research or evaluation endeavor.[13] Based on the examination of research instruments utilized in international studies, 1) the Food Frequency Questionnaire (FFQ) is limited by its reliance on exhaustive food lists, which can hinder respondents' accurate recall of dietary intake. Consequently, validation through supplementary assessment instruments, such as food diaries or urinary sodium excretion assessment, is often necessary,[14,15] 2) the Sodium Food Frequency Questionnaire (FFQ-Sodium)[16] was developed to provide a quick and minimally burdensome method for assessing sodium intake. By focusing on high-sodium foods relevant to high-risk groups in specific communities, it maintains a relatively low error rate.[17] However, it is limited by delays in providing results and the time required to calculate sodium content from the collected data.[18] This tool requires participants to log all meals, food items (including sodium-rich sources), and quantities over at least three days. To ensure data accuracy, simultaneous data collection is needed across all participants within a standardized timeframe.[19] Similarly, the Dietary Record and Discretionary Salt Questionnaire (DSQ) assesses sodium intake through the self-recording of food consumption and added sodium over a minimum of three days. Data are captured electronically via websites, mobile apps, or computer programs to reduce recall errors and facilitate easy recording across different formats.[20]

An analysis of available data reveals a limited range of standardized and accessible tools for assessing sodium intake in Thailand, with reliance mainly on laboratory testing.

To address this gap, this study aimed to develop a self-assessment instrument for sodium consumption, targeting the 30-44 age group in the 4th Health Region, Thailand. This demographic was chosen due to age-related declines in physiological functions, the high prevalence of chronic non-communicable diseases linked to excessive sodium intake, and limited engagement in sodium reduction strategies. National survey data show this group consumes an average of 3,200 mg of sodium per day, exceeding the recommended intake. The region's diverse population allows for a broad representation of dietary behaviors, making it ideal for developing and validating a comprehensive sodium intake assessment tool.[21,22] The instrument development involved two integrated assessment methods, comprising the Food Frequency Questionnaire (FFQ) and food record, allowing for Receiver Operating Characteristic (ROC) analysis to determine an optimal cut-off point based on urinary sodium excretion data as the gold standard. The development process included defining sodium consumption, following a standardized protocol, and rigorously validating the tool's psychometric properties to ensure reliability and ease of self-assessment. The resulting data is expected to inform evidence-based policy changes and the design of targeted community interventions.

Materials and methods

This research obtained ethical approval from the Naresuan University Ethics Committee (IRB No. P3-0013/2565, COA No.215/2022). The research employed a methodological study design[23] and followed a six-step process[24] for the development and validation of the instrument.

Step 1: Definition of sodium intake

To assess sodium consumption patterns, a literature review was conducted using databases related to nutrition, health, and academic theses, with keywords such as "sodium intake," "dietary sodium sources," "high-sodium foods," "Thai diet," and "sodium intake assessment." Key definitions and food categories were synthesized from the literature.

Concurrently, a community-based survey was administered within the 4th Regional Health area using both online (Google Forms) and offline (paper-based) open-ended questionnaires to identify the kinds of sodium-rich foods commonly consumed. For participants without internet access, paper surveys were distributed via community coordinators and returned using a secure, opaque drop box placed in a central community location. This approach facilitated the inclusion of locally relevant foods not captured in existing literature.

The survey targeted 270 individuals aged 30-44 years, selected through simple random sampling based on Cohen's formula for sample size estimation.[25,35] Inclusion criteria required basic literacy and the ability to recall and report past food consumption. Individuals with cognitive or medical conditions affecting data reliability were excluded.

The findings from the literature and field survey were integrated to compile a context-specific food item list. Selected items were analyzed using Inmucal-N software developed by the Institute of Nutrition, Mahidol University, to determine standardized sodium content per serving and per day, supporting accurate estimation of daily sodium intake.[26]

Step 2: Generate an item pool

Based on the findings from Step 1, a preliminary list of food items was categorized into nine culturally specific groups (e.g., Central, Northern, and Northeastern Thai cuisines). A draft Food Frequency Questionnaire (FFQ) was developed to assess sodium intake, incorporating food types, portion sizes, and sodium content. A Case Record Form (CRF) was also designed to facilitate systematic data collection and analysis.

Step 3: Determine the format for the measurement and assessment scale

An ordinal scale was used to assess consumption frequency, ranging from "rarely/never" to "7 times/day." Each respondent reported both the frequency and quantity of food intake, recorded in standardized serving sizes (e.g., teaspoons for sauces, bowls or grams for solid foods, and milliliters for liquids). Serving sizes were defined in the questionnaire with visual aids and examples. Sodium intake estimation was calculated by multiplying consumption frequency with the average sodium content per portion, allowing conversion into milligrams/day.[27]

Step 4: Face validity

The draft instrument underwent preliminary face validity checks by the research team. Data were organized and sequenced by sodium source category. The instrument was reviewed by academic advisors, focusing on content clarity, item relevance, and redundancy. Revisions were made according to the feedback to improve the overall readability and cultural appropriateness of the items.

Step 5: Content validity

The instrument for assessing sodium consumption was submitted to five experts possessing a minimum of five years of experience in research instrument evaluation and nutrition. Subsequently, the item-objective congruence index (IOC) was calculated.[28]

Step 6: Conduct a construct validity of the questionnaire

The construct validity of the instrument used to assess sodium consumption was evaluated. Before this assessment, the instrument underwent evaluations for face validity, content validity, reliability analysis, and concurrent validity. We conducted the evaluation of construct validity in two main steps.

Step 6.1:

Concurrent validity assessment involved collecting self-reported sodium intake data and 24-hour urinary sodium excretion data from a sample of 120 individuals aged 30-44 years from the 4th Health Region. We calculated the sample size using the proportion estimation method.[34] The multi-stage sampling technique commenced with a stratified random selection of geographical areas at the sub-district level. This process yielded Sam Khok subdistrict, Sam Khok district, Pathum Thani province, with a total population of 6,664 individuals. Subsequently, a proportionate allocation based on the population size within this sub-district resulted in a sample of 68 participants. Concurrently, Lam Sam Phung sub-district, Muak Lek district, Saraburi province, with a total population of 5,096 individuals, was also selected. Following a similar proportionate allocation, this sub-district contributed 52 participants to the sample. Finally, a systematic random sampling method was employed, utilizing numerical codes to represent the personal data of individuals aged 30-44 years within each selected area, culminating in a total sample size of 120 participants. Inclusion criteria comprised participants who were 30-44 years old, able to read, listen, and write in Thai, free from serious illnesses (e.g., chronic kidney failure), not undergoing diuretic treatment, not menstruating during data collection, and provided informed consent. Exclusion criteria included individuals with severe chronic non-communicable diseases or those who had lived in the 4th Health Region for less than one year. Participants were contacted to gather information about the collection, timing, and delivery of their 24-hour urine samples. Concurrently, the researchers collected self-reported sodium intake data using a validated sodium intake questionnaire. Prior to the actual collection, participants were invited to attend a briefing session to introduce and explain the proper 24-hour urine collection procedures. Additionally, participants were provided with an illustrated pamphlet outlining the step-by-step urine collection process, which they could review at their convenience to reinforce their understanding. To ensure proper compliance, a direct communication channel was established, allowing participants to contact the research team in case of questions or uncertainty regarding the collection process.[21,24]

The criteria for sample selection included 24-hour urine output data, urine volume greater than 750 ml per day, serum creatinine for males between 0.98-2.2 gm/24 hrs and between 0.72-1.51 g/24 hrs for females, and sodium intake assessment collected at the same time.[29,30] The measured 24-hour urine indices were calculated as follows: 24-hour urinary sodium excretion (mg/day) = concentration of 24-hour urinary sodium excretion (mmol/L) × 24-hour urine volume (L/day) × molecular weight of Na^+ (23 mg/mmol).[21,24,31]

Step 6.2:

The instrument's consistency (concurrent validity) was analyzed by inputting data into a Receiver Operating Characteristic (ROC) curve analysis. This analysis determines the optimal cut-off point for classifying individuals. The ROC curve, using sensitivity and specificity, can help identify groups at risk of consuming more than 2,000 mg/day.[32]

Results

Step 1: The results defined sodium consumption patterns. A literature review yielded 102 food items, and a local food survey added 55 more.

Step 2: The results generated 157 research questions, categorized by food groups, based on defined parameters.

Step 3: Specified measurement and evaluation methods. Food intake was measured in milligrams (mg), milliliters (mL), grams (g), or standard servings. Serving sizes were associated with the measurements specified in Section 1, such as 1, 1.5, and 2 servings. We evaluated the frequency of consumption for each food item (ranging from 1 to 7 times per day) using a checklist.

Step 4: An initial face validity assessment was conducted. The content validity, measurement properties, and evaluation criteria of the instrument for assessing sodium consumption developed by the researchers were reviewed by the advisor. The instrument's content clarity, redundancy, and academic rigor were assessed through a face-to-face validation process. Based on the feedback received, the instrument was revised, and food items were categorized into nine groups: condiments, single-dish meals, processed foods, fruit dips, beverages and snacks, northeastern Thai food, southern Thai food, northern Thai food, and international food.

Step 5: Content validity was evaluated by five experts with over five years of experience in nutrition and research instrument assessment. The Item-Objective Congruence (IOC) index was calculated, with values ranging from 0.6 to 1.0 indicating high consistency between items and objectives. Experts recommended revisions to six questions due to duplicate food items and suggested adjustments to food groupings. The finalized questionnaire included 151 items.

Step 6: Concurrent validity of the research instrument was evaluated via ROC curve analysis, which facilitated the examination of the association between self-reported sodium intake and 24-hour urinary sodium excretion.

Participant characteristics and sodium excretion

The study included 120 participants. The majority (51.7%) were older than 37 years, with a mean age of 37.6 years. Females comprised 55.8% of the sample. Most participants (60.8%) had not completed a bachelor's degree, and 52.5% reported a monthly income below 20,000 baht (mean income: 17,848.25 baht). Regarding health status, 72.5% reported no history of chronic diseases, while the remaining participants indicated having conditions such as hypertension or diabetes.

Importantly, 70.8% of participants had daily urinary sodium excretion levels exceeding 2,000 mg. The mean urinary sodium excretion was 3,020.81 mg/day. Although this study adopted the World Health Organization's (WHO) recommendation of 2,000 mg/day as the threshold for excess sodium intake, it is acknowledged that other literature (e.g., McCarron *et al*., 2013)[33] identifies a broader normal range (2,622–4,830 mg/day). Nonetheless, the 2,000 mg/day benchmark was used for ROC analysis to align with international dietary guidelines and public health goals (Table 1).

Table 1. Demographic and other characteristics of 120 participants for the ROC curve from a population aged 30-44 years in the 4th Health Region, Thailand.

Variable	Frequency	%
Age		
≤37 years	58	48.3
>37 years	62	51.7
(Mean=37.60, SD=3.58, min=30, Max=44)		
Gender		
Male	53	44.2
Female	67	55.8
Education		
Below bachelor's d egree	73	60.8
Higher than a bachelor's degree	47	39.2
Occupation		
Government employee/private sector/trading/private business	63	52.5
General contractor/farmer/animal raising	57	47.5
Monthly income		
≤20,000 Bht	69	57.5
>20,000 Bht	51	42.5
(Mean=17,848.25, SD=10,693.58, min=6,900, Max=61,000)		
Illness/disease		
Not illness	87	72.5
Illness	33	27.5
24-hour sodium excreted in urine		
≤2,000 mg/day	35	29.2
>2,000 mg/day	85	70.8
(Mean=3,020.81, SD=1,689.09, min=1,155.3, Max=10,658.4)		

ROC curve analysis and cut-off determination

Validation of the food frequency questionnaire (FFQ) and food record methods was conducted using ROC curve analysis, with 24-hour urinary sodium excretion as the reference standard. The optimal cut-off point for the FFQ was determined to be 3.46, derived from the ROC analysis using both Youden's index and Euclidean distance methods. These mathematical criteria were used to identify the point with the best balance between sensitivity and specificity.

The analysis of the Receiver Operating Characteristic (ROC) curve demonstrates that Euclidean's Index (i) and Youden's Index (J) are mathematically derived measures, which, while theoretically sound, may not be directly applicable or interpretable in research settings involving practical health science. In contrast, the optimal cutoff point determined by the Weight NMC (C=3, pr=0.3) appears more suitable for applied research, particularly in the context of assessing the concurrent validity of the developed instrument.[15,42] This cutoff effectively distinguishes appropriate sodium intake levels when compared against the 24-hour urinary sodium excretion, using a benchmark of less than 2,000 milligrams per day as the reference standard for adequate sodium elimination (Table 2).

Table 2. Determination of the optimal sodium intake cutoff point by comparing it with 24-hour urinary sodium excretion.

Consideration	Cut-off point for FFQ-Na	Cut-off point for food record (mg/day)
Euclidian's index (i)	2.87	2,052
Youden's index (J)	2.98	2,094
Weight NMC (C=3, pr=0.3)	3.46	2,146

24-hour urinary biomarkers

All urine samples met the inclusion criteria, with a volume of at least 750 mL. Among males, the average 24-hour urinary sodium excretion was 134.8 mEq (SD=68.72), while it was 128.74 mEq (SD=53.16) for females. Average potassium excretion was 86.09 mEq (SD=34.16) for males and 63.46 mEq (SD=24.74) for females. The mean creatinine excretion was 1.34 gm/24 hrs (SD=0.31) in males and 0.98 gm/24 hrs (SD=0.18) in females. These values confirm physiological plausibility and support the reliability of the urine collection method used for validation (Table 3).

The normal reference ranges for laboratory test results are as follows: sodium excretion: 40.00-220.00 mEq/24 hrs, potassium excretion: 25.0-125.0 mEq/24 hrs; creatinine excretion for males was 0.980-2.200 gm/24 hrs, and 0.720-1.510 gm/24 hrs for females.

Table 3. 24-hour urine excretion laboratory test

Urine laboratory test	Male		Female	
	Mean (Min-Max)	SD	Mean (Min-Max)	SD
Volume (mL)	1150.5 (750-2,100)	108.41	890.3 (750-1,750)	86.75
Sodium (mEq)	134.08 (50.0-201.7)	68.72	128.74 (48.5-198.3)	53.16
Potassium (mEq)	86.09 (26.6-105.4)	34.16	63.46 (27.9-86.7)	24.74
Creatinine (gm/24 hrs)	1.34 (1.08-2.16)	0.31	0.98 (0.79-1.43)	0.18

Diagnostic performance of the instruments

The FFQ demonstrated excellent discriminatory performance for detecting sodium intake above 2,000 mg/day. The ROC curve area was 93.3% (95% CI: 0.880-0.987), with a sensitivity of 92.8%, specificity of 95.2%, and overall accuracy of 86.7%. Of the 120 participants, 112 were correctly classified, with only 8 false positives (6.7%) (Figure 1).

In contrast, the food record method showed weaker performance, with a cut-off point of 2,146 mg/day and an AUC of 67.6% (95% CI: 0.577-0.776). Sensitivity and specificity were 63.8% and 70.4%, respectively, while overall accuracy was 60.3%. This method correctly classified only 81 participants and resulted in a higher false positive rate (36.2%, or 39 individuals) (Table 4).

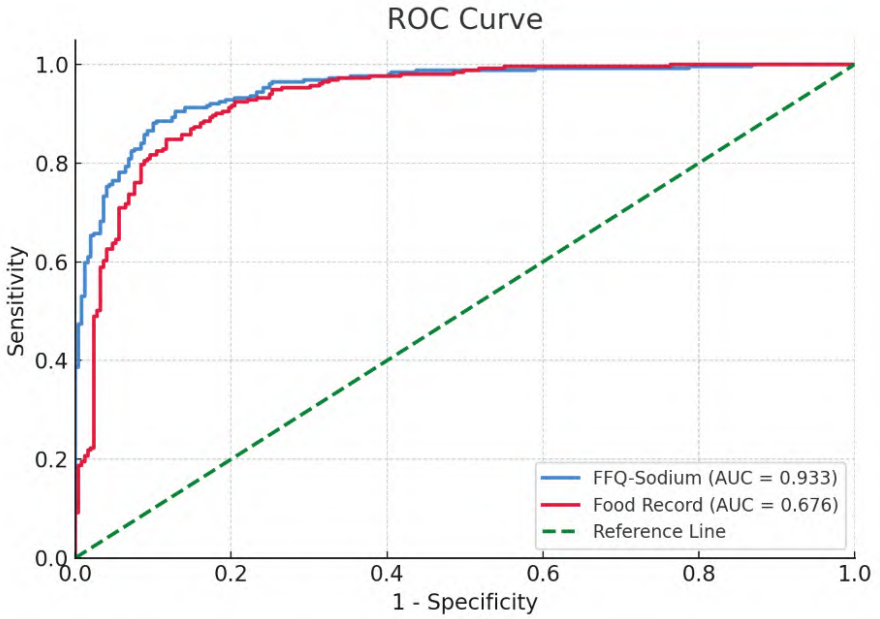


Figure 1. ROC curve of the sodium intake questionnaire.

Table 4. Analysis of the ROC curve for the sodium intake questionnaire.

Parameter	Assessment instrument	
	FFQ-Sodium	Food Record
ROC	0.933	0.676
Cut-off point	3.46	2,146
Number of true positive	112	81
Number of true negative	8	39
Sensitivity	92.8	63.8
Specificity	95.2	70.4
Accuracy	86.7	60.3
95%CI of ROC	0.880-0.987	0.577-0.776
$\chi 2$	6.173	4.807
p value	<0.001	0.001

Discussion

Summary of key findings

This study aimed to develop and validate a culturally tailored sodium intake questionnaire for adults aged 30-44 years in Thailand. The final instrument, based on a food frequency questionnaire (FFQ) format, demonstrated strong diagnostic performance when validated against the 24-hour urinary sodium excretion, a gold standard for sodium intake measurement. Among 120 participants, the instrument achieved a sensitivity of 92.8%, a specificity of 95.2%, and an area under the ROC curve (AUC) of 93.3%, indicating excellent discriminatory ability. In contrast, a food record-based instrument showed lower sensitivity (63.8%), specificity (70.4%), and AUC (67.6%), suggesting relatively poor performance. Furthermore, the FFQ classified 112 participants correctly, with only 8 false positives, compared to 39 false positives for the food record method. The 24-hour urine analysis confirmed the population's sodium and potassium excretion levels, supporting the physiological plausibility of the FFQ's estimates.

Comparison with existing tools

The newly developed Food Frequency Questionnaire (FFQ) offers several advantages over traditional sodium intake assessment tools. Firstly, the FFQ demonstrates a high accuracy with an AUC of 93.3%, strongly correlating with 24-hour urine sodium excretion, compared to only 67.6% AUC for the food record method. The culturally specific food list, comprising 151 items from Thai regional cuisines, enhances accuracy by better capturing the dietary sodium profile.[35] In terms of feasibility, the FFQ significantly reduces participant burden. Unlike the 24-hour urine collection, which is invasive, time-consuming, and logistically challenging, the FFQ can be self-administered or conducted by trained interviewers, minimizing effort and time for both participants and researchers. This makes the FFQ a practical and cost-effective alternative for routine sodium intake monitoring, particularly in resource-limited settings.[36,37] Moreover, the FFQ developed in this study stands out when compared to conventional FFQs in Thailand due to its focus on high-sodium foods and a one-day consumption assessment, eliminating the need for participants to recall their food intake over several days, thus reducing memory bias. This also alleviates the training requirements associated with food record tools, which rely on participants' ability to accurately log and recall food quantities, often leading to inaccurate data. Additionally, the instrument's ROC analysis with 24-hour urinary sodium excretion data adds specificity, offering a more reliable measure of sodium intake than traditional FFQs. In contrast, food record instruments require higher skill levels in writing and understanding food quantities and can be prone to recall bias or errors, making them less efficient for large-scale studies.[12]

In summary, the FFQ developed in this study provides an efficient, reliable, and culturally tailored tool for assessing sodium intake, making it suitable for large-scale public health surveillance and interventions, especially in settings where resources are limited.

Cultural and demographic applicability

This study focused on adults aged 30-44 years in the 4th Health Region, incorporating culturally diverse food categories representative of dietary patterns in the Northern, Central, and Northeastern regions. While this enhances the instrument's internal validity, its generalizability to other cultural or ethnic populations may be limited without adaptation.[38] Given that sodium intake patterns vary by age, ethnicity, and region, future applications of this tool should consider localized validation and modification. For instance, extending the instrument to older adults or adolescent populations would require revising food item lists and frequency categories to align with their distinct consumption behaviors. Additionally, adaptation for use in other countries or cultural contexts would necessitate pilot testing as well as recalibration of the sodium values and cut-off thresholds based on local dietary surveys.[15] The validated FFQ demonstrates strong performance as a culturally tailored, accurate, and cost-effective tool for assessing sodium intake among Thai adults aged 30-44. Its high sensitivity, specificity, and user-

friendliness make it a valuable resource for public health surveillance and dietary interventions. However, broader use across different regions and age groups should be preceded by appropriate cultural and demographic adaptations to maintain its validity and effectiveness.

Cut-off point justification and public health implications

This study validated a sodium intake assessment tool based on consumption frequency and dietary records. The optimal cut-off point of 3.46 in the food frequency questionnaire (FFQ) was identified through ROC curve analysis using 24-hour urinary sodium excretion as the reference standard. This threshold corresponds to approximately 2,000 mg/day, aligning with the WHO's recommended maximum intake for adults. It was chosen to optimize sensitivity (92.8%) and specificity (95.2%), effectively identifying individuals with excessive intake while minimizing false positives. The FFQ is unique in its day-to-day intake scale and was validated through concurrent urinary sodium collection, ensuring consistency between reported behavior and biochemical outcomes. Unlike conventional FFQs or dietary records used internationally, this tool offers enhanced precision for the Thai population. From a public health perspective, the 3.46 threshold enables targeted interventions by accurately identifying individuals exceeding the recommended intake. This supports personalized dietary counseling, risk stratification, and regional monitoring of sodium reduction efforts. As high sodium intake is a modifiable risk factor for hypertension and cardiovascular disease, this validated tool contributes to national strategies for non-communicable disease (NCD) prevention. Those instruments measure general food consumption by asking about frequency (FFQ) and sodium intake (DR) at the same time. In contrast, instruments developed internationally,[40] such as the Discretionary Salt Questionnaire (DSQ) and Food Frequency Questionnaires for Sodium (FFQ-Sodium), focus specifically on sodium intake assessment.[20] This study introduces a specific measurement instrument to assess sodium intake, which is currently unavailable in Thailand. Although its application in Thailand might encounter cultural and dietary consumption contexts that differ from other countries, this instrument offers a unique advantage by measuring both the frequency and quantity of sodium consumption, leading to more precise assessments.

Seasonal variation in food consumption among the Thai population may also influence dietary intake data. In this study, data collection was conducted during the rainy season, a period when access to a wider variety of food sources is more feasible compared to other seasons. This is particularly relevant for local foods, as many ingredients are naturally available and more accessible during the rainy season in each region. We have validated the instrument against 24-hour urine sodium collection.[39] This study validates a newly developed sodium intake assessment tool, which offers an accurate evaluation of sodium intake through urinalysis, comparable to the gold standard laboratory test. Research by DeVellis RF and Matsuno *et al*.[15] supports the importance of contextually appropriate content validity analysis in developing instruments that yield both behavioral and laboratory assessments, ensuring practical applicability, particularly in distinguishing healthy from unhealthy behaviors. The reliability of this tool facilitates the creation of targeted interventions for improved health outcomes. Our study differs from that of Li J. *et al*.,[19] a retrospective 7-day dietary assessment paired with 24-hour urine collection, which showed over 20% misclassification as well as lower sensitivity and specificity due to the retrospective nature and reliance on participant recall. Similarly, our study contrasts with Wong *et al*.'s[20] instrument, which included a 7-day retrospective dietary record, a 24-hour dietary recall, and urinary sodium excretion. Despite an 84.6% discrimination rate, the method by Oria M, Yaktine AL, Strom BL,.[41] had challenges with non-completion rates (over 50%) and required extensive training for accurate data collection, limiting its feasibility and acceptability.[42]

Conclusion

The development of a sodium intake assessment instrument is a process that facilitates easier access to sodium intake results and provides a more accurate understanding of dietary behaviors related to sodium consumption. Health assessment is an integral part of addressing health issues, particularly in identifying specific high-sodium foods. This information can guide more precise dietary choices, ultimately reducing overall sodium intake in daily life.

Conflicts of interest

The authors declare no conflicts of interest.

Suggestions

1. Future studies should validate the instrument using different age groups and cultural contexts to enhance its generalizability and ensure accurate sodium intake assessment across diverse populations.
2. The FFQ-sodium can be incorporated into national health surveys, primary care screenings, and mobile health applications to support public health monitoring, dietary counseling, and salt reduction initiatives.
3. The tool could be improved by integrating digital platforms (e.g., mobile apps with real-time feedback) and combining it with biological markers like repeated 24-hour urine sodium tests to increase precision and user engagement.

Acknowledgements

The authors would like to express their sincere gratitude to all data providers, experts, and funding agencies for their generous support of this research.

References

[1] Podapon J. Online learning management: The way of education. Mahamakut Royal University; 2020.
[2] Gurdil H, Anadol HO, Soguksu YB. The use of artificial intelligence tools in assessing content validity: a comparative study with human experts. arXiv [Preprint]. 2025 Feb 3. Available from: https://arxiv.

org/abs/2503.15525

[3] Ray P, Le Manach Y, Riou B, Houle TT, Warner DS. Statistical evaluation of a biomarker. *Anesthesiology*. 2010;112(4):1023– 40doi: 10.1097ALN.0b013e3181 d47604

[4] Jaafar MK, Rahman AR, Nordin N. Correlation between spot urine sodium, 24-hour urinary sodium, and food frequency questionnaire in estimation of salt intake in young healthy individuals. *J Hypertens*. 2018;36:e288. doi: 10.1097/01.hjh.0000539177.80515.5c

[5] Ministry of Public Health (MOPH). Non-communicable diseases (NCDs) situation in Thailand 2019. [cited 2024 Nov 13]. Available from: http://www.thaincd.com/

[6] World Health Organization. Clinical guidelines for the management of hypertension. Geneva: World Health Organization; 2005 [cited 2021 Feb 19]. Available from: https://www.who.int/iris/handle/10665/119738

[7] Shahar S, Shahril MR, Abdullah N, Borhanuddin B, Kamaruddin MA, Yusuf NAM, et al. Development and relative validity of a semiquantitative food frequency questionnaire to estimate dietary intake among a multi-ethnic population in the Malaysian Cohort Project. Nutrients. 2021;13(4):1163. doi: 10.3390/nu13041163.

[8] Beresford SA, Thompson B, Feng Z, Christianson A, McLerran D, Patrick DL. Development of an assessment tool to measure healthy eating in Navajo children and their families. J Acad Nutr Diet. 2023;123(3):456–65. doi: 10.1016/j.jand.2022.12.005.

[9] McCarron DA, Kazaks AG, Geerling JC, Stern JS, Graudal NA, Dawson-Hughes B. Normal range of human dietary sodium intake: a perspective based on 24-hour urinary sodium excretion worldwide. *Am J Hypertens*. 2013;26(10):1218–23. doi: 10.1093/ajh/hpt139

[10] McLean RM, Farmer VL, Nettleton A, Cameron CM, Cook NR, Campbell NRC, TRUE Consortium. Assessment of dietary sodium intake using a food frequency questionnaire and 24-hour urinary sodium excretion: a systematic literature review. *J Clin Hypertens*. 2017;19(12):1214–30. doi: 10.1111/jch.13136

[11] World Health Organization. WHO Thailand: improving hypertension care cascade with more than 60% control rate through innovation [Internet]. Geneva: World Health Organization; 2023 [cited 2023 Apr 25]. Available from: https://www.who.int/thailand/news/detail/2023/04/25/who-thailand-improving-hypertension-care-cascade-with-more-than-60-percent-control-rate-through-innovation

[12] Chailimpamontree W, Kantachuvesiri S, Aekplakorn W, et al. Estimated dietary sodium intake in Thailand: a nationwide population survey with 24-hour urine collections. *J Clin Hypertens (Greenwich)*. 2021;23(4):744–54. doi: 10.1111/jch.14147

[13] Nephrology Society of Thailand. Nephrology Meeting 2020. [Conference]. 2020 Nov 13–14. Available from: https://www.nephrothai.org/

[14] Saengyo S, Rerkkasem K, Wungrath J. Effectiveness of a Line application together with telephone-based consultation and education program on the dietary knowledge and behavior among caregivers of end-stage renal disease patients on hemodialysis. *Malays J Public Health Med*. 2023;23(2):282–90.doi: 10.37268mjphm/vol.23/no.2/art.1300

[15] Matsuno T, Takachi R, Ishihara J, Ishii Y, Kito K, Maruya S, et al. Validity of the food frequency questionnaire—estimated intakes of sodium, potassium, and sodium-to-potassium ratio for screening at a point of absolute intake among middle-aged and older Japanese adults. *Nutrients*. 2022;14(13):2594. doi: 10.3390/nu14132594

[16] World Health Organization. Global database on body mass index [Internet]. Geneva: World Health Organization; 2021 [cited 2021 Feb 8]. Available from: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html

[17] Huang Y, et al. Measurement error corrected sodium and potassium intake estimation using 24-hour urinary excretion. *Hypertension*. 2014;63(2):238–44. doi: 10.1161/HYPERTENSIONAHA.113.02018

[18] Ummee K, Rerkkasem K, Wungrath J. The effect of fluid overload control program on knowledge and behavior among caregivers of end-stage renal disease patients on hemodialysis. *Kesmas*. 2023;18(4):258–64. doi: 10.21109/kesmas.v18i4.6511

[19] Vanderveen SM, Cohen HW. Sodium intake and mortality follow-up in the Third National Health and Nutrition Examination Survey (NHANES III). *J Clin Hypertens (Greenwich)*. 2009;11(3):85–91. doi: 10.1111/j.1751-7176.2008.00158.x

[20] Wong AT, Munt A, Allman-Farinelli M, Badve SV, Boudville N, Coolican H, et al. Assessment of dietary sodium intake using the scored salt questionnaire in autosomal dominant polycystic kidney disease. *Nutrients*. 2020;12(11):3376. doi: 10.3390 nu12113376

[21] Kato K, Ishigami T, Kobayashi T, et al. Relationship between changes in blood pressure from summer to winter and estimated 24-hour salt excretion using spot urine: the Niigata Wellness Study. Hypertens Res. 2023;46(1):226–30. doi: 10.1038/s41440-022-01049-1.

[22] Li Y, Zhang P, Wu J, Ma J, Xu J, Zhang X, et al. Twenty-four-hour urinary sodium and potassium excretion and their associations with blood pressure among adults in China: baseline survey of Action on Salt China. *Hypertension*. 2020;76(5):1580–8. doi: 10.1161/HYPERTENSIONAHA.120.15809

[23] Polit DF, Beck CT. Nursing research: principles and methods. 7th ed. Philadelphia: Lippincott Williams & Wilkins; 2004.

[24] Tanaka T, Okamura T, Miura K, Kadowaki T, Ueshima H, Nakagawa H, Hashimoto T. A simple method to estimate populational 24-h urinary sodium and potassium excretion using a casual urine specimen. *J Hum Hypertens*. 2002;16(2):97–103. doi: 10.1038/sj.jhh.1001333

[25] Thai Health Promotion Foundation. NCDs group. [cited 2024 Nov 13]. Available from: https://hpfhub.info/fund-management

[26] Fang K, Li Y, Zhang P, Wu J, Xu J, Zhang X, et al. Dietary sodium intake and food sources among Chinese adults: data from the CNNHS 2010–2012. *Nutrients*. 2020;12(2):453. doi: 10.3390/nu12020453

[27] Ji Y, Plourde H, Bouzo V, Kilgour RD, Cohen TR. Validity and usability of a smartphone image-based dietary assessment app compared to 3-day food diaries in assessing dietary intake among Canadian adults: randomized controlled trial. *JMIR Mhealth Uhealth*. 2020;8(9):e16953. doi: 10.2196/16953

[28] Gallani MC, Proulx-Belhumeur A, Almeras N, Després J-P, Doré M, Giguère J-F. Development and validation of a salt food frequency questionnaire (FFQ-Na) and a discretionary salt questionnaire (DSQ) for the evaluation of salt intake among the French-Canadian population. Nutrients. 2020;13(1):105. doi: 10.3390/nu13010105

[29] Gowrishankar M, Blair B, Rieder MJ. Dietary intake of sodium by children: why it matters. Paediatr Child Health. 2020;25(1):47-53. doi: 10.1093/pch/pxz153

[30] Phonsuk P, Vongmongkol V, Satheannoppakao W, Pongutta S, Prasertsom P, Swinburn BA. Sodium intake and socio-demographic determinants of the non-compliance with daily sodium intake recommendations: Thai NHES IV. J Health Popul Nutr. 2020;39(1):10. doi: 10.1186/s41043-020-00213-1

[31] Department of Health, Ministry of Public Health. Annual report; 2019 [cited 2020 Apr 10]. Available from: https://ihri.org/wp-content/uploads/2020/10/Annual-Report-2019-online.pdf

[32] World Health Organization. WHO global sodium benchmarks for different food categories. [Internet]. Geneva: World Health Organization; 2021. Available from: https://www.who.int/

[33] Sari AN, Farapti F, Md Nor N. Salt taste threshold as a detection of salt intake in hypertensive individuals. JBE. 2022 Sep 26;10(3):227-36. doi: 10.20473/jbe.v10i3.25805

[34] Cohen J. Statistical power analysis. Curr Dir Psychol Sci. 1992;1(3):98-101. doi: 10.1111/1467-8721.ep10768783

[35] DeVellis RF. Scale development: theory and applications. 2nd ed. Thousand Oaks, CA: Sage Publications; 2003.

[36] Fielding CL, Langdon E, et al. Application of the sodium dilution principle to calculate extracellular fluid volume changes in horses during dehydration and rehydration. Am J Vet Res. 2008;69(11):1506-11. doi: 10.2460/ajvr.69.11.1506

[37] Freedman LS, Commins JM, Moler JE, Arab L, Baer DJ, Kipnis V, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for potassium and sodium intake. Am J Epidemiol. 2015;181(7):473-87. doi: 10.1093/aje/kwu325

[38] Laffer CL, Scott RC, Titze JM, Luft FC, Elijovich F. Hemodynamics and salt-and-water balance link sodium storage and vascular dysfunction insalt-sensitive subjects. Hypertension. 2016;68(1):195-203. doi:10.1161/HYPERTENSIONAHA.116.07289.

[39] Institute of Nutrition. INMUCAL-Nutrients V2 Program. Nakhon Pathom: Institute of Nutrition, Mahidol University.

[40] Souza DS, Santos BI, Costa BM, Santos DM, Aragão LG, Pires LV, et al. Food frequency questionnaire for foods high in sodium: validation with the triads method. PLoS One. 2023;18(7):e0288123. doi: 10.1371/journal.pone.0288123

[41] Oria M, Yaktine AL, Strom BL, editors. Sodium intake in populations: assessment of evidence. Washington (DC): National Academies Press; 2013.

[42] Sarakarn P, Munpolsri P. Optimal cut-off points for receiver operating characteristic (ROC) curve analysis in developing tools of health innovations: example using Stata. Thai Bull Pharm Sci. 2021;16(1):93-108. doi: 10.69598/tbps.16.1.93-108

Journal of Associated Medical Sciences 2025; 58(3): 43-51

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

Integrated long-term care model for dependent elderly in Mahasarakham Province, Thailand

Weerasak Aneksak[1,2], Chai Meenongwar[2], Choosak Nithikathkul[2*]

[1]Faculty of Medicine, Mahasarakham University, Mahasarakham Province, Thailand.
[2]Tropical Health Innovation Research Unit, Faculty of Medicine, Mahasarakham University, Mahasarakham Province, Thailand.

ARTICLE INFO

Article history:
Received 1 January 2025
Accepted as revised 6 May 2025
Available online 24 May 2025

Keywords:
Elderly care, dependent elderly, non-communicable diseases, integrated care model.

ABSTRACT

Background: Long-term care (LTC) for dependent elderly individuals encompasses a wide range of medical, personal, and social services designed to meet their physical, emotional, and psychological needs over an extended period.

Objective: This study aimed to develop and evaluate an integrated long-term care model for dependent elderly individuals in Mahasarakham Province.

Materials and methods: A mixed-methods research design divided into three phases: Phase 1 focused on identifying the problems and care needs of dependent elderly individuals; Phase 2 involved the development of an integrated care model; and Phase 3 involved the implementation and evaluation of the model. The Geographic Information Systems (GIS) application was crucial for identifying where care support was needed. The study included 95 elderly participants and seven caregivers. Data was collected using questionnaires and structured interviews.

Results: The study included 60% female and 40% male participants, with the majority (34.74%) falling within the age more than 75 years. Hypertension was as the most prevalent underlying condition, affecting 42.11% of the sample. In terms of elderly dependency, 8.42% of participants were severe dependency, while 6.32% were deemed completely dependent. The Barthel ADL Index analysis showed that the average score for daily living activities before care was 6.56 (with a standard deviation of 4.64), and this score went up to 10.22 (with a standard deviation of 8.19) after receiving care. Additionally, caregivers' knowledge regarding elderly dependency care showed significant improvement following a structured training program. The mean score of caregiver knowledge increased after the training, with the change being statistically significant.

Conclusion: This model highlights its potential for broader application, offering a promising approach to improving the quality of life for the elderly by enhancing their ability to perform ADL activities. By integrating this program into local and regional healthcare strategies, it will contribute to the overall healthcare campaign, fostering improved health outcomes and a higher quality of life for dependent elderly people.

Introduction

The global population is undergoing a significant demographic shift, with a rapid increase in the aging population. This trend reflects advancements in healthcare, improved living conditions, and declining fertility rates. By 2022, it was estimated that the global elderly population reached 1.1 billion, accounting for 14% of the total global population of 8 billion.[1] By 2050, this number is expected to double, posing substantial social, economic, and healthcare challenges worldwide.[2]

* **Corresponding contributor.**
Author's Address: *Tropical Health Innovation Research Unit, Faculty of Medicine, Mahasarakham University, Mahasarakham Province, Thailand.*
E-mail address: choosak.l@msu.ac.th
doi: 10.12982/JAMS.2025.071
E-ISSN: 2539-6056

In Thailand, the aging population has grown substantially over the past few decades. In 2022, Thailand's population was recorded at 66 million, based on civil registration data from the Ministry of Interior. Among them, 13 million individuals were aged 60 or older, representing 19% of the total population.[3] By 2040, it is projected that 32.1% of the Thai population will be elderly, marking Thailand as a super-aged society.[4] This rapid demographic change necessitates urgent reforms in healthcare services, particularly in long-term care systems, to support the growing number of elderly individuals with physical, mental, and social vulnerabilities.

The trend is equally evident in Health Region 7 of Thailand, where the elderly population grew from 21.19% in 2021 to 23.79% in 2024.[5] Within this region, Mahasarakham Province, with a total population of 664,651, has 167,528 elderly individuals, constituting 25.2% of its population.[6] Among them, 17,017 are categorized as dependent elderly due to chronic illnesses, disabilities, or other conditions requiring continuous care. Kantharawichai District, one of the province's administrative areas, alone has 1,001 dependent elderly individuals, or 5.88% of the district's population. These individuals often face significant challenges, including social isolation, inadequate access to healthcare services, and difficulties managing chronic diseases.[5]

The study explores the integration of long-term care services for dependent elderly individuals in the Na Sinuan community through the innovative application of Geographic Information Systems (GIS). This approach is intended to optimize resource allocation and improve the efficiency of care delivery systems, thereby addressing the unique challenges faced by dependent elderly populations in rural settings.

Materials and methods

Study design

This study employed a mixed-methods approach, combining research and development (R&D) techniques, to design a care model for the elderly with dependency due to non-communicable chronic diseases. The model was developed using Deming's cycle theory[7] (PDCA: Plan-Do-Check-Act).

Study procedure

***Phase 1**: Identifying problems and care needs for dependent elderly.*

Conduct a comprehensive review of literature and organize focus group discussions to identify problems and care models of dependent elderly in Mahasarakham Province. A questionnaire was applied to 95 purposively sampled patients; this study targeted elderly individuals with dependency who had been diagnosed by physicians with diabetes, hypertension, asthma, and others. The assessment focused on measuring the ability of these dependent elderly individuals to perform ADLs.

***Phase 2**: Model development.*

This phase builds on the findings from Phase 1, which identified policies and issues related to the care needs of dependent elderly with chronic non-communicable diseases in Mahasarakham Province. The goal is to design a care model for the caregivers and elderly, it seems that addresses these needs effectively. The procedure of action research included:

2.1 Conducting a situation analysis of the study area and collect baseline data on elderly individuals with dependency, focusing on their potential, self-care abilities, and care needs. This will be achieved through interviews.

2.2 Employing the action research (AR) methodology based on the framework. The study incorporates empowerment theory and the health belief model to design activities using the quality development process guided by the Deming Cycle (PDCA).

2.3 Focus group discussions and small group meetings (multiple focus group discussions) to understand the context and care needs of elderly individuals with dependency in the context of NCDs in Mahasarakham Province. The project involved collaboration with network partners and stakeholders, including seven caregivers and 95 elderly individuals with dependency. Data will be collected through audio recordings.

2.4 Developing a care model for elderly individuals with dependency in the context of NCDs.

Step 1: Preparation phase (Plan)

This step focused on understanding the process and preparing the target group before starting the intervention.

Step 2: Implementation phase (Do)

This step involved a collaborative learning process, focusing on educating stakeholders about dependent elderly individuals with chronic non-communicable diseases. The development activities follow a structured four-step approach.

D1: Empowerment of elderly individuals and caregivers:

Enhance the capabilities of elderly individuals and caregivers to care for themselves and others by organizing community and family learning activities with the following initiatives: empowering dependent elderly individuals through group sessions held twice a week, each lasting 90-120 minutes, for five weeks; and conducting workshops for caregivers, specifically targeting long-term care managers (CM), to strengthen community-based services that support elderly individuals at the family and community levels.

D2: Development of elderly care systems:

Establish a comprehensive care system that included health screening, medical treatment, referral systems, and community-based health

care services, while conducting educational activities, media exchanges, and interactive health workshops under the program "Healthy Aging Care" to promote health awareness and support.

D3: Testing the developed model:

A quasi-experimental research design, which involved comparing the mean differences in the scores for daily activity performance ability, measured twice—once before and once after participation in the program.

D4: Development of innovations for elderly Care:

Create home-visit initiatives for elderly individuals who remained homebound, ensuring direct interaction and support through consistent visited that strengthen engagement and provide practical assistance in maintaining health and well-being.

Step 3: Observing sustainability (Check)

Researchers observed and documented the implementation of care activities, examining the lifestyle and adherence to the developed care model among elderly individuals with chronic non-communicable diseases.

Step 4: Reflection on model development (Act)

Reflecting on the effectiveness and efficiency of the developed care model in the context of dependent elderly individuals with chronic non-communicable diseases in Mahasarakham Province, evaluating outcomes and areas for improvement.

***Phase 3**: Model implementation and evaluation*

A quasi-experimental research design was employed, comparing the mean differences in the scores of ADL abilities measured twice—before and after participation in the program. The study utilized purposive sampling, focusing on elderly individuals with dependency in the context of NCDs and their caregivers. The 95 dependent elderly and 7 trained caregivers in purposive sampling of, totaled 102 participants.

The target groups were selected based on the following criteria:

Elderly group

Participants included elderly individuals of both genders, aged 60 years and older, with an ADL score of less than 12, as defined by the Guidelines for Screening and Health Assessment of the Elderly.[8] Eligible participants were required to hold Universal Coverage (UC) scheme rights registered with the local administrative organization, reside in the target research area at the time of data collection, demonstrate a willingness to participate voluntarily, and be classified as dependent elderly with chronic conditions such as diabetes or hypertension, etc.

Caregiver group

The primary caregiver was defined as an individual responsible for assisting the elderly with basic self-care activities and actively fulfilling this role at the time of the study. Eligible caregivers provided care for elderly individuals covered under the Universal Coverage Scheme, resided in the same household as the dependent elderly individual, and could be either related (e.g., child, son/daughter-in-law, spouse, grandchild, sibling) or community caregivers. Additionally, caregivers needed to be able to communicate effectively and demonstrate a willingness to participate voluntarily as research informants.

The following steps in Phase 3 were undertaken.

1.1 Enhance the capacity of personnel in providing care for elderly individuals with dependency in the context of NCDs. This was achieved through a hands-on workshop aimed at empowering both the dependent elderly and their family caregivers.

1.2 Conduct focus group discussions and community forums to promote participation and transparency in elderly care through the collaboration of community network partners.

1.3 Organize knowledge-sharing sessions to facilitate the exchange of ideas and the development of innovations for caring for dependent elderly individuals, focusing on knowledge transformation and sustainability.

1.4 Evaluate outcomes using interviews to assess the ability of elderly individuals to perform ADLs.

Data analysis

The demographic data was analyzed using descriptive statistics and presented through frequency, percentage, mean, standard deviation, minimum, and maximum values. The differences in ADL scores were compared using the dependent t-test or paired t-test.

Results

Demographic characteristics

A total of 95 elderly participants, 60% of whom were female, and 40% of whom were male. The age distribution revealed a predominance of individuals aged more than 75 years (34.74%). For the underlying disease found 42.11% of the sample had hypertension, 29.47% had diabetes, 20% experienced a combination of both hypertension and diabetes, as well as asthma (5.26%) and chronic obstructive pulmonary disease (3.16%) Regarding socio-economic factors, a substantial portion of participants were widowed (73.68%) and had received only primary education (76.84%), as shown in Table 1.

Table 1. Socio-demographic data of the elderly with dependency (N=95).

Socio-demographic data	Number (N=95)	Percentage (%)
Sex		
Male	38	40.00
Female	57	60.00
Age		
60-65 years	25	26.31
66-70 years	21	22.11
71-75 years	16	16.84
More than 75 years	33	34.74
Underlying diseases		
Hypertension	40	42.11
Diabetes	28	29.47
Both hypertension and diabetes	19	20.00
Asthma	5	5.26
Chronic obstructive pulmonary disease	3	3.16
Marital status		
Single	5	5.27
Married	18	18.95
Widowed	70	73.68
Divorced	2	2.10
Education		
Not educated	14	14.74
Primary school	73	76.84
Higher than primary school	8	8.42

Elderly dependency classification

The dependency levels among elderly participants were classified based on their ability to perform daily living tasks, assessed using the Barthel Activities of Daily Living (ADL) Index, a standardized tool widely applied in geriatric assessments to determine functional independence.[9] The findings reveal that 6.32% of participants exhibited high dependency, necessitating considerable assistance in their daily activities. In comparison, demonstrated moderate dependency (30.53%), while a smaller group (8.42%), was classified as severely dependent (Table 2). These results presented significant issues as they highlight the varying degrees of dependency within the elderly population and emphasize the necessity for personalized care strategies, with a substantial number of individuals requiring extensive care support, and intervention tailored to their specific needs. These outcomes underscore the necessity of integrating comprehensive care strategies that indicated the varying dependency levels and chronic diseases prevalent among the elderly in the study. Moreover, the high prevalence of chronic diseases such as hypertension and diabetes indicated the importance of targeted healthcare interventions to improve the quality of life for these individuals.

Table 2. Elderly dependency classification.

Symptom group level	Number	%
Mildly dependent	52	54.74
Moderate dependency	29	30.53
Severe dependency	8	8.42
Complete dependency	6	6.32
Total	**95**	**100.00**

Barthel ADL index analysis

The comparison of the mean differences of performing the activities of daily living, with the Barthel ADL Index, among the elderly revealed the following results: The mean score of ADL ability before receiving care was 6.56 (SD=4.640), whereas the mean score after receiving care increased to 10.22 (SD=8.197). This indicates that the ADL ability scores of the elderly after receiving care were significantly higher than those before receiving care, (p=0.026), as shown in Table 3.

The knowledge of elderly dependency care among caregivers

Following the development of caregiving skills was performed among seven caregivers responsible for elderly individuals with dependency and NCDs. Based on the knowledge assessment, this found that the mean score of ten-point knowledge before the skill development training was 54.05 (SD=8.351). After the training, the mean score increased to 65.80 (SD=2.587). This indicated a statistically significant improvement in knowledge levels following the training (p=0.00), as presented in Table 4.

Table 3. Ability to perform activities of daily living (ADL) among the elderly before and after the care intervention (N=95).

Group	Number	Min	Max	Mean	SD	*t*	*p* value
Before	95	0	11	6.56	4.640	2.668	0.026*
After	95	4	15	10.22	8.197		

Table 4. Knowledge of elderly dependency care before and after Intervention in care giver.

Group	Number	Min	Max	Mean	SD	*p* value
Before	7	42	68	54.05	8.351	0.000*
After	7	59	73	65.80	2.587	

Elderly decent care with non-communicable disease model

A total of 95 elderly individuals with dependency conditions within the NCD group was selected. Their LTC conditions were assessed, focusing on health status evaluation using ADL scores and identifying care needed for the elderly in the LTC group with NCD. A GIS was employed to record the location of the homes of the LTC-dependent elderly. The type of elderly individuals in LTC was visualized based on their ADL score groups, represented with different colors: blue, green, orange, and red. Additionally, the locations of trained CG were marked as yellow stars. This mapping supported the planning for appropriate assistance for dependent elderly individuals, as presented in Figure 1.

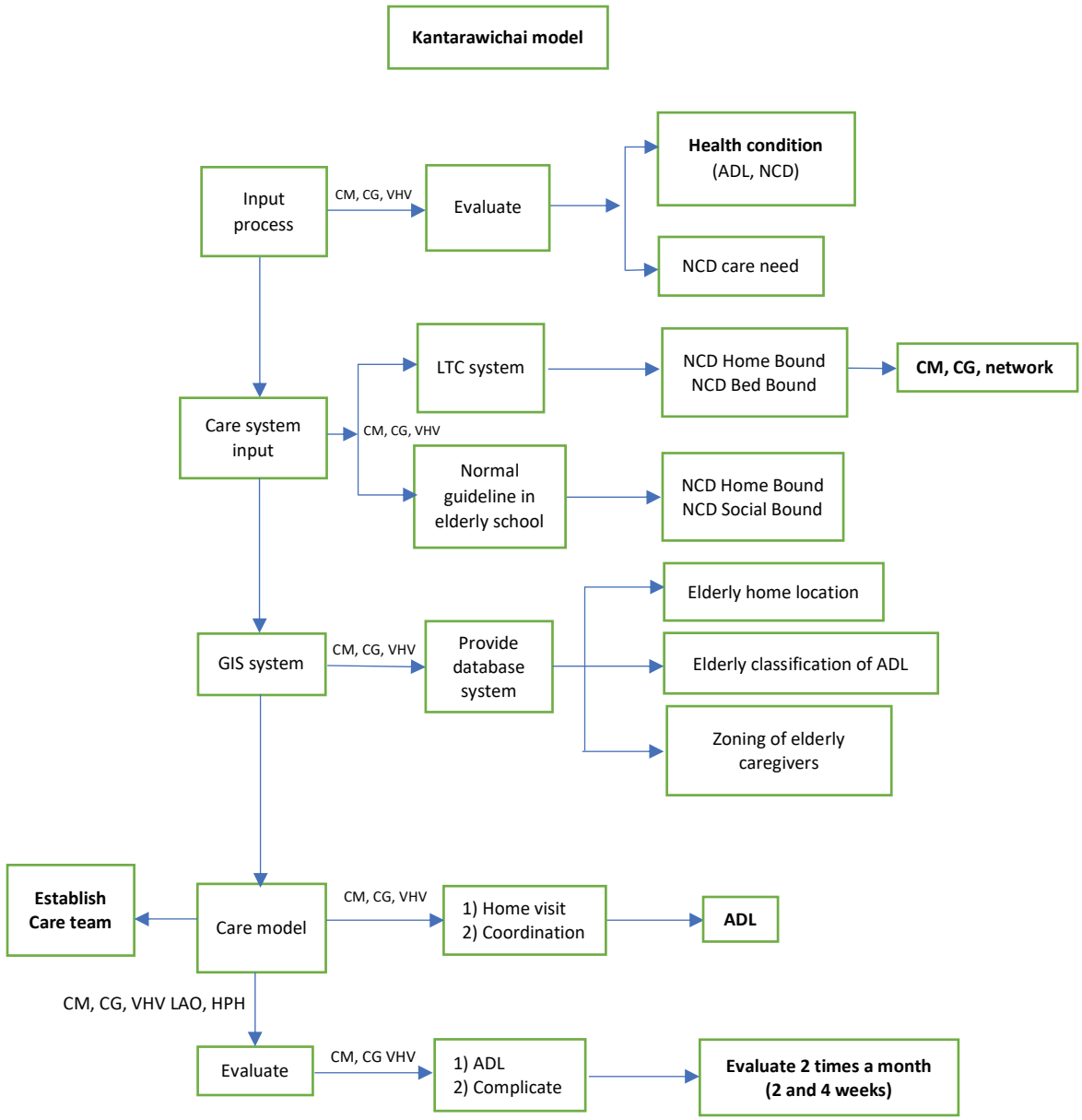


Figure 1. *Kantharawichai elderly dependency care model. CM: care manager, CG: care giver, VHV: village health volunteer, LAO: local administrative organization, HPH: health promoting hospital.*

Distribution of caregivers to support elderly dependency care

The analysis of caregiver distribution for elderly dependency care revealed a strong correlation between the proximity of caregivers' residences to elderly individuals' homes and the frequency of visits. Caregivers residing closer to elderly individuals provided care at a frequency surpassing the government service criteria by 100%. This finding underscores the critical role of geographic proximity in determining accessibility and quality of caregiving interactions. Visual representation of the data (Figure 2) further demonstrated that caregiver proximity significantly enhances care delivery effectiveness for the dependent elderly population. These findings highlighted geographic factors as pivotal in the design of effective caregiving models, particularly for elderly individuals with chronic conditions requiring sustained support. The results align with the overarching objective of this study—to develop improved long-term care strategies tailored to the needs of dependent elderly populations. Additionally, the importance of incorporating spatial considerations into policy and practice frameworks was to optimize caregiving efforts.

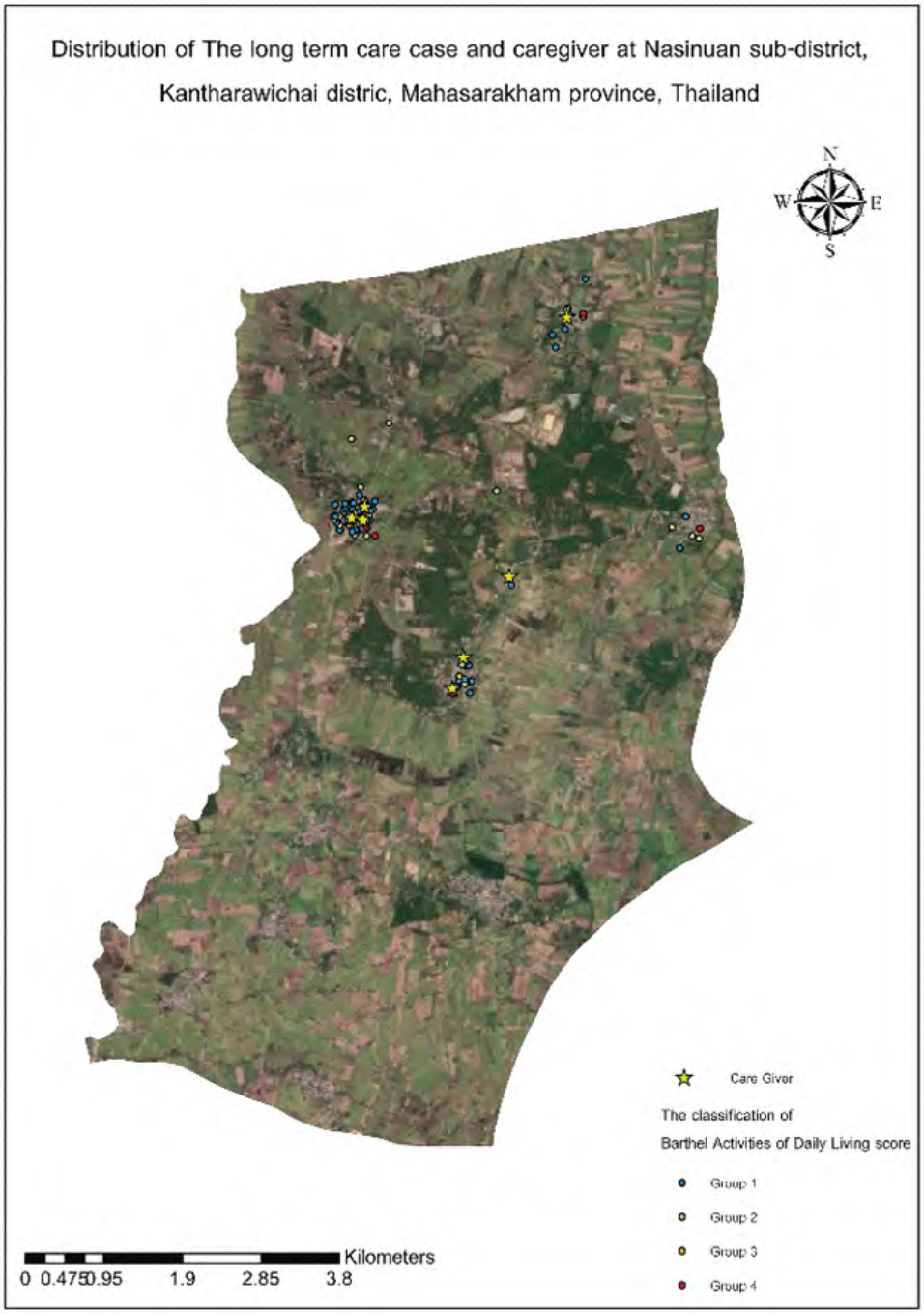


Figure 2. *Geographic proximity of caregivers' residences to elderly individuals' homes and dependency levels.*

Discussion

The current study aimed to develop and assess an integrated long-term care model for dependent elderly individuals with NCDs in Mahasarakham Province, Thailand. The findings suggested that the implemented care model led to significant improvements in the daily living abilities of the elderly participants. These results align with existing literature that demonstrated the efficacy of integrated care models in improving the quality of life for elderly individuals with chronic conditions. In the sometime, there have similarly shown that integrated care approaches, which combine healthcare services, social support, and community resources, contribute significantly to better health outcomes, reduced hospital admissions, and enhanced social engagement among the elderly population.[10,11] Such models provide continuous support and improve communication between caregivers, healthcare professionals, and elderly individuals, thereby fostering trust.[12] The positive outcomes in this study suggested that implementing and scaling such models across other regions could be beneficial, particularly in areas with high elderly populations and limited healthcare infrastructure. It also highlights the importance of local community involvement and caregiver proximity, which are crucial factors in maintaining consistent and accessible care delivery. Continued research and policy support should focus on scaling integrated models while considering community-specific geographic and socio-economic factors to optimize care delivery for elderly populations with NCDs.[13]

In comparison with existing literature, our study provided a more focused examination of elderly individuals living with NCDs, such as diabetes, hypertension, and asthma, while incorporating caregivers as a central component of the care delivery model. While studies have highlighted the importance of comprehensive care systems for elderly populations, our research delves deeper into the specific challenges of managing chronic diseases among dependent elderly individuals.[14,15] Notably, many previous studies have taken a broader approach to elderly care, often overlooking the intricate complexities of chronic disease management and the direct involvement of CG in these contexts. Our study distinguishes itself through the integration of a caregiver training program, which serves as a crucial factor in sustaining care delivery in community-based settings. This approach aligns with the findings that emphasized the necessity of equipping caregivers with the knowledge and skills to address the multifaceted health needs of elderly individuals living with chronic illnesses.[16] Such training programs ensured that caregivers can provide more effective, continuous, and personalized care, ultimately improving health outcomes and quality of life for dependent elderly populations. The model's applying community-based volunteers and leveraging GIS technology for care coordination, has been shown to enhance accessibility and efficiency in care delivery, which could be further expanded and adapted in other regions.

However, this study has several limitations. First, the sample size, although appropriate for the context of this research, is relatively small, limiting the ability to generalize findings on a larger scale. Furthermore, the study was conducted within a single geographical region, which may not fully represent the diverse elderly populations in other parts of Thailand or internationally. This limitation highlights the need for multi-center studies to validate the model's effectiveness across different settings. Additionally, while the model showed promising results, further research is needed to explore its long-term sustainability and the potential impact of ongoing caregiver training and support.

Moreover, while the study measured improvements in daily living abilities, it did not assess the cost-effectiveness of the model. This is an important factor to consider in evaluating the feasibility of scaling the model across other regions or countries. By addressing these limitations in future research endeavors, we can better refine the community-based models, develop scalable interventions, and ultimately enhance the well-being of elderly populations across Thailand and beyond.

In terms of further research, it would be beneficial to explore the impact of this care model on the mental health of elderly individuals, as well as the role of technology in enhancing care delivery. Studies focusing on the mental health outcomes of elderly individuals receiving integrated care have shown promising results,[17] and future research could integrate these aspects into the existing care model. Additionally, expanding to other regions of Thailand or internationally would provide a more comprehensive understanding of the model's scalability and adaptability in different socio-cultural contexts.

Finally, the Long-Term Care (LTC) model for dependent elderly in this study was designed to address their comprehensive care needs by integrating medical, social, and psychological support. This model emphasizes enhancing the ability of elderly individuals to perform activities of daily living (ADL), such as bathing, dressing, and mobility, which are essential for maintaining autonomy and reducing caregiver burden. Improving ADL capabilities is a critical outcome of the LTC model, as it directly correlates with the quality of life (QoL) of dependent elderly individuals. When ADL performance improves, elderly individuals experience greater independence, reduced stress, and enhanced social participation, leading to better physical and mental well-being.[18] Furthermore, the effectiveness of LTC models including caregiver support and community involvement ensures that care is sustainable, personalized, and responsive to the evolving needs of the elderly, and further enhancing their QoL.[19]

Conclusion

This study contributes valuable insights into the development and implementation of an integrated long-term care model for elderly individuals with chronic non-communicable diseases. The results demonstrates that the model can improve daily living abilities, offering

a promising solution for enhancing the quality of life for dependent elderly. Further research, including multi-center trials and cost-effectiveness analyses, is essential for determining the broader applicability and long-term viability of this care model.

Conflict of interest

There was no conflicts of interest in this study.

Funding

This study did not receive funding from any agency and has no conflicts of interest with any funding source.

Acknowledgements

We thank the Faculty of Medicine, Mahasarakham University, the Ministry of Public Health, and local health authorities for their support. Special appreciation goes to the public health volunteers and all contributors who facilitated this study.

References

[1] United Nations. World population prospects 2022: Summary tables [Internet]. 2022 [cited 2024 Dec 6]. Available from: https://www.un-ilibrary.org/content/books/9789210014380/read.

[2] World Health Organization. Global report on aging and health. Geneva: WHO; 2021.

[3] Department of Older Persons. Situation of the Thai Older Persons 2023. Bangkok: Amarin Corporations Public Company Limited; 2023.

[4] Foundation of Thai Gerontology Research and Development Institute. Situation of the Thai elderly 2020. Bangkok: FTGRDI; 2020.

[5] Ministry of Public Health. Health Region 7 Annual Report 2024. Bangkok: MOPH; 2024.

[6] Provincial Health Office of Maha Sarakham. Elderly population report 2024. Maha Sarakham: PHO; 2024.

[7] Deming WE. Out of the crisis. Cambridge: MIT Center for Advanced Educational Services; 1986.

[8] Department of Medical Services, Ministry of Public Health. Guidelines for Screening and Health Assessment of the Elderly, B.E. 2564 (2021). Nonthaburi: Ministry of Public Health; 2021.

[9] Pongratanakul R, Thitisakulchai P, Srinonprasert V, Siriussawakul A, Suraarunsumrit P, Dajpratham P. Development and Psychometric Evaluation of a Short-Form Barthel Index for Older Patients Undergoing Abdominal Surgery. ASEAN J Rehabil Med [internet]. 2023 Sep. 25 [cited 2025 May 9]; 33(3): 135. available from: https://he01.tci-thaijo.org/index.php/aseanjrm/article/view/262902.

[10] Chen TL, Feng YH, Kao SL, Lu JW, Loh CH. Impact of integrated health care on the elderly population: A systematic review of Taiwan's experience. Arch Gerontol Geriatr. 2022; 102: 1-9. doi: 10.1016/j.archger.2022.104746.

[11] Li Y, Luo L, Dong H. Delivering integrated community care for the elderly: A qualitative case study in Southern China. Int J Environ Res Public Health. 2024; 21(6): 1-17. doi: 10.3390/ijerph21060680.

[12] Heggdal K, Mendelsohn JB, Stepanian N, Oftedal BF, Larsen MH. Health-care professionals' assessment of a person-centred intervention to empower self-management and health across chronic illness: Qualitative findings from a process evaluation study. Health Expect. 2021; 24(6): 1367-77. doi: 10.1111/hex.13271.

[13] Liljas, AEM, et al. Impact of Integrated Care on Patient-Related Outcomes Among Older People – A Systematic Review. International Journal of Integrated Care, 2019; 19(3): 1-16. doi: 10.5334/ijic.4632.

[14] Pedersen AKB, Skinner MS, Sogstad M. Needs assessment in long-term care: Expression of national principles for priority setting in service allocation. BMC Health Serv Res. 2024; 24(530): 1-11. doi: 10.1186/s12913-024-10889-1.

[15] Takeda C, Guyonnet S, Sumi Y, Vellas B, Araujo de Carvalho I. Integrated care for older people and the implementation in the INSPIRE care cohort. J Prev Alzheimers Dis. 2020; 7(2): 70-4. doi: 10.14283/jpad.2020.8.

[16] Fulmer T, Mate KS, Berman A. The age-friendly health system imperative. J Am Geriatr Soc. 2018; 66(1): 22-4. doi: 10.1111/jgs.15076.

[17] Chang KT. Introduction to geographic information systems. New York: McGraw-Hill; 2008.

[18] World Health Organization. World report on ageing and health. WHO Press; 2015. Available from: https://www.who.int/publications/i/item/9789241565042.

[19] Kane RL, Kane RA. Assessing older persons: Measures, meaning, and practical applications. Oxford University Press; 2018.

Journal of Associated Medical Sciences 2025; 58(3): 52-59

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 effect of low voltage high frequency electric pulses on the extracellular conductivity, cell permeability, and time-depended manner of MCF7 cell line

Zeinab Shankayi[1,2], S. Mohammad Firoozabadi[1*], Mahmoud Ashouri Mendi[3]

[1]Department of Medical Physics, Tarbiat Modares University, Tehran, Iran.
[2]Neuroscience Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran.
[3]Students' Research Committee, Baqiyatallah University of Medical Sciences, Tehran, Iran.

ARTICLE INFO

Article history:
Received 13 January 2025
Accepted as revised 10 May 2025
Available online 27 May 2025

Keywords:
Low voltage, high frequency, chemotherapy drug, electroporation, quenching.

ABSTRACT

Background: Low voltage, high frequency electrochemotherapy (LVHF ECT) has recently been explored as a method to enhance the permeability of cell membranes to non-permanent chemotherapeutic agents.

Objective: Despite recent advances, it remains unclear whether classical ECT and LVHF ECT (using 50–150 V/cm at pulse frequencies of 4–6 kHz) affect the cell membrane through similar mechanisms.

Materials and methods: We investigated the efficiency of reversible membrane permeabilization in the MCF7 cell line induced by LVHF electric pulses. Specifically, we examined changes in extracellular conductivity, the time-dependent nature of permeabilization, and the effects of this protocol on commonly used permeabilization markers.

Results: LVHF ECT protocols significantly increased the conductivity of the extracellular medium, indicating enhanced membrane permeability in MCF7 cells. This increased permeability was closely associated with elevated membrane conductivity. Notably, most of the membrane permeabilization occurred during pulse application and subsided within one minute after the delivery of LVHF pulses. Experimental data indicate that these electric pulses induce the formation of short-lived pores in the membrane. Furthermore, LVHF pulses did not alter the cytotoxicity of bleomycin; however, this protocol resulted in the quenching of Lucifer yellow fluorescence, a classical marker for membrane permeabilization. These findings suggest that bleomycin is a reliable marker for cell electro-permeabilization under LVHF ECT conditions.

Conclusion: Our results demonstrate that LVHF ECT induces transient, short-lived pores in the cell membrane and increases membrane permeability without affecting bleomycin cytotoxicity. Bleomycin appears to be a suitable marker for assessing electro-permeabilization in this context.

Introduction

Cell poration (CP) relates to the electrical increases of membrane permeability of target cells to the molecules, specifically pharmaceutical compounds and genes.[1-3] This technique, called Electroporation (EP), demonstrates that the theory of the electrical breakdown of cell membrane and pore formation is the most widely accepted.[4,5]

The reversible phenomenon permits the increase of cell membrane permeability and cell suspension conductivity for a limited time.[6,7] During this time of increased membrane permeability, if the high toxicity and non-permeable chemotherapy drug – such as bleomycin– is achievable, then it can enter the cell cytoplasm and

* **Corresponding contributor.**
Author's Address: *Department of Medical Physics, Tarbiat Modares University, Tehran, Iran.*
E-mail address: smohammadfiroozabadi24@chmail.ir
doi: 10.12982/JAMS.2025.072
E-ISSN: 2539-6056

target the intracellular components. This physical process is commonly referred to as Electrochemotherapy (ECT).[7-10] Classical ECT is obtained with a train of eight pulses of 1 Hz or 5 kHz frequency and 1300V/cm amplitude. Such pulses have been successfully applied to treatments of superficial tumors in animals and humans.[8-10] Recently, we focused on the effect of much smaller amplitude pulses. Such protocols are referred to as Low Voltage and High Frequency ECT (LVHF ECT). Permeabilization of the cell membrane and treatment of animal nodules that were induced by LVHF pulses have been reported.[11-15] Whether classical ECT and LVHF ECT act on the membrane in a similar way is unclear. To clarify this ambiguity, we consider the effect of Low Voltage and High Frequency (LVHF) electric pulses on the extra-cellular conductivity, cell permeability, and role of time on cell permeabilization.

The disruption of the cell membrane and its permeability can be indirectly demonstrated using different types of markers added to the cell culture mediums. Many of these markers – such as fluorescence, radiolabels, and bleomycin – have been used in permeabilization research.[13,14,16-18] Bleomycin is a non-permeant molecule which possesses a very high intrinsic cytotoxicity. In vitro studies indicate that a low therapeutic dose of bleomycin is sufficient to generate the breakage of the DNA double-strand and kill the cell by a mitotic and apoptotic cell death process.[17,19,20] Due to previous studies revealing that bleomycin molecules are very sensitive cells to use as a permeabilization marker, we exhibited bleomycin's cytotoxicity effects on MCF7 cells exposed to LVHF electric pulses, and suggested the best LVHF ECT protocol.

In addition, some researchers have studied the changes in the electrical properties of cell suspensions due to EP.[6,7,21,22] These studies concluded that cell suspension conductivity increased, an effect that was observed only for above-the-threshold electric fields.[7,21] Recently, it was suggested that measurements of conductivity could enable observations of cell electroporation.[6,21,23,24] In the current study, we will present results based on the degree of reversible membrane permeabilization as a function of the conductivity and composition of the external medium over a wide range of treatment parameters.

Materials and methods

Cell Line

The adenocarcinoma cell line (MCF7 cell line), found in human breasts, was grown in RPMI containing 10% fetal bovine serum, 160 µg/ml L-glutamine (all from Invitrogen, GIBCO, USA), 100 units/mL penicillin, and 16 µg/mg gentamicin, incubated in 5% CO2 at 37 °C.

Electric pulse exposure

The process of directing electric pulses towards the cells using an ECT-SBDC (designed and made in the Small Business Development Center and Electromagnetic Laboratory of the Medical Physics Department of Tarbiat Modares University, Tehran, Iran) has been described in detail in previous articles.[13,15] The suspended cells were placed between two parallel gold-plated electrodes and LV-HF pulses were applied. 4, 5, and 6 kHz with 4000 electric pulses for 100-µs durations and the high voltage electric pulses were 1000 V/cm in repeated pulse frequencies of 0.001 and 5 with 8 electric pulses for 100-µs durations.

Medium conductivity measurement

After trypsinization and inactivation of trypsin (Bio Idea Group, Tehran, Iran) by the serum factors of the complete medium, the cells were centrifuged for 5 minutes at 500 rpm and resuspended at a density of 500×10^6 cells/mL in RPMI (Invitrogen, GIBCO, USA). 300 µL of the mixture was deposited between the two electrodes and subjected to the selected electric treatment (1: 60 v/cm and 5 kHz repetition frequency, 2: 70 v/cm and 4 kHz repetition frequency, 3: 70 v/cm and 5 kHz repetition frequency, 6:70 v/cm and 6 kHz repetition frequency). A conductometer (CyberScan CON 6000, EUTECH Instruments) was used to measure the conductivity of the cell media. The mean value for all the parameters was calculated from at least three of the measurements.

Time effect

To study the extent of LVHF ECT-induced uptake of bleomycin as a function of time interval after or before exposure of cells to LVHF electric pulses, the chemotherapy drug at 0.1 µM and 1 µM was either present during the electric treatment or added to the cell suspension at various times after the treatment from 0-32 minutes. For determination of cell permeability, the cells were incubated for 48 hours using the same protocols as described before.[25]

The effect of LVHF electric pulses on the permeabilization marker

In the current study, two markers were used. The first marker used bleomycin, the electric pulses (4000 pulses with 70 V/cm and 5 kHz frequency) was applied to bleomycin and immediately this suspension was added to MCF7 cells. The second marker used was Lucifer yellow (Sigma-Aldrich Life Science, USA) which was diluted in phosphate-buffered saline (PBS) with a concentration of 500 µM. The fluorescence emission was measured offline in arbitrary units on a spectrofluorometer (Shimadzu RF-5000, Japan) 40 minutes after the exposure of the LY to the electric pulses. The excitation and emission wavelengths were set at 418 nm and 525 nm, respectively.

Statistical analysis

All results are given as an average of more than three times and are represented in bar graphs. Vertical bars represent the standard deviation of the mean. Statistical analyses were performed using SPSS for windows 16.0 (SPSS Inc., Polar Engineering and Consulting). All data were tested for normality. One-way ANOVA, followed by LSD, was performed and statistical difference analysis was accomplished by a test. The p,0.05 were considered significant for rejection of the null hypothesis.

Results

Medium conductivity

The conductivity changes after low voltage and high frequency for selected protocols compare to standard protocols are displayed in Figure 1. The media conductivity increased considerably for LVHF ECT with 70 V/cm at 4 and 5 kHz frequency. However, based on the synergistic effect and cell suspension connectivity, we found the 70 V/cm amplitude and 5 kHz frequency with 4000 square wave electric pulses to be the best LVHF ECT protocol.

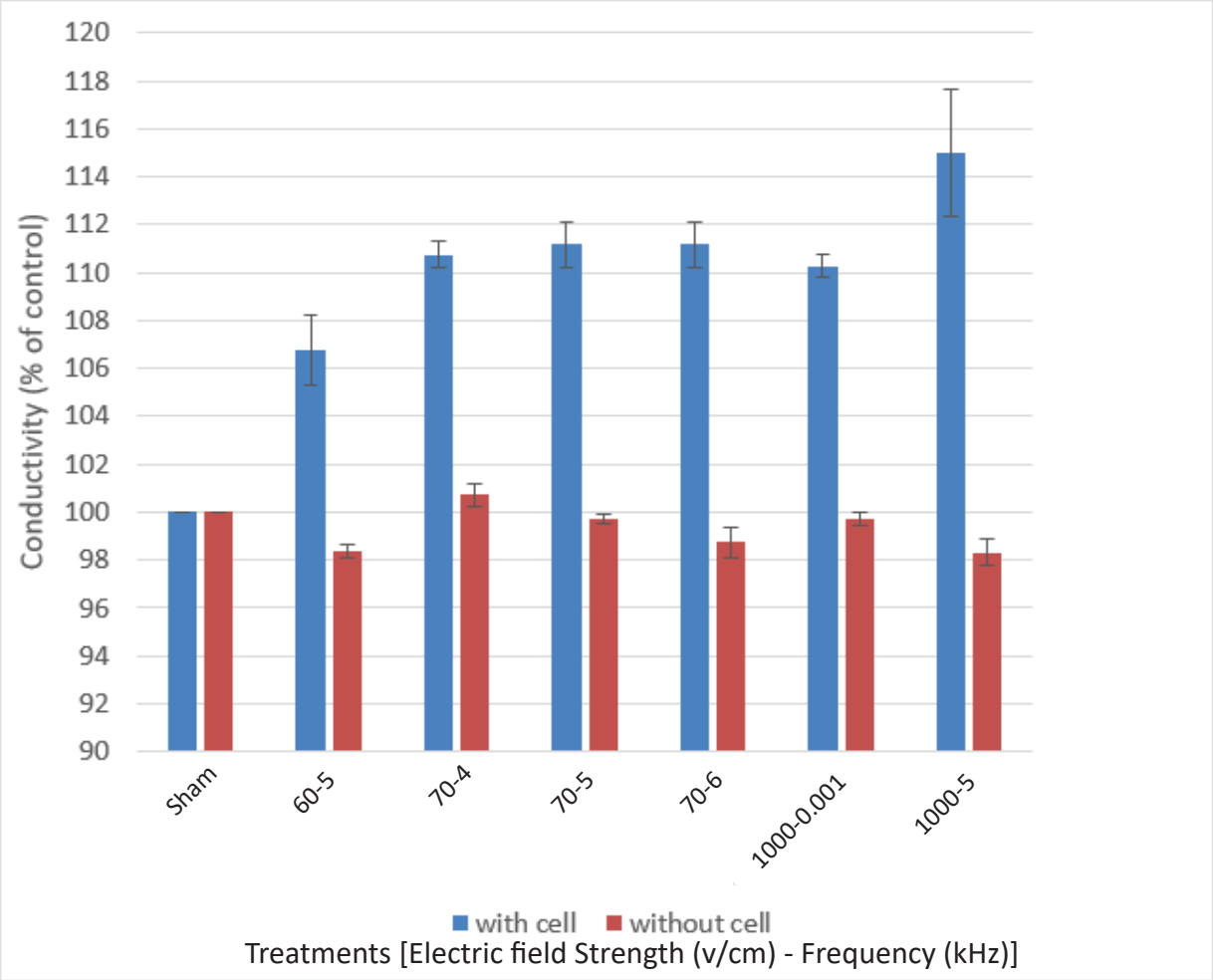


Figure 1*. Conductivity of the cell media changes of attached MCF7 cells to 4 different electric pulse protocols exposures. 1: 60 V/cm with 5 kHz repeated frequency (6-5), 2: 70 V/cm with 4 kHz repeated frequency (70-4), 3: 70 V/cm with 5 kHz repeated frequency (70-5), 4: 70 V/cm with 6 kHz repeated frequency (70-6). The Results are presented as mean±SD. Attributed number to the control group was chosen as 100 and the connectivity of other groups was computed as the percentage of control connectivity.*

Time Effect

To assess the possible reversible pore induction during the pulsation, bleomycin was added to the cell suspension at different times either before or after pulsing. Figure 2 shows that higher uptakes resulted when the chemotherapy drug was added to the suspension before or during the pulsation time. This suggests that a permeabilization occurred during the pulse delivery and the effect disappeared 1 minute after the end of the electric pulse application.

A comparison was made between the smaller and larger doses of the bleomycin. Figure 2 demonstrates that the concentration correlates with the bleomycin uptake only if the bleomycin was added to the cell suspension before or during the pulsation.

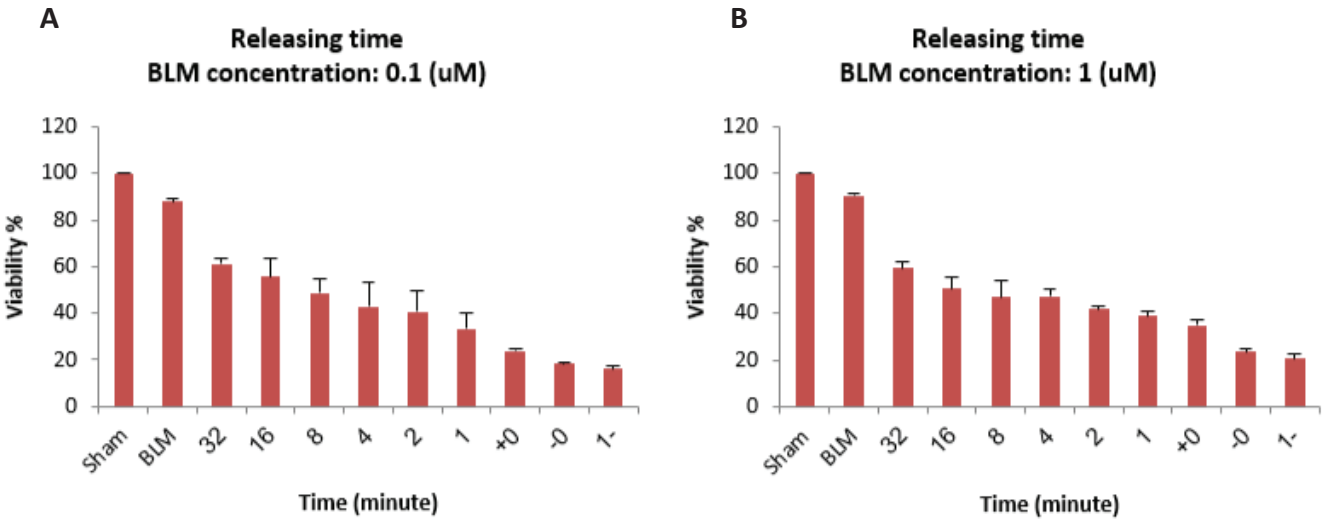


Figure 2*. The temporal decay kinetics of low-voltage, high-frequency electric field-induced bleomycin permeabilization were analyzed. Data are expressed as mean±SD, with control group viability normalized to 100%; all other groups were calculated as a percentage of this baseline. Experimental Groups consist of: Groups 1-32: Bleomycin administration occurred 1-32 minutes post-pulsing after the final electric pulse. Groups 0^- and 0^+: Bleomycin was added immediately before (0^-) or after (0^+) electric pulse delivery. Group -1: Bleomycin was introduced to cell suspensions 1 minute before electroporation.*

The effect of LVHF electric pulses on the permeabilization marker

Our results from the previous part of this research lead us to assume that marker-property changes occurred with electric pulse exposure during the pulsation. Therefore, we set a different experiment. This time, we applied the electric pulses directly to the markers and measured their cytotoxicity and fluorescence intensity.

The changes were not observed between the cytotoxicity of normal bleomycin and that of the bleomycin exposed to the electrical pulses (Figure 3A). Figure 3B indicates the fluorescence intensity for best LVHF ECT protocols (70 V/cm amplitude, 5 kHz frequency with 4000 square wave electric pulses). The comparison between the electric-field-induced fluorescence intensity and the control group is indicative of an enhanced quenching of the fluorescence.

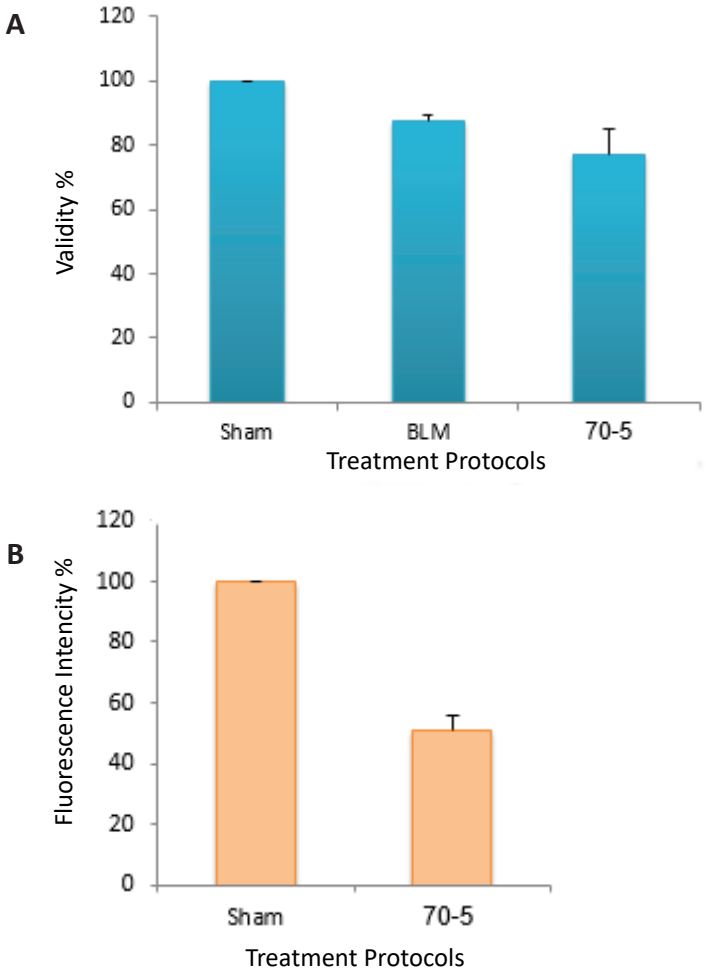


Figure 3. *The effect of pulsed electric field.* ***A:*** *the cytotoxicity of bleomycin. the electric pulses (4000 pulses with 70 V/cm and 5 kHz frequency) were applied to bleomycin and immediately this suspension was added to MCF7 cells and viability was evaluated by an MTT assay,* ***B:*** *Fluorescence intensity of Lucifer Yellow (LY) versus 70 v/cm and 5 kHz electric pulses. The fluorescence emission was measured offline in arbitrary units on a spectrofluorometric 40 minutes after the exposure of the LY to the electric pulses. Results are presented as mean±SD.*

Discussion

As shown in previous studies, electrochemotherapy using low intensity and high frequency increases cell permeability and cures tumors.[13,25,26] Also, in previous studies, we suggested LVHF ECT to reduce the side effect of classical ECT. Using low voltage electric field and high frequencies decrease the number of muscle contractions and reduce the patient's pain during the treatment session.[11-15] This kind of ECT has been shown to be efficient in treatment of animal and human tumors.[11,15,25,26] However, its mechanism is unclear. Theoretical study predicted that cell membrane electro-permeabilization occurred when strong enough electric pulses (more than 300-400 V/cm) are applied and the transmembrane voltage reached to an above threshold transmembrane membrane potential (between 0.2-1.0 V).[6,27] This high electric pulse induced transient aqueous pores, thus enabling the transfer of molecules into cells.[23,28] But, given the results of the in vitro uptake experiments, the membrane permeabilization threshold of cells was less than critical electric pulses amplitude.

Cell suspension conductivity is a physical permeability marker which we confirmed in the current study. We found that our four selected protocols increased media conductivity. These results are in accordance with some previous research results indicating that the cell membrane permeabilization efficiency positively correlated with external conductivity during the application of high-voltage pulses.[28, 29] This increased permeability of a cell membrane was accompanied by the increased membrane conductivity. When the conductivity increased due to electric pulse applying procedure, the membrane partially discharged through the pores as a conductive pathway. This increased membrane conductivity consequently reduced the transmembrane voltage at the regions where electroporation occurred and was accompanied by permeability of a cell membrane. Therefore, our results implied that LVHF electric pulses may extend cell permeabilization immediately after exposure.[6,7,28,29] According to our data, high-intensity ECT was more effective for conductivity than low-intensity ECT, whereas protocols with the same intensity and different frequency revealed

that the medium connectivity may be dependent on cell viability. Dependence of the cell suspension conductivity on the electric field intensity is in accordance with some of the results obtained by previous groups.[23,29] Moreover, results of various studies suggest that concentrations of viable or dead cells after pulsing can influence medium conductivity. Therefore, our results are in agreement with the conclusions of prior literatures.[7,21]

As we noted earlier, measurement of cell medium conductivity is an indirect and convenient method to show cell membrane pore formation.[7,21] Theoretical description predicts that when the pores form, conductivity increased.[21,23,30] Based on this theory and our findings in the previous section, we evaluated the effect of time on cell electro-permeabilization. Pore dynamic theory of electroporation explained that the pore formed in a microsecond, but complete resealing of transient pores occurs in the range of minutes, depending on the electrical parameters used.[31-34] Our observations revealed that the majority of cell membrane permeability occurred during electric pulse application but ended within 1 minute after the delivery of LVHF electric pulses. These results agree with pore dynamic theory.[31-33] Indeed, there are two possible ways of describing the LVHF ECT electro-permeabilization. One way is by the pores formation, which our results support this way very well. The LVHF electric pulses formed a pore with a short life span in the cell membrane which allowed cellular uptake during the electric pulse application. However, pore dynamic theory predicts that small pores grow and expand during pulsation and can survive for a longer period of time after the electroporation.[32,35] If only small pores formed, then the addition of different doses of bleomycin behaved as a function of time, demonstrating that smaller doses killed fewer cells and larger doses killed more cells. Interestingly, between 0.1 µM and 1. 0µM bleomycin, there was a significant difference just when the chemotherapy drug was added to the suspension before or during the pulsation. Therefore, we think this result will require another mechanism. The second way which researcher introduced for electro-permeabilization, is by electro-endocytosis hypothesis. Electro-endocytosis is an endocytosis-like process which induces the cells via low-voltage electric pulses. This process is much longer than the electroporation process and describes a mechanism for uptake of macromolecules or DNA into the cells.[13,24,32,36] Previous studies have reported that plasmid must be added to cells before the electric pulse induction.[32,37] But further studies are needed to maintain the role of electro-endocytosis on the long time electro-permeabilization.

In the current study, we tried to determine the most probable mechanism of LVHF ECT electro-permeabilization. Hence, we examined the effect of LVHF electric pulses on the bleomycin cytotoxicity because the highest number of dead cells was obtained when the bleomycin was added to the cell suspension during or before pulsation. This effect may be due to increased cellular uptake of BLM by electro-permeabilization, or bleomycin cytotoxicity changed during the pulsation. As previously mentioned, the fluorescence spectrum used as a cell membrane permeabilization marker (LY) decreased with the application of an external electric field.[16] Therefore, to clarify the mechanism of LVHF electro chemotherapy, we applied the electric pulses directly to the permeability markers, LY and bleomycin. Our experimental data showed that the fluorescence spectra were reduced after the exposure of the LVHF electric field. This finding is in line with our recent report demonstrating that fluorescence spectra are quenched by electric fields.[16] This is because the electric pulses created oxygen ion radicals which acted as collisional-quencher candidates. Finally, experimental data shown in Figure 3 demonstrated that the changes were not observed between the cytotoxicity of normal bleomycin and that of the bleomycin exposed to the LVHF electric pulses. This result indicated that the electric field has no effect on BLM cytotoxicity and that LVHF electric pulses can induce desirable cell electro-permeabilization.

Conclusion

In the current study, we concluded that LVHF electric pulses could induce electro-permeabilization. The second conclusion that we reached was that the influence of medium conductivity was detected. By measuring electric conductivity and time effect to transport, we suggested that pores with a short life span expanded during the LVHF electric pulse application. Our experimental results exhibit those additional mechanisms, such as the electro-endocytosis process, could be applied to increase molecular uptake during the pulsation.

Ethical approval

The study is approved by ethics committee of Baqiyatallah University of Medical Sciences (IR.BMSU.REC.1402.043).

Acknowledgements

This study was supported by the Shenasa Venture Capital (Pishgaman Amin Sarmayeh Pasargad) and Tarbiat Modares University.

References

[1] Blagus T, Markelc B, Cemazar M, Kosjek T, Preat V, Miklavcic D, et al. In vivo real-time monitoring system of electroporation mediated control of transdermal and topical drug delivery. J Control Rel. 2013; 172: 862-71. doi: 10.1016/j.jconrel.2013.09.030.

[2] Mir LM, Calvet CY, Andre FM, editors. Therapeutic effects of in vivo electroporation: Facilitating drug and gene delivery but not only&#x2026. General Assembly and Scientific Symposium (URSI GASS), 2014 XXXIth URSI; 2014 16-23 Aug. 2014.

[3] Miklavcic D, Kotnik T, Electroporation for electro-chemotherapy and gene therapy. in: PJ. Rosch, MS Markov, Editors. Bioelectromagnetic Medicine, University of Ljubljana, Marcel Dekker, 2004, pp. 637-56.

[4] Chen C, Smye SW, Robinson MP, Evans JA. Membrane electroporation theories: A review. Med Biol Engin Comput. 2006; 44: 5-14. doi: 10.1007/s11517-005-0020-2.

[5] Kotnik T, Kramar P, Pucihar G, Miklavcic D, Tarek M. Cell membrane electroporation- Part 1: The phenomenon. IEEE Electric Insulat Mag. 2012; 28: 14-23. doi: 10.1109/mei.2012.6268438.

[6] Pavlin M, Kandušer M, Reberšek M, Pucihar G, Hart FX, Magjarevićcacute R, et al. Effect of cell electroporation on the conductivity of a cell suspension. Biophysic J. 2005; 88: 4378-90. doi: 10.1529/biophysj.104.048975.

[7] Silve A, Leray I, Poignard C, Mir LM. Impact of external medium conductivity on cell membrane electropermeabilization by microsecond and nanosecond electric pulses. Sci Report. 2016; 6: 19957. doi: 10.1038/srep19957.

[8] Miklavčič D, Pucihar G, Pavlovec M, Ribarič S, Mali M, Maček-Lebar A, et al. The effect of high frequency electric pulses on muscle contractions and antitumor efficiency in vivo for a potential use in clinical electrochemotherapy. Bioelectrochem. 2005; 65: 121-8. doi: 10.1016/j.bioelechem.2004.07.004.

[9] Županič A, Ribarič S, Miklavčič D. Increasing the repetition frequency of electric pulse delivery reducesunpleasant sensations that occur in electrochemotherapy. Neoplasm 2007; 54: 246-50.

[10] Sersa G, Kranjc S, Scancar J, Krzan M, Cemazar M. Electrochemotherapy of mouse sarcoma tumors using electric pulse trains with repetition frequencies of 1 Hz and 5 kHz. J Membrane. 2010; 236: 155-62. doi: 10.1007/s00232-010-9268-z.

[11] Shankayi Z, Firoozabadi SM. Antitumor efficiency of electrochemotherapy by high and low frequencies and repetitive therapy in the treatment of invasive ductal carcinoma in Balb/c mice. Cell J. 2012; 14(2): 110-5. PMID: 23508227, PMCID: PMC3584427.

[12] Shankayi Z, Firoozabadi SMP, Saraf Hassan Z. Comparison of low voltage amplitude electrochemotherapy with 1 Hz and 5 kHz frequency in volume reduction of mouse mammary tumor in Balb/c Mice. Koomesh. 2012; 13: 486-90.

[13] Shankayi Z, Firoozabadi S, Hassan ZS. Optimization of electric pulse amplitude and frequency in vitro for low voltage and high frequency electrochemotherapy. J Membrane Biol. 2014; 247: 147-54. doi: 10.1007/s00232-013-9617-9.

[14] Shankayi Z, Firoozabadi SMP, Saraf HZ. The Endothelial Permeability Increased by Low Voltage and High Frequency Electroporation. J Biomed Physic Engineer. 2013; 3: 87.

[15] Shankayi Z, Firoozabadi S. Tumor growth inhibited by low-voltage amplitude and 5-kHz frequency electrochemotherapy. J Membrane Biol. 2011; 244: 121-8. doi: 10.1007/s00232-011-9405-3.

[16] Tahereh Pourmirjafari Firoozabadi ZS, Azam Izadi, S. Mohammad P. Firoozabadi. Can lucifer yellow indicate correct permeability of biological cell membrane under an electric and magnetic field? Cell J. 2015; 16: 560-3. doi: 10.22074/cellj.2015.501.

[17] Silve A, Leray I, Mir LM. Demonstration of cell membrane permeabilization to medium-sized molecules caused by a single 10ns electric pulse. Bioelectrochem. 2012; 87: 260-4. doi: 10.1016/j.bioelechem.2011.10.002.

[18] Jaroszeski MJ, Gilbert R, Heller R. Electrochemotherapy: an emerging drug delivery method for the treatment of cancer. Advanc Drug Deliver Rev. 1997; 26: 185-97. doi.org/10.1016/S0169-409X(97)00034-3.

[19] Chen J, Stubbe J. Bleomycins: towards better therapeutics. Natur Rev Cancer. 2005; 5: 102-12. doi: 10.1038/nrc1547.

[20] Silve A, Mir LM. Cell electropermeabilization and cellular uptake of small molecules: the electrochemo therapy concept. Clin Aspect Electroporat: Springer; 2011, pp. 69-82.

[21] Pucihar G, Kotnik T, Kandušer M, Miklavčič D. The influence of medium conductivity on electropermeabilization and survival of cells in vitro. Bioelectrochem. 2001; 54: 107-15. doi: 10.1016/S1567-5394(01)00117-7.

[22] Ivorra A, Villemejane J, Mir LM. Electrical modeling of the influence of medium conductivity on electroporation. Physic Chem Chem Physic. 2010; 12: 10055-64. doi: 10.1039/c004419a.

[23] Pavlin M, Miklavčič D. Theoretical and experimental analysis of conductivity, ion diffusion and molecular transport during cell electroporation—relation between short-lived and long-lived pores. Bioelectro chem. 2008; 74: 38-46. doi: 10.1016/j.bioelechem.2008.04.016.

[24] Shankayi Z, Firoozabadi S, Mansurian MG. The effect of pulsed magnetic field on the molecular uptake and medium conductivity of leukemia cell. Cell Biochem Biophysic. 2013; 65: 211-6. doi: 10.1007/s12013-012-9422-6.

[25] Safyari M, Firoozabadi SM, Hassan ZS, Akbari H, Shankayi Z. Optimization of Low Voltage and High Frequency in vitro and in vivo for clinical application. Life Sci Student J. 2023; 1: 89-100.

[26] Mofid B, Shankayi Z, Novin K, Dehghani S, Shankayi M, Haghighatkhah H, et al. Effective treatment of cervical lymph node metastasis of breast cancer by low voltage high-frequency electrochemotherapy. Act Medic Iran. 2017: 55(4):268-271. PMID: 28532140

[27] Neu WK, Neu JC. Theory of Electroporation. In: Efimov IR, Kroll MW, Tchou PJ, (Eds). Cardiac Bioelectric Therapy: Mechanisms and Practical Implications. Boston, MA: Springer US; 2009. pp. 133-61.

[28] Rems L, Miklavčič D. Tutorial: Electroporation of cells in complex materials and tissue. J Appli Physic. 2016; 119: 201101. doi.org/10.1063/1.4949264.

[29] Suzuki DO, Ramos A, Ribeiro MC, Cazarolli LH, Silva FR, Leite LD, et al. Theoretical and experimental analysis of electroporated membrane conductance in cell suspension. IEEE Transact Biomedic Engineer. 2011; 58: 3310-8. doi: 10.1109/TBME.2010.2103074.

[30] Pavlin M, Leben V, Miklavčič D. Electroporation in dense cell suspension—Theoretical and experimental analysis of ion diffusion and cell permeabilization. Biochimic Biophysic Act. 2007; 1770: 12-23. doi: 10.1016/j.bbagen.2006.06.014.

[31] Gissel H, Lee RC, Gehl J. Electroporation and Cellular Physiology. In: Kee ST, Gehl J, Lee EW, Editors. Clin Aspect Electroporat. New York, NY: Springer New York; 2011, pp. 9-17.

[32] Zaharoff DA, Henshaw JW, Mossop B, Yuan F. Mechanistic analysis of electroporation-induced cellular uptake of macromolecules. Experiment Biol Med. 2008; 233: 94-105. doi: 10.3181/0704-RM-113.

[33] Prausnitz MR, Milano CD, Gimm JA, Langer R, Weaver JC. Quantitative study of molecular transport due to electroporation: uptake of bovine serum albumin by erythrocyte ghosts. Biophysic J. 1994; 66: 1522. doi: 10.1016/S0006-3495(94)80943-9.

[34] Ji Z, Kennedy SM, Booske JH, Hagness SC. Experimental studies of persistent poration dynamics of cell membranes induced by electric pulses. IEEE Transact Plasma Sci. 2006; 34: 1416-24. doi: 10.1109/TPS.2006.877250.

[35] Bier M, Hammer SM, Canaday DJ, Lee RC. Kinetics of sealing for transient electropores in isolated mammalian skeletal muscle cells. Bioelectromagnet. 1999; 20: 194-201. doi: 10.1002/(SICI)1521-186X(1999)20:3<194::AID-BEM6>3.0.CO;2-0.

[36] Yadegari-Dehkordi S, Firoozabadi SM, Moghadam MF, Shankayi Z. Role of endocytosis pathways in electropermeablization of MCF7 cells using low voltage and high frequency electrochemotherapy. Cell J. 2021; 23: 445-50. doi: 10.22074/cellj.2021.7203

[37] Pakhomov AG, Miklavcic D, Markov MS. Advanced electroporation techniques in biology and medicine: CRC Press; 2010.

Journal of Associated Medical Sciences 2025; 58(3): 60-64

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 hepatitis B virus infection among pregnant women attending prenatal care at Mae Sai Hospital in Thailand: A retrospective cross-sectional study

Nisarat Saowaros[1†], Phirada Mukphet[1†], Sunarin Inthra[2], Sittiporn Suwannamit[3], Krung Phiwpan[3*]

[1]Student program in Medical Technology, Department of Medical Technology, School of Allied Health Sciences, University of Phayao, Phayao Province, Thailand.
[2]Medical Technology Laboratory, Mae Sai Hospital, Chiang rai Province, Thailand.
[3]Department of Medical Technology, School of Allied Health Sciences, University of Phayao, Phayao Province, Thailand.

ARTICLE INFO

Article history:
Received 25 March 2025
Accepted as revised 21 May 2025
Available online 27 May 2025

Keywords:
Prevalence, hepatitis B virus, pregnancy.

ABSTRACT

Background: Hepatitis B virus (HBV) infection is a serious public health problem worldwide and a major cause of liver diseases. Pregnant women with HBV infection are a reservoir for the virus and can transmit the infection to their children. Most infants (90%) infected with HBV develop chronic infection and 25% die prematurely from liver cancer or cirrhosis.

Objective: To assess the prevalence of HBV infection among pregnant women attending prenatal care at Mae Sai Hospital, a border region in Northern Thailand.

Materials and methods: This retrospective cross-sectional research study collected data from the hospital records database between January 2019 and September 2023. All pregnant women were tested for serum HBsAg marker by the STANDARD Q HBsAg Test. Statistical analyses were conducted using IBM SPSS version 29.0, with logistic regression used to detect the associations between variables at a significance level of $p<0.05$.

Results: The study population comprised 2962 pregnant women, with 29.98% Thai and 70.02% non-Thai divided into Burmese (48.12%) and other ethnic groups (51.88%). The overall prevalence rate of HBV infection among pregnant women was 4.9% (95%CI=4.1-5.7). The HBV prevalence rates of Thai pregnant women and non-Thai pregnant women were 2.9% (95%CI=1.8-4.2) and 5.7% (95%CI=4.7-6.9), respectively. Two variables (nationality and year of birth) were associated with hepatitis B infection. Non-Thai pregnant women had a 2-fold (95%CI=1.998-3.161, $p=0.003$) greater chance of HBV infection than Thai pregnant women, while Thai pregnant women born before the implementation of the national vaccination program had a 4.2-fold (95%CI=1.757-10.037, $p=0.004$) greater chance of HBV infection than those born after the program implementation.

Conclusion: Results showed a lower prevalence of hepatitis B virus infection in pregnant women compared to previous studies. Nationality and vaccine implementation were identified as associated factors with hepatitis B infection. Raising awareness, promoting preventive measures, and implementing screening before pregnancy planning, or receiving the HBV vaccine to prevent infection during pregnancy are essential to stop the transmission of the hepatitis B virus to newborns.

Introduction

Hepatitis B virus (HBV) is an emerging global health problem as either acute or chronic infection. The World Health Organization (WHO) 2024 Global Hepatitis Report stated that the number of deaths from viral hepatitis is increasing. Worldwide, 254 million people live with hepatitis B, while only 3% (7 million people) living with chronic hepatitis B infection had received antiviral therapy

† Co-first author
* Corresponding contributor.
Author's Address: *Department of Medical Technology, School of Allied Health Sciences, University of Phayao, Phayao Province, Thailand.*
E-mail address: krung.ph@up.ac.th
doi: 10.12982/JAMS.2025.073
E-ISSN: 2539-6056

at the end of 2022.[1] Those living with chronic hepatitis B have a high risk of progression in the long term to diseases such as chronic hepatitis, liver cirrhosis, hepatocellular carcinoma, and even death.[2] Chronic hepatitis B (CHB) is defined by the presence of detectable HBsAg in the blood for longer than six months.[3] Age of exposure is a key factor in determining the risk of chronic infection. The probability of developing chronic hepatitis B infection during the perinatal period is high at 70 to 90% compared to 5 to 10% for those acquiring infection during adulthood.[4]

The risk of HBV transmission to newborn infants is 70 to 90% due to the simultaneous positivity of mothers with HBsAg and HBeAg seromarkers. By contrast, there is a 10 to 40% risk of HBV transmission if mothers test positive for only HBsAg.[5] Therefore, pregnant women attending antenatal care should be screened for HBsAg and Anti-HBs. HBV is vaccine-preventable, and immunization offers more than 95% protection against the development of chronic infection.[6] HBV vaccination should be administered at birth for infants exposed to the virus.[7] In Thailand, the Thailand National Strategies (TNS) have implemented a national health benefit package for infants born after 1992 including the viral hepatitis B vaccine. The first stated aim of the TNS was to eliminate viral hepatitis by 2016, while the latest strategy (2022-2030) has the goal of eliminating viral hepatitis B as a major public health threat by 2030.[8] Despite the presence of vaccines, HBV infection remains a major obstacle to public health in Thailand. In 2023, the Department of Disease Control, Ministry of Public Health announced that 2.2 to 3 million Thais were living with chronic HBV infection, with a prevalence of 4 to 5% in those born before 1992. The goals of the national strategy are to reduce viral hepatitis B incidence by 95% and the mortality associated with viral hepatitis B by 65% by 2030.[9]

This cross-sectional study assessed the prevalence of hepatitis B virus infection among pregnant women attending antenatal care (ANC) at Mae Sai Hospital in Chiang Rai. This secondary hospital has a capacity of 90 beds and is in the northernmost part of Thailand near the Myanmar border, serving Thai, Burmese, and other ethnic groups such as Mong and Karen who are living in Thailand without Thai identity documentation. This assessment evaluated the potential development of chronic liver disease and liver cancer.

Materials and methods

All pregnant women attending antenatal care (ANC) at Mae Sai Hospital between January 2019 and December 2023 were included in this study. They underwent a blood test to screen for HBV infection during the initial prenatal visit, and HBsAg screening was conducted by the STANDARD Q HBsAg Test. The sensitivity and specificity of the device were rated at 100% by the company provided and at 98% (95% CI: 94.29-99.63) and 100% (95% CI:98.12-100%), respectively by the Department of Medical Sciences, Thailand. Data were retrospectively collected from the electronic database of Mae Sai Hospital. Inclusion criteria were pregnant women attending ANC at Mae Sai Hospital who had HBsAg serology marker data, while pregnant women without this data were excluded. The research was conducted in accordance with the ethical guidelines outlined in the Declaration of Helsinki, The Belmont Report, the CIOMS guidelines, the International Conference on Harmonization in Good Clinical Practice or ICH-GCP, and 45CFR 46.101(b), and was approved by the University of Phayao Human Ethics Committee No. HREC-UP-HSST1.2/131/67. All the pregnant women were self-identified as indigenous people with de-identified codes.

Statistical analysis

All data were analyzed using SPSS statistical software (IBM SPSS version 29.0). Continuous and categorical variables were presented as numbers and percentages. Categorical data were analyzed and compared using the chi-square test, with logistic regression used to detect the associations between variables at a significance level of $p<0.05$.

Results

Between January 2019 and December 2023, 2,692 pregnant women were admitted to the antenatal care unit at Mai Sai Hospital and included in the analysis. Table 1 shows the demographics of the pregnant women. The sample was divided into the HBsAg positive and HBsAg negative groups, and 131 pregnant women were HBsAg positive with a prevalence rate of 4.9% (95%CI=4.1-5.7). The participants were also segregated by nationality and age. Approximately 30% and 70% were Thais and non-Thais, respectively. The non-Thai group consisted of Burmese (48.1%) and other ethnicities (51.9%). Based on the nationality distribution, the prevalence of HBsAg positive in the Thai group and non-Thai group was 2.9 % (95%CI=1.8-4.2) and 5.7% (95%CI=4.7-6.9), respectively. The prevalence rate of HBsAg positive was highest in the 31-40 age group. The prevalence of HBsAg positivity had the same tendency by age group in the Thai and non-Thai populations. Multivariate analysis showed that nationality not age group was associated with hepatitis B infection; non-Thai pregnant women had a 2-fold (95%CI=1.998-3.161, p=0.003) greater chance of HBV infection than Thai pregnant women. Based on the annual distribution, the number of HBV-infected subjects in each year was 4/184 (2019), 5/131 (2020), 3/151 (2021), 4/166 (2022), and 7/175 (2023) for the Thai group, and 28/445 (2019), 21/350 (2020), 15/258 (2021), 17/283 (2022), and 27/549 (2023) for the non-Thai group. The prevalence rate of HBsAg positive in the Thai group was between 1.99 and 4.0%, while the prevalence rate of HBsAg positive in the non-Thai group was between 4.92 and 6.29% (Figure 1).

Table 1. Prevalence and univariate and multivariate analyses of the factors associated with HBV infection in pregnant women.

	HBsAg		Total N (%)	χ^2 (*p* value)	Univariate analysis			Multivariate analysis		
	Positive N (%)	Negative N (%)			OR	95% CI	*p* value	OR	95% CI	*p* value
Nation group	131 (4.9)	2,561 (95.1)	2,692	0.006**						
Thai	23 (2.9)	784 (97.1)	807 (2.0)		1.00					
Non-Thai	108 (5.7)	1,777 (94.3)	1,885 (70.0)		2.072	1.311-3.275	0.002**	1.998	1.262-3.161	0.003**
Age (years)	131	2561	2,692	0.160						
<20	17 (3.8)	433 (96.2)	450				0.164			
21-30	62 (4.4)	1,334 (95.6)	1,396		1.184	0.685-2.047	0.546			
31-40	49 (6.3)	729 (93.7)	778		1.712	0.947-3.010	0.062			
>40	3 (4.4)	65 (95.6)	68		1.176	0.335-4.123	0.801			

***Note**: **significance level at p<0.01, ***significance level at p=0.001.*

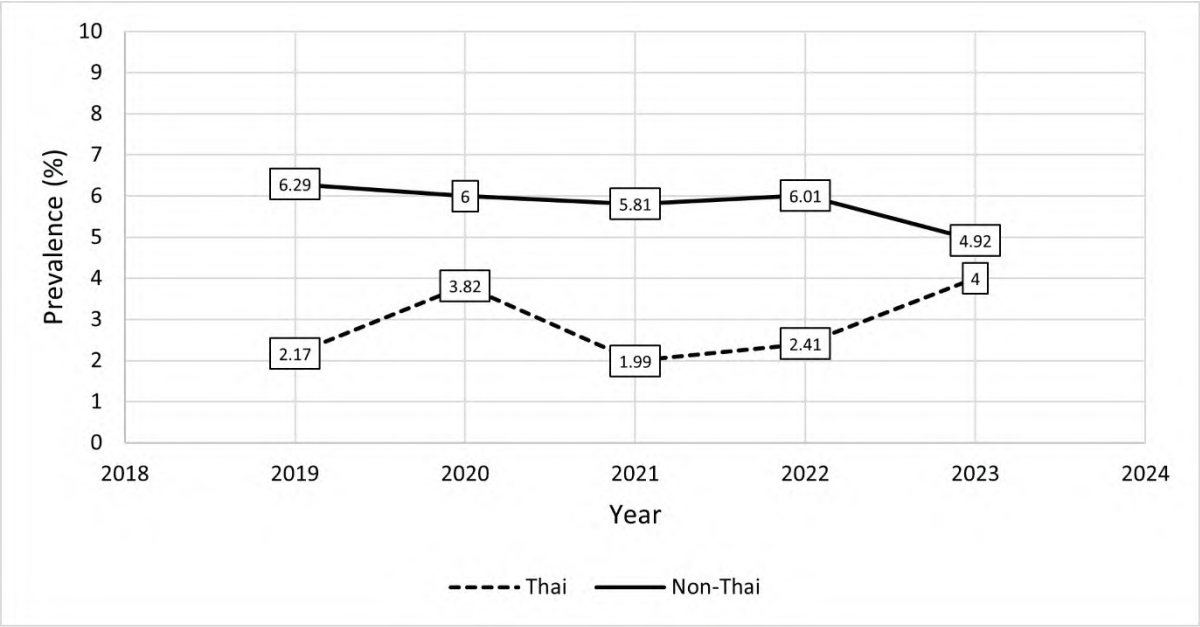


***Figure 1**. Distribution of HBV prevalence in pregnant women between 2019 and 2023. The solid line represents the tendency of the prevalence rate of non-Thai pregnant women, while the dashed line represents Thai pregnant women.*

The vaccine program was implemented in children born after 1992 by the Department of Disease Control, Ministry of Public Health of Thailand, and in pregnant women born before 1992 and born after 1992 at 36.7% and 63.3% respectively. The prevalence rate of HBsAg positive significantly decreased in Thai pregnant women from 1.5% to 0.5% (*p*=0.001), whereas non-Thai pregnant women recorded a slight decrease in the prevalence rate. However, the prevalence rate of HBsAg positive significantly decreased between women born before and after 1992 (*p*=0.002). Multivariate analysis showed that the vaccine implementation program was an associated factor in hepatitis B virus infection; Thai pregnant women born before the implementation of the national vaccination program had a 4.2-fold (95%CI=1.757-10.037, *p*=0.004) greater chance of HBV infection than Thai pregnant women born after the implementation (Table 2).

Table 2. Prevalence and univariate and multivariate analyses identifying vaccine implementation as an associated factor in HBV infection of Thai pregnant women.

Nation group	Prevalence of Hepatitis B viral infection in pregnant women N (%)		Total	χ^2 *p* value	Univariate analysis			Multivariate analysis		
	Born before 1992 (N=988)	Born after 1992 (N=1,704)	2,692		OR	95% CI	*p* value	OR	95% CI	*p* value
Thai	15 (1.5)	8 (0.5)	23	0.001***	1.745	1.227-2.481	0.002**	1.684	1.183-2.396	0.004**
Non-Thai	50 (6.8)	58 (5.0)	108	0.1042	1.384	0.937-2.045	0.102			
Total	65 (6.6)	66 (3.9)	131	0.002**						

***Note**: **significance level at p<0.01, ***significance level at p=0.001.*

Discussion

The hepatitis B virus causes both major and common infectious diseases of the liver. HBsAg positivity indicates active acute or chronic infection and may range from asymptomatic infection or mild disease to severe or rarely, fulminant hepatitis. Acute infection is usually a self-limiting disease with a case fatality rate of 0.5 to 1%, while chronic infection, defined by the presence of detectable HBsAg in the blood for longer than six months, has a high fatality rate. Age of exposure is a key factor in developing the risk of chronic infection. Hepatitis B infection in childhood is responsible for 95% of chronicity. This retrospective cross-sectional study involved 2,692 pregnant women attending antenatal care at Mae Sai Hospital, a facility located in a region of high population mobility and ethnic diversity along the Thai Myanmar border. The overall prevalence of HBV infection in pregnant women in this study was 4.87%, and lower than the rate of hepatitis B infection in pregnant women at Umphang Hospital, also located along the Thai-Myanmar border, with an overall prevalence rate of 6.7%.[10] The rate of hepatitis B infection in pregnant women at the Thai-Myanmar border was 8.3%.[11] Previous studies of HBV infection in pregnant women reported a prevalence rate of 11.40% in a Rohingya camp, 8% in Eastern Ethiopia, and 11.74% in a mountainous region of Southern China, all higher than the rate in our study.[11-13] Based on the nationality distribution, Thai pregnant women showed a lower prevalence than non-Thai pregnant women at 2.9% versus 5.7%. Thai pregnant women had a low prevalence rate compared to a previous study of pregnant women at Chiang Mai University Hospital, located in the northern part of Thailand, and a high prevalence rate compared to Thai pregnant women in the central and northeast areas of Thailand.[14-17] The prevalence rate in these studies tended to increase with age, consistent with our study. This occurred due to the impact of the HBV immunization implementation in Thailand, which was introduced in 1992, and because of increased knowledge in the young population.

However, despite the implementation of the vaccine program in 1992, Thai pregnant women born after 1992 were still infected with the hepatitis B virus, possibly because the HBV-neutralizing antibodies were still below the protective level. The Centre for Disease Control (CDC) considers that people who have received three doses of hepatitis B vaccine and tested positive for anti-HBs (>10 mIU/mL) are immune to hepatitis, while those who do not respond to vaccination or have a low level are thought to remain susceptible to infection. Participant information concerning vaccination medical history or boosting was not included in our study. Shakeri *et al.* revealed that 10 years after having the hepatitis B vaccine, the protection rate was 96.5%.[18] Anti-HBs levels decrease rapidly in the first year after hepatitis B vaccination in children but this decrease slows down in the following years. If the anti-HBs level remains above 10 mIU/mL, protection can continue for up to 24 years.

Adults who received routine hepatitis B vaccination in the newborn period had 71% protection when they reached adulthood.[19] However, a booster dose should be given to women planning for a pregnancy to negate the chance of hepatitis B infection.

Limitations

The findings of this one-center, hospital-based, retrospective study may not accurately represent all pregnant women, and further validation is required in other regions among larger population groups. Follow-up assessments of the severity of infection were not conducted through hepatitis B profiling and comprehensive historical data on boosting or receiving HBV vaccinations in pregnant women were lacking, while the follow-up analyses of maternal and neonatal outcomes encountered problems with data access.

Conclusion

The prevalence of hepatitis B in Thai pregnant women at Mae Sai Hospital was 2.85%, and lower than the prevalence rate in non-Thai pregnant women. The overall prevalence of HBV infection in pregnant women in this study was 4.87%, and lower than the rate of hepatitis B infection in pregnant women compared to previous studies. The overall prevalence rate in this study was less than the goal of the Department of Disease Control, Ministry of Public Health, which aims to eliminate viral hepatitis B as a major public health threat by 2030. The prevalence rate of HBV infection significantly decreased in Thai pregnant women born after the implementation of the national vaccination program. However, steps to prevent maternal HBV infection should be followed

by women planning for pregnancy by undergoing HBV testing, especially those belonging to high-risk non-Thai populations.

Conflict of interest

The authors declare that this research was conducted in the absence of competing interest.

Funding

None

Acknowledgements

The authors thank the entire staffs at Mae Sai Hospital for being cooperative in this project. We also would like to extend our appreciation to the School of Allied Health Sciences, University of Phayao.

References

[1] World Health Organization. Hepatitis B WHO website. 2024 [Cited 2025 Mar6]. Available from: https://www.who.int/publications/i/item/9789240091672.

[2] Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the global burden of disease study 2013. Lancet. 2015; 385(9963): 117-71.

[3] World Health Organization. WHO Guidelines on Hepatitis B and C Testing. Geneva: World Health Organization; 2017 Feb. 4, BACKGROUND – EPIDEMIOLOGY AND NATURAL HISTORY. [cited 2025 Mar6]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK442290/.

[4] Shapiro CN. Epidemiology of hepatitis B. Pediatr Infect Dis J. 1993; 12(5): 433-7. doi: 10.1097/00006454-199305000-00036.

[5] Piratvisuth T. Optimal management of HBV infection during pregnancy. Liver Int. 2013;33(1):188-94. doi: 10.1111/liv.12060.

[6] Chang MH, Chen DS. Prevention of hepatitis B. Cold Spring Harb Perspect Med. 2015; 5(3): 021493. doi: 10.1101/cshperspect.a021493.

[7] World Health Organization. Hepatitis B WHO website. 2024. [cited 2025 Mar6]. Available from:https://www.who.int/news-room/fact-sheets/detail/hepatitis-b.

[8] Posuwan N, Wanlapakorn N, Sintusek P, Wasitthankasem R, Poovorawan K, Vongpunsawad S, Poovorawan Y. Towards the elimination of viral hepatitis in Thailand by the year 2030. J Virus Erad. 2020; 6(3): 100003. doi: 10.1016/j.jve.2020.100003.

[9] Department of Disease Control, Ministry of Public Health. Thailand National Strategies to Eliminate Viral Hepatitis 2022-2030. [cited 2025 Mar6]. Available from:chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.globalhep.org/sites/default/files/content/resource/files/2023-02/Book%20Strategies%20to%20Eliminate%20Viral%20Hepatitis%202022%20%E2%80 %93%202030%20Po7.pdf

[10] Poopaibool N. Prevalence of hepatitis B viral infection in pregnant women on the Thailand-Myanmar border. Sawanpracharak Medical Journal. 2020; 7(2): 151-63 (in Thai).

[11] Banks T, Kang J, Watts I, Tyrosvoutis ME, Min AM, Tun NW, Keereecharoen L, Simmawong W, Wanyatip S, Hanboonkunupakarn B, Nosten F, McGready R. High hepatitis B seroprevalence and risk factors for infection in pregnant women on the Thailand-Myanmar Border. J Infect Dev Ctries. 2016; 10(4): 384-8. doi: 10.3855/jidc.7422.

[12] Ahamed M S, Chaklader T, Hossain S, Farah S, Ali M, Shahnaz F. Prevalence of hepatitis B infection among the pregnant women of forcibly displaced Myanmar Nationals in a selected Rohingya camp. J Dhaka National Med Coll Hos. 2023; 29(02): 32-7. doi: 10.5281/zenodo.10241439.

[13] Umer A, Teklemariam Z, Ayele F, Mengesha MM. Prevalence of hepatitis B infection and its associated factors among pregnant mothers attending antenatal care at public hospitals at Hararghe, Eastern Ethiopia. Front Glob Womens Health. 2023; 27(4): 1056488. doi: 10.3389/fgwh.2023.1056488.

[14] Deng Q, Lin L, Guo W, Deng X, Zhang Q, Hou J. Prevalence of hepatitis B virus infection among pregnant women in the mountainous regions of southern China: A retrospective single-center study. J Clin Lab Anal. 2023; 37(2): e24837. doi: 10.1002/jcla.24837.

[15] Porngasemsart Y, Sirilert S, Tongsong T. Change in prevalence of hepatitis B virus infection in pregnant women in the last two decades in Thailand. Viruses. 2024; 16(2): 314. doi: 10.3390/v16020314.

[16] Harnkuno K, Phukwapee D, Somsorn S, Wattanatorn S. Factors associated of hepatitis B virus infection in pregnant women in Pathum Thani, 2018. Office of Disease Prevention and Control Region 7 Saraburi. 2021; 28(2): 11-20 (in Thai).

[17] Surakan P, Yotha N, Simatan S. Prevalence of hepatitis B virus infection in pregnant women, Health Region 7. Office of Disease Prevention and Control Region 7 Khon Kaen. 2020; 29(1): 24-30 (in Thai).

[18] Shakeri H, Rahmanian V, Shakeri M, Mansoorian E. Study of anti-Hbs antibody titer and associated factors among healthcare staff vaccinated against hepatitis B more than ten years in hospitals of Jahrom In 2016. Pharmacophore. 2018; 9: 156-61.

[19] Sehmen E, Yılmaz EM, Oruç MA. Investigation of hepatitis B surface antibody levels in adults with routine hepatitis B vaccination in childhood. Viral Hepat J. 2023; 29(2): 70-4. doi: 10.4274/vhd.galenos.2023.2023-4-6.

Journal of Associated Medical Sciences 2025; 58(3): 65-71

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Effect of task-oriented approach based activities to improve balance among children with cochlear implant

M. Arun Kumar and Harry Pryya S*

Saveetha College of Occupational Therapy, SIMATS, Thandalam, Chennai, India.

ARTICLE INFO

Article history:
Received 24 January 2025
Accepted as revised 18 May 2025
Available online 30 May 2025

Keywords:
Task-oriented therapy, balance, cochlear implants, occupational therapy, Paediatric Balance Scale (PBS).

ABSTRACT

Background: Paediatric occupational therapy is vital for fostering children's motor, sensory, and coordination skills, particularly for those with cochlear implants who often experience vestibular system disruptions leading to balance impairments.

Objective: This study evaluates the effectiveness of task-oriented occupational therapy in improving balance among children with cochlear implants compared to conventional therapy.

Materials and methods: A quasi-experimental design was conducted with 30 participants aged 4–8 years, recruited using convenience sampling and divided into two groups: experimental (task-oriented therapy) and control (conventional therapy). Both interventions were delivered over 12 weeks with 30 sessions. Balance was assessed pre- and post-intervention using the Paediatric Balance Scale (PBS). Statistical analysis was performed to compare pre- and post-test scores within and between groups.

Results: The experimental group (N=15) had a mean pre-test PBS score of 33.2 (SD=2.88) and post-test score of 38.87 (SD=3.66), showing a statistically significant improvement (Z=-3.429, p=0.001). The control group (N=15) had a mean pre-test score of 34.67 (SD=3.79) and post-test score of 36.07 (SD=3.37), also showing significant improvement (Z=-2.829, p=0.005). Post-test scores between the groups revealed a significantly greater improvement in the experimental group compared to the control group (Z=-2.032, p=0.042).

Conclusion: Task-oriented occupational therapy significantly enhances balance in children with cochlear implants, yielding better outcomes than conventional therapy. These findings highlight the importance of tailored interventions in addressing vestibular deficits and improving functional independence in this population.

Introduction

Paediatric occupational therapy plays a crucial role in helping children to develop essential skills for daily living, including gross and fine motor abilities, sensory-motor integration, and visual-motor coordination. Targeted interventions in occupational therapy enhance a child's independence and social interaction skills, which are central to their overall development. Therapists assess a child's functional abilities in daily activities, academic tasks, and play behaviours, comparing them to age-appropriate developmental milestones.[1]

Hearing is a fundamental aspect of effective communication, which is crucial for a child's cognitive, emotional, and behavioural development. Psychomotor developmentis closely linked to auditory input. However,

* **Corresponding contributor.**
Author's Address: *Saveetha College of Occupational Therapy, SIMATS, Thandalam, Chennai, India.*
E-mail address: pryyaharry03@gmail.com
doi: 10.12982/JAMS.2025.074
E-ISSN: 2539-6056

hearing loss, whether partial or complete, affects approximately 7% of children worldwide, with conductive and sensorineural hearing loss being the most common types.[2] For children with profound hearing loss, cochlear implants provide a viable solution to restore auditory function.

Cochlear implants consist of an external component placed behind the ear and an internal device surgically embedded under the skin. The system includes a speech processor, a microphone, a transmitter, a receiver/stimulator, and an electrode array. These components work together to convert sound into electrical impulses, which stimulate the auditory nerve, thereby enabling hearing.[3,4] Despite their success in restoring auditory function, cochlear implants may disrupt the vestibular system, potentially causing postoperative dizziness or balance impairments. Such complications stem from either surgical trauma or electrical stimulation affecting vestibular pathways. Research highlights that up to 75% of cochlear implant recipients may experience vestibular symptoms like vertigo or imbalance, with 34% identifying these as common postoperative effects.[4] Therefore, assessing and addressing vestibular challenges in cochlear implant patients is crucial for comprehensive rehabilitation.

Children with cochlear implants often face significant balance challenges due to disruptions in their vestibular systems. Previous research, such as Amir *et al.* (2015)[5], examined the balance performance of deaf children with and without cochlear implants, utilizing the Bruninks-Oseretsky motor proficiency test.[5] The study identified notable balance deficits among cochlear implant recipients but emphasized the absence of targeted interventions to address these issues.

While numerous studies have explored balance training in children with cochlear implants, there remains a gap in evidence regarding the effectiveness of task-oriented, occupational therapy-based activities. Task-oriented approaches involve practicing meaningful and real-life tasks to improve motor skills, including balance, coordination, and stability.[6] Such interventions have the potential to enhance functional independence and participation in daily activities for children with cochlear implants.[7]

This study was undertaken to fill the existing gap by investigating the effects of task-oriented approach-based activities on improving balance among children with cochlear implants. The findings aim to provide evidence for integrating task-oriented interventions into occupational therapy programs for this population.

Methodology

Research design

Aquasi-experimental study design was employed to investigate the effectiveness of a task-oriented occupational therapy approach on balance skills in children with cochlear implants.

Sampling technique

The study utilized a convenient sampling technique to select participants.

Sample size

A total of 30 subjects participated in this study. The participants were divided equally into two groups including experimental group: 15 subjects underwent task-oriented therapy interventions, and control group: 15 subjects received conventional occupational therapy.

Sample setting

The study was conducted atMGR Home and Higher Secondary School for the Speech and Hearing Impaired.

Selection criteria

Inclusion criteria

1. Children with cochlear implants.
2. Age group: 4-8 years.
3. Children with a Paediatric Balance Scale (PBS) score between 21-40.
4. Both male and female participants were included

Exclusion criteria

1. Children with visual impairments, other vestibular dysfunctions, or neurological conditions.
2. Children with a PBS score below 21 or above 40.
3. Participants outside the specified age range.

Tools used

Paediatric Balance Scale (PBS)[8]

The PBS, a modified version of the Berg Balance Scale, assesses functional balance skills in school-aged children. The scale comprises 14 items, scored from 0 (lowest function) to 4 (highest function), with a maximum score of 56.

Scoring method

- Each item is scored based on the child's best performance. Multiple trials are permitted for some items.
- Points are deducted for not meeting the time or distance requirements, requiring supervision, or using external support.

Reliability

- Test-Retest Reliability: ICC=0.998, indicating excellent stability of the scale over time
- Interrater Reliability: ICC=0.997, confirming high consistency between different raters.

Validity

- Content Validity: Established through expert reviews and adaptation from the well-validated Berg Balance Scale, ensuring relevance for paediatric populations.
- Construct Validity: Strong correlations with self-care measures and motor control tests in children with motor impairments support the scale's ability to assess balance performance.
- Criterion Validity: Demonstrated high correlation with other validated balance and motor performance assessments, confirming its effectiveness in evaluating functional balance skills.

Intervention procedure

The study was conducted over 12 weeks, with participants attending 30 sessions (3 sessions per week), each lasting 45 minutes. The intervention was structured as follows:

- Warm-up activities: a total of 10 minutes with included gross motor tasks and light exercises.
- Program activities: a total of 30 minutes (task-oriented approach for the experimental group and conventional therapy for the control group). The experimental group engaged in task-oriented occupational therapy involving activities to improve balance, coordination, and motor planning. The control group participated in conventional occupational therapy techniques.
- Wind-up activities: 5 minutes.

Conventional occupational therapy for control group[9,10]

The therapy sessions focused on enhancing the child's ability to adapt to their cochlear implant, improving communication and sensory processing skills, and promoting independence in daily activities. A structured approach was followed to address auditory, motor, and socialchallenges commonly associated with cochlear implantation.

Auditory training: Activities designed to help children interpret and respond to sounds through their cochlear implants. Tasks included identifying environmental sounds, differentiating between tones, and following verbal instructions to improve auditory processing.

Sensory integration therapy: Focused on managing sensory sensitivities often experienced by children with cochlear implants. Activities included exposure to varied sensory stimuli, such as textured materials, swings, or weighted vests, to improve sensory processing and tolerance.

Social skills training: Encouraged interaction with peers through structured group activities, emphasizing turn-taking, active listening, and appropriate social responses. Role-playing scenarios were used to teach effective communication in social settings.

Fine and gross motor skill development: Activities to support motor coordination and balance, which can be affected by auditory processing issues. Included tasks like drawing, writing, catching balls, walking on balance beams, or navigating obstacle courses.

Postural training: Addressed balance and spatial orientation issues that might arise due to cochlear implantation. Exercises included maintaining postural control during sitting, standing, and dynamic activities like walking.

Visual-auditory integration: Designed to improve coordination between visual cues and auditory input. Activities included lip-reading, visual tracking combined with sound cues, and matching sounds to visual objects.

Daily living skills training: Focused on age-appropriate tasks like dressing, eating, and personal hygiene. Activities incorporated auditory instructions to help children follow verbal cues and improve functional independence.

Session Duration: 30-45 minutes, 2-3 times per week Pre- and post-intervention assessments were conducted using PBS.

Table 1. Task oriented occupational therapy for Experimental group.[11,12]

Session	Material used	Activities
1	-	Pre-test assessed using PBS
2	Tape	Walk on a flat surface with one or two-hand support. Walk within a taped box with support. Cruise along the wall using both hands.
3	Tape, footprint	Walk in the taped box without support. Cruise along the wall with one hand. Walk 5 steps on footprint with support.
4	Tape, footprint	Walk 10-15 steps on a flat surface without support. Follow footprints in taped boxes. Walk 15 steps with alternate feet on footprint.
5	Chalk, footprint	Walk on a straight line made with chalk. Follow footprints in 5 circles.
6	Chalk, footprint	Follow footprints on stairs. Follow footprints in 5 chalk circles.
7	Chalk, skipping rope	Jump on alternate legs in chalk shapes. Learn rhythmical movement using skipping rope (step-hop).
8	Obstacle course	Cross 5 obstacles. Carry objects while crossing 10 obstacles.
9	Obstacle course	Cross 10 obstacles. Carry objects while crossing 15 obstacles.
10	Chalk	Jump on a circle while rotating. Jump, hop, and walk in a tandem on chalk patterns.
11	Chalk, chair	Hop on boxes based on claps. Jump on a chair with eyes open and supported.
12	Water bottle, rings	Place rings on a water bottle from a chair. Hop and do animal jumps with rings.

Table 1. Task oriented occupational therapy for Experimental group.[11,12] (Continued)

Session	Material used	Activities
13	Tape	Jump 5 times, walk 10 steps, alternate hop 8 steps, and animal jump to the endpoint.
14	Tape, skipping rope	Tip-toe walk, alternate hopping, and pick up objects.
15	Chair, bench	Stand on one leg on a chair for 10 seconds. Hold quadruped and kneeling positions on a bench.
16	Obstacle	Jump over obstacles of different heights.
17	Footprint, note, pencil	Walk on footprints and draw with eyes closed.
18	Single-word puddles	Jump on puddles with the corresponding word written on the board.
19	Chalk	Jump on both legs and alternate legs on chalk patterns.
20	Rolling	Roll forward or backward with arms and legs extended.
21	Chalk or tape	Jump on hopscotch grid, alternating feet, and skip over a square.
22	Spiral (chalk)	Travel and run through spiral paths, moving backward on both spiral and straight paths.
23	Books	Tap elevated targets with hands, with one foot between each.
24	-	Stand on one leg for 20 seconds, feet heel-to-toe for 15 seconds.
25	-	Walk on a hard surface for 10 seconds and backward for 10 seconds. March in place for 15 seconds with eyes open and closed.
26	-	Walk around a big circle for 15 seconds. Stand with feet shoulder-width apart for 20 seconds, eyes open and closed.
27	-	Stand on one leg, moving head horizontally and vertically for 1 minute while focusing on a target.
28	-	Look at the thumb while rotating on a chair for 30 seconds. Walk while moving head horizontally and vertically for 15 seconds each.
29	-	Sit in front of a target, close eyes and recall it 5 times. Open eyes and recall the target 5 times.
30	-	Post-test assessed.

Wind-up Activities: Included relaxation and joyful tasks such as singing, dancing, and round games.

Results

The effectiveness of a task-oriented occupational therapy approach on balance skills in children with cochlear implants was conducted at MGR Home and Higher Secondary School for the Speech and Hearing Impaired. A total of 30 participant was included in the study. This has been categorized into two groups; control (15) and experimental (15). For control group conventional occupational therapy was given whereas for experimental task orientated occupational therapy was given.

The experimental group consisted of 15 participants, including 10 males (66.7%) and 5 females (33.3%), with ages ranging from 6 to 8 years and a mean age of 6.87 years. The pre-test scores in this group ranged from 29 to 40, with a mean of 33.2, while the post-test scores ranged from 33 to 46, with a mean of 38.87. The control group included 15 participants, comprising 10 males (66.7%) and 5 females (33.3%). The age range for this group was slightly broader, from 6 to 9 years, with a mean age of 7.27 years. Pre-test scores in the control group ranged from 30 to 42, with a mean of 34.67, and post-test scores ranged from 32 to 43, with a mean of 36.07.

Both groups were balanced in terms of gender distribution and comparable in age, ensuring the validity of the comparison. The pre-test and post-test scores provide a basis for evaluating the effectiveness of the PBS intervention in both groups.

On comparison between the pre-test and post-test scores in the control group revealed a statistically significant difference. (Table 2) The mean pre-test score was 34.67 (SD=3.79), while the post-test mean score was 36.07 (SD=3.37). The Z value for the comparison was -2.829, and the p-value was 0.005, which is less than the significance level of 0.05. This idicates a statistically significantimprovement in the PBS scores of the control group. Consequently, the null hypothesis is rejected, and the alternate hypothesis is accepted, suggesting that the intervention received by the control group led to a significant improvement in their scores.

Table 2. Comparative analysis of pre- and post- scores in the control group.

Control group	Mean	SD	N	Z value	*p* value
Pre-test	34.66	3.79	15	-2.829	0.005*
Post-test	36.06	3.36	15		

A significant improvement was observed in the experimental group between the pre-test and post-test scores (Table 3). The mean pre-test score was 33.20 (SD=2.88), while the mean post-test score increased to 38.87 (SD=3.66). The Z value was -3.429, and the p-value was 0.001, which is less than the 0.05 threshold forsignificance. This indicates that the intervention had astatistically significant positive effect on the PBS scores of the experimental group. As with the control group, the null hypothesis is rejected in favour of the alternate hypothesis, confirming that the intervention contributed to a significant improvement in the experimental group.

Table 3. Comparative analysis of Pre and Post test scores in experimental group.

Experimental group	Mean	SD	N	Z value	*p* value
Pre-test	33.2	2.88	15	-3.429	0.001*
Post-test	38.86	3.66	15		

The post-test scores between the control and experimental groups, the results indicated a significant difference (Table 4). The mean post-test score for the control group was 36.07 (SD=3.37), while the experimental group achieved a higher mean post-test score of 38.87 (SD=3.66). The Z value for this comparison was -2.032, and the p-value was 0.042, which is below the significance level of 0.05. This finding suggests that the experimental group experienced a significantly greater improvement in PBS scores compared to the control group. The null hypothesis is rejected, and the alternate hypothesis is accepted, indicating that the intervention in the experimental group was more effective in improving PBS scores than the intervention received by the control group.

Table 4. Comparative analysis of post test scores between control and experimental group.

Post test scores	Mean	SD	N	Z value	p value
Control	36.06	3.36	15	-2.032	0.042*
Experimental	38.86	3.66	15		

The control and experimental groups showed statistically significant improvements in PBS scores from pre-test to post-test. The experimental group exhibited a more substantial improvement compared to the control group, as reflected in the post-test score comparison. These findings highlight the effectiveness of the interventions applied to both groups, with the experimental intervention yielding more pronounced improvements in PBS outcomes.

Discussion

This study evaluated the impact of task-oriented approach-based activities on balance improvement in children with cochlear implants, using the Paediatric Balance Scale (PBS) as the assessment tool. The findings indicate that task-oriented interventions significantly enhanced postural control and balance in the experimental group compared to the control group. Children with cochlear implants often experience challenges in maintaining balance due to vestibular dysfunction associated with hearing impairments. The significant improvement in the experimental group's PBS scores (mean increase from 33.20 to 38.87) demonstrates the effectiveness of task-oriented activities in addressing these deficits. In contrast, the control group showed a smaller improvement (mean increase from 34.67 to 36.07), suggesting that generic interventions are less impactful than targeted approaches.

Franjoine *et al.* highlighted the effectiveness of task-specific training for improving balance in children with motor impairments.[13] Similarly, the present study demonstrates that tailored balance activities, such as dynamic transitions and single-leg stances, can significantly enhance balance in children with cochlear implants, addressing their unique vestibular challenges. Research by Wiener-Vacher *et al.* emphasized that vestibular rehabilitation is critical for children with cochlear implants to compensate for vestibular dysfunction.[16] The task-oriented activities used in this study align with their recommendations, integrating sensory-motor challenges that promote vestibular adaptation and postural stability. Donahoe-Fillmore *et al.* stressed the importance of dynamic and playful balance activities in paediatric therapy.[15] This study's intervention incorporated engaging and child-friendly tasks, which likely contributed to better adherence and outcomes in the experimental group. According to Verbecque *et al.* task-specific balance training promotes neuroplastic changes in the brain, improving motor control and postural adjustments.[16] These principles underpin the significant improvements observed in this study.

The structured activities were tailored to simulate real-life tasks, enhancing sensory-motor integration and

balance control. This approach likely compensated for vestibular deficits caused by cochlear implantation. Task-oriented activities provided opportunities for vestibular rehabilitation, helping children adapt to balance challenges through repeated practice. This aligns with the findings of Wiener-Vacher *et al.* who noted the importance of vestibular compensation in children with cochlear implants.[17] The playful nature of the interventions enhanced motivation and participation, critical factors for achieving significant balance improvements in children. Gradual increases in task complexity fostered neuroplasticity and improved motor learning, as suggested by Wang *et al.*[18] This study aimed to evaluate the baseline balance performance and the impact of task-oriented interventions on children with cochlear implants.

The pre-test scores confirmed that both groups were comparable before interventions, ensuring the validity of comparisons. The experimental group showed significantly greater improvements in balance, validating the efficacy of task-oriented activities for addressing vestibular and balance challenges in children with cochlear implants. The significant difference in post-test scores between the experimental and control groups highlights the superiority of task-oriented interventions over generic exercises.

The study results indicated an enhanced ability to engage in age-appropriate occupations requiring balance, such as dressing, toileting, and classroom participation. Given that children with cochlear implants often experienced vestibular dysfunction, the findings highlighted the importance of sensorimotor-based interventions directly linked to functional goals.[19] This study reinforced that task oriented interventions should not focus solely on improving impairment-level deficits but should aim to enable participation in meaningful daily routines, an essential principle of the occupational therapy paradigm.[20]

The use of real-life, play-based, and dynamic balance tasks likely contributed to better engagement and adherence in the experimental group. These elements aligned with core paediatric occupational therapy principles, particularly the concept of providing "just-right challenges" and using play to enhance motivation. Contextualizing therapy tasks allowed children to practice motor strategies in environments resembling their everyday settings, which facilitated the generalization of skills beyond the clinical setting.[20,21]

Overall, the findings emphasized the need for task-oriented, child-cantered balance training programs for children with cochlear implants. Such programs addressed vestibular deficits, enhanced sensory-motor integration, and improved postural control. Therapists were encouraged to incorporate real-life tasks and dynamic challenges into therapy sessions to optimize outcomes.

Limitations

The study has small sample size and short intervention duration limit its generalizability. Future research should include larger populations, long-term follow-up assessments, and evaluations of task-oriented approaches in diverse settings. Additionally, exploring the role of multisensory integration in balance rehabilitation for children with cochlear implants could provide further insights.

Conclusion

This study demonstrates that task-oriented approach-based activities significantly improve balance in children with cochlear implants. The findings underscore the importance of targeted, progressive, and engaging interventions for addressing vestibular and postural deficits in this population, contributing to better functional outcomes and quality of life.

Funding

This study was self-funded by the authors.

Acknowledgements

The authors sincerely thank all the children and their families who participated in this study for their valuable time and cooperation. We express our gratitude to the audiology and rehabilitation team for their support during the intervention process.

References

[1] American Occupational Therapy Association. Occupational therapy practice framework: Domain and process. Bethesda (MD): AOTA Press; 2021. doi:10.1000/aota.2021.001.

[2] World Health Organization. World report on hearing. Geneva: WHO Press; 2023. doi:10.1000/who.2023.002.

[3] Loizou PC. Introduction to cochlear implants. Hoboken (NJ): IEEE Press; 2013. doi:10.1000/ieee.2013.003.

[4] Ramirez A, Vázquez R, González M. Vestibular dysfunction in cochlear implant recipients: A systematic review. J Vestib Res. 2020; 30(3): 123-36. doi:10.1000/jvr.2020.004.

[5] Amir S, Eleonora G, Zohar S. Balance performance of deaf children with or without cochlear implants. Int J PediatrOtorhinolaryngol. 2015; 79(3): 330-4. doi:10.1000/ijporl.2015.005.

[6] Shumway-Cook A, Woollacott MH. Motor control: Translating research into clinical practice. Philadelphia (PA): Lippincott Williams and Wilkins; 2017. doi: 10.1000/lww.2017.006.

[7] 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(2): 114-28. doi:10.1000/ppt.2003.007.

[8] Deng J, Zhu Q, Zhang K, Xie D and Wu W . Vestibular function in children with cochlear implants: Impact and Evaluation. Front Neurol. 2022; 13: 938751. doi.org/10.3389/fneur.2022.938751

[9] Donahoe-Fillmore B, Erdman E, Spencer J, Ward M, Rowe L. Improving balance and stability in typically developing children through dynamic balance activities. J Phys Ther Sci. 2015; 27(2): 353-7. doi:10.1000/jpts.2015.009.

[10] Rahman S, Begum T, Shamsuzzaman M. Effects of sensory integration therapy on improving balance in children with developmental delays. Int J Pediatr. 2018; 6(12): 9231-8. doi:10.1000/ijp.2018.010.

[11] Wang Z, Yang , Ren L. Effects of task-specific balance training on postural control in children: A systematic review. Phys Ther Rev. 2016; 21(4): 240-7. doi:10.1000/ptr.2016.011.

[12] Verbecque E, Vereeck L, Truijen S. Motor learning principles in pediatric balance rehabilitation: A systematic review. Pediatr Phys Ther. 2019; 31(3): 209-18. doi:10.1000/ppt.2019.012.

[13] Gabbard C, Rodrigues LP. Balance and postural control in children: Influences, assessment, and interventions. Phys TherRehabil J. 2021; 101(1): 1-8. doi:10.1000/ptrj.2021.013.

[14] Shepherd RK, Hardie NA. Auditory system and cochlear implant function in children: Implications for balance and motor skills. Cochlear Implants Int. 2022; 18(Suppl 1): S12-S20. doi:10.1000/cii.2022.014.

[15] Schafer EC, Thibodeau LM. Auditory training for children with hearing loss using cochlear implants. J Speech Lang Hear Res. 2006; 49(4): 644-59. doi:10.1000/jslhr.2006.015.

[16] Gordon AM, Charles JR. Task-oriented training in pediatric rehabilitation. In: Campbell JM, editor. Pediatric Rehabilitation. New York: Springer; 2014. p. 192–211. doi:10.1000/springer.2014.016.

[17] Winstein CJ, Kay DB. Translating the science of task-oriented training to clinical practice: Implications for rehabilitation professionals. J Neurol Phys Ther. 2005; 29(3): 109-18. doi:10.1000/jnpt.2005.017.

[18] Chisari D, Vitkovic J, Clark R, Rance G. Vestibular function and balance performance in children with sensorineural hearing loss. Int J Audiol. 2023; 62(2): 1-9. doi:10.1000/ija.2023.018.

[19] Janky KL, Patterson J, Thomas M, Al-Salim S, Robinson S. The effects of vestibular dysfunction on balance and self-concept in children with cochlear implants. Int J Pediatr Otorhinolaryngol. 2023; 171: 111642. doi:10.1016/j.ijporl.2023.111642.

[20] Benjamin RS, Cushing SL, Blakeman AW, Campos JL, Papsin BC, Gordon KA. Evaluating the use of a balance prosthesis during balance perturbations in children and young adults with cochleovestibular dysfunction. Sci Rep. 2023; 13(1): 9721. doi:10.1000/sr.2023.020.

[21] Akbari M. Effect of task-oriented occupational therapy interventions on balance performance of children with cochlear implants. J Clin Res Paramed Sci. 2021; 6(1): 75-83. doi:10.1000/jcrps.2021.021.

Journal of Associated Medical Sciences 2025; 58(3): 72-84

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Effect of postural breath technique (PBT) compared with postural stretching technique (PST) on body flexibility in healthy participants with poor flexibility: A randomized controlled trial

Wut Yi Nway[1], Settapong Nongharnpitak[2], Yodchai Boonprakob[2,3*]

[1]Graduate School, Khon Kaen University, Khon Kaen Province, Thailand.
[2]School of Physical Therapy, Faculty of Associated Medical Science, Khon Kaen University, Khon Kaen Province, Thailand.
[3]Human High Performance and Health Promotion Research Institute, Khon Kaen University, Khon Kaen Province, Thailand.

ARTICLE INFO

Article history:
Received 3 April 2024
Accepted as revised 27 May 2025
Available online 4 June 2025

Keywords:
Body flexibility, hamstring tightness, postural breath technique, musculoskeletal health.

ABSTRACT

Background: Body flexibility is essential in sports science, enhancing athletic performance by improving agility, balance, and coordination. Postural muscles serve postural control and breathing and are crucial in biomechanics and neurophysiology. The postural breath technique, an exercise combining breathing and postural muscle engagement, was designed to enhance flexibility.

Objective: This study compared the effects of postural breath technique (PBT) and postural stretching technique (PST) on body flexibility in individuals with poor flexibility.

Materials and methods: Fifty-eight participants with poor flexibility were randomly divided into two groups. They received a supervised exercise program of either PBT or PST twice weekly for four weeks. Outcome measurements, including Sit and Reach Test (SRT) as the primary outcome, Modified-modified Schober Test (MMST), Passive Knee Extension Test (PKE), and Craniovertebral Angle Test (CVA) as secondary outcomes, were conducted at baseline, immediately after the first session, and after the last session.

Results: Within each group, both PBT and PST significantly increased SRT scores, decreased PKE angle on both sides and improved MMST and CVA. The mixed-model ANOVA revealed significant differences within groups at each assessment point. PBT resulted in immediate and after four weeks improvements in SRT, MMST, PKE, and CVA, while PST showed increased of all outcomes after four weeks. However, there was no significant difference between the PBT and PST groups.

Conclusion: Both PBT and PST demonstrated effectiveness in improving general body and hamstring flexibility. The study's results suggest that clinicians determine the most effective approach for everyone, such as immediate flexibility improvements (PBT) or sustained flexibility gains over time (PST).

Introduction

Flexibility is essential to health and physical fitness, beside muscle strength and body stability.[1] It significantly influences one's ability to perform daily, athletic, and leisure activities. Optimal flexibility contributes to improved performance, while poor flexibility increases the risk of muscle injuries.[2] Many previous literatures emphasize the various aspects of flexibility, particularly in the spine and hamstring muscles, which are related to overall musculoskeletal health.[3,4] The spine's flexibility directly influences the performance of the abdominal muscles.[5]

∗ **Corresponding contributor.**
Author's Address: *School of Physical Therapy, Faculty of Associated Medical Science, Khon Kaen University, Khon Kaen Province, Thailand.*
E-mail address: yodchai@kku.ac.th
doi: 10.12982/JAMS.2025.075
E-ISSN: 2539-6056

Deep breathing enhances the mobility of the thoracic spine and rib cage, thereby indirectly improving overall flexibility.[5] Proper breathing patterns can facilitate better oxygenation of the muscles, reduce tension and enhancing muscle pliability, which is essential for flexibility. The correlation between breathing and flexibility is also evident that controlled breathing can lead to greater muscle relaxation and an increased range of motion.

Postural muscles, which maintain an upright body position against gravity, play in postural control and breathing.[6] These muscles divided into local and global groups, provide stability to lumbopelvic hip complex and support the spine and pelvic during functional activities.[7] Poor flexibility and posture, often due to the inadequate functioning of these muscles, can negatively impact of joint structure, especially in the spine and knee. [8]

The connection between postural muscle and breathing is crucial, as these muscles support the integrity of the breathing mechanism. The diaphragm, a core stabilizer muscle, is closely linked to both postural control and breathing.[6] Ineffective diaphragm function can disrupt breathing patterns, leading to poor balance and compromised postural control.[9] When the diaphragm's function is impaired, chest mobility decreases, limiting deep breathing[10] and reducing flexibility, particularly in the upper body.[11]

Core stability impacts flexibility by providing a solid foundation for movement. When the core muscles are strong and stable, they allow for better movement patterns and alignment, reducing the strain on other muscles and joints.[8] This results in an increased range of motion and flexibility, as the body can move more freely without compensatory tension in the muscles.

Breathing exercises are recognized for their positive impact on flexibility and posture in healthy adults.[12] Additionally, core stabilizing exercises effectively strengthen muscles, improve balance, and increase overall body flexibility.[13,14] The PBT is designed based on the understanding that postural muscles serve postural control and breathing.[15] It simultaneously stimulates postural muscle contraction, Golgi tendon reflex inhibition, and breathing.

Stretching exercises have long been recognized as a fundamental method for improving flexibility, reducing muscle stiffness, and enhancing overall range of motion.[16] In this study, the Postural Stretching Technique (PST) is specifically designed to focus on the key postural muscles including the abdominal muscles, back muscles, pectoral muscles, trunk and hip muscles.

Previous studies have demonstrated positive outcomes of various exercise techniques on flexibility, including breathing exercises, Pilates techniques, and active stretching exercises.[16,17,18] However, there remains a significant gap regarding the combined effects of postural muscle strengthening and breathing exercises on body flexibility. Although most research has focused on these elements in isolation, the synergistic impact of these integrating has yet to be thoroughly investigated. It is essential to approach might enhance flexibility more effectively than any one technique alone. This study aimed to address this gap by comparing the effects of PBT and PST on flexibility. Due to the comprehensive approach of PBT, it was hypothesized that PBT could lead to a more significant improvement in body flexibility compared to PST in individuals with poor flexibility.

Materials and methods

Participants

The study was designed as a single-blinded, randomized controlled trial conducted at the Faculty of Associated Medical Sciences at Khon Kaen University, Thailand. The study protocol was registered in the RCT registry (TCTR20240902002). Participant recruitment focused on students who are studying in Khon Kaen University. Inclusion criteria included both genders, aged 18 to 24 years, with unilateral or bilateral hamstring tightness (PKE angle more than 20 degrees),[19,20] a BMI of 18.5-24.9 kg/m², and a sedentary lifestyle. Exclusion criteria were regular stretching exercises, muscle relaxant usage, fractures of the spine and/or lower extremities, hamstring muscle tendinitis or strain, hypermobility, severe orthopedic or neurological conditions, abdominal surgery within one year, seizures, respiratory problems, and uncontrolled hypotension or hypertension. The sample size was determined using the mean and standard deviation from the flexibility results from a previous study involving the Sit and Reach Test,[21] to calculated in a formula of n/group = $\frac{2\sigma^2(Z\alpha+Z\beta)^2(1-p^2)}{(\mu 1-\mu 2)^2}$,[22] 10% drop-out, therefore in 28 participants per group for a total of 56 participants.

The study was approved by the Khon Kaen University Human Ethics Committee (HE 662057). Participants were recruited through public posters and online information. A physical therapist assessed participants through the recruitment criteria, and written informed consent was obtained. Participants meeting the inclusion criteria were randomly assigned to the experimental (PBT) or comparison group (PST) by another physical therapist using the stratified block randomization with block size of 4 and 6. Gender (male and female), and degree of hamstring tightness ranges (group 1: male, 20-35 degrees, group 2: male, 36-50 degrees, group 3: female, 20-35 degrees, and group 4: female, 36-50 degrees) were used as the stratifying variables.[19] Each stratifying used block randomization with block sizes of 4 and 6 to achieve an approximate balance of important characteristics and equal participants in both groups. The study was single-blind, meaning that the participants were blinded to the group assignments. Participants performed exercises twice a week for 40 minutes per session for four weeks. The decision to conduct the training for four weeks was supported by literature suggesting that this duration is sufficient to observe measurable changes in flexibility.[21] The flow of the study is shown in Figure 1.

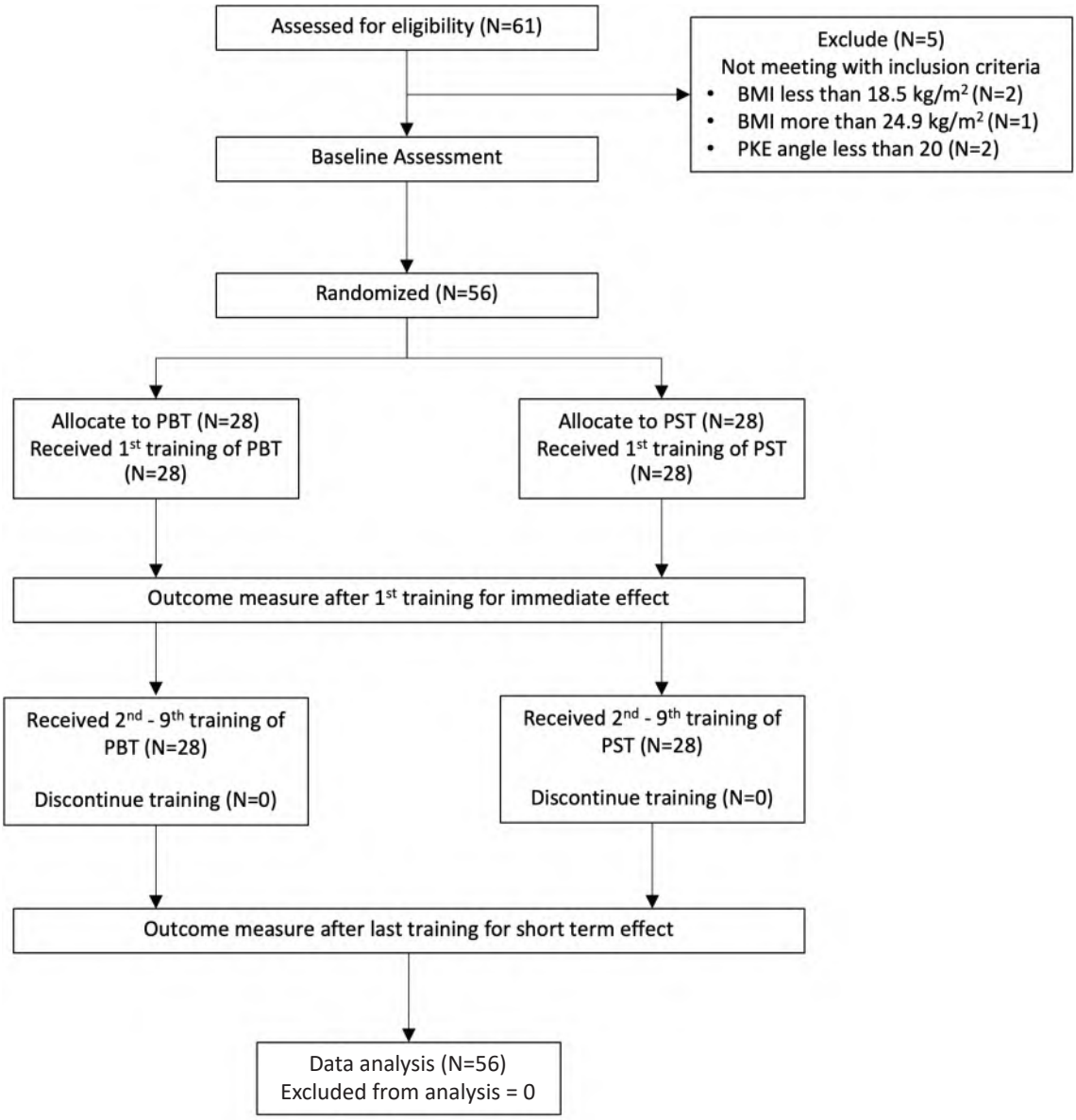


Figure 1. *Flow of the study.*

Intervention

Postural breath technique (PBT)

The experimental group practiced PBT consisting of 7 exercises. All the exercises involved inhaling for 7 seconds and exhaling for 14 seconds, which had to be repeated six times per set, for two sets, with a rest of 30 seconds in between exercises (Figure 2).[15] The design of the exercise intensity in this study was related to the previous study.[18] Detailed exercise procedures are described in supplementary sheet.

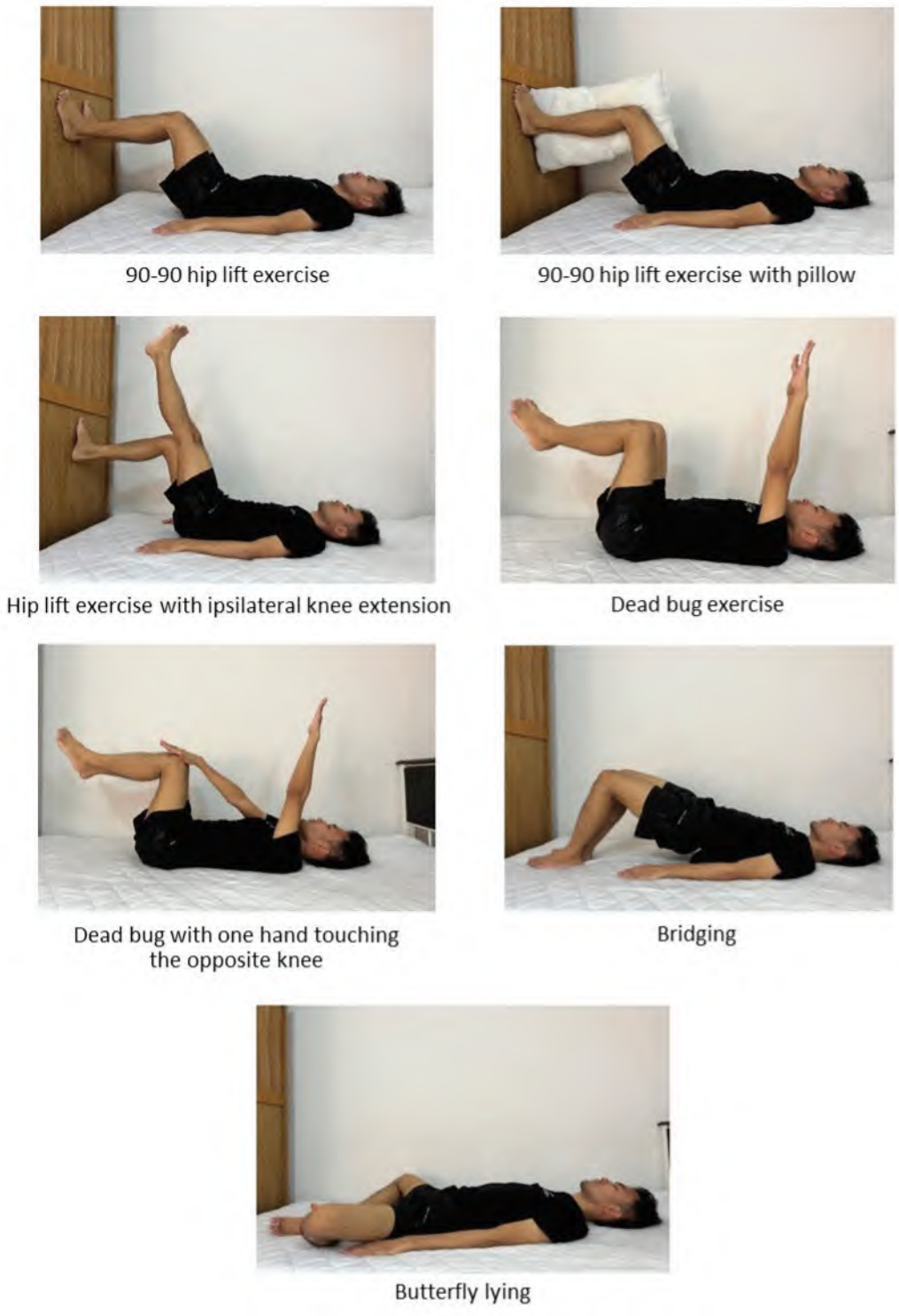


Figure 2. *Exercises positions of PBT.*

Postural stretching technique (PST)

The comparison group practiced PST consisting of 7 exercises. Each stretching exercise was performed holding for 30 seconds, repeated three times for each side, with a 30-second rest between exercises (Figure 3). The design of the exercise intensity was related to the previous study.[18] Detailed exercise procedures are described in the supplementary sheet.

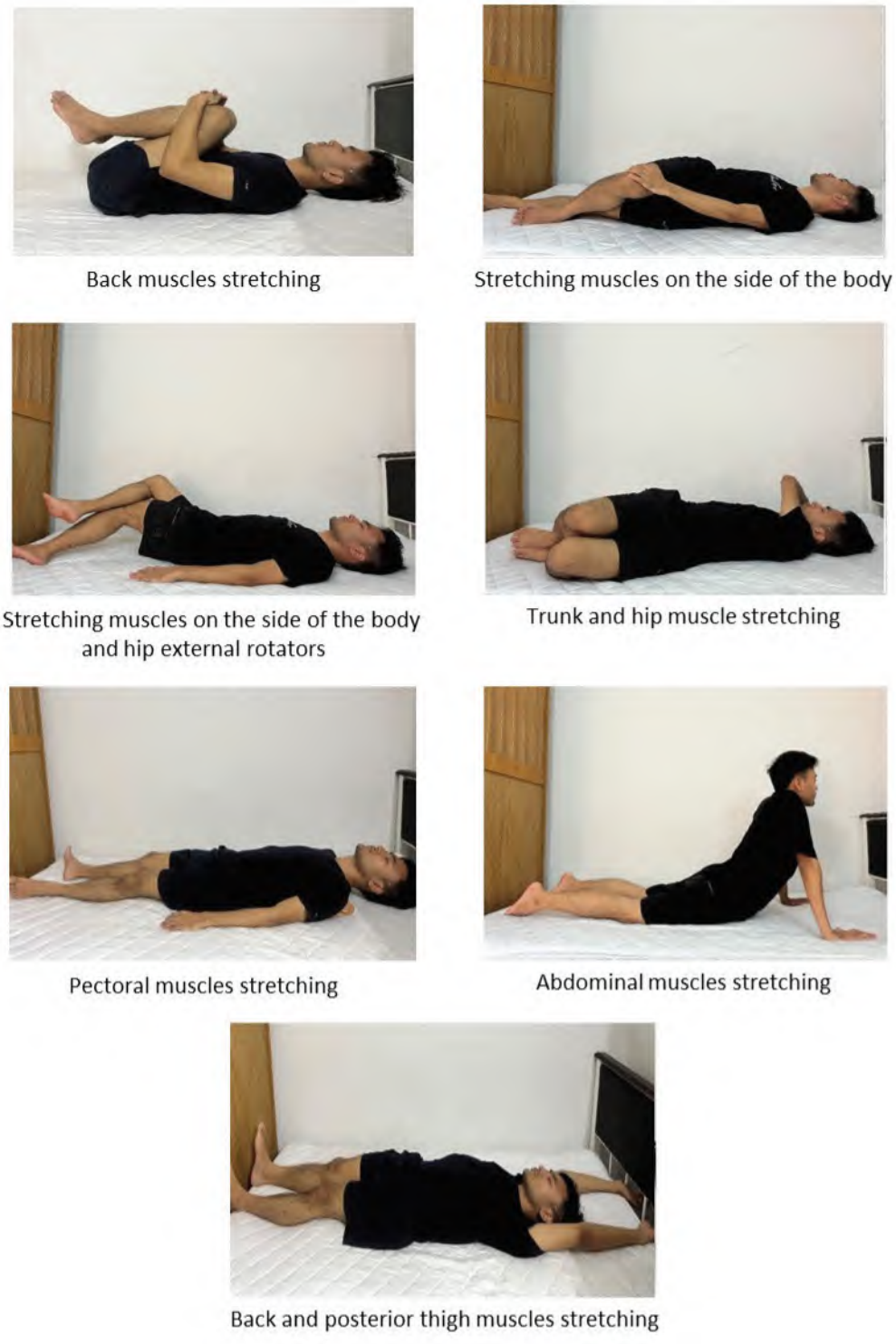


Figure 3. *Exercises positions of PST.*

Outcome parameters

In the primary outcome, the Sit and Reach Test (SRT) is a general flexibility test that relates to several body parts. To conduct the SRT, the participant sits on the floor with legs straight and feet pressed against the measuring box. The individual was then instructed to reach forward over the box, holding that position for two seconds. This test was repeated twice, and the maximum reach in centimeters was recorded as the flexibility value. Reliability of data obtained with this test (Figure 4) was reported to be high (ICC=0.92-0.99). [23]

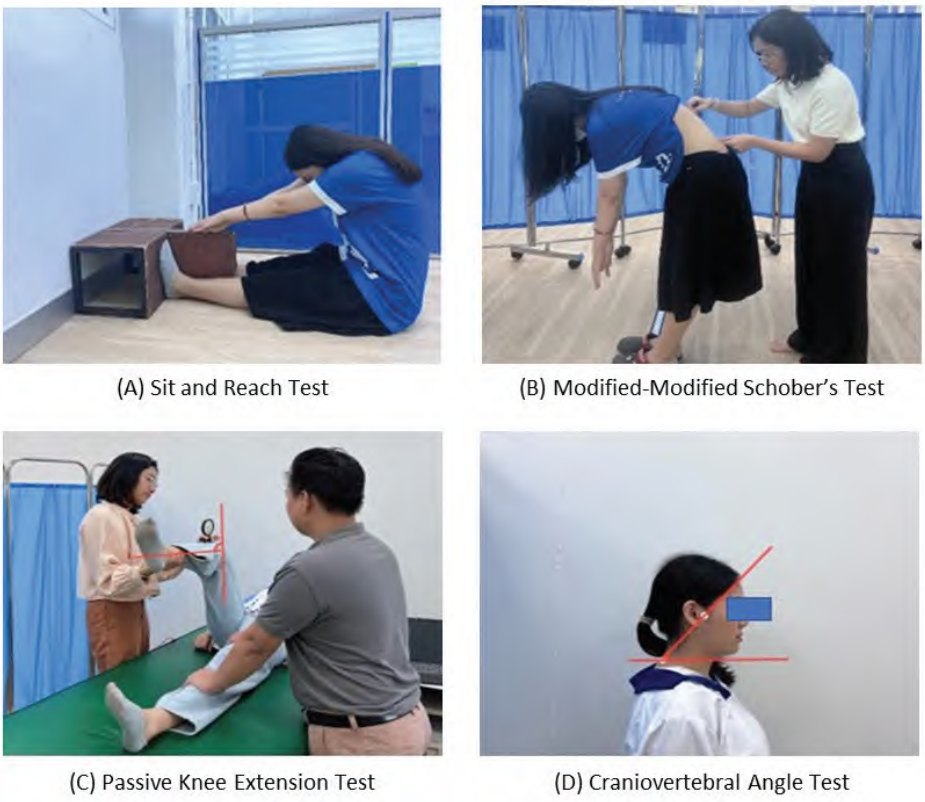
Modified-Modified Schober's test provides information about the range of motion of flexion of the lumbar spine and deep back muscles flexibility. For this procedure, the participants stood with their backs facing the examiner. The examiner palpated the lower edge of the backbone (PSIS) using their thumbs and marked this point with a horizontal line. Another line was drawn 15 centimeters above the midpoint of the first line. The difference between measurements in standing and bending forward was recorded as lumbar flexion. This measurement was repeated three times, and an average value was calculated. The reliability of data obtained with this test was reported to be high (ICC=0.91-0.95) as shown in Figure 4. [24]

The passive knee extension angle test is to determine hamstring flexibility. It is the gold standard for represent hamstrings flexibility. In this procedure, the participant lay comfortably on their back while the assessor bent one hip to 90 degrees, placing the inclinometer on the shin while the opposite leg remained straight on the table. The assessor then slowly extended the bent knee until a tolerable stretch was felt in the posterior thigh, recording the angle. This process was repeated three times, and the average score was calculated. Reliability of data obtained with this test (Figure 4) was reported to be high (ICC=0.92-0.93). [25]

The Craniovertebral Angle Test for forward head posture examination involved the participant sitting in a relaxed position without back support. The examiner marked the bony prominence at C7 and the tragus of the ear. A horizontal line was drawn through C7, and the angle between the line from C7 to the tragus and the horizontal line was measured using a goniometer. Reliability of data obtained with this test (Figure 4) was reported to be high (ICC=0.91-0.99).[26]

Statistical analysis

Data were analyzed using SPSS Version 28.0. The demographic data were described with descriptive statistics such as mean and standard deviation after checking for normal distribution by the Shapiro-Wilk test. The independent t-test was used to compared baseline characteristics variables between groups. Mixed Model Analysis of Variance was used in this study since the dependent variables were assessed at three-time points: baseline, after 1st week, and after 4th week of training. Post-hoc analysis (Tukey's Honest Significant Difference; HSD) was conducted to evaluate specific differences between means in both groups. The statistical significance was set at an alpha level of <0.05.



(A) Sit and Reach Test (B) Modified-Modified Schober's Test

(C) Passive Knee Extension Test (D) Craniovertebral Angle Test

Figure 4. *Tests for outcome measurements.*

Results

Demographic data and baseline characteristics

Of 61 interested volunteers, 56 participants met the inclusion criteria and were randomized into two groups. Both groups had a balanced gender distribution. No dropouts occurred, and no adverse events were reported. Baseline characteristics were comparable, including age, university affiliation, passive knee extension (PKE), and Body Mass Index (BMI). The demographic data and baseline characteristics were found to be equally balanced between groups (p>0.05) as shown in Table 1.

Sit and Reach Test

A mixed-model ANOVA test revealed a non-significant interaction between groups and time [F(1.364, 54.636) = 0.565, p=0.506, partial Ƞp2 =0.010]. There was a difference in the SR at different times for both groups. Within-group comparisons showed statistical significance across various time points in both PBT and PST groups (Table 2). However, no significant differences were found between the groups at any time (p>0.05) (Table 3).

Modified-Modified Schober's Test (MMST)

In the mixed-model ANOVA test, the MMST outcome accepted an assumption of sphericity (Mauchly's test was nonsignificant p>0.05). There was significant interaction between groups and time [F(2,54)=23.575, p<0.001, partial Ƞ2=0.304]. There was a difference in the MMST at different times for both groups. Within-group analyses revealed significant improvement in MMST for the PBT group (p<0.001), while the PST group exhibited substantial improvement from baseline to the short-term effect (p<0.001) (Table 2). Between-group comparison indicated no significant difference in MMST between the two groups (Table 3).

Passive Knee Extension Test (PKE)

The PKE outcome in the mixed-model ANOVA showed a non-significant interaction between groups and time [F(1.725, 54.275)=0.438, p=0.617, partial Ƞp2=0.008] for left-sided PKE values and [F(2,54)=0.375, p=0.688, partial Ƞp2=0.007] for right-sided PKE values, respectively. There was a difference in the PKE at different times for both groups. Within-group comparisons showed improvements in PKE values across various time points (p<0.001) (Table 2). However, there were no significant differences between the groups at any time (p>0.05) (Table 3).

Craniovertebral Angle Test (CVA)

In the mixed-model ANOVA test, the CVA outcome accepted an assumption of sphericity (Mauchly's test was non-significant p>0.05), with a non-significant interaction effect between groups and time [F(2,54)=0.527, p=0.592, partial Ƞp2=0.010]. Within-group comparisons resulted in a significant change in CVA for the PBT group, while the PST group showed a significant difference between baseline and immediate effect (p=0.017) (Table 2). Between-group comparisons indicated no significant differences at any point (Table 3).

Table 1. Demographic data and participant characteristics.

Variable	PBT group (N=28)	PST group (N=28)
Age (years)	19.43±1.03 (18-22)	19.71±1.18 (18-23)
Gender (N, %)		
Male	15 (53.57%)	14 (50.00%)
Female	13 (46.43%)	14 (50.00%)
Weight (kg)	58.33±8.00 (43-74.5)	58.57±9.29 (43-76)
Height (cm)	165.36±9.50 (150-180)	165.61±8.99 (152-180)
Body mass index (kg/m^2)	21.47±1.88 (18.6-24.5)	21.21±2.20 (18.5-24.8)
PKE (Lt) (degree)	31.02±8.81 (20.4-51.4)	31.36±10.18 (21.5-51.4)
PKE (Rt) (degree)	31.03±8.64 (21.0-49.0)	29.93±8.97 (20.2-49.0)

***Note**: PKE(Lt): passive knee extension(left), PKE(Rt): passive knee extension(right). Data presents with mean and min-max and no significance of baseline values between groups.*

Table 2. Comparison of outcomes within each group among baseline, immediate, and the fourth week after training in the PBT and PST group.

Outcomes	PBT					PST				
	Baseline	Immediate effect	Effect size (Cohen's d)	Short term effect	Effect size (Cohen's d)	Baseline	Immediate effect	Effect size (Cohen's d)	Short term effect	Effect size (Cohen's d)
SRT	6.44 ± 6.59	7.61 ± 6.09*	3.95	12.79 ± 5.41**	5.55	5.61 ± 5.35	7.13 ± 5.51*	3.87	11.39 ±5.00**	5.28
MMST	6.33± 0.73	7.34±1.30*	9.59	9.33±0.83**	10.99	6.18±1.10	6.66±1.58	7.84	7.03±1.09**	8.93
Lt. PKE	31.39 ± 8.75	25.11 ± 3.34*	16.25	13.36 ± 6.90**	11.31	31.36 ± 10.18	27.19 ±11.87*	12.47	15.85 ± 7.37**	11.27
Rt. PKE	31.03 ± 8.64	24.31 ± 10.08*	12.79	13.30 ± 6.71**	11.31	29.92± 8.96	25.08 ±12.43*	11.89	16.14 ± 7.67**	11.29
CVA	47.79± 6.03	49.54±5.37*	28.83	52.02±5.19**	29.80	49.04±6.19	50.36±6.08	28.38	52.22±5.88**	29.15

***Note**: SRT: sit and reach test, MMST: modified-modified Schober's test, Rt.PKE: the passive knee extension angle test (right side), Lt.PKE: the passive knee extension angle test (left side), CVA: the craniovertebral angle test. *Statistically significant differences between immediate and baseline (p<0.05). **statistically significant differences between fourth week measurement and baseline (p<0.05).*

Table 3. Comparison of outcomes between PBT and PST group among baseline, immediate, and the fourth week after training.

Outcome measurements	Immediate effects			Short term effects		
	PBT	PST	Mean difference (95% CI)	PBT	PST	Mean difference (95% CI)
SRT	7.61±6.09	7.13±5.51	0.49 (-2.63 to 3.60)	12.79±5.41	11.39±5.00	1.40 (-1.39 to 4.19)
MMST	7.34±1.30	6.66±1.58	0.68 (-0.16 to 1.52)	9.33±0.83	7.03±1.09	2.30 (-1.78 to 2.82)
Lt. PKE	25.11±3.34	27.19±11.87	-2.08 (-7.87 to 3.71)	13.36±6.90	15.85±7.37	-2.49 (-6.36 to 1.37)
Rt. PKE	24.31±10.08	25.08±12.43	-0.78 (-6.84 to 5.29)	13.30±6.71	16.14±7.67	-2.84 (-6.70 to 1.02)
CVA	49.54±5.37	50.36±6.08	-0.821 (-3.89 to 2.25)	52.02±5.19	52.22±5.88	-0.20 (-3.17 to 2.77)

***Note**: SRT: sit and reach test, MMST: modified-modified Schober's test, Rt.PKE: the passive knee extension angle test (right side), Lt.PKE: the passive knee extension angle test (left side), CVA: the craniovertebral angle test, no significance between groups.*

Discussion

The aim of the study was to compare the effects of PBT and PST on body flexibility in individuals with poor flexibility. Within each group, both PBT and PST significantly increased SRT scores, decreased PKE angle on both sides and improved MMST and CVA. The mixed-model ANOVA revealed significant differences within groups at each assessment point. PBT resulted in immediate and after four weeks improvements in SRT, MMST, PKE, and CVA, while PST showed increased of all outcomes after four weeks. However, there was no significant difference between the PBT and PST groups.

Effect on General Flexibility

PBT significantly increased SRT at all time points, showing improved trunk flexibility. These findings align with previous studies using a different technique but a similar style of core movement with breathing control.[12,16] Several mechanisms may have contributed to the increase in SRT scores. PBT consists of active movements of core muscles with breathing. This keeps the main muscles, such as the rectus abdominis, obliques, and deep core stabilizers like the transversus abdominis, in an eccentric contraction and stretching rhythm during exhalation. Eccentric contraction occurs when the muscle lengthens under tension, which is beneficial for increasing flexibility as it promotes muscle elongation and enhances the neurophysiological properties of contractile tissues.

When PBT is applied, muscles activate the golgi tendon organ (GTO), a proprioceptive sensory receptor located in the tendons that can detect changes in muscle tension. The activation of the GTO inhibits alpha motor neuron activity, leading to a decrease in muscle tension, which allows the sarcomeres (the basic unit of a muscle's contraction) to lengthen.[27] In contrast, the PST also contributes to increased flexibility, its primary mechanism might be related to the viscoelastic properties of the muscles and tendons. Shrier and Gossal and Kubo *et al*. have reported that stretching exercises increase the range of motion by either decreasing viscoelasticity or increasing stretch tolerance.[28,29]

Effect on hamstring muscles

PBT significantly decreased the degree of passive knee extension angle, indicating improved hamstring flexibility. This technique utilizes proprioceptive neuromuscular inhibition to release contraction knots of the sarcomere and enhance blood circulation and oxygen supply. The muscle movements combined with slow, deep breathing produce an excellent oxygen supply.[30,31] Increasing blood flow removes pain and proinflammatory mediators,[30] which can lead to increased flexibility.

In addition to PBT, the PST also improved hamstring flexibility. The stress-strain curve explains this effect. When force is gently applied perpendicular to the tissue, the wavy collagen fibers straighten. With increased tension, a recoverable deformation occurs within the elastic range, resulting in new length and enhanced flexibility and range of motion.[32]

Effect on back extensor muscles

Both groups showed statistically significant improvements in short-term effects, with the PBT group being significantly superior, consistent with previous studies.[33] In those studies, researchers found that the increasing flexibility of university students could be measured using the modified Schober test after applying pilates techniques. PBT likely activates core muscles such as the iliocostalis, erector spinae, and multifidi, improving these muscles' flexibility.[34] The eccentric contractions of these muscles during PBT contribute to the gradual lengthening of muscle fibers, which enhances flexibility.

Effect on neck extensor muscles

The PBT significantly increased CVA with both immediate and short-term effects. In contrast, Byung-Sun Kong *et al*. found no significant interaction effects on CVA in straight-necked women because of breathing exercises.[35] The significant increased CVA of the present study may be due to the role of postural muscles in PBT.

The PST group showed an increase in short-term effects, consistent with a previous study that reported static stretching of hamstring muscles increases cervical range of motion.[36] This is likely due to the concept of myofascial meridians, where loosening the musculofascial component in one area can increase the flexibility of the same element in a distant location on the same meridian.[37] Additionally, reducing neural tension may also play a role, as stretching can alleviate nerve compression, thereby improving flexibility in connected regions.

Both PBT and PST exercises can explain the same effect of force on the Golgi tendon organ. From neuroscience knowledge, when an agonist muscle is contracted long enough or stretched with an appropriate force, the Golgi tendon organ of the muscle sends an electrical signal through the 1b afferent fiber into the spinal cord. It stimulates inhibitory interneurons to inhibit the alpha motor neuron of the agonist muscle for relaxation, a mechanism called autogenic inhibition.[38] Based on biomechanics, the postural muscles have two important functions: postural control and breathing.[39] Therefore, the design of exercises in the PBT group includes breathing exercises.

Conclusion

A recent study investigated the effect of PBT and PST on body flexibility over four weeks. Results indicate that both PBT and PST similarly improve general body and hamstring flexibility, with immediate improvements observed in back and neck extensor muscles flexibility with PBT, while PST shows increased flexibility after four weeks. However, there is no significant difference between the two groups. Both PBT and PST are recommended in clinical practice for improving flexibility, with PBT offering immediate benefits and PST providing sustained gains. The choice between these techniques should be based on patient needs, clinical goals, and practitioner expertise. Limitations of the recent study include focusing on young,

healthy individuals and a short duration, suggesting future studies with broader participant demographics and more extended training periods to explore long-term effects comprehensively. Moreover, we could not entirely control the participants' daily activities, which should be monitored for specific results in the future study.

Ethical approval

The study protocol was registered in the RCT registry (TCTR20240902002).

Conflict of interest

The authors declared no potential conflicts of interest to the research, authorship, and/or publication of this article.

Funding

Khon Kaen University Scholarship for ASEAN and GMS Countries' Personnel, 2021.

Acknowledgements

We thank all volunteers for participating in this the research and the research assistants for communicating Thai language and data collection for this study.

References

[1] Pfeifer CE, Ross LM, Weber SR, Sui X, Blair SN. Are flexibility and muscle-strengthening activities associated with functional limitation? Sports Med Health Sci. 2022; 4(2): 95-100. doi: 10.1016/j.smhs.2022.03.001.

[2] Mocanu GD, Dobrescu T. Improving upper body flexibility in students through various types of stretching during physical education lessons. J Phys Educ Sport. 2021; 21(3): 1533-43. doi: 10.7752/jpes.2021.03195.

[3] Raftry SM, Marshall PWM. Does a 'tight' hamstring predict low back pain reporting during prolonged standing? J Electromyogr Kinesiol. 2012; 22(3): 407-11. doi: 10.1016/j.jelekin.2012.02.008.

[4] Reis FJ, Macedo AR. Influence of hamstring tightness in pelvic, lumbar and trunk range of motion in low back pain and asymptomatic volunteers during forward bending. Asian Spine J. 2015; 9(4): 535-40. doi: 10.4184/asj.2015.9.4.535.

[5] Lee K. The relationship of trunk muscle activation and core stability: a biomechanical analysis of pilates-based stabilization exercise. Int J Environ Res Public Health. 2021; 18(12804): 1-11. doi: 10.3390/ijerph 182312804.

[6] Srijessadarak T, Arayawichanon P, Kanpittaya J, Boonprakob Y. Diaphragmatic Mobility and Chest Expansion in Patients with Scapulocostal Syndrome: A Cross-Sectional Study. Healthcare. 2022; 10(950): 1-9. doi: 10.3390/healthcare10050950.

[7] Shankar G, Chaurasia V. Comparative study of core stability exercise with Swiss ball in improving trunk endurance. Int J Health Sci Res. 2012; 2(5): 57-61.

[8] Swann J. Good positioning: the importance of posture. NRC. 2009; 11 (9): 467-9. doi: 10.12968/nrec.2009.11.9.43734.

[9] Kocjan J, Gzik-Zroska B, Nowakowska K, Burkacki M, Suchoń S, Michnik R, et al. Impact of diaphragm function parameters on balance maintenance. PLoS One. 2018; 13(12): 1-14. doi: 10.1371/journal.pone.0208697.

[10] Magee D. Thoracic (dorsal) spine, Lumbar spine. Orthopedic Physical Assessment, 6th Ed., Elsevier: Amsterdam, The Netherlands, 2013, 508-648.

[11] Tomaszewska A, Pawlicka-Lisowska A. Evaluation of an influence of systematic motor activity on the body posture of young people. Pol Merkur Lek. 2014; 36(215): 336-40.

[12] Csepregi É, Gyurcsik Z, Veres-Balajti I, Nagy AC, Szekanecz Z, Szántó S. Effects of classical breathing exercises on pPosture, spinal and chest mobility among female university students compared to currently popular training programs. Int J Environ Res Public Health. 2022; 19(3728): 1-22. doi: 10.3390/ijerph19063728.

[13] Granacher U, Schellbach J, Klein K, Prieske O, Baeyens JP, Muehlbauer T. Effects of core strength training using stable versus unstable surfaces on physical fitness in adolescents: a randomized controlled trial. BMC Sports Sci Med Rehabil. 2014; 6(1): 1-11. doi: 10.1186/2052-1847-6-40.

[14] Kalaycioglu T, Apostolopoulos NC, Goldere S, Duger T, Baltaci G. Effect of a core stabilization training program on performance of ballet and modern dancers. J Strength Cond Res. 2020; 34(4): 1166-75. doi: 10.1519/JSC.0000000000002916.

[15] Hruska R, Coughlin KJ, Masek J. Cough-variant asthma: Responsive to integrative management and postural restoration. Explore. 2005; 1(5): 377-9. doi: 10. 1016/j.explore.2005.06.008.

[16] Batista LH, Vilar AC, de Almeida Ferreira JJ, Rebelatto JR, Salvini TF. Active stretching improves flexibility, joint torque, and functional mobility in older women. Am J Phys Med Rehabil. 2009; 88(10): 815-22. doi: 10.1097/PHM.0b013e3181b72149.

[17] Phrompaet S, Paungmali A, Pirunsan U, Sitilertpisan P. Effects of pilates training on lumbo-pelvic stability and flexibility. Asian J Sports Med. 2011;2(1): 16-22. doi: 10.5812/asjsm.34822.

[18] Charoenwipas P, Buabucha P, Arayawichanon P, Hansawong T, Matho L, Puangmali A, et al. Theraputic effects of posture breath technique in patients with scapulocostal syndrome. Thai J Phys Ther. 2023; 45(1): 66-81 (in Thai).

[19] Yıldırım MŞ, Tuna F, Demirbağ Kabayel D, Süt N. The cut-off values for the diagnosis of hamstring shortness and related factors. Balk Med J. 2018; 35(5): 388-93. doi: 10.4274/balkanmedj.2017.1517.

[20] Joshi DG, Balthillaya G, Prabhu A. Effect of remote myofascial release on hamstring flexibility in asymptomatic individuals - A randomized clinical trial. J Bodyw Mov Ther. 2018; 22(3): 832-7. doi: 10.1016/j.jbmt.2018.01.008.

[21] Na'ima AL, Sari GM, Utomo DN. Combination effect of core stability exercise and contract relax exercise on hamstring flexibility. J Phys Conf Ser. 2019; 1146 (012035): 1-5. doi: 10.1088/1742-6596/1146/1/012035.

[22] Frison L, Pockok SJ. Repeated measures in clinical trials: analysis using mean summary statistics and its implication for design. Statist Med. 1992; 11(13): 1685-704.

[23] Ayala F, de Baranda PS, Croix MD, Santonja F. Reproducibility and criterion-related validity of the sit and reach test and toe touch test for estimating hamstring flexibility in recreationally active young adults. Phys Ther Sport. 2012; 13(4): 219-26. doi: 10.1016/j.ptsp.2011.11.001.

[24] Rezvani A, Ergin O, Karacan I, Oncu M. Validity and reliability of the metric measurements in the assessment of lumbar spine motion in patients with ankylosing spondylitis. Spine. 2012; 37(19): 1189-96. doi: 10.1097/BRS.0b013e31825ef954.

[25] Liu H, Shen Y, Xiong Y, Zhou H, Mao Y, Shen Q, Hong W, et al. Psychometric properties of four common clinical tests for assessing hamstring flexibility in young adults. Front Physiol. 2022: 13(911240): 1-8. doi: 10.3389/fphys.2022.911240.

[26] Gadotti IC, Armijo-Olivo S, Silveira A, Magee D. Reliability of the cranio-cervical posture assessment: visual and angular measurements using photographs and radiographs. J Manipulative Physiol Ther. 2013; 36(9): 619-25. doi: 10.1016/j.jmpt.2013.09.002.

[27] McArdle WD, Katch FI, Katch VL. Exercise Physiology: Energy, Nutrition, and Human Performance. 6th Ed. Philadephia: Lippincott Williams & Wilkins; 2007.

[28] Shrier I, Gossal K. Myths and truths of stretching: individualized recommendations for healthy muscles. Phys Sportsmed. 2000; 28(8): 57-63. doi: 10.3810/psm.2000.08.1159.

[29] Kubo K, Kanehisa H, Fukunaga T. Effect of stretching training on the viscoelastic properties of human tendon structures in vivo. J Appl Physiol. 2002; 92(2): 595-601. doi: 10.1152/japplphysiol.00658.2001.

[30] Simons DG, Travell JG, Simons LS. Myofascial pain and dysfunction: The trigger point manual. Volume 1, 2nd Ed. Baltimore: Williams & Wilkins; 1999.

[31] Boonprakob Y, Phadungkit S, Nongharnpitak S, Srijessadarak T, Supasatean W, Nakhengrit C. Trigger point: Curable or palliative symptoms. Bull Chiang Mai Assoc Med Sci. 2016; 49(1): 155-66 (in Thai).

[32] Liemohn W. Exercise Prescription and the Back. New York: McGraw-Hill; 2001.

[33] Atilgan E, Tarakci D, Mutluay F. Examining the postural awareness and flexibility changes in physical therapy students who took clinical pilates class. Pak J Med Sci. 2017; 33(3): 640-4. doi: 10.12669/pjms.333.12808.

[34] Brumitt J, Matheson JW, Meira EP. Core stabilization exercise prescription, part I: current concepts in assessment and intervention. Sports Health. 2013; 5(6): 504-9. doi: 10.1177/1941738113502451.

[35] Byung-Sun Kong, Beom-Cheol Jeong, Kyung-Tae Yoo. Effect of breathing exercises via joint mobilization on the lung function and spinal alignment of straight-necked women. J Korean Soc Phys Med. 2020; 15(4): 55-65. doi: 10.13066/kspm.2020.15.4.55.

[36] Hyong IH, Kang JH. The immediate effects of passive hamstring stretching exercises on the cervical spine range of motion and balance. J Phys Ther Sci. 2013; 25(1): 113-6. doi: 10.1589/jpts.25.113.

[37] Khunkitti C, Mato L, Hunsawong T, Paungmali A, Boonprakob Y. Effects of occipito-frontal release with dorsiflexion technique (OFRDF) for improvement of superficial back line flexibility. J Assoc Med Sci. 2020; 54(1): 26-34. doi: 10.14456/jams.2021.4.

[38] Bentivoglio M, Cotrufo T, Ferrari S, Tesoriero C, Mariotto S, Bertini G, Berzero A, Mazzarello P. The original histological slides of Camillo golgi and his discoveries on neuronal Structure. Front Neuroanat. 2019; 13:3. doi: 10.3389/fnana.2019.00003.

[39] Aramaki Y. Relationship between postural control and respiratory movement during one-leg standing in healthy males. J Phys Ther Sci. 2023; 35(6): 479-82. doi: 10.1589/jpts.35.479.

Supplementary

Exercise methods for participants in the experimental group.

Postural breath technique				
No.	**Exercise**	**Position**	**Training method**	**Number of trainings**
1	90-90 hip lift exercise	The participant lies supine comfortably.	The volunteer flexes hips and knees 90 degrees so that the soles of feet touch the wall with their buttocks lifted (lift tail bone 1-inch from the floor). Inhales for 7 seconds and then exhales for 14 seconds.	Repeat six times per one set, rest and relax for 30 seconds, performs two sets.
2	90-90 hip lift exercise with pillow	The participant lies supine comfortably.	The participant flexes hips and knees 90 degrees and pins the pillow so that the soles of feet touch the wall with their buttocks lifted (lift tail bone 1-inch from the floor). Inhales for 7 seconds and then exhales for 14 seconds. During exhalation, the participant compresses the pillow slowly.	Repeat six times per one set, rest and relax for 30 seconds, performs two sets.
3	90-90 hip lift exercise with ipsilateral knee extension	The participant lies supine comfortably.	The participant flexes hips and knees 90 degrees so that the soles of feet touch the wall with their buttocks lifted (lift tail bone 1-inch from the floor). Inhales for 7 seconds and then exhales for 14 seconds. During exhalation, the participant extends one side of the knee slowly. On the next exhalation, the volunteer extends the other knee alternatively.	Repeat six times per one set, rest and relax for 30 seconds, performs two sets.
4	Dead bug exercise	The participant lies supine comfortably	The participant flexes hips and knees 90 degrees with both arms raising. Inhales for 7 seconds and then exhales for 14 seconds.	Repeat six times per one set, rest and relax for 30 seconds, performs two sets.
5	Dead bug exercise with one hand touching the opposite knee	The participant lies supine comfortably.	The participant flexes hips and knees 90 degrees with both arms raising. Inhales for 7 seconds and then exhales for 14 seconds. During exhalation, the participant uses one hand to touch the opposite knee. On the next exhalation, the volunteer uses the other hand to touch the opposite knee alternatively.	Repeat six times per one set, rest and relax for 30 seconds, performs two sets.
6	Bridging	The participant lies supine comfortably.	The participant lies in a crook and raises their hips above the floor. Inhales for 7 seconds and then exhales for 14 seconds.	Repeat six times per one set, rest and relax for 30 seconds, performs two sets.
7	Butterfly lying	The participant lies supine comfortably.	The participant lies in a crook position. Inhales for 7 seconds and then exhales for 14 seconds. The participant folds their hips with inhalation and spreads their hips with exhalation.	Repeat six times per one set, rest and relax for 30 seconds, performs two sets

Exercise methods for participants in the control group

Postural muscle stretching exercises				
No.	**Exercise**	**Position**	**Training method**	**Number of trainings**
1	Back muscles stretching	The participant lies supine comfortably	The participant holds his or her knees to the anterior chest wall for 30 seconds and then returns to starting position. Rest and relax for 30 seconds, and performs the next session	3 times
2	Back and posterior thigh muscles stretching	The participant lies supine comfortably.	The participant dorsiflex the ankle and rise his or her arm over head for 30 seconds. Rest and relax for 30 seconds, and performs the next session.	3 times
3	Stretching the muscles on the sides of the body	The participant lies supine comfortably.	The participant lifts their right leg across the left leg. Using the left hand, hold the right knee for 30 seconds. Rest and relax for 30 seconds, and performs the next session	3 times on the right and 3 times on the left sides
4	Stretching the muscles on the sides of the body and hip external rotators	The participant lies supine comfortably.	The participant lifts the right leg cross and holds the left leg for 30 seconds. Rest and relax for 30 seconds and performs the next session.	3 times on the right and 3 times on the left sides
5	Trunk and hip muscles stretching	The participant lies supine comfortably.	The participant changes from supine lying to crook lying and bends shoulders forwardly with hands interlocking. Twist the upper and lower body in opposite directions during exhale and hold for 30 seconds. Inhale and return to starting position. Rest and relax for 30 seconds and performs the next session.	3 times on the right and 3 times on the left sides.
6	Pectoral muscles stretching	The participant lies supine comfortably.	Roll two towels in a circle and insert it under the scapulae for dropping shoulders downwardly. Holding 60 seconds and return to starting position. Rest and relax for 30 seconds and performs the next session.	3 times
7	Abdominal muscles stretching	The participant lies in a prone position	The participant pushes the body backward with the upper limbs and holds for 30 seconds. Return to starting position. Rest and relax for 30 seconds, and performs the next session	3 times

Journal of Associated Medical Sciences 2025; 58(3): 85-91

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Speech therapy for children with cleft lip and palate: Telepractice combined with face-to-face via Application for Articulation Therapy-Thai (AAT-T) in the pandemic of Covid-19

Sasalaksamon Chanachai[1] and Benjamas Prathanee[1,2*]

[1]Department of Otorhinolaryngology, Faculty of Medicine, Khon Kaen University, Khon Kaen Province, Thailand.
[2]Mekong Health Science Research Institute, Khon Kaen University, Khon Kaen province, Thailand.

ARTICLE INFO

Article history:
Received 15 March 2025
Accepted as revised 27 May 2025
Available online 4 June 2025

Keywords:
Application, cleft palate, speech therapy, telemedicine.

ABSTRACT

Background: In developing countries, access to speech therapy is restricted due to limited resources, lack of trained professionals, and geographical barriers. The COVID-19 pandemic further exacerbated these issues, making in-person therapy sessions difficult or impossible for many patients. In response, combining telepractice and mobile applications can help bridge this gap in accessibility to speech therapy services for children with cleft lip and palate.

Objective: To compare the percentage of correct consonants (PCC) before and after implementing telepractice combined with face-to-face speech therapy using the Application for Articulation Therapy-Thai (AAT-T) for children with cleft lip and palate and to evaluate the effectiveness of implementing the combination of AAT-T with telepractice regarding improvement in pronunciation.

Materials and methods: A one-group pretest-posttest design was employed to evaluate a cohort of 18 children aged 5 to 13 years who have undergone surgery for cleft lip and palate, including 12 boys and 6 girls. During pre- and post-implementation of telepractice and face-to-face speech therapy using the Application for Articulation Therapy-Thai (AAT-T), the percentage of correct consonants (PCC) for each individual was elicited via the Thai Connected Speech Standard Test by qualified speech and language pathologists. The speech therapy program consisted of five 45-minute face-to-face training sessions using AAT-T held once a month, along with 15 weekly telepractice sessions, each lasting 30 minutes, conducted over a period of five months.

Results: After telepractice and face-to-face speech therapy with AAT-T, the numbers of PCC scores significantly increased (mean difference=8.64, SD=8.89; 95% confidence interval=4.22-13.06).

Conclusion: Telepractice combined with face-to-face speech therapy using AAT-T effectively corrected the articulation disorders in patients with cleft lip and cleft palate. This strategy can be applied in any areas that have limited resources, geographical barriers, or a pandemic situation that requires social distancing in Thailand. The application can also be developed in any language for use in enhancing speech accessibility in other developing countries.

Introduction

Cleft palate with or without cleft lip (CP±L) is a birth defect that occurs when the roof of the mouth and the upper lip do not completely join during the early months of pregnancy. A cleft lip occurs because parts of the face do not fuse correctly, which is often linked to a problem with the movement of certain cells known as neural crest cells. A cleft palate results when the two sides of the

* **Corresponding contributor.**
Author's Address: *Department of Otorhinolaryngology, Faculty of Medicine, Khon Kaen University, Khon Kaen Province, Thailand.*
E-mail address: bprathanee@gmail.com
doi: 10.12982/JAMS.2025.076
E-ISSN: 2539-6056

palate do not grow together properly. These two types of clefts can occur on their own (non-syndromic) or as part of a larger set of health issues (syndromic). Worldwide, CP±L affects roughly 0.57 to 1.57 out of every 1,000 babies born alive;[1] however, the prevalence of 2.14 per 1,000 live births in Thailand is higher than the global average (95% CI=2.08-2.20).[2]

Even after receiving surgery to repair a cleft lip and/or palate (CP±L), many children continue to have problems with how their soft palate and throat work together (velopharyngeal function). These chronic issues can cause difficulties with speech and language development, as it makes it hard for them to pronounce words clearly. They might also have resonance defects (hypernasality/hyponasality) and voice problems. Furthermore, velopharyngeal insufficiency (VPI) often develops after surgery and can result in compensatory articulation disorders (CAD) that may actually make their speech more difficult for others to understand.[3,4] CAD commonly affects sounds produced far back in the oral cavity, including glottal stops, pharyngeal fricatives, nasal fricatives, mid-dorsum palatal stops, and pharyngeal plosives, among others. These articulation disorders are quite common, ranging from 71.18% to 83.8% of individuals with this condition.[4-7] Subsequently, speech interventions should begin at a young age as the primary objective is to facilitate typical speech and language skills and give children opportunities for academic and social success.[8]

Access to speech therapy for Thai children with CP±L is significantly restricted due to a lack of qualified therapists and the challenges of distance, a common issue in developing nations. Consequently, these children often receive speech therapy very infrequently (once every 1-3 months) which results in extremely slow progress. Previous efforts to address this issue focused on training based on community resources, infrastructure, and the availability of personnel. These approaches have included various training methods, such as intensive speech summer camps,[9-12] parent-led speech interventions,[13-15] and community-based speech therapy models and networking.[16,17] Telepractice (TP) is currently an appropriate model of service delivery for communication disorders that may be the primary mode of service delivery or may be used to supplement in-person services, which is referred to as hybrid service delivery.[18] Moreover, the need for social distancing during the COVID-19 pandemic compounded the existing challenges by limiting the accessibility of speech therapy services. Thus, TP offers a promising solution for delivering specialized speech therapy to children with cleft palate during crisis situations.[19]

Facilitated by information and communication technology, which provides an efficient means of delivering speech therapy to children with CP±L, accessibility has improved[20-22] and expenses can be reduced compared to traditional center-based services. A combination of TP, remote patient monitoring, and/or in-person services, for example an online or mobile app used to share asynchronous information between the client and the clinician between synchronous in-person or virtual sessions, can be implemented to meet the needs of individual clients.[18] In addition, applications can also support speech training by increasing children's attention and positive response, and parents can use them to help correct errors and improve their children's pronunciation skills.[23]

The COVID-19 pandemic, coupled with existing geographical limitations and a shortage of speech therapists, made traditional face-to-face therapy difficult to access. To address this issue, telepractice (TP) and mobile apps offer a method to deliver speech therapy services remotely. As a result, the Application for Articulation Therapy-Thai (AA-T), which focuses on 27 initial and final Thai consonant sounds, was created by the corresponding author via collaboration with the Digix Technology Company.[24] The AAT-T is composed of eight levels of articulation exercises: Level 1: Isolated sounds; 20 exercises focusing on individual sounds, accompanied by videos demonstrating the motion of the speech organs for each target sound, Level 2: Nonsense syllables (1-syllable); 20 exercises involving single nonsense syllables, Level 3: Nonsense syllables (2-syllable); 20 exercises featuring two nonsense syllables that share the same vowel, Level 4: Nonsense syllables (3-syllable); 20 exercises consisting of three nonsense syllables with the same vowel, Level 5: Nonsense syllables (2-3-syllable); 20 exercises that include two to three nonsense syllables with different vowels in each syllable, Level 6: Word level; 35 exercises featuring words, including 15 pictures and a reading list of 20 words, Level 7: Short phrases/sentences; 35 exercises consisting of short phrases or sentences accompanied by 15 pictures and a reading list of 20 short phrases or sentences containing 3 to 5 syllables, and Level 8: Long phrases/sentences; 35 exercises involving longer phrases or sentences featuring 15 pictures and a reading list of 20 phrases or sentences containing 4 to 8 syllables. This systematic method was designed to gradually improve the pronunciation of children with CP±L in order to achieve correct articulation.[24] Using the application, children can practice by naming pictures of words, short phrases, long phrases, and sentences, or by repeating what they hear after pressing a button.

The objectives of this study were to compare the percentage of correct consonants (PCC) before and after implementing a combination of TP and face-to-face speech therapy using the Application for Articulation Therapy-Thai (AA-T) in children with CP±L and to assess its effectiveness.

Materials and methods

Study design and setting

A one-group pretest-posttest design was used. Sample size was calculated with an alfa set to a 95% confidence interval and a beta power of 80% for pre- and post-intervention, based on a previous study. [25] According to these calculations with a 20% dropout, a minimum of 19 patients was required. This research was approved by

the Ethics Committee of Human Research of Khon Kaen University. The project number was HE654002 (approved on April 22, 2022)**.**

Inclusion Criteria

Children with CP±L, aged from 5 to 13, who had caregivers and articulation disorders were enrolled in the study.

Exclusion Criteria

Children with cleft palate, with or without cleft lip (CP±L), could not participate in this study if they had any of the following: 1) hearing loss in either ear ≥40 decibels, 2) global developmental delay, 3) disabilities impacting their speech and language, 4) fewer than two articulation errors (excluding errors with the /r/ sound), or 5) their caregivers were unable to follow the study plan or read and write in Thai.

Of the twenty-one children with CP±L who were initially enrolled, two were excluded from the study. One child had an insufficient number of articulation errors, and the other had attention deficit hyperactivity disorder and delayed language development, making assessment participation impossible. Following the second visit for the in-person speech therapy sessions, another participant withdrew from the study due to the caregiver's inability to adhere to the protocol**.**

Assessment

Participants were assessed on site at the Speech and Swallowing clinic, Srinagarind Hospital as follows: Oral examination and physical development were individually investigated by a speech and language pathologist (SLP). Physical development was evaluated by reviewing Srinagarind Hospital's Health Object System (electronic medical records) and health notebook with physical screening by the researcher.

The hearing test was administered by a qualified audiologist using conditional play audiometry via audiometer (Interacoustics AC 40),[26] which is a standard test used in the Speech and Hearing clinic, Srinagarind Hospital, Khon Kaen University.

Articulation test: Children with CP±L were evaluated using the Thai Connected Speech Articulation Test [27] by consensus of two independent SLPs with over 10 years of experience with cleft palate. If there was any disagreement, a third senior SLP who has more than 35 years of experience with cleft palate investigated and provided a consensus among the evaluators. The Thai Connected Speech Articulation Test has been utilized in previous studies to elicit and evaluate speech production outcomes.[12,28,29] Participants were asked to identify pictures by name. If a child was unable to read the name or identify the picture, the researcher asked her/him to repeat the name after the researcher's saying.

Intervention

Before the study began, caregivers received training on the AAT-T. This included hands-on instruction covering the various training stages, such as how to use the application, score progress, and document the results, in order to ensure they were well-prepared. Then, instruction on providing service was individually conducted for caregivers by SLPs. Moreover, caregivers were given home exercises designed to address articulation errors in children with CP±L. Based on articulation errors, phonological, traditional, and specific strategies for compensatory articulation disorders were individually practiced. This training followed a hierarchical approach, starting with isolated sounds, moving through nonsense syllables, and then progressing to words, phrases, sentences, reading, and storytelling, while focusing on the target sounds. Advancement was considered upon the child achieving 90% accuracy in producing the target sounds at each stage.

This treatment protocol combined monthly in-person therapy with regular telehealth sessions and structured home practice. Each month began with a 45-minute face-to-face session using AAT-T, followed by 30 minutes weekly remote therapy delivered by SLPs through video platforms in the remaining weeks. Parents conducted daily home practice following professional guidelines and documented the progress, with advancement to higher difficulty levels occurring when children demonstrated 90% accuracy with the target sounds.[30] A Daily Home Record was used to track compliance. The program delivered 20 total speech therapy sessions (five sessions onsite for face-to-face speech therapy and 15 TP).

The exercises were structured into 8 main levels, targeting 21 initial sounds and 6 final sounds. These sounds were composed of:

- Level 1: 20 isolated sounds (video: motion of speech organs for target sounds).
- Level 2: 20 1-syllable nonsense sounds.
- Level 3: 20 2-syllable nonsense sounds (with the same vowel in each syllable).
- Level 4: 20 3-syllable nonsense sounds (with the same vowel in each syllable).
- Level 5: 20 2-3-syllable nonsense sounds (with a different vowel in each syllable).
- Level 6: 15 pictures of words and a 20-word list.
- Level 7: 15 pictures of 3-5 syllable short phrases/ sentences and a list of 20 3-5 syllable short phrases/ sentences.
- Level 8: 15 pictures of 4-8-syllable long phrases/ sentences and a list of 20 4-8-syllable long phrases/ sentences. There was also list of reading exercises that provided 35 speech samples in each level for children who possessed literacy skills.

Statistical analysis

The main outcome of this study was the articulation errors that were scored as 0: correct or normal, or 1: incorrect. The percentage of correct consonants (PCC) was calculated, and normal distribution was assessed with Kolmogorov-Smirnov tests. The scores of pre- and post-intervention PCC were analyzed as mean and standard deviation. Resonance was evaluated via perceptual assessment with scores classified based on standard

protocols.[31,32] The mean differences of the pre- and post-intervention PCC scores numbers were compared using independent t-tests. All statistical analyses were performed using SPSS software [IBM SPSS Statistics; version: 28.0.1.0(142)], which is a licensed program available for the Department of Epidemiology, Faculty of Medicine, Khon Kaen University.[33]

Results

Initially, nineteen children with CP±L, aged 5-13 years (mean 7.69, standard deviation 2.13) participated in this study. Unfortunately, one child left the study because their caregiver could not continue following the study plan; therefore, 18 participants, including 6 girls and 12 boys, completed the study. Cleft types were divided into cleft palate (3 children), bilateral cleft lip and palate (6 children), unilateral left cleft lip and palate (5 children), unilateral right cleft lip and palate (4 children). Perceptual and baseline characteristics are presented in Table 1.

Individual percentage of correct consonants (PCC) from pre- and post-speech therapy are presented in Figure 1. Comparison of the percentages of correct consonants from pre- and post-speech therapy is displayed in Table 2. The results reveal that the combination of TP via AAT-T was a strategy that effectively increased PCC scores in children with CP±L.

Table 1. Characteristics of children with CP±L.

Code	Diagnosis	Resonance	Voice	Intelligibility
A01	URCLP	2	0	1
A02	BCLP	1	1	1
A03	ULCLP	2	0	1
A04	CP	2	0	2
A05	ULCLP	0	0	1
A07	URCLP	2	0	1
A08	BCLP	2	0	2
A09	BCLP	3	0	2
A11	ULCLP	3	0	1
A12	BCLP	1	1	2
A13	CP	0	0	1
A14	BCLP	1	0	1
A16	BCLP	2	0	1
A17	URCLP	2	1	1
A18	URCLP	0	0	1
A19	ULCLP	3	1	3
A20	ULCLP	0	0	1
A21	CP	3	0	3

***Note**: CP: cleft palate, ULCLP: unilateral left cleft lip and palate, URCLP: unilateral right cleft lip and palate, BCLP: bilateral cleft lip and palate, **Resonance score**; -1: hyponasality, 0: normal, +1: mild hypernasality, +2: moderate hypernasality, +3: severe hypernasality, **Voice**; 0: normal voice, 1: abnormal voice, **Intelligibility**; 1: intelligibility (>75%), 2: intelligibility if topic known (50-75%), 3: unintelligibility (<75%).*

Table 2. Comparison pre- and post-percentage of corrected consonants.

Assessment	Pre - test				Post - test				Paired differences			
	Min	Max	Mean	SD	Min	Max	Mean	SD	Mean difference	SD	95%CI	
											Lower	Upper
PCC	51.85	85.19	72.63	9.60	66.67	92.59	81.27	9.46	8.64	8.89	4.22	13.06

***Note**: PCC: percentage of corrected consonants.*

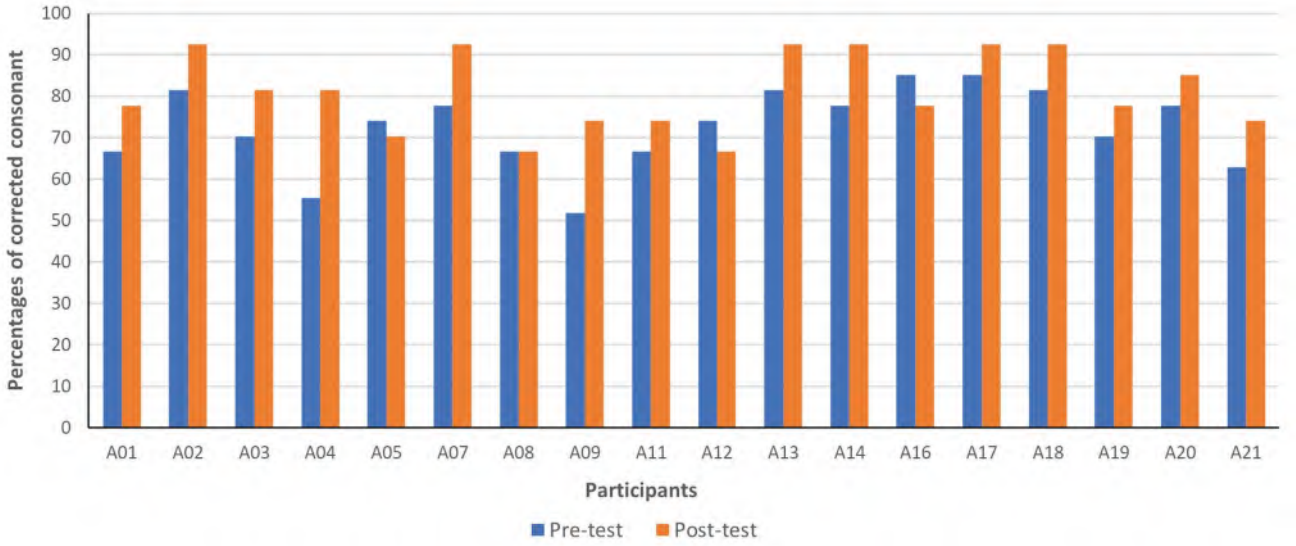


***Figure 1**. Pre- and post- number percentages of corrected consonant.*

Discussion

TP and telehealth are commonly used today.[34] TP has been used in recent decades for speech therapy,[19-21,35,36] particularly during the pandemic, and telehealth applications are currently gaining popularity in the management of cleft lip and/or palate, assisting in areas such as diagnosis and treatment planning, especially for speech therapy.[23] To aid in speech skills practice, a dedicated application is required for use by professionals, parents, and caregivers. Subsequently, AAT-T was designed as a software program to facilitate performance of articulation exercises for speech therapy, which is a complex and highly individualized intervention, which means that AAT-T can provide a tailored approach. TP for children with CP±L has not previously been available for speech therapy in Thailand. Therefore, this research was undertaken as a pioneer study in this field in order to determine the benefits and the effectiveness of conducting speech therapy via AAT-T in conjunction with TP.

The effectiveness of the combination of AAT-T with TP in this study was demonstrated by significant improvement in the PCC scores within 5 months (mean difference=8.64; 95% confidence interval=4.22-13.06) with a total of five sessions of face-to-face speech therapy and 15 sessions of TP for 18 children with CP±L, and it can be used in other areas of Thailand. This finding is in agreement with a previous study, in which one month of weekly, 45-minute TP sessions, incorporating storybook reading and virtual choir singing with the targeted sounds, significantly reduced the severity of compensatory articulation disorders in children with cleft lip and palate aged 4 to 12.[20] In addition, a teletherapy program known as Enhanced Milieu Teaching with Phonological Emphasis (EMT+PE), which was implemented by the children's parents, improved the talking rate, speech, and vocabulary of very young children with cleft lip and/or palate, ranging from 21 to 28 months old.[21] Moreover, during the COVID-19 pandemic in Japan, TP was successfully used as an alternative therapy for children with cleft lip and palate, with parents reporting complete satisfaction.[35] A review of multiple studies also found that TP for speech therapy was effective in decreasing articulation errors during the COVID-19 pandemic[19] and suggested that audiovisual materials such as applications are useful for families receiving and practicing the interventions.

The strength of this study is that, based on the overall data (Figure 1 and Table 2), AAT-T with TP was shown to provide benefits and effectiveness. For the participants who demonstrated the highest improvements in training performance (A04, A07, A09, A14), their individual daily record of homework revealed that caregivers had good compliance regarding homework and consistently provided home-based exercises. They also cooperated very well in both the face-to-face and online training sessions, which was particularly pronounced in participants A07 and A09, whose caregivers demonstrated consistent adherence to the home-based speech practice protocol, following both the prescribed procedures and time frames. Additionally, participants with normal resonance or mild hypernasality demonstrated greater improvements in post-training scores, particularly those with speech intelligibility levels above 75%, which especially affected the training outcomes (A04, A14), suggesting that resonance and intelligibility levels also significantly influenced the outcomes.

This stands in contrast to the outcomes observed in participants who exhibited unexpected findings or no measurable improvement (A05, A08, A12, A16) in performance following the intervention. These three children and caregivers had poor cooperation during therapy sessions, both in-person and remote, as the children struggled with short attention spans, making it difficult to focus on the target sounds. Caregivers had difficulty managing their children's behavior, hindering home practice. Despite "short attention span" being an exclusion criterion in the screening by observation,

this lack of cooperation during assessment may have resulted in underestimated scores. Thus, future studies should implement more formal screening procedures to accurately identify and exclude participants with attention deficits, particularly in the case of the caregiver of A08, who did not conduct any home practice as recommended by SLPs, as the caregiver was unable to manage the child's behavior during the TP-based practice sessions. In summary, the outcome scores from these children may have had an impact on the wide confidence interval of the mean difference.

Despite the unexpected findings potentially arising due to these reasons, the results still support the hypothesis that suggests that AAT-T combined with TP is beneficial for reducing articulation errors. Consequently, future research should prioritize controlling these influencing factors. For the clinical implications, the findings support previous studies, as AAT-T with TP provided benefits and effectiveness for correcting articulation errors of children with CP±L where there were geographical barriers, the pandemic situation, and a shortage of trained professionals in Thailand. Thus, AAT-T may be applied to use with any language for further development.

Regarding the limitations of the study, although the sample size was calculated based on the previous study, the 95% confident interval is still a wide range. This might be due to the excessively small number of participants. Moreover, several participants had a short attention span, which resulted in negative outcomes. To accurately assess the effectiveness of the interventions, a larger number of participants and a formal investigation of attention span should be included in further research.

Conclusion

Twenty sessions of speech therapy via AAT-T combined with TP (five sessions of face-to-face speech therapy and 15 sessions of TP) significantly improved PCC percentages in patients with CP±L. Thus, this approach can be applied in regions with limited resources, geographical barriers, or in situations requiring social distancing, such as during a pandemic, particularly within the context of Thailand, and can also be adapted for use with other languages to enhance speech therapy accessibility in any area that has similar contexts.

Ethical approval

The research was approved by the Ethics Committee of Human Research of Khon Kaen University. The project number was HE654002 (approved on April 22, 2022).

Funding

The study was supported by the Operation Smile Thailand.

Conflict of interest

All authors declare no conflicts of interest.

Acknowledgements

The authors sincerely express gratitude to the children with cleft lip and/or palate (CL±P) and their families for their valuable cooperation throughout the project. The authors appreciate to the staff of the Speech and Swallowing Clinic at Srinagarind Hospital and Research Center of cleft Lip – Palate and Craniofacial Deformities, Khon Kaen University, for their assistance with data collection.

References

[1] Panamonta V, Pradubwong S, Panamonta M, Chowchuen B. Global birth prevalence of orofacial clefts: A systematic review. J Med Assoc Thai. 2015; 98(Suppl 7): S11-21.

[2] Fuangtharnthip P, Chonnapasatid W, Thiradilok S, Manopatanakul S, Jaruratanasirikul S. Registry-based study of prevalence of cleft lip/palate in Thailand from 2012 to 2015. Cleft Palate Craniofac J. 2021; 58(11): 1430-7. doi:10.1177/1055665620987677.

[3] Scherer NJ, D'Antonio LL, McGahey H. Early intervention for speech impairment in children with cleft palate. Cleft Palate Craniofac J. 2008; 45(1): 18-31. doi: 10.1597/06-085.1.

[4] Rezaei P, Poorjavad M, Abdali H. Speech outcomes after palatal closure in 3-7-year-old children. Braz J Otorhinolaryngol. 2022; 88(4): 594-601. doi: 10.1016/j.bjorl.2020.08.005.

[5] Albustanji YM, Albustanji MM, Hegazi MM, Amayreh MM. Prevalence and types of articulation errors in Saudi Arabic-speaking children with repaired cleft lip and palate. Int J Pediatr Otorhinolaryngol. 2014; 78(10): 1707-15. doi: 10.1016/j.ijporl.2014.07.025.

[6] Prandini EL, Pegoraro-Krook MI, Dutka Jde C, Marino VC. Occurrence of consonant production errors in liquid phonemes in children with operated cleft lip and palate. J Appl Oral Sci. 2011; 19(6): 579-85. doi: 10.1590/s1678-77572011000600007.

[7] Prathanee B, Thanawirattananit P, Thanaviratananich S. Speech, language, voice, resonance and hearing disorders in patients with cleft lip and palate. J Med Assoc Thai. 2013; 96 (Suppl. 4): S71-80.

[8] Kuehn DP, Moller KT. Speech and language issues in the cleft palate population: the state of the art. Cleft Palate-Craniofac J. 2000; 37 (348): 1-35. doi: 10.1597/1545-1569_2000_037_0348_saliit_2.3.co_2.

[9] Prathanee B, Lorwatanapongsa P, Makarabhirom K, Suphawatjariyakul R, Wattanawongsawang W, Prohmtong S, Thanaviratananita P. Speech camp for children with cleft lip and/or palate in Thailand. Asian Biomed. 2011; 5(1): 111-8. doi:10.5372/1905-7415.0501.013.

[10] Suphawatjariyakul R, Lorwatanapongsa P, Makarabhirom K, Prathanee B, Manochiopinig S, Wattanawongsawang W. Speech camp: community-based speech therapy model for Thai children with cleft lip/palate in Amnatchareon Province. Saraburi Hosp Med J. 2007; 33(2): 118-25.

[11] Pamplona C, Ysunza A, Patiño C, Ramírez E, Drucker M, Mazón JJ. Speech summer camp for treating articulation disorders in cleft palate patients. Int J Pediatr Otorhinolaryngol. 2005; 69(3): 351-9. doi:

10.1016/j.ijporl.2004.10.012.

[12] Prathanee B, Thanawirattananit P, Surit P, Mitkitti R, Makarabhirom K. Speech task force and quality of life after surgery in children with cleft lip and palate: Limitation of professionals. Arch Plast Surg. 2024; 51(3): 275-83. doi: 10.1055/s-0043-1776738.

[13] Pamplona MC, Ysunza A, Jimenez-Murat Y. Mothers of children with cleft palate undergoing speech intervention change communicative interaction. Int J Pediatr Otorhinolaryngol. 2001; 59(3): 173-9. doi: 10.1016/s0165-5876(01)00476-1.

[14] Sweeney T, Hegarty F, Powell K, Deasy L, Regan MO, Sell D. Randomized controlled trial comparing parent led therapist supervised articulation therapy (PLAT) with routine intervention for children with speech disorders associated with cleft palate. Int J Lang Commun Disord. 2020; 55(5): 639-60. doi: 10.1111/1460-6984.12542.

[15] Pushpavathi M, Kavya V, Akshatha V. Efficacy of focused stimulation in early language intervention program for toddlers with repaired cleft palate. Glob J Oto. 2017; 9: 5-12. doi: 10.19080/GJO.2017.09.555752.

[16] Prathanee B, Dechongkit S, Manochiopinig S. Development of community-based speech therapy model: for children with cleft lip/palate in northeast Thailand. J Med Assoc Thai. 2006; 89(4): 500-8.

[17] Prathanee B, Makarabhirom K, Jaiyong P, Pradubwong S. Khon Kaen: a community-based speech therapy model for an area lacking in speech services for clefts. The Southeast Asian journal of tropical medicine and public health. 2014; 45(5): 1182-95.

[18] ASHA. Telepractice 2025 [cited 2025 18 April 2025]. Available from: https://www.asha.org/practice-portal/professional-issues/telepractice/.

[19] Palomares-Aguilera M, Inostroza-Allende F, Solar LR. Speech pathology telepractice intervention during the COVID-19 pandemic for Spanish-speaking children with cleft palate: A systematic review. Int J Pediatr Otorhinolaryngol. 2021; 144: 110700. doi: 10.1016/j.ijporl.2021.110700.

[20] Pamplona MDC, Ysunza PA. Speech pathology telepractice for children with cleft palate in the times of COVID-19 pandemic. Int J Pediatr Otorhinolaryngol. 2020; 138: 110318. doi: 10.1016/j.ijporl.2020.110318.

[21] Philp J, Ellis PK, Scherer NJ, Lien KM. Enhanced milieu teaching with phonological emphasis: A pilot telepractice parent training study for toddlers with clefts. Children (Basel). 2021; 8(9). doi: 10.3390/children8090736.

[22] Nakarmi KK, Mehta K, Shakya P, Rai SK, Bhattarai Gurung K, Koirala R, Pradhan B, Rai SM. Online speech therapy for cleft palate patients in rural Nepal: Innovations in providing essential care during COVID-19 Pandemic. J Nepal Health Res Counc. 2022; 20(1): 154-9. doi: 10.33314/jnhrc.v20i01.3781.

[23] Dhillon H, Chaudhari PK, Dhingra K, Kuo RF, Sokhi RK, Alam MK, Ahmad S. Current applications of artificial intelligence in cleft care: A scoping review. Front Med (Lausanne). 2021; 8: 676490. doi: 10.3389/fmed.2021.676490.

[24] Prathanee B. Application for Articulation Therapy-Thai: AAT-T. Thailand, 2023.

[25] Prathanee B, Pumnum T, Yoodee P, Makarabhirom K. Speech therapy model for patients with cleft palate in Lao People's Democratic Republic: Lack of speech services. Int J Pediatr Otorhinolaryngol. 2020; 138: 110366. doi: 10.1016/j.ijporl.2020.110366.

[26] Interacoustics. AC40 2025 [cited 2025 11 May]. Available from: https://www.interacoustics.com/audiometers/ac40.

[27] Prathanee B. Assessment and intervention of articulation disorders in children with cleft lip and palate. In: Prathanee B, editor. Cleft lip and palate: speech problem and interdisciplinary approaches Vol II. Khon Kaen: Khon Kaen University Printing; 2014. p. 393-442.

[28] Ooppanasak N, Makarabhirom K, Chowchuen B, Prathanee B. Speech outcomes in children with cleft and palate: Srinagarind Hospital, Khon Kaen University, Thailand. J Med Assoc Thai. 2019; 102(6): 10-8.

[29] Makarabhirom K, Prathanee B, Suphawatjariyakul R, Yoodee P. Speech therapy for children with cleft Lip and palate using a community-based speech therapy mModel with speech assistants. J Med Assoc Thai. 2015; 98 (Suppl. 7): S140-50.

[30] Prathanee B. Record book of speech therapy for children with cleft lip and palate. KhonKaen: Klangnanawittaya Press; 2010.

[31] Prathanee B, Lorwatanapongsa P, Anantapong D, Buakanok N. Thai speech parameters for patients with cleft palate in a universal reporting system. Asia Pac J Speech Lang Hear. 2011; 14(1): 31-49. doi: 10.1179/136132811805334902.

[32] Henningsson G, Kuehn DP, Sell D, Sweeney T, Trost-Cardamone JE, Whitehill TL. Universal parameters for reporting speech outcomes in individuals with cleft palate. Cleft Palate Craniofac J. 2008; 45(1): 1-17. doi: 10.1597/06-086.1.

[33] IBM. IBM SPSS Statistics Desktop 28.0.1.0 2025 [cited 2025 11 May]. Available from: https://www.ibm.com/software/reports/compatibility/clarity/product.html?id=0B6294300A3F11ECADF816E2F-D273C71.

[34] Burke BL, Jr., Hall RW. Telemedicine: Pediatric applications. Pediatrics. 2015; 136(1): e293-308. doi: 10.1542/peds.2015-1517.

[35] Hayakawa T, Imura H, Inoue C, Mori T, Aihara Y, Tsujiuchi S, Niimi T, Natsume N. Efficacy of telepractice, an alternative therapy tool during the coronavirus disease 2019 pandemic, for speech disorders related to congenital anomalies. Congenit Anom (Kyoto). 2023; 63(6): 206-10. doi: 10.1111/cga.12543.

[36] Rangarathnam B, McCullough GH, Pickett H, Zraick RI, Tulunay-Ugur O, McCullough KC. Telepractice versus in-person delivery of voice therapy for primary muscle tension dysphonia. Am J Speech Lang Pathol. 2015; 24(3): 386-99. doi: 10.1044/2015_ajslp-14-0017.

Journal of Associated Medical Sciences 2025; 58(3): 92-101

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 effects of executive function interventions on executive functions and language skills in school-age children with specific language impairment

Natwipa Wanicharoen, Thanasak Kalaysak*, Supaporn Chinchai

Communication Sciences and Disorders Division, Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 15 January 2025
Accepted as revised 26 May 2025
Available online 4 June 2025

Keywords:
Executive function intervention, language, specific language impairment, systematic review.

ABSTRACT

Background: Children with specific language impairment (SLI) have problems with non-linguistic areas such as executive function (EF) skills, in addition to language skills. EF skills are crucial for language development and processing. Conversely, language acquisition can also enhance EF skills, suggesting a bidirectional relationship. Speech therapists (STs) play a crucial role in providing interventions focused on EF skills for school-age children with SLI. Nevertheless, there is a scarcity of comprehensive evaluations regarding the effects of EF interventions for school-age children with SLI that enhance EF and language skills.

Objective: The purpose of this systematic review (SR) was to investigate which EF interventions have affected EF skills for school-age children with SLI and to investigate how improvements in EF skills could improve language skills in these children with SLI.

Materials and methods: This SR followed the Preferred Reporting Items for Systematic Reviews (PRISMA) guidelines with a descriptive-analytical approach. The protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO; CRD42024545361). Searched databases included ERIC, PubMed, APA PsycINFO, and ProQuest Dissertations and Theses. The authors used the Single-Case Experimental Design (SCED) and the Joanna Briggs Institute (JBI) critical appraisal tool for quality assessment.

Results: A total of 5,737 studies were retrieved, of which 4 studies were included in this review. The evidence supports the notion that EF interventions could improve EF skills (i.e., visuospatial WM, attention, inhibition, and cognitive flexibility) as well as language skills (i.e., language comprehension and production, particularly grammatical skills) in children with SLI aged 6 to 12 years. This study also indicates that while there are promising outcomes, the effects can be inconsistent and vary depending on the type of intervention and the specific skills targeted. EFs offer a cognitive framework that facilitates language acquisition, comprehension, and production throughout development and across different contexts.

Conclusion: EF interventions could have the potential to improve both EF skills and language skills. EF skills are essential for language development and processing, and vice versa. Therefore, STs could integrate EF interventions with traditional language interventions.

* **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: thanasak.k@cmu.ac.th
doi: 10.12982/JAMS.2025.077
E-ISSN: 2539-6056

Introduction

Executive functions (EFs) are a set of cognitive capacities that enable individuals to regulate goal-directed behavior, planning, mental flexibility, and self-monitoring. The term EF encompasses three main cognitive capacities: inhibition (deliberately overcoming a dominant response), working

memory (WM; involving maintenance, monitoring, updating, and manipulation of short-term memory contents, as well as interference), and attentional set shifting.[1-2] EFs gradually develop and change across an individual's lifespan and can be improved at any time.[1]

The relationship between EF skills and language acquisition is a complex and multifaceted area of study, with research indicating a significant interplay between these cognitive domains. EF skills are crucial for language development and processing. Conversely, language acquisition can also enhance EF skills, suggesting a bidirectional relationship.[3-5] For instance, Filipe *et al.*[4] highlight that WM and cognitive flexibility have been shown to predict language abilities significantly in preschool children, explaining a substantial portion of the variance in language outcomes beyond age, gender, and nonverbal intelligence. WM is particularly important for tasks that require holding and manipulating information, such as understanding complex sentences. While inhibition is a critical EF component, its direct impact on language outcomes is less pronounced compared to working memory and cognitive flexibility. However, it plays a role in managing distractions and focusing on relevant linguistic information.

Children with specific language impairment (SLI) have a communication impairment in language skills that is markedly below age-appropriate and does not result from other developmental abnormalities such as hearing loss, cognitive impairment, or a clear neurological diagnosis.[6] Not only do children with SLI have problems with language, such as limited vocabulary, morphological errors, grammar and syntax challenges, and pragmatic difficulties,[7] but they also have problems with non-linguistic areas such as EF skills that might lead to social, literacy, and working memory challenges.[8-11]

In the past decade, there have been several systematic reviews and meta-analyses investigating EF skills in school-age children with SLI. Many children with SLI have difficulty with EFs compared to typically developing (TD) children. [8-11] A SR of Cama and Leon-Rojas[8] found that children with SLI have problems with WM, which includes phonological, auditory, and visual/verbal memory. Furthermore, attention deficits are prevalent, affecting their ability to focus and process information efficiently, and also the ability to plan and internalize speech is defective, contributing to difficulties in expressive skills both verbally and non-verbally.

A meta-analysis by Estes *et al.*[9] also found that children with SLI experience word learning and sentence processing difficulties due to lower performance in phonological short-term memory (pSTM) compared to their peers on average. Meanwhile, a meta-analysis by Vugs *et al.*[10] found that children with SLI performed lower in visuospatial working memory (VSWM) than their peers on average, leading to nonverbal learning challenges. Additionally, a meta-analysis by Pauls & Archibald[11] found that children with SLI performed lower in inhibition and in cognitive flexibility than their peers on average, which might increase distractibility, reduce listening comprehensibility, and affect multiple elements of pragmatic skills. As a result, children with SLI had EF deficits that exceeded those of TD children and related to their language skills.

Speech therapists (STs) play a pivotal role in addressing EFs for school-age children with SLI.[12] EF interventions are essential as they address the cognitive and communicative challenges faced by children with SLI.[12-15] However, the effects can be inconsistent and vary depending on the type of intervention and the specific skills targeted. For example, EF interventions including WM training programs, including interactive and computerized,[12] WM training interventions including listening recall training task and odd one out span training task,[13] a computer-based EF training (Braingame Brian),[14] and Cogmed WM training.[15] Additionally, the results varied, encompassing both WM and language, with a particular focus on WM[12-15] and language.[12-13,15] Furthermore, some studies examined behavioral problem[14] and IQ.[15] Nevertheless, there is a scarcity of comprehensive evaluations regarding the effects of EF interventions for school-age children with SLI that enhance EF skills and language skills and how improvements in EF skills could improve language skills in these children.

The purpose of this SR was to investigate which EF interventions have affected EF skills for school-age children with SLI and to investigate how improvements in EF skills could improve language skills in these children with SLI. In the present SR, we sought to answer the following questions:

1) What EF interventions have affected EF skills for school-age children with SLI?
2) How could improvements in EF skills improve language skills in school-age children with SLI?

Materials and methods

This SR followed the PRISMA guidelines with a descriptive-analytical approach.[16] The protocol was registered in the PROSPERO (CRD42024545361).

Inclusion and exclusion criteria

This SR followed the criteria for the inclusion and exclusion of studies in this review based on population, intervention, comparison, outcome, and study design (PICOS) principles.

Types of participants

The target participants included children aged 6-12 years old who were diagnosed with SLI. We excluded participants with other diagnoses or mixed diagnoses in their developmental and medical histories, such as attention-deficit/hyperactivity disorder (ADHD), autism spectrum disorder (ASD), intellectual disability (ID), sensory disorders, brain damage, or seizures.

Types of intervention

The target interventions included EF interventions, programs, principles, trainings, or strategies used with school-age children with SLI. We did not consider programs

that were not primarily focused on EF skills.

Types of comparators

We compared the intervention to other types of interventions, such as waiting lists, traditional language interventions, control groups, or no control group.

Types of outcomes

The target outcomes included EF skills, such as attention control, behavioral inhibition, and working memory, as a primary outcome and related to language skills. We excluded studies that did not measure EFs as primary outcomes.

Types of studies

We considered randomized or non-randomized controlled trials. Other types of research designs, such as expert opinions, case reports, and qualitative studies, were not considered.

Search strategy

The authors searched for the studies between 1 July and 31 October 2024. Searched databases included ERIC (3,492), PubMed (15), APA PsycINFO (4), and ProQuest Dissertations and Theses (1,782). To ensure that all existing literature on the research questions addressed by this review was included, we used reference tracking and hand searching. We also searched Brain Sciences (93), Child Language Teaching and Therapy (242), American Journal of Speech-Language Pathology (44), and Clinical Linguistics & Phonetics (69). The SR included studies that were published from 2014 to 2024. Full text was in English.

The searched terms were using booleans, truncations, and other operators: ("specific language impairment" OR "developmental language disorder" OR "primary language impairment") AND ("executive function" OR "attention" OR "inhibition" OR "shifting" OR "working memory") AND ("executive function intervention" OR "program" OR "principle" OR "training" OR "strategy").

Study selection and data extraction

The two review authors (N.W. and T.K.) searched and eliminated duplicates, including documents indexed in two or more databases as well as nonoriginal publications like books, book chapters, and journals. We were then responsible for study selection during the first screening phase, independently assessing the title and abstract of all documents to find those of potential relevance. The selection process includes documents with sufficient information for screening (title and abstract). In a second screening process, using full texts, we independently determined which studies met the inclusion criteria for this review. When we could not reach a consensus, a third review author (S.C.) was consulted as necessary, whose decision was final. In the last process, we independently extracted the main characteristics of the included studies using a designed template.

Data analysis

This SR used a descriptive analysis approach. We did not analyze subgroups or subsets or carry out a meta-analysis of the studies due to heterogeneous studies.

Quality assessment

To assess the quality of the studies, we used the SCED[17] for single-subject studies (Table 1) and the JBI critical appraisal tool[17] for quasi-experimental studies (Table 2).

Table 1. Risk of bias for single subject study using the SCED.

Study	Clinical history	Target behaviors	3 phases	Baseline	Treatment phase	Data record	Inter-rater reliability	Assessor independence	Statistical analysis	Replication	Generalization	Total score
Shahmahmood Toktam *et al*. 2018[12]	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	8 (moderate)

***Note**: Y: yes, N: no.*

Table 2. Risk of bias for quasi-experimental studies using the JBI quasi-experimental studies checklist.

JBI quasi-experimental studies checklist	Henry *et al.* (2022)[13]	Vugs *et al.* (2017)[14]	Holmes *et al.* (2015)[15]
1. 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	Y	Y
2. Were the participants included in any comparisons similar?	Y	NA	Y
3. Were the participants included in any comparisons receiving similar treatment/care, other than the exposure or intervention of interest?	N	NA	N
4. Was there a control group?	Y	N	Y
5. Were there multiple measurements of the outcome both pre- and post-intervention/ exposure?	Y	Y	Y
6. Was follow-up complete and if not, were differences between groups in terms of their follow-up adequately described and analyzed?	Y	Y	N
7. Were the outcomes of participants included in any comparisons measured in the same way?	Y	Y	Y
8. Were outcomes measured in a reliable way?	Y	Y	Y
9. Was appropriate statistical analysis used?	Y	Y	Y
Overall appraisal	8/9 (Low risk)	6/9 (Moderate risk)	7/9 (Moderate risk)

***Note**: Y: yes, N: no, NA: not applicable*

The SCED comprises 11 items, with 10 of them focusing on assessing the methodology and statistical analysis quality. It classifies studies scoring between 9-11 as good quality, those scoring 6-8 as moderate quality, and any scoring under 5 as poor quality.[17]

The JBI quasi-experimental studies checklist comprises 9 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. The JBI checklist assesses each item by assigning a score of 1 for yes, 0 for no, and 0 for unclear or not applicable items. There is a high risk of bias if 20-50% of items score yes, a moderate risk if 50-80% of items score yes, and a low risk if 80-100% of items score yes.[18]

The two review authors (N.W. and T.K.) independently assessed the studies. Any disagreement was resolved by consensus, and whenever this was not possible, a third review author was consulted (S.C.), whose decision was final.

Results

Figure 1 illustrates a methodology flowchart. The preliminary database and hand search in relevant journals yielded a total of 5,737 articles relevant to the topic. The authors removed the duplicates from these articles, leaving a total of 1,330. During the initial screening, we excluded 1,320 studies from these publications. We systematically filtered the full-text articles for eligibility assessment using the PICOS design. After a thorough review of the 10 articles, the search yielded four matches for the inclusion/ exclusion criteria defined in the methods.[12-15]

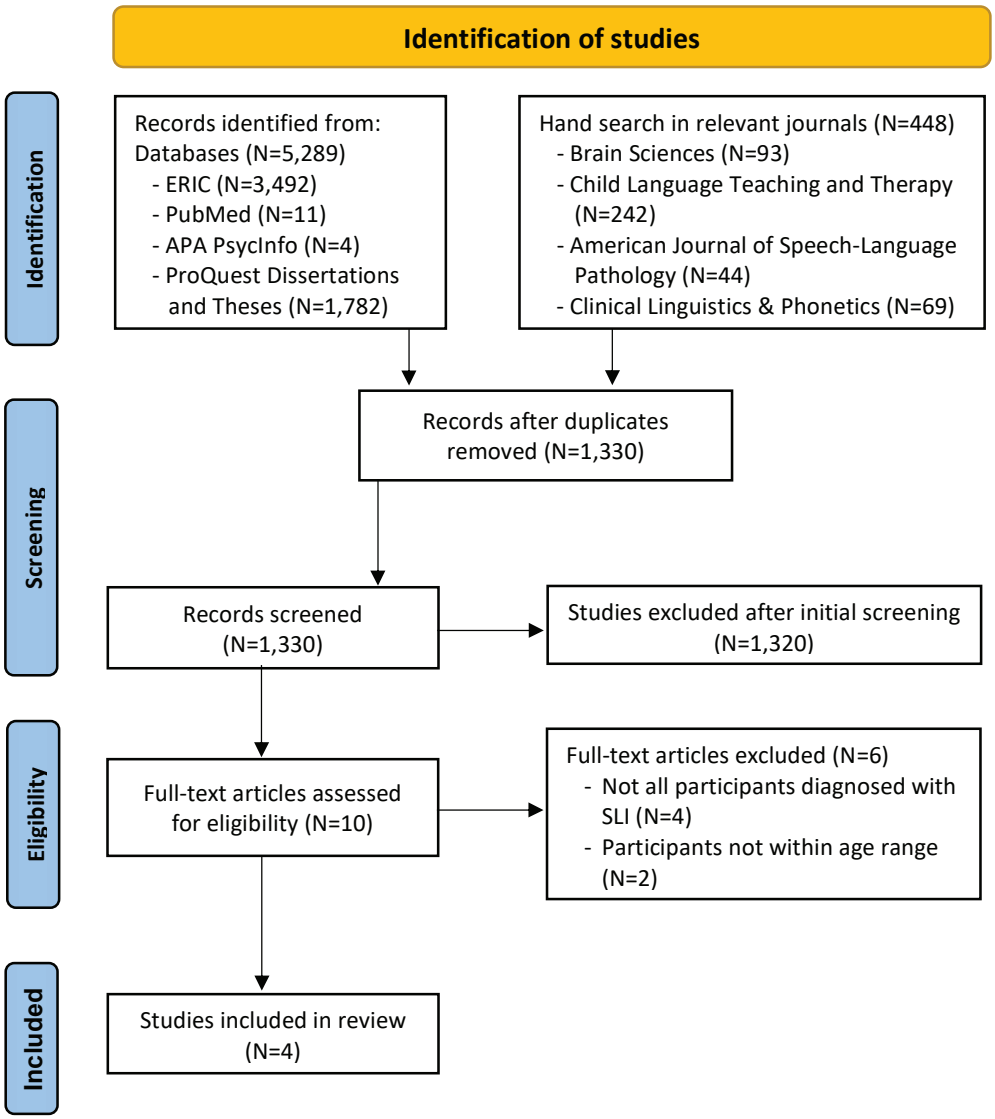


Figure 1. *Methodology flowchart.*

Table 3 demonstrates the included studies. Three studies were conducted in Europe, i.e., the UK,[13-14] and the Netherlands.[15] One study was conducted in Asia, in Iran.[12] According to levels of evidence at Johns Hopkins Nursing Evidence-Based Practice,[19] our review identified three studies using quasi-experimental design (Level II)[12-14] and one study using single case design (Level III).[12]

The present SR involved a total of 94 participants, with sample sizes ranging from 10 to 47. All studies included only children aged between 6 and 12 years. All participants had received a diagnosis of SLI (which can be used interchangeably with developmental language disorder (DLD) or primary language impairment (PLI), which identified by a psychological assessment or a psychiatric diagnosis.[12-15] Severity levels were reported in only one study, ranging from mild to severe SLI.[12]

The studies included in this review examined EF interventions, including WM training programs, including interactive and computerized,[12] WM training interventions, including listening recall training task and odd one out span training task,[13] a computer-based EF training (Braingame Brian),[14] and Cogmed WM training.[15] Speech-language pathologist was clearly reported as an interventionist in one study.[12] However, the researcher and researcher assistant were reported in of Henry *et al.*,[13] Vugs *et al.*,[14] and Holmes *et al.*[15] The duration of sessions ranged from 5 weeks[12] to 8 weeks.[15]

Shahmahmood Toktam *et al.* focused on WM, pSTM, morpho-syntax skills, and executive WM.[12] Henry *et al.* focused on WM tasks and language comprehension.[13] Vugs *et al.* focused on tasks of three trained EFs (visuospatial WM, inhibition, and cognitive flexibility), other untrained neurocognitive functions (verbal WM, attention, planning, and fluency), and parents' and teachers' ratings of EF and behavioral problems.[14] Holmes *et al.* focused on STM, WM, language, and IQ.[15]

Table 3. Included studies.

Author(s) Year Country	Design	Participants	Intervention methods	Interventionist	Duration	Outcomes	Results
Shahmahmood Toktam *et al.* (2018)[12] Iran	A single-subject experimental design	Total: 10 participants Age: 6-8 years olds Persian-speaking children Diagnosis: mixed receptive and expressive language impairment Severity: moderate to severe	1. WM training (interactive and computerized (auditory components of WM training software, persian version of robomemo, and a WM training program)) 2. Language intervention	A speech-language pathologist	15 sessions (three 60-minute sessions per week for five weeks)	- WM - pSTM - Morpho-syntax skills	Direct WM remediation can lead to beneficial changes in WM skills including NWR and BDS and morpho-syntactic language skills.
Henry *et al.* (2022)[13] United Kingdom	A pre-test/ intervention/ post-test/ 9-month follow-up design	Total: 47 participants (the experimental group =24 and the control group =23) Age: 6-10 years olds English-speaking children	WM training interventions (listening recall training task and odd one out span training task)	Researcher	18 sessions over of around 10 min three times a week (total 3 h)	- EWM - WM tasks - Language comprehension	Children who received WM training interventions showed significantly higher WM scores and language comprehension at both time points than children in the active control group.
Vugs *et al.* (2017)[14] Netherlands	A pre-test/ intervention/ post-test/ 6-month follow-up design	Total: 10 participants Age: 8-12 years olds Dutch-speaking children	A computer-based EF training (Braingame Brian)	Research assistant	25 sessions over 6 weeks	Tasks of three trained EFs (visuospatial WM, inhibition, cognitive flexibility) - Other untrained neurocognitive functions (verbal WM, attention, planning and fluency) - Parents and teachers' ratings of EF and behavioral problems	Children with SLI showed significant improvement on a task of cognitive flexibility directly after training, as well as a positive trend for improvement in the visuospatial storage component of WM and inhibition.
Holmes *et al.* (2015)[15] United Kingdom	A pre-test/ intervention/ post-test design	Total: 27 participants (the experimental group =12 and the control group =15) Age: 8-11 years olds English-speaking children	Cogmed working memory Training	Research assistant (Cogmed trainer)	20 45-min sessions for 8 weeks	- STM - WM - Language - IQ	Both children who received the Cogmed WM interventions and children in the control group showed significant post-training gains on visuospatial storage (or visuospatial short-term memory).

The results Shahmahmood Toktam *et al.* showed that direct WM remediation can lead to beneficial changes in WM skills including non-word repetition (NWR) and backward digit span (BDS) and morpho-syntactic language skills.[12] Focusing on WM as the primary target of intervention can improve the children's participation in treatment (attention), as well as their linguistic comprehension and production skills, especially their grammatical skills. Furthermore, the findings showed that language intervention also led to an improved grammatical receptive and expressive functioning of all participants, though it did not notably change the participant's performances in WM tasks, including the word list recall and backward word span tests.

The results of Henry *et al.* found that children who received WM training interventions showed significantly higher WM scores and language comprehension at both time points than children in the active control group.[13] Improvements not only in two trained tasks (direct effects) and six untrained working-memory tasks (near-transfer effects) but also sentence comprehension and receptive grammar assessments. The researchers chose face-to-face delivery because it can enhance children's motivation and focus their attention on input, thereby creating an ideal environment for training participation. This may be especially crucial for the executive-load component of the executive working memory activities assigned to the trained group. This shows that the focus and/or enjoyment of the activity, as well as motivation, maybe more important than the mode of delivery in achieving favorable outcomes. The intervention program demonstrated advantages: it involved short sessions over a short period, caused little disruption in the school day, and was enjoyed by children.

The results of Vugs *et al.* found that children with SLI showed significant improvement on a task of cognitive flexibility directly after training, as well as a positive trend for improvement in the visuospatial storage component of WM and inhibition.[14] At 6-month follow-up, the children performed significantly better on the visuospatial storage component of WM, inhibition, and cognitive flexibility. The study observed significant improvement at the 6-month follow-up in two neurocognitive functions, sustained attention and attention control, which were not part of the program's training. Regarding the behavioral ratings, both parents and teachers reported significantly fewer attention problems. Moreover, parents reported significantly fewer problems with WM and metacognition, thought problems, externalizing behavioral problems, and overall behavioral problems.

The results of Holmes *et al.* found that both children who received the Cogmed WM interventions and children in the control group showed significant post-training gains on visuospatial storage (or visuospatial short-term memory).[15] Children in the LLA (low language ability) group improved significantly on one of the two verbal STM measures (digit span but not word span), although group interaction training was not significant. Low verbal IQ scores were strongly and specifically associated with greater gains in verbal STM. One possibility is that these children's daily practice on several Cogmed tasks that require them to remember the order of spoken information (letters and numbers) may help them come up with simple ways to practice, which in turn improves their verbal STM performance. Following training, children with higher verbal IQs made greater gains in visuospatial short-term memory. Verbal abilities may be critical in developing new strategies to meet the complex demands of visuo-spatial working memory.

Discussion

The first purpose of this SR was to investigate which EF interventions have affected EF skills for school-age children with SLI. The evidence supports the notion that EF interventions could improve EF skills (i.e., visuospatial WM, attention, inhibition, and cognitive flexibility) as well as language skills (i.e., language comprehension and production, particularly grammatical skills) in children with SLI aged 6 to 12 years. This study also indicates that while there are promising outcomes, the effects can be inconsistent and vary depending on the type of intervention and the specific skills targeted.

Two studies found that computer-based EF training, including Braingame Brian[14] and Cogmed WM training[15] improved the trained visuospatial WM of school-age children with SLI. Klingberg *et al.* investigated the effect of computer-based WM training in children with ADHD.[20,21] The results indicated that these children improved in visuospatial WM tasks. Therefore, our SR supports the use of computerized interventions, which could be used in clinical settings for STs to enhance WM.

Furthermore, Holmes *et al.* found that visuospatial WM improved, especially in individuals with higher baseline verbal IQs and non-word repetition scores.[15] This result demonstrated that the complex demands of visuospatial WM tasks required strong verbal skills. Children with strong language abilities may find it simpler to employ verbal labels to recode stimuli such as spatial locations or colors, giving them additional ways to preserve memory items. A meta-analysis of Vugs *et al.* found that there was a correlation between more severe language impairment and greater impairment in visuospatial storage.[4] However, Vugs *et al.* did not report verbal IQs or non-word repetition scores.[14]

Three studies also found that computer-based EF training, including interactive and computerized,[12] Braingame,[14] as well as training interventions including listening recall training tasks and odd one-out span training tasks on paper-based materials, improved attention skills in school-age children with SLI.[13] There were several explanations for this finding. For example, Shahmahmood Toktam *et al.*[12] found that WM training programs indirectly improved attention skills by improving one of the WM-related tasks (i.e., backward digit span, BDS). Likewise, Vugs *et al.* found that EF intervention enhanced two tasks related to neurocognitive functions that were not addressed in the program, such as sustained attention and attention control.[14] Besides, Henry *et al.* found that face-

to-face delivery was used for social engagement, which may increase children's motivation and focus attention on input, creating an environment that is conducive to training adoption.[13]

Interesting, the study of Vugs *et al.* was only one study that investigated the trainability of inhibition and cognitive flexibility.[14] Following the training and during the 6-month follow-up, they noticed a significant enhancement in cognitive flexibility performance, as well as a positive tendency toward improved inhibition. Like Van der Oord *et al.* this study investigated the effect of computerized EF training focused on improving three EFs: WM, inhibition, and cognitive flexibility in children with ADHD.[22] EF intervention utilizing games could enhance children's motivation and cognitive performance. Leading to neuroplastic changes, especially in regions such as the prefrontal cortex, which is essential for inhibition.[23]

Also, Van der Oord *et al.* incorporated gamification aspects to augment children's motivation and potentially amplify their cognitive performance throughout training.[22] Gaming enhances the release of striatal dopamine, which is believed to enhance arousal and cognitive control functions, particularly during EF training.[24] It should be noted that the average performance of the children with SLI in inhibition and cognitive flexibility was lower than that of their peers.[5] Hence, the EF intervention should enhance the capacity for inhibition and cognitive flexibility.

On balance, the findings indicate that several interventions effectively improved EF skills, including computer-based programs (Braingame Brian, Cogmed, and other interactive computerized training) and paper-based materials with listening recall and odd-one-out span tasks. These interventions demonstrated improvements in visuospatial working memory, attention, inhibition, and cognitive flexibility, though with varying consistency.

The second purpose of this SR was to investigate how improvements in EF skills could improve language skills in school-age children with SLI. The evidence presented in the four studies also supports the notion that EF skills are essential for language development and processing, and vice versa.

Henry *et al.* found an improvement in sentence comprehension. Also, it should be noted that this study found an improvement in attention skills in children with SLI.[13] It is possible that EF interventions, such as listening recall training tasks and odd one-out span training tasks on paper-based materials, which improve attention control, can lead to better concentration during sentence processing. This ensures that the individual remains focused on the meaning of the sentence without being sidetracked by irrelevant stimuli, which enhances overall comprehension. In line with Gillam *et al.*[25] and Montgomery *et al.*,[26] the relationship between WM and sentence comprehension is a complex interplay involving various cognitive functions such as fluid reasoning, controlled attention, and long-term memory for language knowledge. In this association, WM acts as a mediator, facilitating the integration and processing of information necessary for understanding sentences. However, Montgomery *et al.*[26] encouraged interventionists to address underlying language skills such as syntax and sentence comprehension directly rather than trying to improve WM through training. They suggested methods for implicit and explicit interventions to reduce WM demands and improve language, respectively.

Shahmahmood Toktam *et al.* found an improvement in linguistic comprehension and production skills, especially their grammatical skills.[12] Also, it should be noted that this study found an improvement in attention skills in children with SLI. Attention, as part of the broader EF system, helps individuals focus on relevant linguistic information, manage distractions, and process language with precision. It is probable that WM training, which improves attention, is the underlying mechanism for language skill growth in response to a range of linguistic intervention activities.[27]

Furthermore, the findings of Shahmahmood Toktam *et al.*[12] are consistent with the findings of Holmes *et al.*[14] Their WM training method for a sample of low-WM children resulted in near transfer to other WM activities and far transfer to mathematics and English. This indicated that there are cross-domain interactions between language and WM, which is consistent with the findings of Ebert and Kohnert and supports cognitive theories of language.[28] However, the study by Henry *et al.*[29] found no significant far-transfer effects on mathematics and reading skills following WM training, except for a notable improvement in reading comprehension. Positive results could suggest that EF skills, such as cognitive flexibility and inhibitory control, play significant roles in language development. Inhibitory control may be crucial for language development because it enables children to concentrate on interpretations of a message.[30] Additionally, it could be necessary for communicative perspective-taking.[31] Cognitive flexibility was a predictor of the narrative structure.[32] It is crucial to consider how these EF skills affect language development. Children may therefore be able to utilize language more flexibly if they have cognitive flexibility. Therefore, the question of whether low-WM children experience near transfer to other WM activities and far transfer to mathematics and English is a complex one, with research providing mixed results. While some studies suggest potential benefits of WM training, others highlight limitations in the transfer effects, particularly in far transfer to academic skills such as mathematics and English.

Overall, the results of the current research, which are consistent with previous reports, indicate that EF interventions have shown potential in improving EF skills in these children. EFs and language are closely intertwined in cognitive processes. It should be the notion that EFs provide the cognitive framework that supports language acquisition, comprehension, and production throughout development and across different contexts. EF skills such as the ability to retain information (working memory), focus and process relevant linguistic information precisely without being distracted (attention), filter out irrelevant information (inhibition), and shift focus between activities (cognitive flexibility) were found to be crucial foundations

for enabling more effective communication.

For school-age children with SLI, STs could integrate EF interventions with traditional language interventions, as language skills have little bearing on EFs. EF skills are essential for language development and processing. Conversely, language acquisition may also improve EFs, indicating a bidirectional relationship.[9-11] Therefore, our finding suggests that STs should not neglect EFs because EF interventions may not only successfully enhance children with SLI but also indirectly affect their language skills. However, more research is needed to better understand the bidirectional relationship between EF and language skills in children with SLI, as well as to provide more tailored speech therapy to these individuals.

Limitations and future research

This present SR has several limitations. First, we only selected articles in English. Other languages might yield complementary results. Secondly, there have only been four recent studies published to date, which included one single-subject design[11] and three quasi-experimental studies.[13-15] The randomized controlled trial study, which is widely regarded as the gold standard of experimental research, was notably absent. Therefore, it should be concerned about the limited sample size, making findings difficult to generalize to the population. Thirdly, the evidence from the four included studies in this review focuses on school-age children with SLI, meaning that the findings are not generalizable to preschool children or adults with SLI. Further research is required to determine the effects of EF interventions on individuals with SLI under 6 or above 12, as well as how improvements in EFs could improve language skills in young and adult children with SLI. Finally, we must recall that STs implemented the interventions,[12] a fact not explicitly stated in some studies.[13-15] Therefore, further research is necessary to define the interventionist clearly for STs to know their role in intervening both in language and EF areas. There should be an understanding of STs' experiences remediating developmental EF impairments with SLI, as well as their confidence and expertise in this domain.

Conclusion

In conclusion, this present SR indicates that EF interventions could enhance both EF and language skills in school-age children with SLI. The research suggests that, although there are encouraging results, the benefits may be variable and depend on the type of intervention and the specific skills addressed. Cognitive processes intricately connect EFs and language. EF skills are assumed to provide the cognitive framework that underpins language acquisition, comprehension, and production throughout development and across various contexts. School-age children with SLI could benefit from EF interventions in conjunction with traditional language interventions. STs must not overlook EFs, as interventions targeting them may not only effectively improve EF skills but also indirectly influence language skills. Further research is needed to better understand the bidirectional relationship between EFs and language skills in children with SLI and to provide more individualized speech therapy for these individuals.

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

None

Ethical approval

The study was approved for exemption review by the Ethics Committee, Faculty of Associated Medical Sciences, Chiang Mai University (CMU), Thailand (AMSEC-66EM-009).

References

[1] Diamond A. Executive functions. Annu Rev Psychol. 2013; 64: 135-68. doi:10.1146/annurev-psych-113011-143750.

[2] Gilbert SJ, Burgess PW. Executive function. Curr Biol. 2008; 18(3): R110-R114. doi: 10.1016/j.cub.2007.12.014.

[3] Shokrkon A, Nicoladis E. The directionality of the relationship between executive functions and language skills: A literature review. Front Psychol. 2022; 13: 848696. doi: 10.3389/fpsyg.2022.848696.

[4] Filipe MG, Veloso AS, Frota S. Executive functions and language skills in preschool children: The unique contribution of verbal working memory and cognitive flexibility. Brain Sci. 2023; 13(3): 470. doi: 10.3390/brainsci13030470.

[5] Veraksa AN, Bukhalenkova DA, Kovyazina MS. Language proficiency in preschool children with different levels of executive function. Psychol. Russ. State Art. 2018; 11(4): 115-29. doi: 10.11621/pir.2018.0408.

[6] Leonard LB. Specific language impairment across languages. Child Dev Perspect. 2014; 8(1): 1-5. doi: 10.1111/cdep.12053.

[7] van der Lely HK. Domain-specific cognitive systems: insight from Grammatical-SLI. Trends Cogn Sci. 2005; 9(2): 53-9. doi: 10.1016/j.tics.2004.12.002. PMID: 15668097.

[8] Flores Camas R, Leon-Rojas JE. Specific language impairment and executive functions in school-age children: A systematic review. Cureus. 2023; 15(8): e43163. doi: 10.7759/cureus.43163.

[9] Estes KG, Evans JL, Else-Quest NM. Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. J Speech Lang Hear Res. 2007; 50(1): 177-95. doi: 10.1044/1092-4388(2007/015).

[10] Vugs B, Cuperus J, Hendriks M, Verhoeven L. Visuospatial working memory in specific language impairment: A meta-analysis. Res Dev Disabil. 2013; 34(9): 2586-97. doi: 10.1016/j.ridd.2013.05.014.

[11] Pauls LJ, Archibald LM. Executive functions in children with specific language impairment: A meta-analysis. J Speech Lang Hear Res. 2016; 59(5): 1074-86. doi:

10.1044/2016_JSLHR-L-15-0174.

[12] Shahmahmood Toktam M, Zahra S, AliPasha M, Ali M, Shahin N. Cognitive and language intervention in primary language impairment: Studying the effectiveness of working memory training and direct language intervention on expansion of grammar and working memory capacities. Child Lang. Teach. Ther. 2018; 34(3): 235-68. doi: 10.1177/0265659018793696.

[13] Henry LA, Christopher E, Chiat S, Messer DJ. A short and engaging adaptive working-memory intervention for children with developmental language disorder: effects on language and working memory. Brain Sci. 2022; 12(5): 642. doi: 10.3390/brainsci12050642.

[14] Vugs B, Knoors H, Cuperus J, Hendriks M, Verhoeven L. Executive function training in children with SLI: A pilot study. Child Lang. Teach. Ther. 2017; 33(1): 47-66. doi: 10.1177/0265659016667772.

[15] Holmes J, Butterfield S, Cormack F, Loenhoud AV, Ruggero L, Kashikar L, Gathercole S. Improving working memory in children with low language abilities. Front. Psychol. 2015; 6: 519. doi: 10.3389/fpsyg.2015.00519.

[16] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. BMJ. 2009; 339: b2700. doi: 10.1136/bmj.b2700.

[17] L Tate R, McDonald S, Perdices M, Togher L, Schultz R, Savage S. Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing the Single-Case Experimental Design (SCED) Scale. Neuropsychol Rehabil. 2008; 18(4): 385-401. doi: 10.1080/09602010802009201.

[18] Barker TH, Habibi N, Aromataris E, Stone JC, Leonardi-Bee J, Sears K, Hasanoff S, Klugar M, Tufanaru C, Moola S, Munn Z. 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.

[19] Newhouse RP, Dearholt SL, Poe SS, Pugh LC, White KM. John Hopkins Nursing Evidence-Based Practice Model and Guidelines. Indianapolis, IN: Sigma Theta Tau International Honor Society of Nursing, 2007.

[20] Klingberg T, Forssberg H, Westerberg H. Training of working memory in children with ADHD. J Clin Exp Neuropsychol. 2002; 24(6): 781-91. doi: 10.1076/jcen.24.6.781.8395.

[21] Klingberg T, Fernell E, Olesen PJ, Johnson M, Gustafsson P, Dahlström K, Gillberg CG, Forssberg H, Westerberg H. Computerized training of working memory in children with ADHD-a randomized, controlled trial. J Am Acad Child Adolesc Psychiatry. 2005; 44(2): 177-86. doi: 10.1097/00004583-200502000-00010.

[22] Van der Oord S, Ponsioen AJ, Geurts HM, Brink ET, Prins PJ. A pilot study of the efficacy of a computerized executive functioning remediation training with game elements for children with ADHD in an outpatient setting: outcome on parent-and teacher-rated executive functioning and ADHD behavior. J Atten Disord. 2014; 18(8): 699-712. doi: 10.1177/1087054712453167.

[23] Knutson B, Fong GW, Bennett SM, Adams CM, Hommer D. A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. Neuroimage. 2003; 18(2): 263-72. doi: 10.1016/s1053-8119(02)00057-5.

[24] Houghton S, Milner N, West J, Douglas G, Lawrence V, Whiting K, Tannock R, Durkin K. Motor control and sequencing of boys with Attention-Deficit/Hyperactivity Disorder (ADHD) during computer game play. Brit J Educational Tech. 2004; 35(1): 21-34. doi: 10.1111/j.1467-8535.2004.00365.x.

[25] Gillam RB, Montgomery JW, Evans JL, Gillam SL. Cognitive predictors of sentence comprehension in children with and without developmental language disorder: Implications for assessment and treatment. Int J Speech Lang Pathol. 2019; 21(3): 240-51. doi: 10.1080/17549507.2018.1559883.

[26] Montgomery JW, Gillam RB, Evans JL. A new memory perspective on the sentence comprehension deficits of school-age children with developmental language disorder: Implications for theory, assessment, and intervention. Lang Speech Hear Serv Sch. 2021; 52(2): 449-66. doi: 10.1044/2021_LSHSS-20-00128.

[27] Gillam RB, Loeb DF, Hoffman LM, Bohman T, Champlin CA, Thibodeau L, Widen J, Brandel J, Friel-Patti S. The efficacy of Fast ForWord language intervention in school-age children with language impairment: A randomized controlled trial. J Speech Lang Hear Res. 2008; 51(1): 97-119. doi: 10.1044/1092-4388(2008/007).

[28] Ebert KD, Kohnert K. Non-linguistic cognitive treatment for primary language impairment. Clinical Linguistics & Phonetics. 2009; 23(9): 647-64. doi: 10.1080/02699200902998770.

[29] Henry LA, Messer DJ, Nash G. Testing for near and far transfer effects with a short, face-to-face adaptive working memory training intervention in typical children. Infant Child Dev. 2014; 23(1): 84-103. doi: 10.1002/icd.1816.

[30] Ye Z, Zhou X. Involvement of cognitive control in sentence comprehension: Evidence from ERPs. Brain Res. 2008; 1203: 103-115. doi: 10.1016/j.brainres.2008.01.090.

[31] Brown-Schmidt S. The role of executive function in perspective taking during online language comprehension. Psychon Bull Rev. 2009; 16(5): 893-900. doi: 10.3758/PBR.16.5.893.

[32] Oshchepkova E, Bukhalenkova D, Veraksa A. The relation between cognitive flexibility and language production in preschool children. InInternational Conference on Cognitive Sciences 2020 Oct 10 (pp. 44-55). Cham: Springer International Publishing. doi:10.1007/978-3-030-71637-0_5.

Journal of Associated Medical Sciences 2025; 58(3): 102-110

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

Medial longitudinal arch collapse of flexible flatfoot during running

Supoj Tangwanicharoensuk and Siriporn Sasimontonkul*

Department of Health and Movement Sciences, Faculty of Sports and Health Science, Kasetsart University, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 26 November 2024
Accepted as revised 6 May 2025
Available online 10 June 2025

Keywords:
Flexible flatfoot; navicular drop; medial longitudinal arch; muscle activation

ABSTRACT

Background: Uncertainty regarding an actual amount of the medial longitudinal arch (MLA) collapse during running among the flexible flatfoot and normal foot individuals has been existed. Additionally, it was unclear whether the alterations of muscle activation and lower limb motion enhanced the MLA collapse in the flexible flatfoot individuals.

Objective: This cross-sectional study aimed to compare the dynamic MLA collapse during running between the flexible flatfoot and the normal foot individuals. It also sought to investigate whether the greater MLA collapse in the flexible flatfoot individuals was related to the alteration of kinematics, kinetics, and muscle activation.

Materials and methods: Sixteen-volunteered males participated in this study. Based on their static navicular drop test, they were assigned into either the flexible flatfoot or the control groups totally 8 participants in each group. They performed barefoot running at a speed of 3.5±0.25 m/s while the three-dimensional motion of the lower body, electromyography and ground reaction force were recorded synchronously. A subject-specific foot model was constructed from the additional reflective markers attached on their feet to determine the maximal dynamic navicular drop (ND_{Dmax}) during the stance phase.

Results: The flexible flatfoot group had greater MLA collapse during midstance and propulsion which coincided with greater hip and knee flexions, and ankle dorsiflexion. The 95% CI of their ND_{Dmax} was in the ranges of 27.18 to 37.68 mm while that of the control group ranged between 20.71 mm and 27.82 mm. The flexible flatfoot group had lesser activations of the hip, knee, and ankle extensors to counteract the flexion moments. They also had larger ankle inversion ($p=0.04$) without greater activation of peroneus longus.

Conclusion: Greater MLA collapse in the flexible flatfoot individuals at midstance and propulsion was associated with greater hip and knee flexions, and ankle dorsiflexion. It might be the consequence of the lower activation of their leg extensors.

Introduction

Flexible flatfoot is a less rigid medial longitudinal arch (MLA) which collapses from weight bearing and recoils back after unloading.[1] The MLA is formed by the calcaneus, talus, navicular, three cuneiform bones, and the first, second, and third metatarsals. The role of MLA is shock absorption which attenuates ground reaction forces transferred to the lower limb.[2,3] Malfunction of plantar intrinsic and extrinsic foot muscles leads to unstable and malalignment of MLA.[4] Reduction of the abductor hallucis muscle activation from nerve blockage[5] and fatigue exercise also resulted in MLA collapse.[6] Greater MLA collapse may

* **Corresponding contributor.**
Author's Address: *Department of Health and Movement Sciences, Faculty of Sports and Health Science, Kasetsart University, Bangkok, Thailand.*
E-mail address: siriporn.s@ku.th
doi: 10.12982/JAMS.2025.078
E-ISSN: 2539-6056

lead to the exercise-induced leg pain[7] and plantar fasciitis in runners.[8]

Nevertheless, the evidence regarding the contribution of muscle function to MLA collapse has been unclarified. There was reported that runners with greater MLA collapse and foot pronation have larger plantar intrinsic foot muscle, the flexor digitorum brevis and flexor digitorum longus.[9] However, the size of their peroneus longus (PL)[9,10] did not differ from that of normal foot runners. Hence, it was likely that other factors besides foot muscle imbalance or overactivation of PL relate to the MLA collapse during running.

Abnormal motion of the proximal joint may enhance the MLA collapse. Nevertheless, there were controversial data[11] from the difference in the study interventions such as footwear condition,[11] locomotion speed,[12-16] foot classification,[17] and foot modelling.[9,12-14] During walking, the FF adolescents[13] and adults[14] had greater plantarflexion, smaller dorsiflexion,[12] and eversion[12] of the tibiotalar joint. In contrast, the low, compared to the high, MLA runners exhibited greater ankle dorsiflexion[15] and knee flexion[16] without the difference in rearfoot eversion during barefoot running. Moreover, the similar static pronated feet presented different dynamic arch heights at the stance phase.[17,18]

Therefore, this study aimed to determine the amount of MLA collapse that occurred during running in the flexible flatfoot and normal foot individuals. Moreover, it also investigated whether the difference in this dynamic MLA collapse between groups was related to the alteration of kinematics, kinetics, and muscle activation.

In this study, the MLA collapse was determined from the navicular height relative to the plantar surface of foot during the stance phase of running using the subject-specific foot model. The difference between groups was analyzed across the entire stance phase of running using Statistical Parametric Mapping. We hypothesized that the flexible flatfoot individuals had greater dynamics MLA collapse compared to normal foot individuals. Furthermore, the alteration of the lower limb motion and muscle activation related to the difference in dynamics MLA collapse between groups.

Materials and methods

Participants

Based on the peak hindfoot inversion from walking trials of flat and normal foot individuals,[12] the effect size of 1.94 and 95% power to detect the difference between group at 0.05 alpha level, the sample size of each group should be at least 7 (G*power 3.1.9.7).[19] Totally 16 male volunteers, aged 18-22 years to account for the foot growth, participated in the study. Inclusion criteria were (1) the right foot dominant and (2) engaging in 10 km running or less, 2-3 times per week for at least 1 year. Exclusion criteria were having (1) any lower extremity injuries or neurological conditions within 6 months prior to the study and (2) leg length discrepancy greater than 10 mm. They were allocated to either the flexible flatfoot (FF) or the control (CON) groups based on their static navicular drop (ND_S). ND_S was the difference in navicular heights measured in sitting and standing positions on the first visit for group allocation. The most prominent part of navicular tuberosity was marked by a trained researcher and used for the navicular height and ND_S evaluations. The test-retest intraclass correlation, ICC [3, 1], from the repeated ND_S equaled 0.895 and its 95% CI was in the range of 0.780-0.955. The standard error of the measurements was 1.15 mm.

ND_S of young adults (24.13±3.41 years of age) with low MLA (ND_S>10) had an excellent correlation with the foot posture index-6, the clinical assessment for foot posture, at 0.818 (Pearson's product moment correlation coefficient, r).[20] The test-retest intraclass correlation (ICC) of ND_S was reported in the range of 0.950-0.955.[20] ND_S of the middle-aged adults (42.7±7.3 yrs) also had moderate to strong correlation with the radiography (r=0.61-0.89).[21] The average ND_S of the healthy male (age 19.6±2.8 yrs; body height of 172.9±5.3 cm) detected from radiography equaled 0.7±0.2 cm.[22] ND_S of the 5-8-year-old children, who were identified as having flexible flatfoot from abnormal radiological findings, equaled 15.56±1.93 mm.[23] However, MLA of children mature with age. ND_S of adults with low MLA (mean 11.83±1.68 mm) strongly correlated with the foot posture index-6 score ≥6 which indicated foot pronation.[20] Moreover, the 2-D dynamic navicular drop, occurred during walking, was in the range of 1.7 to 8.7 mm (95% CI).[24] After the initial ND_S screening of 31 volunteers, therefore, participants who had ND_S greater than 10 mm were assigned to the FF group and those who had ND_S less than 9 mm were assigned to the CON group, total 8 participants in each group. The Research Ethics Committee approved (COA64/044) the study, and the informed consents were obtained from all participants.

Measuring muscle activation

A wireless electromyogram (EMG) system (Noraxon Ultium®, Noraxon USA Inc., Scottsdale, AZ) was synchronized with infrared cameras and force plates to record muscle activation during running at a sampling rate of 2000 Hz. Based on SENIAM (www.seniam.org), adhesive Ag/AgCl surface electrodes (BlueSensor P, Ambu Inc., Malaysia) were attached to skin, with the interelectrode distance of 3.4 cm, above peroneus longus (PL), gastrocnemius medialis (GM), vastus lateralis (VL), vastus medialis oblique (VM), semitendinosus (ST), tensor fasciae latae (TFL), and gluteus medius (GLM) muscles. For tibialis anterior (TA), electrodes were placed at 47.5% between the tip of fibular head and the medial malleolus.[25] Before running, EMGs of each muscle were recorded during the maximal voluntary contraction (MVC) following protocol in table 1. After setting up position, participants were encouraged to perform movement against the resistance from the researcher with their maximal effort for 3-5 seconds, total 3 trials for each muscle testing. A 2-minute rest was allowed between the tests.

Table 1. MVC protocol and testing position for each muscle.

Muscle(s)	Movement	Position
TA	Ankle dorsiflexion with inversion	Seated, ankle plantarflexed to 10 degrees.[25]
PL	Ankle eversion with plantarflexion	Supine, hip, and knee flexed to 0 degrees.
GM	Plantarflexion	Prone, hip and knee flexed to 0 degrees, ankle dorsiflexed to 90 degrees.
VL&VM	Knee extension	Seated, hip flexed to 90 degrees and knee flexed to 60 degrees.[26]
ST	Knee flexion	Prone, knee flexed to 60 degrees.[27]
TFL	Hip abduction with flexion and internal rotation	Semi-prone, hip flexed to 45 degrees, abducted to 30 degrees and knee flexed to 90 degrees.
GLM	Hip abduction with extension and external rotation	Side-lying, hip flexed to 0 degree, abducted to 25 degrees and knee flexed to 90 degrees.

Three-dimensional motion capture and foot modelling

Seven near infrared cameras (Vicon Motion System Ltd., Oxford, UK) captured running motion of participants at 200 Hz. They were synchronized with three force plates (Kistler Corporation, Winterthur, Switzerland) which collected ground reaction forces (GRFs) at 1000 Hz. Retro-reflective markers were attached to the skin of lower extremities following the lower body Plug-in-Gait marker set. An additional foot markers of the foot model were attached at the medial aspect of the 1st metatarsal head (1MH), the lateral aspect of the 5th metatarsal head (5MH), the navicular tuberosity (NAV) and the distal of the calcaneus (DCA) to compute navicular height during running.[28] The root mean square errors of this model for detecting the navicular height of healthy adults over the stance phase of barefoot walking was 0.25 mm compared to wearing a ski boot with rigid marker set.[28] After static calibration, participants performed a 20-meter barefoot running on synthetic track at the speed of 3.5±0.25 m/s. This running speed range was chosen because it was in the average running speed range of amateur runners[29] and the preferred running speed range of recreational active runners.[30] Running speed was monitored by the infrared timers (Fitness Technology, KMS, Australia). After practice, three running trials which the foot landed in the middle of the force plate were captured. Next, kinematic, and kinetic variables were computed using inverse dynamics (Vicon Nexus®, version 2.8.2). The foot model was constructed to determine the dynamic navicular height (NH_D), which was defined as the perpendicular distance from NAV to the plane formed between 1MH, 5MH, and DCA (Figure 1), using customized software (Matlab R2022a, The Mathworks Inc., Natick, USA). The origin of the foot model was located at the calcaneus and the directions of the local X, Y, Z axes were shown (Figure 1). The local coordinate of NAV was determined from its global coordinate and the rotational matric between the global and the local axes. The maximum dynamic navicular drops (ND_{Dmax}), representing maximal MLA collapse, was defined as the difference between NH_D at heel strike and the minimal NH_D at stance phase of running. The percentage change of ND_{Dmax} was the difference between the ND_{Dmax} and ND_S as the percent change from ND_S.

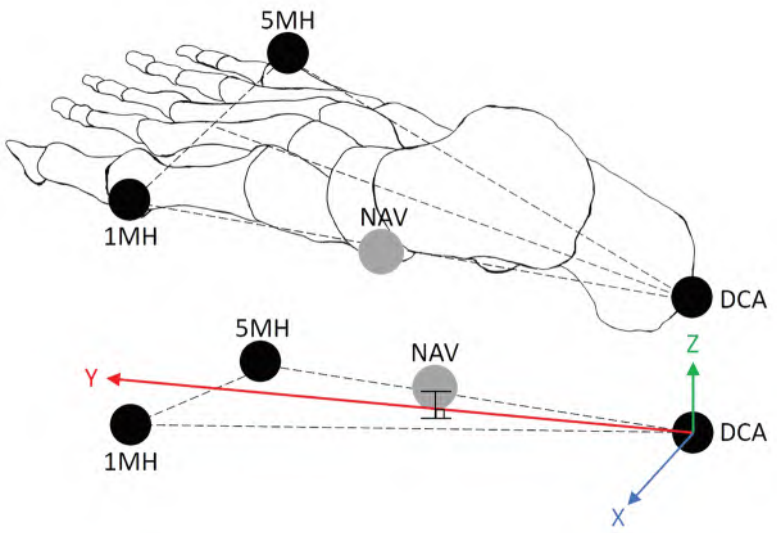


Figure 1. *The dynamic foot model and location of model's markers.*

Measuring strengths of dorsiflexor and plantarflexor

The peak torques of dorsiflexor and plantarflexor were measured using an isokinetic dynamometer (Biodex System 4 Pro, Biodex medical systems, USA) to detect the difference in muscle strength among the groups. The participants sat on the dynamometer chair with backward tilt around 70 degrees. Their right hip and knee were flexed at 110 degrees and 70-80 degrees, respectively, until the tibia was parallel to the ground. Then, the right foot was fixed with the tested footplate, and the lateral malleolus was aligned with the dynamometer's axis of rotation. Next, participants performed 5 repetitions of plantarflexion and dorsiflexion with their maximal effort using the concentric-concentric mode at the speed of 30°/sec. The maximal plantarflexion and dorsiflexion torque were normalized by body weight (%BW) and reported as the plantarflexion and dorsiflexion peak torque. The test-retest intraclass correlation coefficients (ICC) of ankle dorsiflexion and plantarflexion were 0.93 and 0.91, respectively.[31]

Data processing

The EMG signals collected during running were rectified and normalized to their maximal EMG amplitude obtained from MVC. Next, the smoothing technique, a moving root mean square with an exponential weight method was applied to the normalized EMG. Thereafter, it was filtered by a low-pass infinite impulse response filter with the passband frequency below the signal's mean frequency (Matlab R2022a, The Mathworks Inc., Natick, USA). The normalized EMG, NH_D, kinematics and kinetics at each percentage of the stance phase were determined from cubic spline interpolation. In addition, the normalized EMGs at each percent stance were summed up throughout the stance phase to represent total muscle activation, iEMG.

Statistical analysis

Normal distributions of variables were observed using the Shapiro–Wilk test. For the normal distributed variables, the between-group difference and the association between variables were evaluated using the unpaired T-test and Pearson's product moment correlation, respectively. The Mann-Whitney U test was applied to the non-normal distributed variables, the ST iEMG (SPSS for Windows, version 25, IBM Inc., Armonk, NY, USA). Classical statistical parametric mapping (SPM), the unpaired T-test from the open-source package spm1d version M.0.4.11(spm1d.org, © T. Pataky), was deployed on all trials of the time domain data to assess the between group difference at each percent stance using MATLAB (Matlab R2022a, The Mathworks Inc., Natick, USA).[32] The statistical significance was set at $p<0.05$.

Results

There were no between-group differences in running speed, dorsiflexion and plantarflexion peak torques (Table 2).

Table 2. Characteristics of the participants, static navicular drop, running speed and peak torques.

	CON (N=8)		FF (N=8)		*p* value
	Mean±SD	95% CI	Mean±SD	95% CI	
Age (year)	19.75±0.70	19.15-20.34	19.12±1.12	18.18-20.06	0.205
Weight (kg)	61.62±5.68	56.87-66.37	62.16±4.13	58.70-65.61	0.832
Height (cm)	1.70±0.04	1.67-1.74	1.72±0.07	1.66-1.79	0.570
ND_S (mm)	7.13±1.31	6.04-8.23	13.54±1.43	12.34-14.74	0.000**
Running speed (m/s)	3.53±0.11	3.44-3.63	3.47±0.10	3.39-3.56	0.267
T_{DF} (%BW)	48.51±9.74	38.36-54.66	47.68±7.46	41.44-53.92	0.791
T_{PF} (%BW)	103.50±24.03	83.40-123.59	124.81±32.79	97.39-152.22	0.160

*Note: CON: control group, FF: flexible flatfoot, ND_S: static navicular drop, kg: kilogram, cm: centimeter, mm: millimeter, m/s: meter/second, T_{DF}: dorsiflexion peak torque, T_{PF}: plantarflexion peak torque, BW: body weight, ** statistically significant difference at p<0.01 using unpaired T-test.*

There was a negative correlation between the percentage change of ND_{Dmax} and plantarflexion peak torque (r=-0.438, p=0.045, N=16) from the pooled data of both groups (Figure 2A). Comparing between groups, the FF group had lower dynamic navicular height (Figure 2B) which resulted in greater ND_{Dmax} (Figure 2C). The 95% confidence interval (CI) of ND_{Dmax} of the FF group was in the range of 27.18-37.68 mm while that of the CON group was in the range of 20.71-27.82 mm.

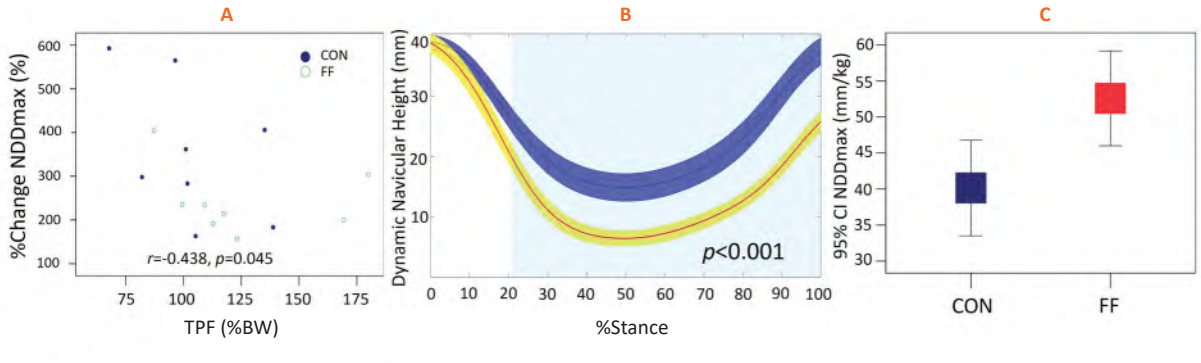


Figure 2. *The maximum dynamic navicular drops (ND_{Dmax}) and the dynamic navicular height (NH_D) during the stance phase. A: correlation between the percentage change of ND_{Dmax} and plantar flexion peak torque (TPF) from the pooled data, B: Ensemble graph of NH_D at each percent of the stance phase, C: The 95% CI of ND_{Dmax}, solid line: mean, error bar: SEM, blue line: control group, red line: FF group, CON: control, FF: flexible flatfoot. The shaded area shows the significant difference between groups using the SPM analysis with unpaired T-test.*

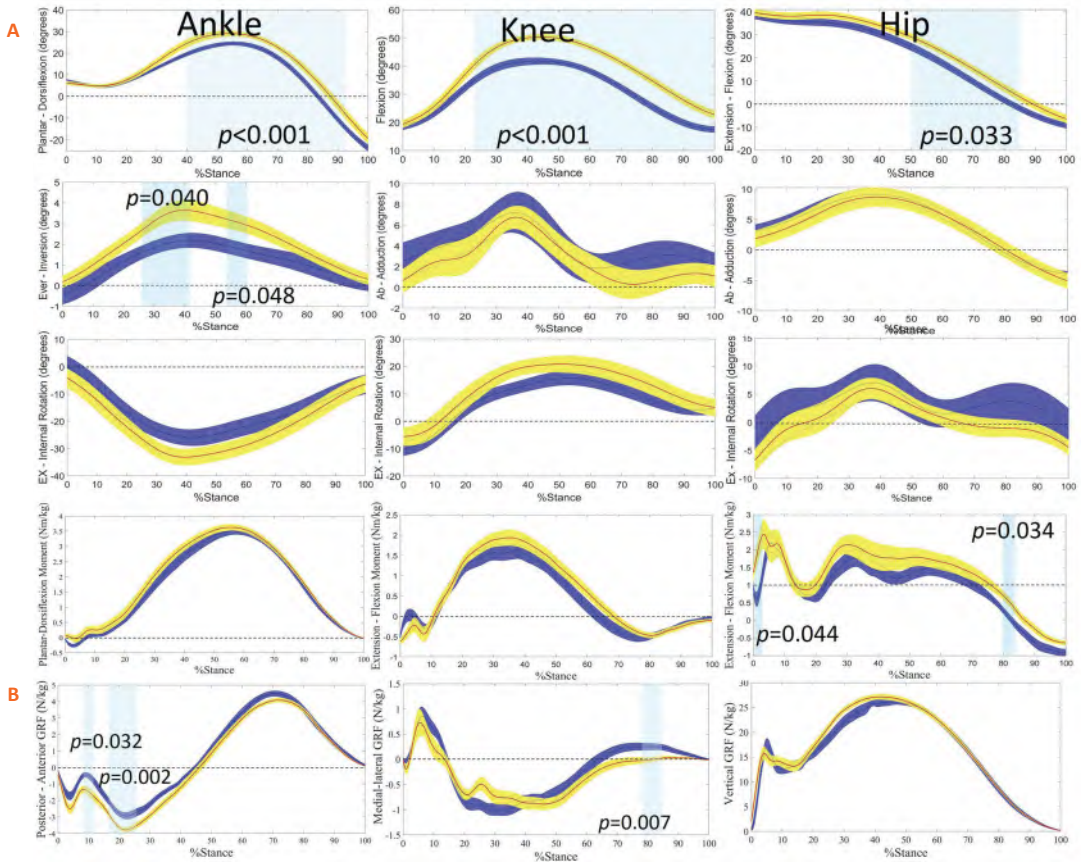


Figure 3. *Ensemble graphs of joint angles, moments, and GRF during the stance phase. A: joint angles and moments, B: GRF were plotted against the stance phase of running. Positive values of joint angle and moment are dorsiflexion/flexion, inversion/adduction, and internal rotation. Positive values of GRF are anterior, lateral and superior, solid lines: mean, error bars: SEM, blue lines: control group, red lines: FF group. Shaded areas show significant difference between groups using the SPM analysis with unpaired T-test.*

The dynamic navicular drop of the FF group was greater than that of the CON group (Figure 2B) during midstance and propulsion (20-100% stance). This incidence coincided with their greater knee flexion angle (Figure 3A). The greater ankle dorsiflexion and hip flexion angles also existed in the FF group during propulsion. Moreover, the FF participants plantarflexed their ankle with a smaller angle during push-off. The FF participants also had a greater hip flexion moment at initial impact and a smaller hip extension moment during push-off. A larger ankle inversion existed in the FF group during the initial and mid propulsions. Nevertheless, there were no between group differences in the adduction-abduction and the internal-external rotation angles of the knee and the hip. Compared to the CON participants, the larger GRF applied to the FF participants posteriorly during the midstance while the smaller GRF applied laterally during push-off (Figure 3B).

The lesser VL, VM, GLM activations (Figure 4) occurred in the FF group at initial propulsion (45-50% stance) and their lesser GM activation also presented at push-off (75% stance). Moreover, TA, GM, GLM and ST iEMGs of the FF group were lesser than those of the normal foot group (Figure 5).

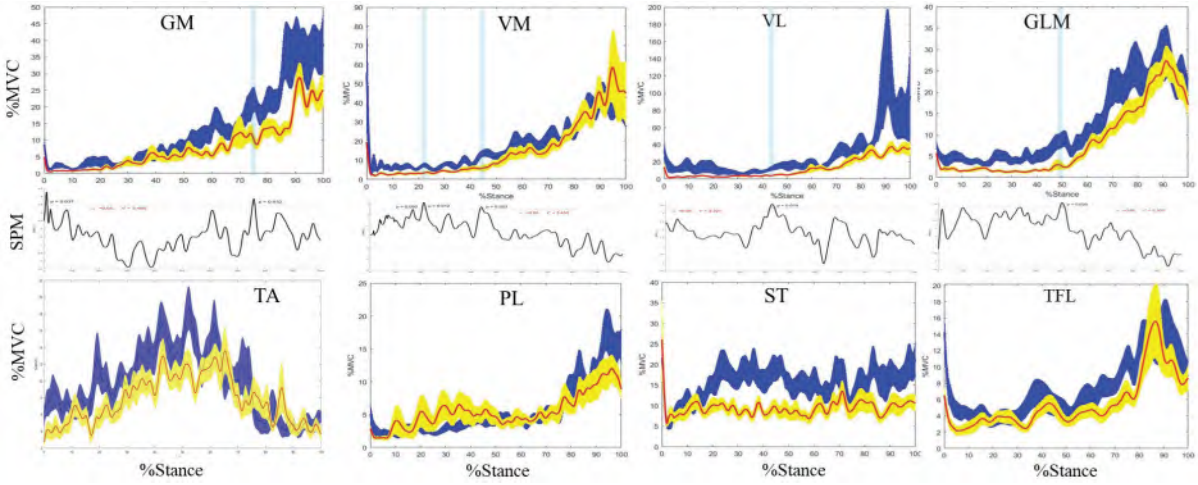


Figure 4. *Ensemble graphs of muscle activation during the stance phase of running. Solid lines: mean, error bars: SEM, blue lines: control group, red lines: FF group. Shaded areas show the significant difference between groups using SPM analysis with unpaired T-test.*

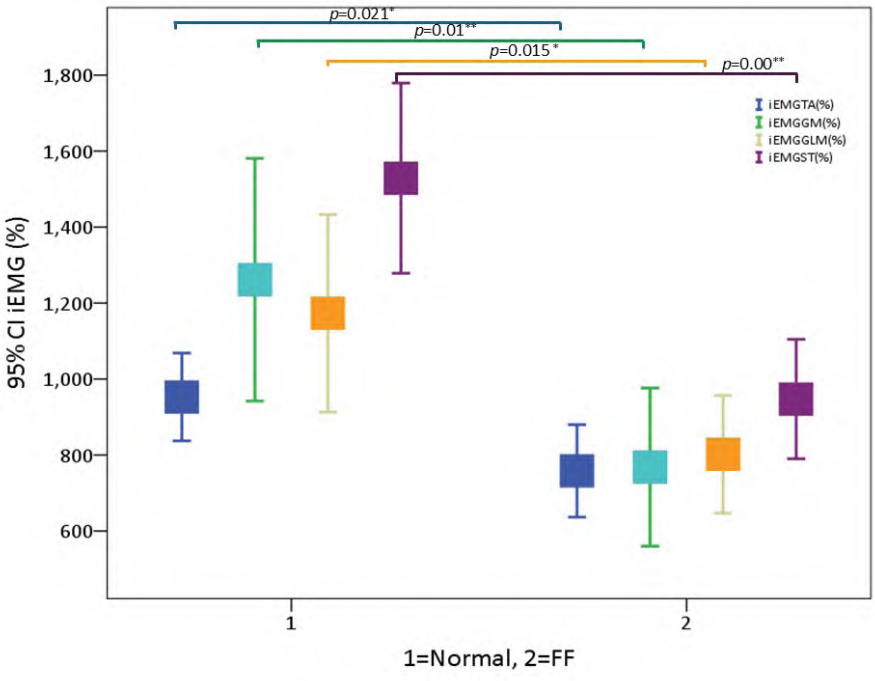


Figure 5. *The 95% CI of iEMG. The differences between groups of TA, GM and GLM iEMGs were assessed by the unpaired T-test while that of ST was detected from the Mann-Whitney U test.*

Discussion

This study investigated whether the alterations of muscle activation and limb motion in the FF individuals associated with the greater MLA collapse during running. Our 3-D foot model supported the hypothesis that the FF participants had greater dynamic MLA collapse during running compared to the normal foot participants (27.18-37.68 mm VS 20.71-27.82 mm, the 95% CI of ND_{Dmax}). Moreover, this ND_{Dmax} from the normal foot's running was greater than the ND_{Dmax} range of 1.7-8.7 mm from 2-D walking analysis.[24] However, running at the same speed as walking also increased NH_D by 3.5 mm.[33] The alteration of lower limb kinematics,[33] an elevation of the foot arch from greater hallux dorsiflexion,[34] the stiffer arch and an increase in MLA angle during propulsion,[35] and a larger fall of MLA with running may relate to the greater NH_D during running.[36] Moreover, our study showed that the greater percentage change from static to dynamic navicular drops during running was correlated with lower plantarflexion peak torque. Therefore, the weakness of plantar flexor could increase MLA collapse during running.

The FF group had greater MLA collapse during midstance and propulsion which coincided with greater ankle dorsiflexion, knee, and hip flexions. It supported the previous finding[15] that FF individuals exhibited greater ankle dorsiflexion during propulsion. However, neither excessive ankle eversion nor substantial PL activation existed in the FF group. In contrast, they had larger ankle inversion during midstance and propulsion. Greater activation of tibialis posterior and lesser activation of PL were also found in the flat-arched, compared to the normal-arched, groups during the propulsion of walking.[37] Therefore, PL may not be a major contributor to the dynamic MLA collapse in FF individuals.

Greater dynamic MLA collapse in the FF group could be the consequence of greater lower extremities collapse as well as an insufficient activation of leg extensors. Their larger hip flexion during propulsion and smaller hip extension moment during push-off resulted from the greater hip flexion moment at initial impact and lesser activation of hip extensors, GLM and ST, respectively. Similarly, the larger posterior GRF at midstance accompanying with the lesser activation of knee extensors, VL and VM also led to the greater knee flexion. The former study also reported the increasing of peak knee flexion from the reduction of knee stiffness in the low arch runners, whereas the activation of knee extensor, VL, raised knee stiffness.[38] In addition, the knee and hip stiffnesses had positive relationship with ankle stiffness.[39] Moreover, the greater dorsiflexion angle occurred in the FF group during propulsion should be the consequence of their lesser plantar flexor activation, GM, because the anterior and vertical GRFs were not significantly different between groups.

The limitation of this study was that skin artifacts could induce marker vibration during running. The skin marker set of the similar MLA model underestimated the MLA deformation about 2 degrees compared to the radiography-based MLA deformation.[40] Hence, the exact amount of navicular drop occurs during running could be a bit greater than the computed navicular drop of our study.

Conclusion

The FF participants had greater MLA collapse during running which associated with the lower activation of leg extensors to counteract the hip, knee, and ankle flexions at midstance and propulsion. Greater hip flexion moment at the initial impact and the lesser activations of GLM, and ST enhanced larger hip flexion in the FF group. The larger posterior GRF and lesser activations of VL and VM resulted in their greater knee flexion. Lesser GM activation in the FF participants related to greater dorsiflexion even though both FF and normal foot groups had comparable plantar flexor strength. Nevertheless, greater percentage change of dynamic from static MLA collapse was correlated with lower plantarflexion peak torque.

Funding

This research is supported by the Graduate School, Kasetsart University

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

I wish to acknowledge the Department of Health and Movement Sciences, Kasetsart University, for providing all facilities.

References

[1] Atik A, Ozyurek S. Flexible flatfoot. North Clin Istanb. 2014; 1(1): 57-64. doi: 10.14744/nci.2014.29292.

[2] McPoil TG, Knecht HG. Biomechanics of the foot in walking: A function approach. J. Orthop Sports Phys Ther. 1985; 7(2): 69-72. doi: 10.2519 jospt.1985.7.2.69.

[3] Franco AH. Pes Cavus and Pes Planus: Analyses and treatment. Phys Ther. 1987; 67(5): 688-94. doi: 10.1093/ptj/67.5.688.

[4] McKeon PO, Hertel J, Bramble D, Davis I. The foot core system: a new paradigm for understanding intrinsic foot muscle function. Br J Sports Med. 2015; 49(5): 290. doi: 10.1136/bjsports-2013-092690.

[5] Fiolkowski P, Brunt D, Bishop M, Woo R, Horodyski M. Intrinsic pedal musculature support of the medial longitudinal arch: an electromyography study. J Foot Ankle Surg. 2003; 42(6): 327-33. doi: 10.1053/j.jfas.2003.10.003.

[6] Headlee DL, Leonard JL, Hart JM, Ingersoll CD, Hertel J. Fatigue of the plantar intrinsic foot muscles increases navicular drop. J Electromyogr Kinesiol. 2008; 18(3): 420-5. doi: 10.1016/j.jelekin.2006.11.004.

[7] Reinking MF. Exercise-related leg pain in female collegiate athletes: the influence of intrinsic and extrinsic factors. Am J Sports Med. 2006; 34(9): 1500-7.doi: 10.1177/0363546506287298.

[8] Wiegand K, Tandy R, Freedman Silvernail J. Plantar fasciitis injury status influences foot mechanics during running. Clin Biomech (Bristol). 2022; 97:

105712.doi: 10.1016/j.clinbiomech.2022.105712.

[9] Zhang X, Aeles J, Vanwanseele B. Comparison of foot muscle morphology and foot kinematics between recreational runners with normal feet and with asymptomatic over-pronated feet. Gait Posture. 2017; 54: 290-4. doi: 10.1016/j.gaitpost.2017.03.030

[10] Okamura K, Hasegawa M, Ikeda T, Fukuda K, Egawa K, Kanai S. Classification of medial longitudinal arch kinematics during running and characteristics of foot muscle morphology in novice runners with pronated foot. Gait Posture. 2022; 93: 20-5. doi: 10.1016/j.gaitpost.2022.01.006.

[11] Hollander K, Zech A, Rahlf AL, Orendurff MS, Stebbins J, Heidt C. The relationship between static and dynamic foot posture and running biomechanics: A systematic review and meta-analysis. Gait Posture. 2019; 72: 109-22. doi: 10.1016/j.gaitpost.2019.05.031.

[12] Kruger KM, Graf A, Flanagan A, McHenry BD, Altiok H, Smith PA, *et al.* Segmental foot and ankle kinematic differences between rectus, planus, and cavus foot types. J Biomech. 2019; 94: 180-6. doi: 10.1016/j.jbiomech.2019.07.032.

[13] Caravaggi P, Sforza C, Leardini A, Portinaro N, Panou A. Effect of plano-valgus foot posture on midfoot kinematics during barefoot walking in an adolescent population. J Foot Ankle Res. 2018; 11: 55. doi: 10.1186/s13047-018-0297-7.

[14] Phan CB, Lee KM, Kwon SS, Koo S. Kinematic instability in the joints of flatfoot subjects during walking: A biplanar fluoroscopic study. J Biomech. 2021; 127: 110681. doi: 10.1016/j.jbiomech.2021.110681.

[15] Powell DW, Williams DS 3rd, Windsor B, Butler RJ, Zhang S. Ankle work and dynamic joint stiffness in high-compared to low-arched athletes during a barefoot running task. Hum Mov Sci. 2014; 34: 147-156. doi: 10.1016/j.humov.2014.01.007.

[16] Williams DS, McClay IS, Hamill J, Buchanan TS. Lower extremity kinematic and kinetic differences in runners with high and low arches. J Appl Biomech. 2001; 17(2): 153-63. doi: 10.1123/jab.17.2.153.

[17] Tweed JL, Campbell JA, Avil SJ. Biomechanical risk factors in the development of medial tibial stress syndrome in distance runners. J Am Podiatr Med Assoc. 2008; 98(6): 436-44. doi: 10.7547/0980436.

[18] Langley B, Cramp M, Morrison SC. Clinical measures of static foot posture do not agree. J Foot Ankle Res. 2016; 9: 45. doi: 10.1186/s13047-016-0180-3.

[19] Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007; 39(2): 175-91. doi: 10.3758/bf03193146.

[20] Zuil-Escobar JC, Martínez-Cepa CB, Martín-Urrialde JA, Gómez-Conesa A. Evaluating the medial longitudinal arch of the foot: Correlations, reliability, and accuracy in people with a low arch. Phys Ther. 2019; 99(3): 364-72. doi: 10.1093/ptj/pzy149.

[21] Hannigan-Downs K, Harter R, Smith G. Radiographic validation and reliability of selected clinical measures of pronation. J Ath Tr. 2000; 35: S12.

[22] Akiyama K, Noh B, Fukano M, Miyakawa S, Hirose N, Fukubayashi T. Analysis of the talocrural and subtalar joint motions in patients with medial tibial stress syndrome. J Foot Ankle Res. 2015; 8: 25. doi: 10.1186/s13047-015-0084-7.

[23] Žukauskas S, Barauskas V, Čekanauskas E. Comparison of multiple flatfoot indicators in 5-8-year-old children. Open Med (Wars). 2021; 16(1): 246-56. doi: 10.1515/med-2021-0227.

[24] Nielsen RG, Rathleff MS, Simonsen OH, Langberg H. Determination of normal values for navicular drop during walking: a new model correcting for foot length and gender. J Foot Ankle Res. 2009; 2: 12. doi: 10.1186/1757-1146-2-12.

[25] Sacco IC, Gomes AA, Otuzi ME, Pripas D, Onodera AN. A method for better positioning bipolar electrodes for lower limb EMG recordings during dynamic contractions. J Neurosci Methods. 2009; 180(1): 133-7. doi: 10.1016/j.jneumeth.2009.02.017.

[26] Harput G, Ulusoy B, Akmese R, Ergun N. Comparison of muscle activation levels and knee valgus between individuals with medial patellofemoral ligament reconstruction and healthy individuals during fatiguing stepdown task. Clin Biomech (Bristol). 2020; 78: 105067. doi: 10.1016/j.clinbiomech.2020.105067.

[27] Worrell TW, Karst G, Adamczyk D, Moore R, Stanley C, Steimel B, *et al.* Influence of joint position on electromyographic and torque generation during maximal voluntary isometric contractions of the hamstrings and gluteus maximus muscles. J Orthop Sports Phys Ther. 2001; 31(12): 730-40. doi: 10.2519/jospt.2001.31.12.730.

[28] Eichelberger P, Blasimann A, Lutz N, Krause F, Baur H. A minimal marker set for three-dimensional foot function assessment: measuring navicular drop and drift under dynamic conditions. J Foot Ankle Res. 2018; 11: 15. doi: 10.1186/s13047-018-0257-2.

[29] Lourenço TF, Nunes LAS, Martins LEB, Brenzikofer R, Macedo DV. The performance in 10 km races depends on blood buffering capacity. J Athl Enhanc. 2019; 8: 1. doi: 10.4172/2324-9080.1000312.

[30] Orendurff MS, Kobayashi T, Tulchin-Francis K, Tullock AMH, Villarosa C, Chan C, *et al.* A little bit faster: Lower extremity joint kinematics and kinetics as recreational runners achieve faster speeds. J Biomech. 2018; 71: 167-75. doi: 10.1016/j.jbiomech.2018.02.010.

[31] Tuominen J, Leppänen M, Jarske H, Pasanen K, Vasankari T, Parkkari J. Test-retest reliability of isokinetic ankle, knee and hip strength in physically active adults using Biodex System 4 Pro. Methods Protoc. 2023; 6(2): 26. doi: 10.3390/mps6020026.

[32] Serrien B, Goossens M, Baeyens JP. Statistical parametric mapping of biomechanical one-dimensional data with Bayesian inference. Int Biomech. 2019; 6(1): 9-18. doi: 10.1080/23335432.2019.1597643.

[33] Pohl J, Jaspers T, Ferraro M, Krause F, Baur H, Eichelberger P. The influence of gait and speed on the dynamic navicular drop - A cross-sectional

study on healthy subjects. Foot (Edinb). 2018; 36: 67-73. doi: 10.1016/j.foot.2018.04.001.

[34] Shono H, Matsumoto Y, Kokubun T, Tsuruta A, Miyazawa T, Kobayashi A, *et al.* Determination of relationship between foot arch, hindfoot, and hallux motion using Oxford foot model: Comparison between walking and running. Gait Posture. 2022; 92: 96-102. doi: 10.1016/j.gaitpost.2021.10.043.

[35] Stolwijk NM, Koenraadt KL, Louwerens JW, Grim D, Duysens J, Keijsers NL. Foot lengthening and shortening during gait: a parameter to investigate foot function?. Gait Posture.2014; 39(2): 773-7. doi: 10.1016/j.gaitpost.2013.10.014.

[36] Caravaggi P, Pataky T, Günther M, Savage R, Crompton R. Dynamics of longitudinal arch support in relation to walking speed: contribution of the plantar aponeurosis. J Anat. 2010; 217(3): 254-61. doi: 10.1111/j.1469-7580.2010.01261.x.

[37] Murley GS, Menz HB, Landorf KB. Foot posture influences the electromyographic activity of selected lower limb muscles during gait. J Foot Ankle Res. 2009; 2: 35. doi: 10.1186/1757-1146-2-35.

[38] Williams DS 3rd, Davis IM, Scholz JP, Hamill J, Buchanan TS. High-arched runners exhibit increased leg stiffness compared to low-arched runners. Gait Posture. 2004; 19(3): 263-9. doi: 10.1016/S0966-6362(03)00087-0.

[39] Akl AR, Baca A, Richards J, Conceição F. Leg and lower limb dynamic joint stiffness during different walking speeds in healthy adults. Gait Posture. 2020; 82: 294-300. doi: 10.1016/j.gaitpost.2020.09.023.

[40] Caravaggi P, Rogati G, Leardini A, Ortolani M, Barbieri M, Spasiano C. Accuracy and correlation between skin-marker based and radiographic measurements of medial longitudinal arch deformation. J Biomech. 2021; 128: 110711. doi:10.1016/j.jbiomech.2021.110711.

Journal of Associated Medical Sciences 2025; 58(3): 111-121

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

Edge-based AI approach for blood vessel segmentation in coronary X-ray angiography

Mohd Osama[1,2] and Rajesh Kumar[1*]

[1]*Department of Electronics and Communication, University of Allahabad, India.*
[2]*United Institute of Technology, Prayagraj, India.*

ARTICLE INFO

Article history:
Received 3 March 2025
Accepted as revised 26 May 2025
Available online 6 June 2025

Keywords:
Computer-aided diagnosis (CAD), machine learning for medical imaging, feature selection, cardiovascular disease, Sorbel edge detection, Prewitt edge detection, Robert's edge detection, Canny edge detection.

ABSTRACT

Background: An organizational report indicates that heart attacks lead to seventy percent of human fatalities. Heart-related diseases strike people in India who range between the ages of 30-60 years. X-ray coronary angiography functions as the key procedure for detecting these conditions. The manual process of heart vessel segmentation by cardiologists becomes slow and needs significant effort because different professional skill levels affect the consistency of their output results.

Objective: A study proposes automatic coronary angiography segmentation through artificial intelligence analysis of edge features to accurately detect the main cardiovascular artery system edges.

Materials and methods: The Mendeley public database contained 100 patient images for training purposes and 34 images for validation purposes. The VGG Image Annotator tool served to create binary masks for annotation purposes. The analysis incorporated traditional edge detection methods that included Sobel, Prewitt, and Roberts along with Canny.

Results: The tested model obtained 99% accuracy alongside a positive predictive value (PPV) of 96% and Sensitivity of 94% and Dice Coefficient of 95%. The upcoming research will focus on developing soft computing approaches for detecting stenosis in segmented images.

Conclusion: The method demonstrates better performance metrics that show superior capability to previous techniques implemented in this field. New studies are needed to analyze soft computing techniques to identify vascular structures in coronary angiographic images.

Introduction

Cardiovascular Disease (CVD) is a situation in which arteries that supply blood to the heart muscles become narrowed or blocked due to fatty deposits of plaque. Cholesterol and fat-like substances may be responsible for forming plaques. Plaques reduce blood flow or cause complete blockage of arteries. When the heart does not receive enough blood that contains oxygen due to narrowed or blocked arteries, it may result in chest pain or serious complications such as heart attacks. Cardiovascular disease (CVD) is to be more responsible for a significant amount of mortality globally and is considered the cause of fatalities.[1] Invasive coronary angiography (ICA) remains the gold standard for diagnosing coronary vessel disease.

However, its effectiveness can be influenced by patient-specific anatomical variations and image quality dependency. Despite its widespread use, previous studies have struggled to achieve optimal performance and accuracy in predictive analysis. Accurate assessment

* **Corresponding contributor.**
Author's Address: *Department of Electronics and Communication, University of Allahabad, India.*
E-mail address: rajeshkumariitbhu@gmail.com
doi: 10.12982/JAMS.2025.079
E-ISSN: 2539-6056

of the degree of vessel stenosis plays a crucial role in guiding treatment management and improving patient outcomes. The formation of coronary artery plaque leads to the development of the condition. Plaque is a substance composed the cholesterol, calcium, fatty deposits, and other materials present in the human blood. Over time, these fatty deposits lead to the narrowing (stenosis) and hardening (sclerosis) of the artery walls, a condition known as atherosclerosis.

This process reduces the elasticity of the arteries and restricts blood flow, which can limit the supply of oxygen and nutrients to the heart muscle. If the plaque ruptures, it can trigger the formation of blood clots, further obstruct the blood flow and potentially lead to serious conditions such as heart attack, or other types of coronary disease (CAD).[2] The plaques are detected by some physical observation of heart abnormalities by medical practice. Abnormalities of the coronary arteries can lead to severe consequences, including myocardial infarction, heart failure, or sudden cardiac arrest. In recent years, substantial efforts have been directed towards enhancing diagnostic methods in cardiology to facilitate early detection and treatment, ultimately improving patient survival rates and quality of life.[3] The medical term stenosis defines a lumen constriction that reduces coronary artery blood flow to the heart tissue. The accumulation of plaque leads to arterial narrowing that produces an irritated condition known as myocardial infarction or heart attack. The spread of arterial plaque causes inflammation in the proximal side wall which raises the heart attack danger.

Figure 1A shows the first image describes the catheter injected from the leg and guided to aorta, that means catheter is typically inserted into a blood vessel through a small incision, most commonly in the femoral artery located in the leg. This minimally invasive procedure is known as catheterization. The catheter is a thin, flexible tube that is carefully guided through the vascular system using real-time imaging techniques, such as fluoroscopy (X-ray guidance). Figure 1B shows catheter tip stopped at artery Positioning the catheter at this location is critical for identifying any blockages, narrowing (stenosis), or other abnormalities that may impair blood flow to the heart muscle and for planning potential interventions like angioplasty or stenting. Figure 1C shows the die injected into the arteries for stenosis in left coronary artery. During the procedure, a contrast dye is often injected through the catheter to enhance the visibility of blood vessels on imaging using coronary angiography, helping physicians assess the extent of disease or abnormalities and guide the intervention precisely. There are some popular approaches for imaging tests such as X-ray angiography, CT scan, and MRI scan systems. The manual observation and detection of blood vessels in coronary angiography images leads to subjectivity, time consumption, and variability depending on the observer's expertise and experience. Therefore, automatic detection of coronary vessels is needed for accurate, fast, and objective detection and diagnosis.

Nowadays, people rely more on machines than on human beings. Therefore, artificial intelligence plays a vital role in medical diagnosis. There are various AI techniques that integrate medical image analysis using digital image processing and computer vision. At this stage, X-ray angiography imaging is taken of a segmented part of the arteries. Segmentation techniques performed by experienced human professionals often yield better results than machines. Research institutions utilize artificial intelligence approaches with computer-aided systems and digital image processing to boost blood vessel identification in coronary angiography picture segmentation practices. The two main approaches for segmentation methods are divided into supervised methods and unsupervised procedures. The image structure serves as the sole foundation for segmentation methods because the analysis depends on intensity and gradient information. The methods edge detection, thresholding, deformation and graph cuts become tools to recognize image boundaries for object definition.[5,6] Multiple machine-learning models exist for both segmenting and reconstructing the structure of coronary arteries. The techniques used for this purpose include edge

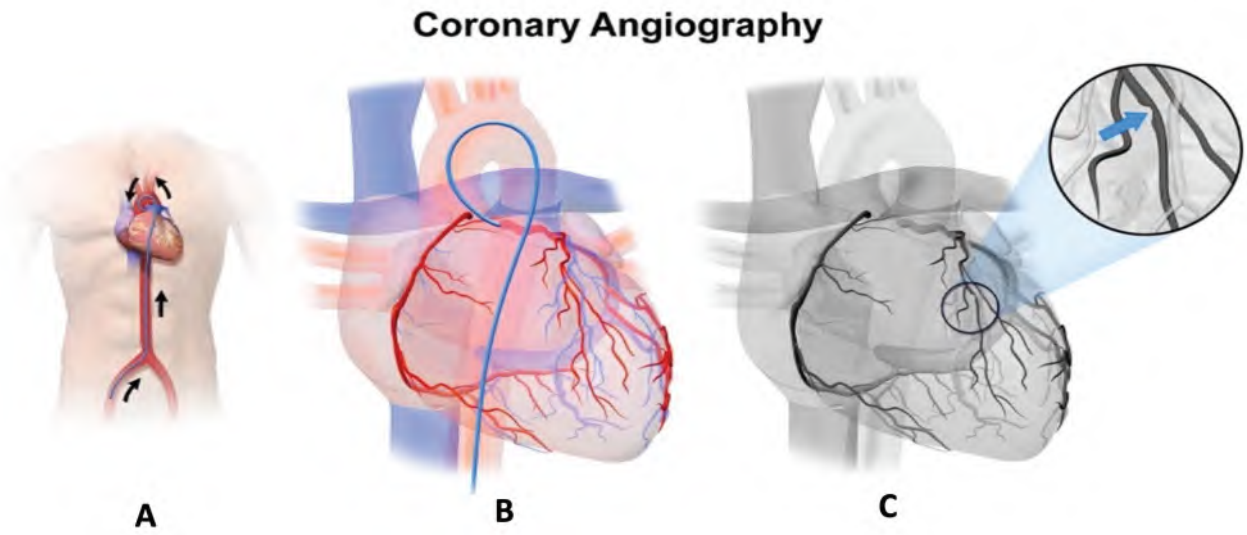


***Figure 1**. Coronary angiography depicting blood vessels.[4]*

detection, region-based analysis, tracking techniques, feature extraction and learning-based methods as well as others.[7,8] In this study, we concentrate on feature-based machine learning techniques to detect blood vessels in coronary angiography images. The methods use self-learning abilities to extract information from historical data to enhance the precision of segmentation results.[9,10]

The algorithms enhance the segmentation process while shortening the computation time required. Especially in X-ray heart angiography image segmentation, vessel boundary detection is obscured by the low contrast of images and noisiness, besides the overlap of structures.[11] The main challenges are the necessity of many annotated images and variability in imaging conditions that impede generalization ability. Most previous methods failed to handle the information on a fine level with handcrafted features and were prone to overfitting.[12] The various contributions to this manuscript are as follows:

1. **Problem identification:** Outlined the problems of manual segmentation of coronary angiography images, which is affected by the variation in expertise of different cardiologists and time-consuming processes.
2. **Integration of traditional edge detection algorithms:** Traditional edge detection algorithms like as Sobel edge detector, Prewitt edge detector, Robert's edge detector, and Canny edge detector have been employed in extracting dominating edges of a cardiovascular artery system from angiography images.
3. **Artificial intelligence application**: It incorporated a Random Forest algorithm in selecting and refining the strong edges detected by the traditional edge detection methods, increasing the accuracy of segmentation.

The hypothesis of coronary angiography images segmentation using automatic techniques of edge-based feature analysis using Random Forest, would offer a means to achieve such an effective result as more prominence and consistent results, while at the same time fewer man-hours spent on manual segmentations.[13] These methods are expected to provide higher consistency and overcome the variability and time burden of cardiologists performing the segmentation.

The article focuses on study of coronary angiography and section I is an introduction to coronary angiography. In section II we review published research on segmentation of cardiovascular images, specifically, feature based machine learning approaches. In section III we describe our system and spend a great deal of time discussing our machine learning model. In section IV there is detail of the edge detection technique which is used to recognize edges. section V presents experimental results and performance evaluation of the model. In the final section both research and findings of the study are summarized.

Literature review

The study proposes Early diagnosis becomes vital for preventing significant health risks associated with cardiovascular diseases at a global level because of their broad prevalence. The ECG tool functions as a main diagnostic instrument, yet existing diagnosis processes miss accurate interpretations. The combination of BiAE technology that unites BiLSTM cells with autoencoders makes ECG signal features easier to extract. The new algorithm uses its abilities to detect important elements that guide accurate classification. Among the ten machine learning models, SVM obtained 96% accuracy for classifying ECG signal types while categorizing them into five groups. Using AI for ECG diagnosis helps medical staff diagnose patients more accurately while also speeding up the process through improved efficiency and decreased operator mistakes. New advancements make it possible to discover CVDs at an earlier stage and identify them better. The application of artificial intelligence in ECG signal analysis will transform patient diagnosis as well as cardiac medical care procedures.[14]

Diagnosis of atherosclerosis and stroke risk heavily relies on accurate carotid plaque segmentation together with classification methods in ultrasound imaging context. Current techniques maintain either complex procedures or neglect task correlations which results in diminished performance levels. RCCM-Net represents a new multi-task learning framework which unites RCM for region confidence analysis with CCM for sample category confidence management to boost segmentation together with classification results. RCM provides probabilistic assessment of plaques for classification purposes followed by CCM which creates categorical sample weights to improve segmentation. The performance of RCCM-Net surpassed both standalone and multi-task approaches when analyzing 1270 ultrasound images because it reached 85.82% accuracy in classification plus 84.92% Dice-similarity coefficient. An ablation study demonstrated that RCM together with CCM achieved their target purpose. The results demonstrate that RCCM-Net shows promise for medical examination of carotid artery plaques. Through this method doctors would achieve better early identification of atherosclerosis and stroke prevention treatment opportunities.[15]

Medical staff must perform manual corrections to traditional edge detection algorithms in quantitative coronary angiography, resulting in reduced accuracy. AngioPy improves segmentation precision through added features that let users mark ground-truth points which minimizes the requirement for manual input. AngioPy demonstrated excellent performance with an F1 score of 0.927 after its evaluation of 2455 images from FAME 2 before validating its results on 580 images. Measurements of vessel dimensions between AngioPys and Medis QFR® revealed outstanding agreement when r was 0.96 at $p<0.001$ level. Minimum luminal diameter analysis through the model produced an excellent correlation score of r=0.93. The reliability of AngioPy as a coronary segmentation system is supported by these research results. Manual corrections are not required because this system improves the efficiency of QCA. The technology

demonstrates the capability to transform diagnostic and clinical operational models related to coronary disease diagnosis.[16]

Cardiovascular diseases act as the main death-causing factors which require prompt diagnosis together with precise identification. The process of feature engineering required by traditional machine learning hinders effectiveness when dealing with complex datasets. DL models feature automatic feature abstraction while delivering better efficiency along with accuracy levels. The research describes the creation of EnsCVDD-Net and BlCVDD-Net as DL models for CVD classification through the combination of LeNet with GRU and MLP structures. The selection of optimal data balancing methods together with feature selection algorithms helps improve model operational efficiency. The EnsCVDD-Net model reached a performance mark of 88% accuracy; however, BlCVDD-Net exceeded the competitive models with a remarkable 91% accuracy. SHAP Additive exPlanations provides interpretation through its capability to assess feature contribution levels. The proposed frameworks represent a beneficial technique for enhancing cardiovascular disease screening accuracy.[17]

Personalized medical interventions become necessary because cardiovascular diseases exist with multiple complex characteristics. Predictive analysis is improved through AI and ML techniques to enhance diagnosis and treatment methods. The researchers applied statistical approaches together with AI/ML techniques to discover vital transcriptomic markers for CVD predictive purposes. Genes were evaluated using three statistical methods that ended in feature prioritization procedures before employing ML methodology. Soft voting ensemble was used as the ML model framework to achieve 96% accuracy performance. The research identified eighteen important biomarkers which were verified through clinical record comparisons. Biomarkers function as initial warning signals for detecting CVD. The AI-based prediction system provides healthcare professionals with a dependable methodology for patient-specific care.[18]

Progressive diagnostic tools are crucial for medical care because cardiovascular disease remains a primary global health challenge. The necessity of automated cardiac MRI analysis involves supervised learning approaches that require significant labeled data. The research describes a partially supervised Strong-Teacher Consistency Network which segments multi-class cardiac MRI structures by utilizing unlabeled data. The model employs a multi-teacher framework, hybrid loss functions, and virtual adversarial training for robust learning. The model achieves advanced accuracy in its performance evaluation using the MM-WHS and ACDC datasets relative to other existing models. The system obtains a 90.14% success rate on MM-WHS tasks and 78.45% accuracy on ACDC while using limited labeled data. The methodology provides better performance than both fully supervised and single-teacher systems. The innovation increases automated cardiovascular diagnostic accuracy when working with small, labeled datasets.[19]

The authors proposed solving the traditional boundaries encountered when performing vascular hemodynamic evaluations manually in clinical settings. The research team created a deep learning method to process CT images through automatic segmentation followed by vessel reconstruction and prediction of computational fluid dynamics (CFD). A CNN design update performed segmentation, and the MC algorithm did 3D reconstruction tasks. The Res2-CD-UNet model demonstrated superiority over previous methods because it achieved a 92.76% accuracy rate on aortic-artery data while attaining 94.57% accuracy on lower-limb artery data sets. The model achieved better relative geometric errors and computed results at a much faster pace. The software used OpenFOAM to conduct CFD simulations that delivered improved efficiency. The method enables physicians to view accurate arterial hemodynamics while maintaining fast speeds and operability. The application of this approach demonstrates high promise to enhance medical diagnostic systems for blood vessels.[20]

The left ventricular wall identification process in myocardial perfusion scintigraphy images gets improved thanks to U-Net convolutional neural networks for detecting coronary artery diseases. An analysis of 83 clinical exams with 4,150 images showed that the AI segmentation yielded an 87% dice coefficient (DC) together with an intersection over Union of 0.8 compared to conventional methods. Communication science validated the model both internally and externally to meet clinical practice requirements. The integration of artificial intelligence results in improved accuracy and efficiency when performing diagnoses in myocardial imaging procedures. This development will help researchers conduct future studies with the focus on artifact correction.[21]

The research introduces SCS-SLSP as a semi-supervised cardiac image segmentation framework which combines SPG modules with SL theoretical principles. The soft labeling mechanism of SL identifies in-distribution pixels as well as uncertain areas and SPG enhances the extraction of vital features through its modules. The Hard Uncertain Pixel Mining (HM) module achieves higher accuracy through two methods which include trustworthy pseudo-label allocation and hard case processing using contrastive learning. The SCS-SLSP system demonstrated superior performance than current methods when working with restricted amounts of labeled data through tests on three open-source datasets. The dice score increased by 0.41% for the ACDC dataset which used 10% labeled data while reducing segmentation errors substantially. By using this approach, both hospital operations can get improved segmentation results with better accuracy and efficiency.[22]

This work presents CapNet which represents a light deep-learning system designed for cardiac MRI segmentation to reduce the excessive operational costs of transformer-based models. The model uses both convolutional structures and mixing mechanics to perform efficient training operations using fewer computational parameters. The adoption of attention mechanisms allows the model to adapt to different cardiac shape variations.

Tversky Shape Power Distance loss serves as a new loss function which boosts segmental accuracy in the process. Open-source assessments conducted on three public datasets reveal that CapNet achieves exceptional dice similarity scores better than present-day standard models. A statistical review verifies its successful performance even with decreased computing demands.[23]

The research presents an AI-based system for cardiac ultrasound image segmentation which lowers requirements for large, annotated dataset dependency. The system makes use of artificial intelligence generative capabilities to generate multi-class RGB masks which directly segment heart structures. The novel implementation of conditional generative adversarial networks (CGAN) produces better accuracy because it includes conditional inputs and paired RGB masks. The methodology achieves superior performance than state-of-the-art models during evaluations on three separate datasets. This technique allows superior segmentation while being less sensitive to noise. The method helps automate cardiac imaging operations while minimizing expenses alongside dependence on human experts. [24]

Materials and methods

1. **Dataset**: Our research utilized a dataset comprising images of heart blood vessels along with their corresponding annotated vessel masks. This publicly available dataset was sourced from Mendeley and included a dataset consisting of 134 angiography images out of which 100 angiography images were used for training and 34 images for validation purpose.[25] The binary masks were annotated using the VGG Image Annotator tool. All images underwent a preprocessing step, including cleaning, before being used for segmentation.
2. **Methods**: In the field of medical imaging, identifying and analyzing targets can be highly challenging, particularly when dealing with low-quality images of blood vessels in coronary angiography, which often suffer from poor pixel resolution and inadequate color schemes. The proposed technique involves several steps, including data preprocessing and feature-based segmentation. Segmentation is performed by selecting features through machine learning techniques. In this study, we propose a segmentation model that applies machine learning algorithms to segment blood vessels in X-ray coronary angiography images.

This model begins by processing the data using Random Forest methodology while performing feature selection activities. The results undergo machine learning model analysis to conduct the assessment. This methodology provides a more efficient solution for diagnosing coronary artery blockages, reducing the time required by cardiologists or physicians. The segmentation of heart blood vessels involves several steps using different types of filters. As shown in Figure 2, the process starts by taking coronary angiography images of blood vessels as input. These images undergo preprocessing using a Gabor filter, which smooths the images while minimizing noise interference. The following image demonstrates different edge detection techniques (Canny, Sobel, Prewitt and Roberts) used after preprocessing. After appropriate segmentation of coronary angiography blood vessels, the Random Forest algorithm selects essential heart image features. The system performs training subsequently and its output results are assessed alongside visualization processing. Edge detection techniques serve as fundamental tools for accurate heart vessel segmentation processes.

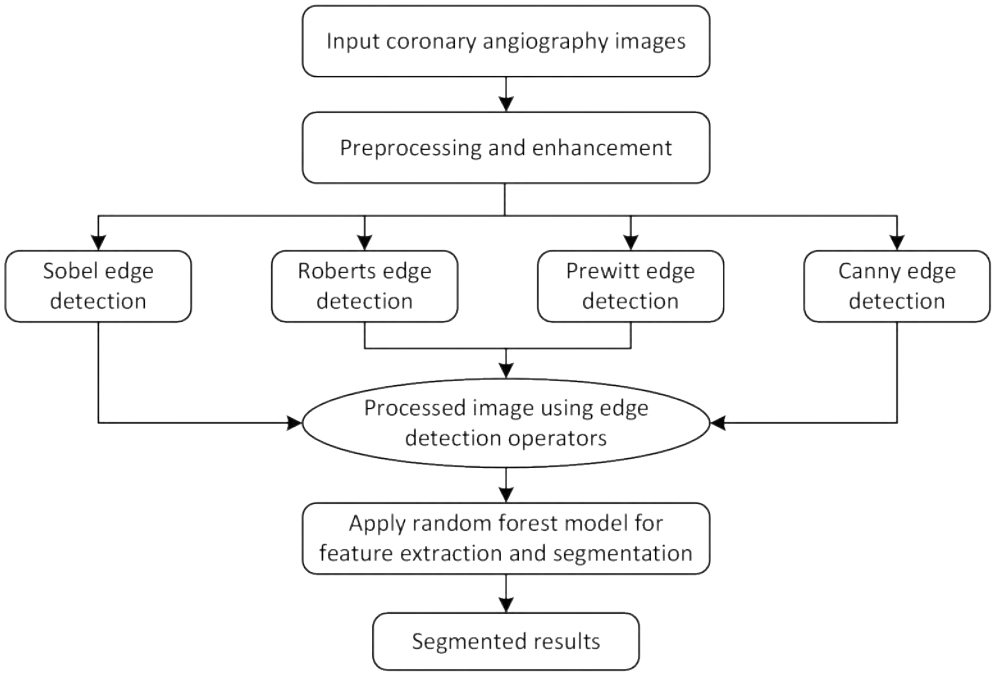


Figure 2. *Proposed model for vessel segmentation.*

Canny edge detection algorithm

The Canny edge detection technique is a technique in the literature for detecting edges in an image, and it is applied to different tasks such as image segmentation. This technique balances two major performance criteria. First is effective edge detection and the second is precise edge localization, which means that the edge points are as close as possible to the true edges, and each edge has only one response. The algorithm processes edges through gradient analysis which examines intensity changes and directions of gradient alterations.[26] There is various process in the algorithm will be as follows as smoothing of image that reduces noise in the image for better accuracy in detecting edges. Computation of gradient based on calculation of the magnitude and direction of change in intensity. Lastly, double thresholding which classifies the edges into strong and weak concerning the level of intensity. Strong edges are of utmost importance in creating a continuous chain of edges, which is the main factor in segmenting regions of interest in an image. The algorithm connects the pixels along the strong edges to help in the separation and segmentation of meaningful areas in the image.

Sobel edge detection

The detection of edges in images, commonly known as Sobel edge detection, is one of the techniques designed for such purposes. The procedure basically works by convolving an image with what is called a Sobel kernel-special matrices that enhance the difference in intensity between neighboring pixels. It yields a map of gradient magnitudes that then show the edges of an image.[27] This output is generally used as the base for further image segmentation processes in which the edges outline the borders that demarcate regions within the image.

Roberts edge detection

The Roberts edge detection algorithm is the simplest method for detecting edges in an image. This is like the 2-D gradient operator on grayscale images. The algorithm processes edges through gradient analysis which examines intensity changes and directions of gradient alterations.

$$[G] = [G_x] + [G_y] \quad (1)$$

[Gx] and [G_y] describe the gradients in the x and y directions, respectively. These outcomes after magnitude map, highlighting edges of an image. This information about the gradient can easily be used for segmentation using the detected edges to delineate different regions within the image.[28] However, care must be taken that in comparison with other operators such as Sobel or Canny, the Roberts edge-detection algorithm is less reliable. Due to its smaller kernel size, which has a limited detailed information of image.

Prewitt's edge detection

The basic yet effective edge detection method called Prewitt's enhances image edge detection for analysis purposes. The Prewitt Kernel works as a 3×3 matrix component of this algorithm to detect intensity changes between adjacent image pixels. The edge detection process determines partial derivative values through gradient values that extend in horizontal (Gx) as well as vertical (Gy) directions. Equations allow the determination of these gradient values.

$$G_x = [A_2+cA_3+A_4] - [A_0+cA_7+A_6] \quad (2)$$

$$G_y = [A_6+cA_5+A_4] - [A_0+cA_1+A_2] \quad (3)$$

Here, c is a constant that assigns greater importance to pixels nearer to the center of the mask. The operation produces a gradient magnitude, which helps in identifying edges within the image. Further, it can be used in edge information on image segmentation for which the edges define regions between different portions within the image.[29] Owing to its larger kernel size and better capability in capturing edge details of an image, the Prewitt method usually would be more reliable compared to the Roberts method. However, it still will not be considered as robust as other advanced algorithms in edge detection techniques, whereas Sobel and Canny methods, use advanced techniques for the same operation in images.

Random Forest classifier (RF)

The Random Forest classifier serves the dual purpose of feature selection and classification. For blood vessel segmentation, it selects important features such as color, texture, etc., from the input data, which are then used by a segmentation algorithm to do the actual segmentation.[30] Following is a sequence of actions when a classifier is applied:

Feature selection: Running Random Forest algorithm on the training data with extracted features; assessing the importance of features depending on how decision trees have been split.

Segmentation: The selected features become the input for a segmentation algorithm that performs the segmentation, either by traditional image processing methods or by machine learning techniques.

Evaluation: Measure the outcomes of the segmentation algorithm in terms of performance metrics such as confusion matrices.

The classification performance will improve when Random Forest classifier feature selection methods work together with appropriate segmentation methods.

Performance metrics

The performance assessment of the classification model includes accuracy measurements taken during training and testing operations. A confusion matrix provides an evaluation of the predicted data against real dataset labels. The classification outcomes consist of four cases where correct classifications receive labels true positive and true negative but incorrect ones become false positive or false negative. The assessment of the classification model performance relies mainly on the

positive predictive value (PPV) together with Sensitivity as the commonly utilized metrics in these experiments.

***Positive Prediction Value (PPV)*:** the model, in essence, quantifying how sure it is that such a prediction of positive is actually positive - that is, the number of true positives compared against the total sum of actual true and false positives.[31] PPV can be calculated in Equation (4).

$$\text{Positive Prediction Value} = \frac{TP}{TP+FP} \quad (4)$$

***Sensitivity*:** Sensitivity determines whether the model detects genuine positive situations accurately. The ratio between authentic positive outcomes and all actual positive results yields sensitivity measurement. This calculation excludes only true positive (TP) and false negative (FN) from the count.[32] Equation (5) states the mathematical formula for sensitivity calculation.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (5)$$

***Dice Coefficient (DC)*:** The dice coefficient functions as an evaluation measure for binary classification models. This measure tracks how well the model identifies positive cases without excessive numbers of incorrect positive or negative classifications. The model's performance effectiveness receives a balanced assessment through a harmonic mean calculation involving accuracy and sensitivity results.[33] The performance evaluation of the DC occurs through Equation (6).

$$\text{Dice Coefficient} = 2 \times \frac{(\text{Positive Prediction Value} * \text{Sensitivity})}{(\text{Positive Prediction Value} + \text{Sensitivity})} \quad (6)$$

***Accuracy*:** A model achieves accuracy when it predicts correctly for every output it generates. Equation (7) provides a method to calculate accuracy.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

Confusion metrics used to check the performance of the generated model through machine learning on various levels of confidence.

Results

This section showcases the model's experimental results. In this study, the predicted masks were assessed by comparing them with the annotated masks for evaluation. The model is based on a segmentation approach that utilizes feature extraction from coronary angiography blood vessel images. Different edge detection methods, involving Sobel, Roberts, Prewitt, and Canny, play an important role in enhancing the features of an image. Each operator may be used in different ways to extract edges or boundaries in coronary angiography images to ensure that all details concerning the feature under study are captured. These processed images highlight blood vessels and other critical structures for highly informative features toward which the subsequent machine learning model shall be working. The use of multiple operators in the system enhances diversity among the extracted edge information, promoting the robustness of segmentation. The edges of vessels in coronary angiography were identified using Sobel, Roberts, Prewitt, and Canny edge detection methods. Then, an AI-driven Random Forest model was applied to identify and select the most optimal edges from the results of these algorithms. In this process, an image is analyzed using a Random Forest model-based method of feature extraction concerning segmentation. In this feature selection, only the most relevant information of the image content, like vessel or other boundary edges, is kept while noise is reduced and segmentation accuracy improved. Such features make the Random Forest model highly effective for non-linear relationships and handling high-dimensional data. The model then segments the image into meaningful regions once it extracts the relevant features, hence isolating main elements like blood vessels while discarding less important background regions.

An AI-based technique demonstrates its ability to segment coronary angiography blood vessels in heart patients as presented in Figure 2. The segmented blood vessels demonstrate high correspondence to their corresponding annotated ground truth images. The performance evaluation of this proposed method utilizes accuracy alongside DC, PPV, and sensitivity which stem from confusion matrix calculations.

Figure 3 indicates the segmented image using an angiography image and the ground truth images. Performance of blood vessel segmentation from coronary angiography images using the Random Forest-based model. The three columns represent, respectively, the original grayscale images, the annotated ground truth masks by experts, and the predicted segmentation by the model. The accuracy is 99% with a dice coefficient of 95%, indicating a large overlap between the vessel areas predicted by the model and the ground truth images. In fact, a PPV of 96% depicts that most of the vessels predicted by the network are correctly classified with only a tiny number of false positives, while a sensitivity of 94% underlines how strong the model is in terms of detecting most vessel areas. As illustrated in the figure, though there are slight mistakes, in general the predicted results match the annotated masks very well, A confusion matrix developed with segmentation results enables researchers to determine the total counts of true positive, false positive, true negative and false negative outcomes. The collected information aids both verification and accuracy enhancement for segmentation procedures.

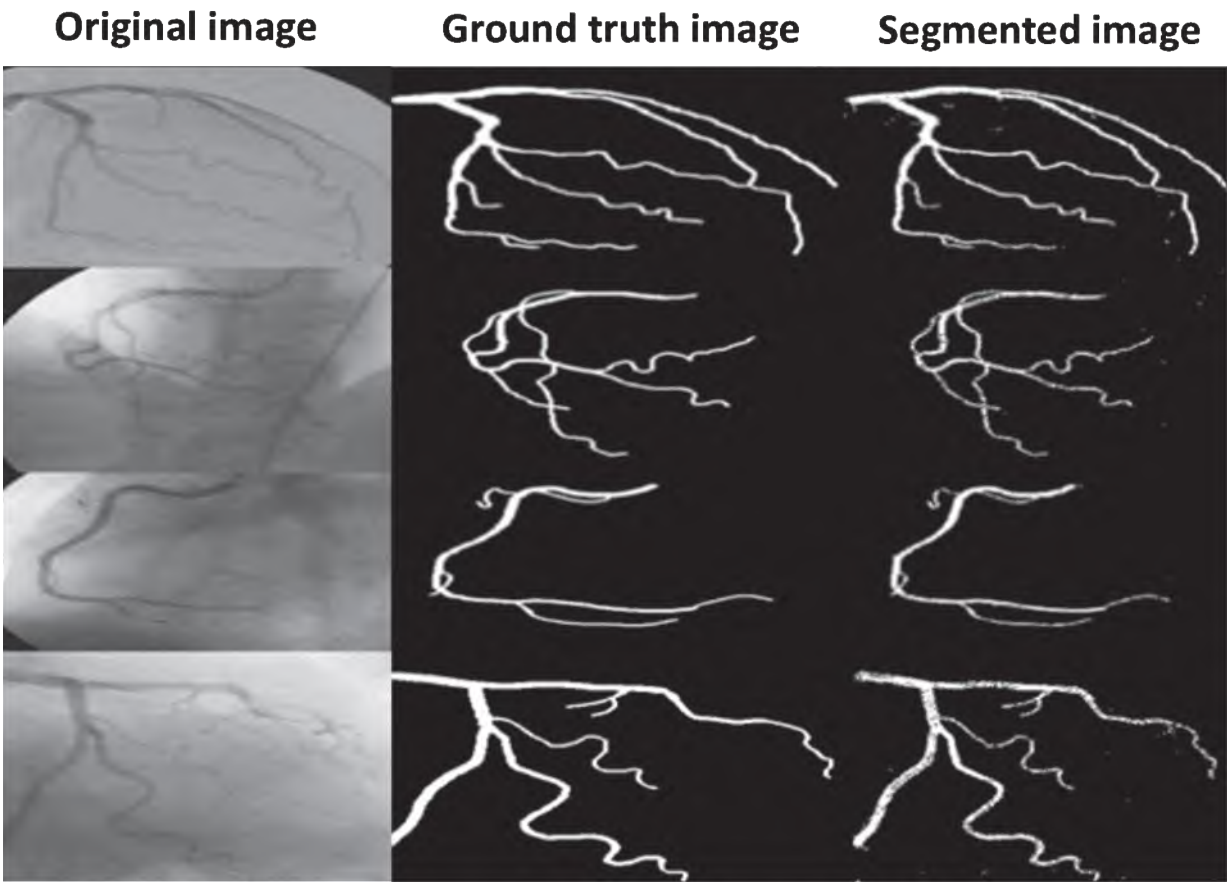


Figure 3. *Outcome of segmented blood vessels in coronary angiography.*

Table 1 clearly demonstrates that various performance metrics of the Artificial Intelligence based model, such as accuracy, PPV, DC, and sensitivity, have been analyzed in the context of the current work involving edge-based feature selection and the AI model. Several evaluation metrics have been computed and compared against other existing models with datasets.

Table 1 highlights the use of edge detection techniques such as Sobel, Roberts, Prewitt, and Canny algorithms for vessel segmentation. The experimental results demonstrate the performance of the proposed model in terms of PPV, sensitivity, DC, and accuracy, comparing favorably to other existing models. From Table 1, it is evident that the predicted results for blood vessel segmentation in coronary angiography images deliver promising outcomes. In the proposed model, a total of 134 images were utilized for training and testing. The research designated 80% of images for training purposes and reserved 20% for validation. Specialized AI features were integrated into the Random Forest model by the research team to obtain precise target region localization.

Table 1. Comparison table of predicted results.

Methods	Authors	Material	PPV	Sensitivity	DC	Accuracy
Deep learning (MTNet)	Ruochen Liu, Song Gao *et.al*.[35]	Retinal blood vessels	80%	98%	88%	96%
Machine learning (supervised method)	Hyungjoo Cho,June-Goo Lee, *et al*.[36]	Angiography images	84%	80%	82%	82%
DL (PSPNet)	Xiliang Zhu, Zhaoyun Cheng, *et. al*.[37]	Coronary angiography image segmentation	86%	94%	89%	95%
Deep learning (U-Net)	Dongxue Lian *et.al*.[38]	Coronary angiography images	84%	87%	83%	98%
Edged based feature using AI	Our's	Coronary angiography Images	96%	94%	95%	99%

Note*: PPV: positive predictive value, DC: dice coefficient.*

The following table compares various approaches applied to image-based datasets in medical imaging with respect to metrics such as PPV, sensitivity, DC, and accuracy. A performance study of each technique adopted using different datasets is conducted—retinal blood vessels, angiography images, and coronary angiography images. The proposed methodology, marked as "Our's," is the edge-based feature extraction approach which has been empowered by a Random Forest model on coronary angiography images. Conclusively, the proposed approach outperforms other methods on most metrics. Having identified the image of retinal blood vessels with a deep learning model, the first approach of MTNet reached 80% PPV, with 98% sensitivity, 88% DC, and accuracy at 96%. MTNet, while performing exceptionally in sensitivity to identify true positive cases, performed lower in PPV, probably due to increased false positives[34]. This trend underlines the need for a model characterized by better balance in terms of all metrics. In the case of angiography images, regarding the overall performance of machine learning supervised techniques, the PPV is 84%, sensitivity is 80%, DC is 82%, and accuracy is 82%. This approach has given a moderately good performance; however, the sensitivity resulting from this approach is relatively low, which means failing to detect some positive cases.[35] Moreover, the DC illustrating the overlap between the predicted and ground truth segmentations is low, which suggests possible drawbacks in the accuracy of segmentation. The deep learning-based image segmentation approach in coronary angiography, PSPNet, gives a PPV of 86%, sensitivity of 94%, DC of 89%, and accuracy of 95%. Compared to all the earlier-described techniques, this technique gives much better sensitivity and overall segmentation accuracy.[36] Hence, it signifies higher capacity for detection and classification of features in medical images. However, a slightly lower PPV against "Our's" indicates that there is still room for improvement in this technique to get better precision. using U-Net as the architectural framework for examining images from coronary angiography. The accuracy, precision, recall and specificity of their approach were 84%, 87%, 83% and 98%, respectively.[37] Edge-based feature extraction using AI shows much better performance than their methods.

The method in the last row, "Our's," represents an edge-based feature extraction technique by employing a Random Forest model on coronary angiography images. It outperformed well with a PPV of 96%, sensitivity of 94%, DC of 95%, and accuracy of 99%. Compared to all other techniques, the "Our's" method was superior for most metrics with the highest PPV, DC, and accuracy. With the high PPV, the method seems to minimize false positives effectively, making it more accurate. Its sensitivity is also compared with PSPNet at 94%, hence showing very good capability for the detection of true positives. Moreover, the DC has indicated excellent overlap between predicted and ground truth segmentations at 95%, hence highly reliable for correct analysis of images. Thus, at the highest accuracy of 99%, it is also quite robust and has the potential for clinical medical use.

Data availability statement

The dataset used in this experiment is publicly available on the Mendeley website.

Conclusion

Based on the results and analysis, the experimental findings prove that the developed protocol executes segmentation tasks in coronary angiography images with superior effectiveness. The proposed method produced segmentation results with accuracy reaching 99% and dice coefficient at 95% and positive predictive value at 96% and Sensitivity at 94%. The method's superior scores confirm that it ranks above former techniques known to work in this domain. Research in this field should investigate potential soft computing methods for classifying coronary angiography vascular structures.

Ethical approval

There was no human or animal ethics violated in this experiment.

Funding

None

Conflict of interest

There is no conflict of interest for this manuscript.

Credit authorship contribution statement

The first two authors designed and implement the core idea, while the third author provide expert insights on the experimental results and analysis.

References

[1] Di Cesare M, Perel P, Taylor S, Kabudula C, Bixby H, Gaziano TA, et al. The heart of the world. Global Heart. 2024; 19(1): 11. doi: 10.5334/gh.1288.

[2] Abubakar M, Irfan U, Abdelkhalek A, Javed I, Khokhar MI, Shakil F, et al. Comprehensive quality analysis of conventional and novel biomarkers in diagnosing and predicting prognosis of coronary artery disease, acute coronary syndrome, and heart failure, a comprehensive literature review. J Cardiovasc Transl Res. 2024; 17(6): 1258-85. doi: 10.1007/s12265-024-10540-8.

[3] Ren P, He Y, Guo N, Luo N, Li F, Wang Z, et al. A deep learning-based automated algorithm for labeling coronary arteries in computed tomography angiography images. BMC Med Inform Decis Mak. 2023; 23(1): 249. doi: 10.1186/s12911-023-02332-y.

[4] Modric J. Coronary Artery Disease Evidence: Diet, Exercise - eHealthStar [Internet]. eHealthStar - Evidence Based Health Articles. 2016. Available from: https://www.ehealthstar.com/conditions/coronary-heart-disease.

[5] Shukur BS, Mijwil MM. Involving machine learning techniques in heart disease diagnosis: a performance analysis. International J Electr Comput Eng. 2023; 13(2): 2177. doi: 10.11591/ijece.v13i2.pp2177-2185.

[6] Vij R, Arora S. A hybrid evolutionary weighted ensemble of deep transfer learning models for retinal vessel

segmentation and diabetic retinopathy detection. Comput Electr Eng. 2024; 115: 109107. doi: 10.1016/j.compeleceng.2024.109107.

[7] Duan X, Sun Y, Wang J. ECA-UNet for coronary artery segmentation and three-dimensional reconstruction. Signal Image Video Process. 2023; 17(3): 783-9. doi: 10.1007/s11760-022-02288-y.

[8] Nobre Menezes M, Silva JL, Silva B, Rodrigues T, Guerreiro C, Guedes JP, et al. Coronary X-ray angiography segmentation using Artificial Intelligence: a multi-centric validation study of a deep learning model. Int J Cardiovasc Imaging. 2023; 39(7): 1385-96. doi: 10.1007/s10554-023-02839-5.

[9] Wang G, Zhou P, Gao H, Qin Z, Wang S, Sun J, et al. Coronary vessel segmentation in coronary angiography with a multi-scale U-shaped transformer incorporating boundary aggregation and topology preservation. Phys Med Biol. 2024; 69(2): 025012. doi: 10.1088/1361-6560/ad0b63.

[10] Arefinia F, Aria M, Rabiei R, Hosseini A, Ghaemian A, Roshanpoor A. Non-invasive fractional flow reserve estimation using deep learning on intermediate left anterior descending coronary artery lesion angiography images. Sci Rep. 2024; 14(1): 1818. doi: 10.1038/s41598-024-52360-5.

[11] Zou Z, Li D, Guo H, Yao Y, Yin J, Tao C, et al. Enhancement of structural and functional photoacoustic imaging based on a reference-inputted convolutional neural network. Opt Express. 2025; 33(1): 1260-70. doi: 10.1364/OE.541906.

[12] Rezaee K, Zhu M. Diagnose Alzheimer's disease and mild cognitive impairment using deep CascadeNet and handcrafted features from EEG signals. Biomed Signal Process Control. 2025; 99: 106895. doi: 10.1016/j.bspc.2024.106895.

[13] Nande SB, Patwardhan SD. Automated Reservoir Characterization of Carbonate Rocks using Deep Learning Image Segmentation Approach. SPE J. 2024; 29(08): 4356-75. doi: 10.2118/219769-PA.

[14] Ren H, Sun Q, Xiao Z, Yu M, Wang S, Yuan L, et al. Heterogeneous feature fusion based machine learning strategy for ECG diagnosis. Expert Syst Appl. 2025: 271: 126714. doi: 10.1016/j.eswa.2025.126714.

[15] Gan H, Zhou R, Ou Y, Wang F, Cheng X, Fu L, et al. A region and category confidence-based multi-task network for carotid ultrasound image segmentation and classification. IEEE J Biomed Health Inform. 2025: online ahead of print. doi: 10.1109/JBHI.2025.3529483.

[16] Mahendiran T, Thanou D, Senouf O, Jamaa Y, Fournier S, De Bruyne B, et al. AngioPy Segmentation: An open-source, user-guided deep learning tool for coronary artery segmentation. Int J Cardiol. 2025; 418: 132598. doi: 10.1016/j.ijcard.2024.132598.

[17] Khan H, Javaid N, Bashir T, Akbar M, Alrajeh N, Aslam S. Heart disease prediction using novel ensemble and blending based cardiovascular disease detection networks: EnsCVDD-Net and BlCVDD-Net. IEEE Access. 2024; 12: 109230-54. doi: 10.1109/ACCESS.2024.3421241.

[18] DeGroat W, Abdelhalim H, Patel K, Mendhe D, Zeeshan S, Ahmed Z. Discovering biomarkers associated and predicting cardiovascular disease with high accuracy using a novel nexus of machine learning techniques for precision medicine. Sci Rep. 2024; 14(1): 1. doi: 10.1038/s41598-023-50600-8.

[19] Qiu Y, Meng J, Li B. Semi-supervised Strong-Teacher Consistency Learning for few-shot cardiac MRI image segmentation. Comput Methods Programs Biomed. 2025; 261: 108613. doi: 10.1016/j.cmpb.2025.108613.

[20] Shi L, Guo H, Liu J. Rapid and automatic hemodynamic assessment: integration of deep learning-based image segmentation, vessel reconstruction, and CFD prediction. Quant Imaging Med Surg. 2025; 15(2): 1358-70. doi: 10.21037/qims-24-1721

[21] Nogueira SA, Luz FA, Camargo TF, Oliveira JC, Campos Neto GC, Carvalhaes FB, et al. Artificial intelligence applied in identifying left ventricular walls in myocardial perfusion scintigraphy images: pilot study. Plos One. 2025; 20(1): e0312257. doi: 10.1371/journal.pone.0312257.

[22] Yu X, Zhu H, Huang B, Hou T, Lu W, Chen N, et al. SCS-SLSP: Hard uncertain pixels mining and utilization for semi-supervised cardiac image segmentation using subjective logic theory and subset prototype generation. Biomed Signal Process Control. 2024; 92: 106145. doi: 10.1016/j.bspc.2024.106145.

[23] Pham TV, Vu TN, Le HM, Pham VT, Tran TT. CapNet: An Automatic Attention-Based with Mixer Model for Cardiovascular Magnetic Resonance Image Segmentation. J Imaging Inform Med. 2025; 38(1): 94-123. doi: 10.1007/s10278-024-01191-x.

[24] Wang G, Zhou M, Ning X, Tiwari P, Zhu H, Yang G, et al. US2Mask: Image-to-mask generation learning via a conditional GAN for cardiac ultrasound image segmentation. Comput Bio Med. 2024; 172: 108282. doi: 10.1016/j.compbiomed.2024.108282.

[25] Danilov V, Klyshnikov K, Kutikhin A, Gerget O, Frangi A, Ovcharenko E. Angiographic dataset for stenosis detection. Mendeley Data. 2021, Version 1. doi:10.17632/ydrm75xywg.1.

[26] Li H, Xu K. Innovative adaptive edge detection for noisy images using wavelet and Gaussian method. Sci Rep. 2025; 15(1): 5838. doi: 10.1038/s41598-025-86860-9.

[27] Priyanka V, Rama YS, Sravani K, Kavya B. Implementation of Sobel Edge Detection with Image Processing on FPGA. In: 2024 2nd World Conference on Communication & Computing (WCONF) 2024 Jul 12 (pp. 1-5). IEEE. doi: 10.1109/WCONF61366.2024.10692301.

[28] Pritha A, Fathima G. A Detailed Description on Various Techniques of Edge Detection Algorithms. In: Shubham Mahajan, Kapil Joshi, Amit Kant Pandit, Nitish Pathak (Editors). Integrating Metaheuristics in Computer Vision for Real-World Optimization Problems. Crivener Publishing LLC; 2024: pp. 193-205. doi: 10.1002/9781394230952.ch11.

[29] Al Rawahi S. A Comparison of Sobel and Prewitt Edge Detection Operators. EJCS. 2025; 1(1): 49-58. doi:

10.63496/ejcs.Vol1.Iss1.16.

[30] Zhang R, Jiang G. Exploring a multi-path U-net with probability distribution attention and cascade dilated convolution for precise retinal vessel segmentation in fundus images. Sci Rep. 2025 Apr 18; 15(1): 13428. doi: 10.1038/s41598-025-98021-z

[31] Ylescupidez A, Speake C, Pietropaolo SL, Wilson DM, Steck AK, Sherr JL, et al. OGTT metrics surpass continuous glucose monitoring data for T1D prediction in multiple-autoantibody–positive individuals. J Clin Endocrinol Metab. 2024; 109(1): 57-67. doi: 10.1210/clinem/dgad574.

[32] Challoumis-Κωνσταντίνος Χαλλουμής C. Comparative analysis between capital and liability-Sensitivity Method. OJRE. 2024; 7(2): 63-76. doi: 10.32591/coas.ojre.0702.02063c.

[33] Das N, Das S. Attention-UNet architectures with pretrained backbones for multi-class cardiac MR image segmentation. Curr Prob Cardiol. 2024; 49(1): 102129. doi: 10.1016/j.cpcardiol.2023.102129.

[34] Liu R, Gao S, Zhang H, Wang S, Zhou L, Liu J. MTNet: A combined diagnosis algorithm of vessel segmentation and diabetic retinopathy for retinal images. Plos One. 2022; 17(11): e0278126. doi: 10.1371/journal.pone.0278126.

[35] Cho H, Lee JG, Kang SJ, Kim WJ, Choi SY, Ko J, et al. Angiography-based machine learning for predicting fractional flow reserve in intermediate coronary artery lesions. J Am Heart Assoc. 2019; 8(4): e011685. doi: 10.1161/JAHA.118.011685.

[36] Zhu X, Cheng Z, Wang S, Chen X, Lu G. Coronary angiography image segmentation based on PSPNet. Comput Methods Programs Biomed. 2021; 200: 105897. doi: 10.1016/j.cmpb.2020.105897.

[37] Wang L, Liang D, Yin X, Qiu J, Yang Z, Xing J, et al. Coronary artery segmentation in angiographic videos utilizing spatial-temporal information. BMC Med imaging. 2020; 20: 1-0. doi: 10.1186/s12880-020-00509-9.

Journal of Associated Medical Sciences 2025; 58(3): 122-128

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

Airway clearance techniques for pneumonia patients: A survey of Thai physical therapists

Chatchai Phimphasak[1,2], Hathaichanok Makphin[3], Dhissanuvach Chaikhot[3], Kittipun Aronpharungsunti[3], Jaturong Chimpalee[4], Arunrat Srithawong[5], Sahachat Aueyingsak[2,3*]

[1]Department of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kean Province, Thailand.
[2]Innovation to Improve Cardiopulmonary & Physical Performances Research Group, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kean Province, Thailand.
[3]Department of Physical Therapy, College of Health Sciences, Christian University of Thailand, Nakhon Pathom Province, Thailand.
[4]Respiratory Physical Therapy Unit, Central Chest Institute of Thailand, Nonthaburi Province, Thailand
[5]Department of Physical Therapy, School of Allied Health Sciences, University of Phayao, Phayao Province, Thailand.

ARTICLE INFO

Article history:
Received 31 January 2025
Accepted as revised 13 June 2025
Available online 16 June 2025

Keywords:
Pneumonia, physical therapy, airway clearance techniques.

ABSTRACT

Background: Pneumonia remains a significant global health challenge, necessitating effective management strategies, including airway clearance techniques (ACTs) utilized by physical therapists (PTs).

Objective: This study aims to investigate the selection and usage of ACTs among Thai PTs and examine the influence of work characteristics on these choices.

Materials and methods: A descriptive survey design was employed, targeting licensed PTs in Thailand with at least one year of experience in chest physical therapy. The survey, distributed via online forms and physical therapy units, gathered data on personal demographics, commonly used ACTs, and work characteristics. Out of 178 responses, 137 were included in the study.

Results: The results indicated a predominant reliance on manual techniques, particularly percussion and postural drainage, regardless of experience level. However, PTs specializing exclusively in chest physical therapy showed a higher propensity for using advanced techniques and devices such as positive expiratory pressure (PEP) and oscillating positive expiratory pressure (OPEP) devices. The study highlights significant gaps in training and access to advanced airway clearance devices, suggesting a need for targeted continuing education programs.

Conclusion: The findings underscore the importance of promoting a broader range of ACTs in clinical practice to improve patient outcomes. By addressing the barriers to adopting advanced techniques, PTs can enhance their therapeutic interventions and provide more comprehensive care for pneumonia patients. This study provides valuable insights into the current landscape of airway clearance therapy in Thailand, emphasizing the need for improved training and access to necessary equipment to optimize patient care.

Introduction

Pneumonia remains a significant global health challenge, particularly among vulnerable populations such as the elderly and immunocompromised individuals.[1] Effective management of pneumonia encompasses various therapeutic approaches, with physical therapy playing a vital role in improving patient outcomes.[2] A key aspect of this management is addressing decreased lung volume due to atelectasis, which is crucial for improving ventilation and ventilation/perfusion matching.[3] Among the array of interventions, airway clearance techniques (ACTs) stand out as essential tools, particularly in the late

∗ **Corresponding contributor.**
Author's Address: *Department of Physical Therapy, College of Health Sciences, Christian University of Thailand, Nakhon Pathom Province, Thailand.*
E-mail address: sahachata@christian.ac.th
doi: 10.12982/JAMS.2025.080
E-ISSN: 2539-6056

stages of pneumonia or for patients using mechanical ventilation.[4] These techniques are pivotal for respiratory physical therapists (PTs), aiding in the removal of mucus from the airways and enhancing lung function.

Understanding the pathophysiology and prognosis of diseases like pneumonia is essential for PTs who treat these patients. However, in the context of Thailand, PTs have diverse work characteristics in hospitals, and there is no standardized course for a Diploma of the Thai Board of Physical Therapy in each specialty.[5] As a result, the choice of ACTs is often based on undergraduate education and may be further supplemented by personal experiences in treating this patient group. This may lead to a variety of treatment approaches for the same condition.

The lack of information on the selection of ACTs makes it challenging to analyze the problem or create appropriate curricula in Thailand. This study aims to investigate the ACTs selected by Thai physical therapists and do different work characteristics influence their selection of ACTs.

Materials and methods

This study employs a descriptive survey design to examine the ACTs used by PTs for pneumonia patients across Thailand. Ethical approval was obtained from the Research Ethics Committee of Christian University of Thailand, under the reference number 07/2566. The target participants were licensed physical therapy in Thailand with at least one year of experience in chest physical therapy, aged between 22 and 60 years.

The survey was distributed using Google Forms through PT groups on social media platforms. It was also sent to physical therapy units in hospitals across Thailand. Data collection spanned two three-month periods, with reminder invitations sent after two weeks to encourage participation.

The survey questionnaire consisted of 15 items divided into three sections. The first section gathered personal information, including age, gender, education level, region of healthcare facility, and type of healthcare facility. The second section collected data on the most frequently used airway clearance techniques. To ensure validity and relevance, the questionnaire was reviewed and evaluated by two qualified PTs with over five years of experience in chest physical therapy.

Data analysis

Data analysis for this study involved chi-square tests to compare categorical variables, expressed as number and percentage. Statistical analysis was performed using the IBM SPSS statistical package (version 28.0), with a significance level set at $p<0.05$

Results

The questionnaires were distributed to physical therapy units and administered online, yielding a total of 178 responses. However, 41 questionnaires were excluded due to respondents having not worked in chest physical therapy for over a year. Consequently, 137 participants were included in this survey study.

The general characteristic of Thai PTs is shown in Table 1 An overview of the use of airway clearance techniques by Thai PTs for pneumonia patients is presented in Figure 1, while details on routine practice subgroups are shown in Figure 2 and Table 2.

Table 1. General characteristics of Thai physical therapists with varying levels of experience.

Characteristics	All group (N=137)	Exclusive chest PT (N=26)	All fields alongside chest PT (N=74)	All fields without chest PT (N=37)
Age (years)				
22-30 (%)	57 (42)	6 (23)	36 (49)	15 (41)
31-40 (%)	69 (50)	16 (62)	34 (46)	19 (51)
>40 (%)	11 (8)	4 (15)	4 (5)	3 (8)
Female (%)	86 (63)	11 (42)	49 (66)	26 (70)
Level of educations				
Bachelor's degrees (%)	105 (77)	15 (58)	64 (86)	26 (70)
Master's degrees (%)	26 (19)	8 (30)	9 (13)	9 (25)
Doctoral degrees (%)	6 (4)	3 (12)	1 (1)	2 (5)
Workplace locations				
Bangkok metropolitan region (%)	31 (22)	7 (28)	16 (22)	8 (22)
Southern region (%)	18 (13)	4 (15)	10 (14)	4 (11)
Central region (%)	27 (20)	4 (15)	11 (15)	12 (32)
Northern region (%)	19 (14)	5 (19)	8 (11)	6 (16)
Western region (%)	5 (4)	0	3 (4)	2 (5)
Eastern region (%)	8 (6)	2 (8)	2 (3)	4 (11)
Northeastern region (%)	29 (21)	4 (15)	24 (31)	1 (3)
Types of healthcare facilities				
Primary care hospitals (%)	18 (13)	0	8 (11)	10 (27)
Secondary care hospitals (%)	38 (28)	3 (11)	25 (34)	10 (27)
Tertiary care hospitals (%)	27 (20)	7 (27)	14 (19)	6 (16)
Advanced tertiary care hospitals (%)	15 (11)	9 (35)	4 (5)	2 (5)
Private hospitals (%)	36 (26)	7 (27)	21 (28)	8 (22)
Municipality (%)	3 (2)	0	2 (3)	1 (3)

***Note**: Data are presented as the number (%).*

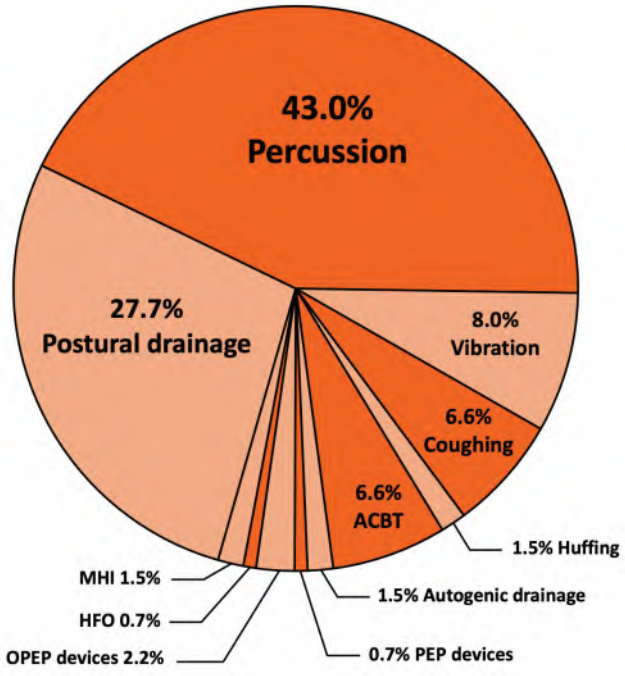


***Figure 1**. Overview of popular airway clearance techniques for pneumonia patients among Thai physical therapists. ACBT: active cycle of breathing technique, HFO: high frequency oscillation, MHI: manual hyperinflation, OPEP: oscillating positive expiratory pressure devices, PEP devices: positive expiratory pressure devices.*

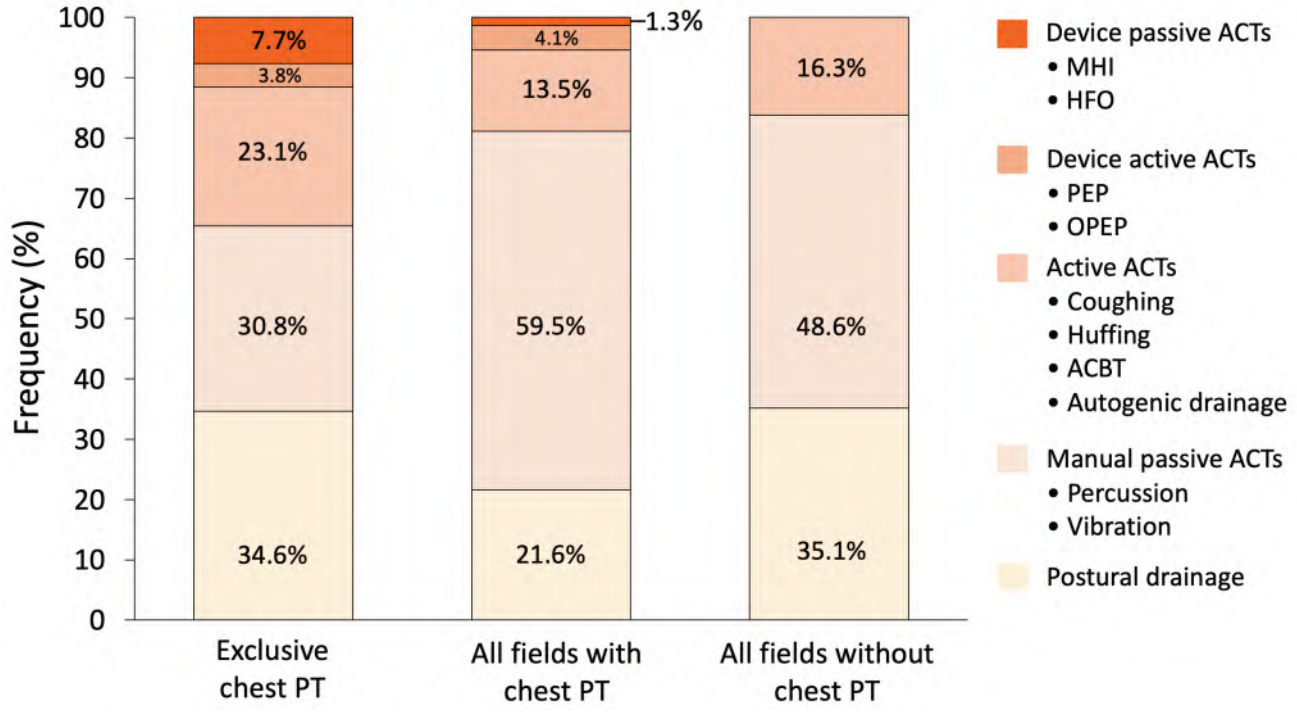


Figure 2. *Popular airway clearance techniques for pneumonia patients among Thai physical therapists, presented by subgroup. ACTs are categorized based on their techniques or mechanisms for airway clearance techniques, ACBT: active cycle of breathing technique, HFO: high frequency oscillation, MHI: manual hyperinflation, OPEP: oscillating positive expiratory pressure devices, PEP devices: positive expiratory pressure devices.*

Table 2. Comparison of Airway Clearance Technique Usage Among Thai Physical Therapists with Varying Levels of Experience (N=137).

Airway clearance techniques	Exclusive chest PT (N=26)	All fields with chest PT (N=74)	All fields without chest PT (N=37)
Postural drainage (%)	9 (34)	16 (22)	13 (35)
Manual passive ACTs			
- Percussion (%)	6 (22)	37 (51)*	16 (43)
- Vibration (%)	2 (8)	7 (9)	2 (5)
Active ACTs			
- Huffing (%)	1 (4)	1 (1)	0
- Coughing training (%)	2 (8)	5 (7)	2 (5)
- Autogenic drainage (%)	0	1 (1)	1 (3)
- ACBT (%)	3 (12)	3 (4)	3 (9)
Device active ACTs			
- PEP devices (%)	0	1 (1)	0
- OPEP devices (%)	1 (4)	2 (3)	0
Device passive ACTs			
- HFO (%)	1 (4)	0	0
- Manual hyperinflation (%)	1 (4)	1 (1)	0

***Note**: Data are presented as number (%). ACBT: active cycle of breathing technique, HFO: high frequency oscillation, MHI: manual hyperinflation, OPEP: oscillating positive expiratory pressure devices, PEP devices; positive expiratory pressure devices. *p<0.05 (Exclusive chest PT vs All fields with chest PT).*

Discussion

This study represents the first survey conducted among Thai PTs concerning ACTs for pneumonia patients. The predominant techniques identified in this survey were percussion and postural drainage. Interestingly, regardless of experience, PTs typically rely on these manual techniques. However, PTs who specialize exclusively in chest physical therapy showed a greater propensity for utilizing active and passive device-aided ACTs.

Pneumonia leads to decreased lung volume, increased airway resistance, and reduced compliance, ultimately causing hypoxemia.[1] Late-stage inflammatory processes can result in sputum secretion, further increasing airway resistance.[6] Effective ACTs are crucial, yet no universally effective long-term methods exist. Techniques such as the active cycle of breathing (ACBT) and positive expiratory pressure (PEP) devices have demonstrated benefits, including improved cure rates and reduced hospital stays in pneumonia patient.[2]

Despite the variety of available ACTs, our survey showed no significant preference among Thai PTs across different experience routine practice. However, percussion and postural drainage were the most used techniques. Percussion is a non-invasive, cost-effective method widely practiced in Thailand and other regions such as Turkey.[7] Its limitations include dependency on external factors such as frequency and force of application, with studies indicating challenges in maintaining effective percussion due to muscle fatigue.[8] Moreover, manual passive ACTs show varying efficiency in airway clearance depending on the individual user and the mechanism of action of each technique.[9,10] Postural drainage involves body positioning to use gravity for mucus movement towards larger airways, improving lung volume.[11] Though effective, it requires specific positions, which can be challenging to maintain and implement.

On the other hand, airflow-dependent clearance techniques, such as active techniques, using devices, and MHI, are not popular among Thai PTs. These techniques directly alter airflow in the airways, providing immediate relief by facilitating the removal of mucus and improving ventilation, which may be appropriate for low lung volume in pneumonia patients.[12] However, based on the available evidence, it is difficult to definitively determine the efficacy of each technique in pneumonia patients[2,13] and cystic fibrosis patients.[14] Therefore, prior research suggests similar effectiveness between active and manual techniques. When choosing a method, PTs might consider their effort, favoring techniques that provide comparable results with a lower workload.

In Thailand, the absence of specific guidelines for PT management in pneumonia leads PTs to rely on personal experience and undergraduate training. Comparatively, a retrospective study in Australia highlighted both similarities and differences in practice. For instance, while postural drainage is similarly popular, MHI is more commonly used in Australia.[15] An expert consensus from eight countries also suggests the benefits of combining manual techniques with MHI when sputum volume or viscosity is high.[13] Recent clinical guidelines recommend MHI for mechanically ventilated patients due to its advantages in lung volume improvement and secretion clearance.[16] Active treatments such as deep breathing exercises and early mobilization are well-supported by evidence, whereas manual techniques rely more on expert opinion.[17,18,19]

The limited use of techniques such as vibration, huffing, and autogenic drainage, as well as devices like PEP and OPEP, highlights potential areas for improvement in training and practice. Factors such as lack of familiarity, limited access to equipment, and institutional protocols may contribute to the underutilization of these methods. This gap suggests a need for targeted continuing education programs to broaden the skill set of PTs and enhance patient care. However, while exclusive chest PTs demonstrated a greater inclination to use devices and MHI, these techniques have not been widely adopted overall.

Implications for practice and training

Given the reliance on manual techniques and postural drainage, it is crucial to promote the adoption of a wider range of ACTs. Training programs should emphasize the benefits and proper use of advanced techniques and devices to ensure that therapists are well-equipped to deliver comprehensive care. Additionally, enhancing access to essential equipment and fostering interdisciplinary collaboration can facilitate the integration of these techniques into routine practice.

Our findings highlight potential barriers to the adoption of newer ACTs, such as limited training access, equipment availability, and institutional protocols. Addressing these barriers is essential to ensure equitable access to evidence-based respiratory care. It is important to provide PTs with the necessary resources and support to effectively implement these techniques.

Limitation

This study has several limitations, including a small sample size, the absence of patient condition such as consciousness and intubation status in the survey, which may influence decision-making, reliance on self-reported data, and the measurement of technique preference rather than actual clinical practice. These factors may affect the generalizability of the findings and suggest the need for cautious interpretation.

Further research is necessary to explore the efficacy of various airway clearance techniques in different settings and patient populations. Long-term studies could provide more robust evidence to support the development of clinical guidelines and enhance therapeutic decision-making. Additionally, investigating the impact of specific training programs on technique utilization and patient outcomes would provide valuable insights into the best practices for respiratory care.

Conclusion

This study highlights the current landscape of airway clearance therapy for pneumonia patients in Thailand, revealing a predominant reliance on manual techniques and postural drainage. By understanding the preferences, challenges, and opportunities in clinical practice, we can work towards optimizing therapeutic interventions and improving patient outcomes in respiratory care. Addressing the identified barriers and promoting continuing education will be crucial in advancing the quality of care provided by Thai PTs.

Ethic approval

This study received ethical approval from the Research Ethics Committee of Christian University of Thailand (reference number 07/2566).

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research received funding support from the Christian University of Thailand [grant number 003/2566].

Conflict of interest

The authors have disclosed no competing financial or non-financial interests related to the research, authorship, and/or publication of this article.

Acknowledgements

The authors gratefully acknowledge the support of the College of Health Sciences, Christian University of Thailand, Nakhon Pathom, Thailand. We extend our sincere thanks to all Thai physical therapists for their generous contribution to this research.

References

[1] Mandell LA, Wunderink RG, Anzueto A, Bartlett JG, Campbell GD, Dean NC, et al. Infectious Diseases Society of America/American Thoracic Society Consensus Guidelines on the Management of Community-Acquired Pneumonia in Adults. CID. 2007; 44: S27-S72. doi:10.1086/511159.

[2] Chen X, Jiang J, Wang R, Fu H, Lu J, Yang M. Chest physiotherapy for pneumonia in adults. Cochrane Database Syst Rev. 2022; 9: CD006338. PMID: 36066373, PMCID: PMC9447368.

[3] Kacmarek RM, Stoller JK, Heuer AJ, Chatburn RL, Kallet RH, editors. Egan's fundamentals of respiratory care. 12th Edition. St Louis, Missouri: Elsevier; 2021.

[4] Van Der Lee L, Hill A-M, Jacques A, Patman S. Efficacy of Respiratory Physiotherapy Interventions for Intubated and Mechanically Ventilated Adults with Pneumonia: A Systematic Review and Meta-Analysis. Physiother Can. 2021; 73: 6-18. doi: 10.3138/ptc-2019-0025.

[5] Physical Therapy Council. Regulations of the Physical Therapy Council Concerning Restrictions and Conditions in the Practice of Physical Therapy, B.E. 2551. 2008. Available: https://pt.or.th/PTCouncil/law/1-2/10.pdf.

[6] Pahal P, Rajasurya V, Sharma S. Typical Bacterial Pneumonia. StatPearls. Treasure Island (FL): StatPearls Publishing; 2023. Available: http://www.ncbi.nlm.nih.gov/books/NBK534295/.

[7] Çakmak A, İnce Dİ, Sağlam M, Savcı S, Yağlı NV, Kütükcü EÇ, et al. Physiotherapy and rehabilitation implementation in intensive care units: A Survey Study. Turk Thorac J. 2019; 20: 114-9. doi: 10.5152/TurkThoracJ.2018.18107.

[8] Audsavachulamanee B, Aueyingsak S, Ubolsaka-Jones C, Kosura N, Chanapon J, Yubonpan N, et al. Thai physiotherapists' performance of manual chest wall percussion on an artificial lung: frequency, force, and fatigue perception. Arch AHS. 2023; 35: 23-34.

[9] Audsavachulamanee B, Nwe AA, Ubolsaka-Jones C, Kosura N, Hongratana G, Aueyingsak S, et al. The flow-pressure and cardiopulmonary responses of manual chest percussion with different ranges of oscillation frequency in healthy subjects. Trends Sci. 2024; 21: 7345. doi:10.48048/tis.2024.7345.

[10] Audsavachulamanee B, Aueyingsak S, Ubolsakka-Jones C, Bansri A, Chaiyasang J, Srilasalai S, et al. Physiotherapist performance during manual chest vibration in simulated adult lung: frequency, force and pattern. Asia Pac J Sci Technol. 2024; 29(3): 14-21. doi:10.14456/apst.2024.48.

[11] Paz JC, West MP. Acute care handbook for physical therapists. 4th Edition. St. Louis, Missouri: Elsevier Saunders; 2014.

[12] Langenderfer B. Alternatives to percussion and postural drainage. A review of mucus clearance therapies: percussion and postural drainage, autogenic drainage, positive expiratory pressure, flutter valve, intrapulmonary percussive ventilation, and high-frequency chest compression with the ThAIRapy Vest. J Cardiopulm Rehabil. 1998; 18: 283-9. doi: 10.1097/00008483-199807000-00005.

[13] Van Der Lee L, Hill A, Patman S. Expert consensus for respiratory physiotherapy management of mechanically ventilated adults with community-acquired pneumonia: A D elphi study. J Eval Clin Pract. 2019; 25: 230-43. doi: 10.1111/jep.13077.

[14] Mckoy NA, Wilson LM, Saldanha IJ, Odelola OA, Robinson KA. Active cycle of breathing technique for cystic fibrosis. Cochrane Database Syst Rev. 2016; 7: CD007862. doi:10.1002/14651858.CD007862.pub4.

[15] Kwan BPM, Hill A-M, Elliott M, Van Der Lee L. A retrospective study of physiotherapy management for patients with pneumonia requiring invasive ventilation in a single-center Australian ICU. Hong Kong Physiother J. 2022; 42: 55-64. doi: 10.1142/S1013702522500068.

[16] Van Der Lee L, Patman S, Hill A-M. Development of a clinical practice guideline for physiotherapy management of adults invasively ventilated with community-acquired pneumonia. Physiotherapy. 2024; 122: 57-67. doi: 10.1016/j.physio.2023.12.003.

[17] Schweickert WD, Pohlman MC, Pohlman AS, Nigos C, Pawlik AJ, Esbrook CL, et al. Early physical and occupational therapy in mechanically ventilated, critically ill patients: a randomised controlled trial. The Lancet. 2009; 373: 1874-82. doi:10.1016/S0140-6736(09)60658-9.

[18] Larsen T, Lee A, Brooks D, Michieli S, Robson M, Veens J, et al. Effect of early mobility as a physiotherapy treatment for pneumonia: A systematic review and meta-analysis. Physiother Can. 2019; 71: 82-9. doi: 10.3138/ptc.2017-51.ep.

[19] Van Der Lee L, Hill A-M, Patman S. After-hours respiratory physiotherapy for intubated and mechanically ventilated patients with community-acquired pneumonia: An Australian perspective. Aust Crit Care. 2018; 31: 349-54. doi: 10.1016/j.aucc.2017.10.001.

Journal of Associated Medical Sciences 2025; 58(3): 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

Determining the optimal frequency and current intensity of microcurrent stimulation for postural sway induction: A pilot study

Wonjong Yu[1] and Haneul Lee[2]*

[1]*Department of Physical Therapy, Eulji University, Seongnam, Korea.*
[2]*Department of Physical Therapy, College of Medical Science, Gachon University, Incheon, Korea.*

ARTICLE INFO

Article history:
Received 4 October 2024
Accepted as revised 12 June 2025
Available online 18 June 2025

Keywords:
Electromyography, microcurrent stimulation, postural sway, vestibular system, rehabilitation

ABSTRACT

Background: Maintaining postural balance is crucial for various daily activities, and the vestibular system plays a significant role in this process. Dysfunction in the vestibular system can result in impaired postural control and balance. Microcurrent therapy, which uses lower current intensities than traditional electrical stimulation methods, may offer a less painful alternative for vestibular system training.

Objective: This pilot study aimed to investigate the effects of microcurrent stimulation on vestibular system training for posture control, with an emphasis on sway velocity and muscle activity, and to determine its optimal frequency and electric current intensity.

Materials and methods: Twenty healthy young adults (mean age, 22.35 years; 76.5% female; mean body mass index, 24.5±3.5 kg/m²) participated in this study. Surface electromyography (EMG) of the tibialis anterior muscle and postural sway were measured during microcurrent stimulation at varying frequencies (0.5 Hz to 2.0 Hz) and intensities (50 μA to 400 μA) on the bilateral mastoid processes of 15 seconds each. The center of pressure displacement was recorded using a pressure platform, and muscle activity was assessed through EMG data.

Results: The results showed that frequency did not significantly affect postural sway or muscle activity, whereas intensity did. Muscle activity was greatest at 1.5 Hz and 400 μA, and postural sway was most pronounced at 2.0 Hz and 400 μA. No adverse reactions were reported during or after the stimulation sessions.

Conclusion: Microcurrent stimulation with optimal current and intensity may be a useful, non-invasive option for vestibular system training, particularly in populations where traditional electrical stimulation may cause discomfort. Further studies involving older adults and individuals with vestibular dysfunction are necessary to validate these findings.

Introduction

To maintain one's balance, the human body's center of gravity should fall within the weight support center of mass, for which continuous adjustment is required from the positioning of joints and muscle activity.[1] Sustaining standing balance is among the most important functions of the human body in various activities of daily living.[2] The central nervous system (CNS) controls one's postural balance through the vestibular, visual, and proprioceptive systems,[3,4] as well as the motor system of the human body.[1,2,5] Among the three systems, the vestibular system is the most unique because it is instantly multisensory and multimodal. For instance, it works with proprioception and

* **Corresponding contributor.**
Author's Address: *Department of Physical Therapy, College of Medical Science, Gachon University,Incheon, Korea.*
E-mail address: leehaneul84@gachon.ac.kr
doi: 10.12982/JAMS.2025.081
E-ISSN: 2539-6056

corollary discharge of the motor system/plan, enabling the cortex to differentiate activities caused by passive head movements.[6] During various movements of the head and body, the vestibular system maintains stability.[4,7] The peripheral vestibular system detects physical stimuli, such as head rotation, acceleration, and gravity, and converts them to electrical signals for transmission to the CNS.[8] The CNS uses information from the vestibular system to determine the head's position and speed, integrates it with information from the proprioceptive and visual systems, and maintains the body's posture and balance.[8]

Vestibular system dysfunction has gained clinical recognition as an important differential diagnosis for decreased postural control. People with balance disorders, vestibular system defects, or both tend to abstain from physical activity and social gatherings to avoid unexpected falls, which affect their quality of life.[9] Vestibular system training was developed to gradually increase eye, head, and body movements to stimulate the vestibular system and promote the CNS to adapt to new information.[10] This training can reduce the fear of dizziness during daily activities and enhance the balance sensation, making it ideal for patients with peripheral vestibular dysfunction.[10] Four types of vestibular training are available: adaptation, habituation, substitution, and balance exercise.[11] The recovery of vestibular function mainly relies on CNS plasticity, and natural recovery, adaptation, and substitution can promote recovery through these mechanisms.

Additionally, the approach widely used to improve postural control involves inducing perturbations and enhancing the effects of balance training by repeating the process of recovery using motor learning through repetition and adaptation.[12,13] It is possible that postural sway caused by electrical stimulation could serve as a form of balance training, as the individual would need to use their muscles to destabilize the stimulation and maintain their balance.[14] This is similar to other forms of balance training such as standing on a wobble board or performing exercises on an unstable surface. However, the effectiveness of this training type may depend on several factors, such as the frequency and intensity of electrical stimulation. In addition, it is important to ensure that electrical stimulation during training is safe and does not pose a risk of injury.

Microcurrent stimulation has been hypothesized to enhance neural activity through mechanisms such as increased ATP production, which supports cellular energy metabolism and facilitates neural adaptation.[15] Additionally, microcurrent has been reported to improve blood flow and tissue repair, which may aid in rehabilitative processes involving the vestibular system.[16] On a neural level, electrical stimulation, including microcurrent, has the potential to enhance synaptic plasticity, a critical factor in the central nervous system's ability to adapt to balance perturbations.[17] These mechanisms highlight its potential as a therapeutic tool for vestibular system rehabilitation.

Sinusoidal or galvanic currents have been used as direct electrical stimulation methods for the vestibular systems to improve postural control. Such currents are applied to the mastoid process to stimulate the vestibular nerve directly;[18] however, most patients complain of unpleasant pain due to strong electrical currents.[19] Previous studies have demonstrated the potential of vestibular stimulation to improve postural sway and muscle activation in different age groups, highlighting its utility in balance control interventions.[20] Despite this, the discomfort associated with galvanic stimulation often limits its clinical application, especially for individuals with heightened pain sensitivity. In contrast, microcurrent therapy stimulates muscle contraction without discomfort or pain by using a microampere (μA) rather than milliampere (mA) current.[21] Research has shown that microcurrent stimulation can improve balance performance in the 100-300 μA,[22] suggesting its potential to painlessly stimulate the vestibular system. Furthermore, microcurrents have been reported to alleviate muscle symptoms, supporting their use in clinical rehabilitation.[16,23]

Direct current galvanic stimulation has been shown to effectively stimulate the vestibular system but is often accompanied by discomfort and pain, limiting its clinical application.[24] To address this limitation, this study explored the use of microcurrent stimulation, which delivers significantly lower intensities, as an alternative method for vestibular stimulation. However, the optimal intensity of microcurrent stimulation for effective vestibular activation remains unclear. Therefore, this study systematically varied microcurrent intensity (50 μA to 400 μA) and frequency (0.5 Hz to 2.0 Hz) to identify parameters that could maximize the effects on postural control while minimizing discomfort.

The primary objective was to determine the most effective microcurrent intensity for stimulating the vestibular system, with the potential to develop a safer and more tolerable intervention for clinical rehabilitation.

Materials and methods

Ethical approval

The study was conducted in accordance with the guidelines of the Declaration of Helsinki and was approved by the Institutional Review Board of Eulji University (EU19-68). All participants signed an informed consent form before starting the experiment. The manuscript was prepared in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology guidelines.

Participants

Twenty healthy young adults aged ≥18 years with no problem receiving microcurrent stimulation to the vestibular system volunteered to participate. Three participants were excluded because they did not meet the inclusion criteria. The participants were excluded if they had a history of recent lower-extremity injury or disability, visual or neurological deficits, cardiopulmonary diseases, or mental problems. We also excluded subjects

if their one-leg standing test time was <27.5 sec, the mean of young healthy adults aged 20-30 years to prevent falls during the balance tasks. We also excluded those who showed an unpleasant reaction to electrical stimulation during the screening process.

Procedure

When the participants arrived at the laboratory, their general anthropometric characteristics such as height, weight, and body mass index were obtained. Next, electromyography (EMG) signals of the tibialis anterior (TA) muscle and postural sway were recorded on the pressure platform while the subjects received various microcurrent stimulations on the bilateral mastoid processes for 15 s each. To minimize head movement, participants were instructed to stare at the "+" sign at eye level on the wall and maintain a standing posture.

Outcome measurements

Measurement of postural control (or center of pressure)

A Zebris FDM-S pressure platform (Zebirs Medical GmbH, Isny im Allgau, Germany) was used to assess center of pressure (CoP) displacement. The pressure platform-calibrated capacitive force sensors underneath the platform at a sampling rate of 60 Hz enable the assessment of the distribution of static forces. MR 3.8 software (Noraxon Inc., Scottsdale, AZ, USA) was used to integrate the force signals and graphic representation of the CoP. The area of the CoP path (mm^2), length of the CoP path (mm), and sway velocity (mm/s), defined as the mean CoP movement speed, were used to determine postural sway. The participants were instructed to maintain their standing posture on a pressure platform during various microcurrent stimulations.

Measurement of muscle activity on EMG

A wireless surface Trigno EMG sensor (DELSYS, Natick, MA, USA) was used at a sampling rate of 2000 Hz to measure the muscle activity. Prior to the EMG measurement, the skin was shaved and wiped with alcohol to reduce impedance. Disposable pre-gelled bipolar Ag/AgC1 surface electrodes were placed over the muscle belly parallel to the muscle fibers. The electrodes were placed on the muscle belly of the TA of both legs. Surface EMG electrodes were secured with a bandwidth set between 10 Hz and 350 Hz, and the notch filter was set at 60 Hz to minimize noise distortion. The signal of each muscle was normalized to the reference maximal voluntary isometric contraction (MVIC). The participants were seated on chairs, and maximum dorsiflexion forces were measured against the exertional force toward the floor on the participant's ankle during lifting to measure the MVIC of the TA muscle. MVIC measurements were taken three times with a 1-min rest period between contractions, and the average middle 3 s of the MVIC was analyzed.[25]

The TA muscle was selected for its key role in maintaining postural balance during anterior-posterior sway, which is commonly induced by vestibular stimulation. Previous studies have highlighted the involvement of the TA, soleus, and gastrocnemius muscles in stabilizing postural sway.[16] The TA muscle, in particular, is essential for controlling dorsiflexion and anterior stability, making it a suitable indicator of balance control.[26] Future research could include other muscle groups, such as the soleus and gastrocnemius, to provide a broader understanding of balance mechanisms.

EMG data of the TA were collected for 15 s while participants were standing on the pressure platform during microcurrent stimulation of the vestibular system at varying frequencies and current intensities, but only for 10 s. The root mean square value was used to quantify the electric signal of the amplitude of the EMG signal for a given period because it estimates the muscular exertion intensity.[27,28]

Microcurrent electrical stimulation

Microcurrent electrical stimulation was performed using an Acutron Mentor (Eastwestmed, Inc., Phoenix, AZ, USA), with HRTC 32AP stimus hydrogel electrodes were placed on the mastoid process.

The range of 50–400 microamperes (μA) for current intensity was selected based on a combination of prior research and preliminary testing. Previous studies using galvanic vestibular stimulation (GVS) to stimulate the vestibular system reported increased postural sway at current intensities of 0.2 mA, 0.5 mA,[27,28] and 0.7 mA.[29] To identify the minimum and maximum microcurrent intensities capable of stimulating the vestibular nerve without causing discomfort, preliminary testing was conducted in this study. The range of 50-400 μA was established as the threshold where participants reported no unpleasant sensations while maintaining sufficient stimulation of the vestibular nerve. Additionally, because this study targeted healthy young adults with intact vestibular and somatosensory systems, it was hypothesized that effective stimulation could be achieved using relatively low microcurrent intensities.

Based on the previous evidence and preliminary testing, microcurrent stimulation began with 0.5 Hz and 50 μA intensity while the participants maintaineda standing position. Subsequently, 0.5 Hz and 100 μA, 200 μA, 300 μA, and 400 μA were applied, respectively, with a 15-sec break between experiments. Similarly, frequencies of 1.0 Hz, 1.5 Hz, and 2.0 Hz were applied in combination with current intensities of 50 μA, 100 μA, 200 μA, 300 μA, and 400 μA, respectively. If the participant experienced nausea, a tickling sensation, unbearable pain, or an unpleasant feeling due to electrode attachment during the experiment, the trial was immediately discontinued (Figure 1).

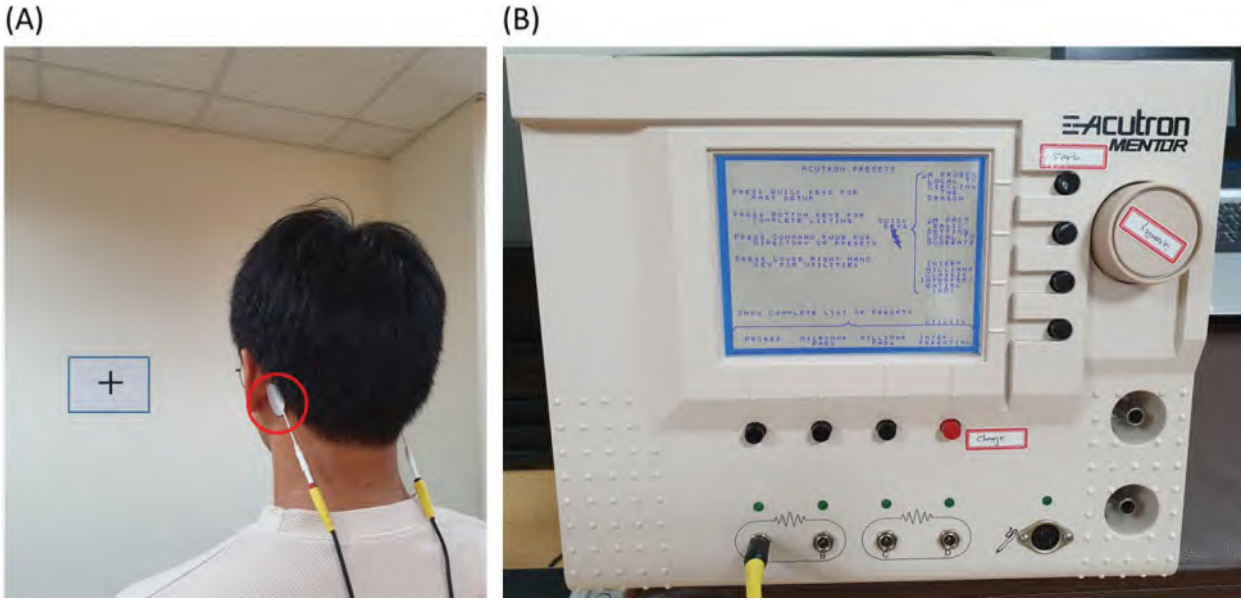


Figure 1. *Application of microcurrent stimulation on bilateral mastoid processes.*

Statistical analysis

SAS version 9.4 (SAS Institute, Inc. Cary, NC, USA) and R software version 4.0.4 (R Project for Statistical Computing, Vienna, Austria) were used to analyze the data, which were summarized as means and standard deviations (SD) or medians with maximum and minimum. The Shapiro-Wilk test was used to confirm the assumption of normality of the continuous variables. Non-parametric repeat-measures analysis of variance was conducted to test the data. The level of significance was set at α=0.05.

Results

None of the participants reported unpleasant symptoms or irritation under the electrodes during or after stimulation trials. Additionally, none of the participants noticed the stimulation applied during the trials. The patients' general characteristics are described in Table 1.

Table 2 and Figure 2 show the TA muscle activity and balance parameters such as CoP area, CoP length, and CoP velocity. There was no significant interaction between the frequency and intensity postural balance parameters (CoP area, CoP sway length, and CoP sway velocity) and EMG. In addition, muscle activity and postural balance parameters were not statistically significant among the different frequencies (p>0.05) but were significantly different among intensities (p<0.001). Muscle activity was maximum at a combination of 1.5 Hz and 400 µA and minimal at 2.0 Hz and 50 µA. The CoP area, sway length, and velocity were the greatest at the combination of 2.0 Hz and 400 µA, while CoP area was the lowest at 0.5 Hz and 50 µA, while CoP length and velocity were the lowest at 1.5 Hz 50 µA.

The findings demonstrated a significant increase in muscle activity and postural sway with higher intensity levels, particularly at 300 µA and 400 µA. This trend suggests that intensity serves as a critical determinant of vestibular stimulation effectiveness, likely due to its direct impact on vestibular nerve excitation thresholds.

Table 1. General characteristics of participants (N=20)

Characteristics	
Sex, female [N (%)]	13 (76.5%)
Age (years; mean±SD)	22.35±2.12
Height (cm; mean±SD)	171.5±15.2
Weight (kg; mean±SD)	65.2±10.5
BMI (kg/m^2; mean±SD)	24.5±3.5

Note: BMI: body mass index, SD: standard deviation.

Table 2. Muscle activity and postural sway at different frequencies and intensities.

Frequency (Hz)	Amplitude (μA)	EMG (%)	p†	CoP area (mm^2)	p†	CoP sway length (mm)	p†	CoP sway velocity (mm/sec)	p†
0.5	50	32.5±9.7	0.466	130.0±144. 2	0.562	96.3±87.7	0.686	9.5±8.6	0.485
	100	32.5±9.4		139.8±179.9		100.6±80.3		10.0±7.9	
	200	33.1±9.4		157.8±147.2		121.6±111.1		12.0±9.6	
	300	33.5±8.9		161.2±108.8		123.6±96.7		11.9±11.0	
	400	34.1±9.1		202.3±161.2		129.1±134.0		12.7±9.3	
1.0	50	32.6±9.5		137.8±112.5		99.3±73.2		9.6±7.2	
	100	32.8 ±9.7		141.9±201.3		112.7±94.91		11.1±9.42	
	200	33.1± 9.4		162.2±171.9		124.9±137.8		12.1±8.6	
	300	33.2±9.3		170.6±213.4		126.0±121.3		12.3±11.9	
	400	33.9±8.5		209.0±194.5		150.6±190.6		13.7±19.0	
1.5	50	32.3±9.3		133.4±77.8		87.9±71.3		9.4±6.0	
	100	32.4±9.9		136.8±106.6		96.4±60.9		10.8±10.0	
	200	32.6±9.6		158.8±111.9		110.3±102.1		11.5±7.0	
	300	33.0±9.7		180.2±115.9		106.2±97.1		13.5±9.5	
	400	34.1±9.3		207.4±165.6		129.0±135.2		14.8±13.2	
2.0	50	31.3±8.8		136.6±81.2		151.2±107.9		11.8±10.4	
	100	31.7±9.0		143.5±110.3		156.4±121.8		12.4±11.7	
	200	32.1±9.0		163.0±167.8		156.8±99.3		14.4±9.4	
	300	33.5±8.8		188.0±120.1		160.4±108.5		14.9±10.3	
	400	34.0±٨,٩		212.8±183.5		172.3±154.1		17.0±12.0	

Note: EMG: electromyography, CoP: center of pressure, † Interaction between frequency and intensity.

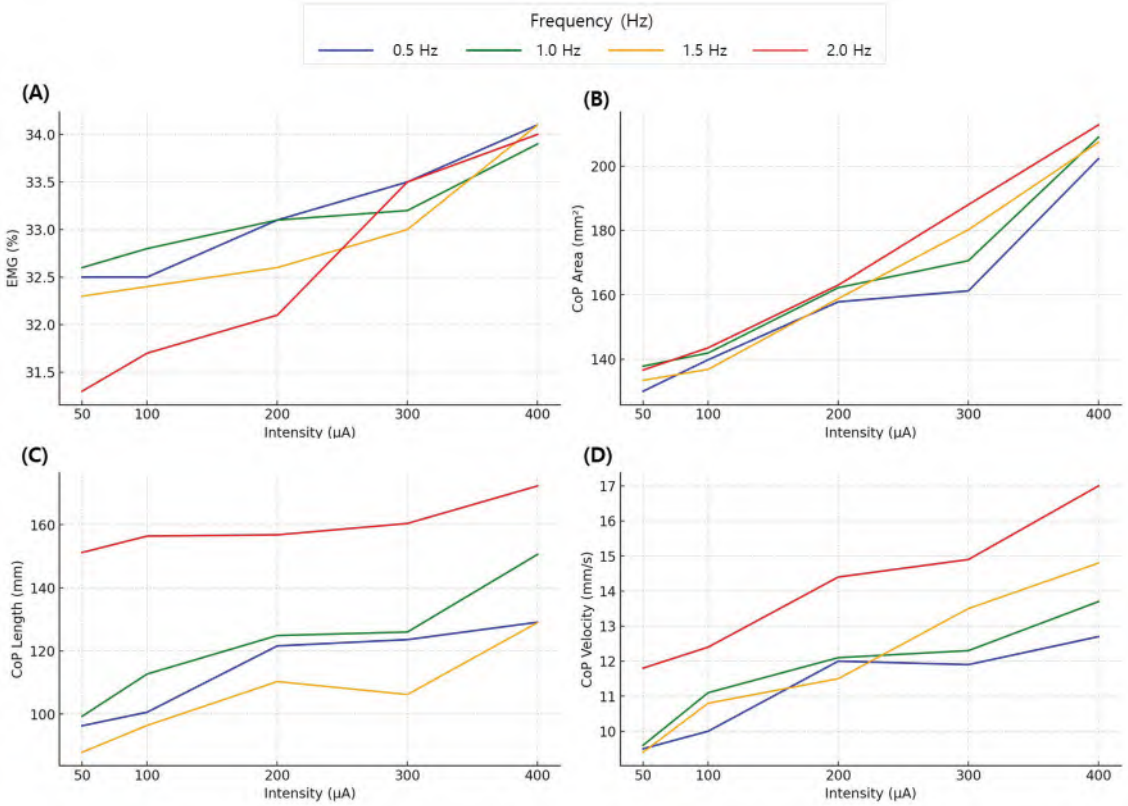


Figure 2. *Frequency and Intensity on muscle activity (A), CoP area (B), CoP Length (C), and CoP velocity (D).*

Discussion

This pilot study aimed to investigate the effect of microcurrent stimulation on vestibular system training for posture control and determine the optimal frequency and electric current intensity for training. The microcurrent stimulation used in this study involved applying electrical stimulation directly to the mastoid process, which stimulated the vestibular nerve directly without passing through the peripheral vestibular receptors. The resulting sway was recorded using a force plate and EMG without involving the peripheral vestibular receptors. Our results showed that the frequency of microcurrent stimulation did not significantly affect postural sway and muscle activity, whereas intensity did. The lack of significant impact from frequency variations on postural sway and muscle activity could be attributed to the lower sensitivity of the vestibular system to frequency changes compared to intensity. Previous studies suggest that muscle responses and postural adjustments are more closely tied to the magnitude of the stimulation rather than its temporal characteristics.[30,31] This indicates that intensity plays a more dominant role in eliciting physiological responses during vestibular stimulation. Muscle activity was greatest at a combination of 1.5 Hz and 400 μA, while the postural sway parameters were greatest at a combination of 2.0 Hz and 400 μA, supporting previous evidence that optimal stimulation amplitude for balance improvement is within the rage of ±100 to ±400 μA.[22] Our results also suggest that when individuals with reduced postural balance control are stimulated, TA muscle activity increases along with postural balance parameters.

The findings of this study align with previous research on GVS, demonstrating that increased current intensity induces greater postural sway due to heightened vestibular nerve activation.[24,29] Similarly, our findings indicate that microcurrent stimulation, despite using significantly lower current levels (1/1000th of GVS intensity), can effectively stimulate the vestibular system. This highlights its potential as a safer and more tolerable alternative for balance training. Compared to GVS, microcurrent stimulation utilizes significantly lower intensities, offering a safer and more tolerable alternative, particularly for healthy individuals and populations with heightened sensitivity.

Galvanic and microcurrent differ in current intensity, but they are both direct currents. Microcurrents can stimulate the vestibular system with minimal stimulation with minimal discomfort.[29] This reduced pain or unpleasant sensations increases user satisfaction compared to conventional galvanic stimulators. While galvanic current intensity is often limited by pain, microcurrent stimulation allows greater flexibility in adjusting intensity and frequency, making it beneficial for personalized treatments. This approach is particularly advantageous for elderly individuals with decreased balance ability, where a reduced pain burden lowers the risk of falling and enhances overall safety.

The existing exercise for the vestibular system involves physical stimulation of the hair cells in the peripheral vestibular receptors, which generates electrical signals and triggers the nystagmus reflex.[32] However, it is unclear which parts of the peripheral vestibular receptors and vestibular nerves should be targeted. However, in the current study,

electrical stimulation was applied directly to the ampulla, which can create a postural sway or nystagmus response without passing through the peripheral vestibular organs. This type of stimulation may be useful in cases of vestibular nerve problems since it bypasses the nerve and directly stimulates the vestibular end organs. Furthermore, microcurrent stimulation of the vestibular system can lead to exercise-like effects because it initiates a process of self-recovery in which the body repeatedly adjusts and stabilizes itself after each stimulation. This process can strengthen the vestibular system, similar to strengthening the muscles.[33] In addition, a microcurrent stimulator was proven to raise the tissue temperature and generate sedation, decreasing symptomatic pain.[34] It also effectively improves muscle function and promotes recovery by enhancing blood flow through the skin;[16] thus, it has been widely used as an effective, easy-to-use, noninvasive, and time-efficient treatment method.

Limitation

This pilot study of healthy adults aimed to determine the most effective frequency and electric current intensity of microcurrent stimulation for training the vestibular system. However, as this is a pilot study, a formal sample size calculation was not performed, and the findings are based on a small sample of 20 participants. While this sample size was sufficient for a preliminary investigation, future studies with larger and more diverse populations are necessary to validate these findings and ensure their generalizability. It is also necessary to test this system in elderly individuals or patients who require vestibular rehabilitation to confirm the appropriateness of its parameters. Additionally, further research combined with various balance training methods is needed, as it might have stronger synergistic effects with balance training. Finally, studies investigating the prolonged effects of microcurrent stimulation over time are required to establish its long-term efficacy and safety.

Conclusion

This pilot study found that the optimal frequency and current intensity for microcurrent vestibular stimulation for improving postural control is 2.0 Hz and 400 μA for postural sway and 1.5 Hz and 400 μA for muscle activity. These findings suggest that microcurrent stimulation may be a valuable tool for vestibular system training, particularly in cases of vestibular dysfunction. Its low-intensity, non-invasive nature makes it a more suitable and tolerable option for elderly individuals or those with heightened sensitivity, as it minimizes pain and discomfort. This approach has potential for enhancing balance training and supporting vestibular rehabilitation. Microcurrent stimulation could be tailored for diverse patient populations, but further studies with larger samples and longitudinal designs are needed to validate its efficacy and broader clinical applications

Ethical approval

The study adhered to the guidelines of the Declaration of Helsinki and was approved by the Institutional Review Board of Eulji University (EU19-6). Informed consent was obtained from all study subjects.

Funding

This study was supported by the Researchers Supporting Project of Eulji University in 2020 and the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2018R1C1B5041814).

Conflicts of Interest

The authors declare no conflicts of interest.

CRediT authorship contribution statement

Conceptualization: WY and HL; **methodology:** WY; **formal analysis:** HL; **investigation:** WY; **data curation:** WY and HL; **writing-original draft preparation:** WY and HL; **writing-review and editing:** WY and HL, **funding acquisition:** WY. All authors have read and agreed to the published version of the manuscript.

References

[1] Alonso AC, Brech GC, Bourquin AM, Greve JM. The influence of lower-limb dominance on postural balance. Sao Paulo Med J. 2011; 129(6): 410-3. doi: 10.1590/s1516-31802011000600007.

[2] Cherng RJ, Chen JJ, Su FC. Vestibular system in performance of standing balance of children and young adults under altered sensory conditions. Percept Mot Skills. 2001; 92(3 Pt 2): 1167-79. doi: 10.2466/pms.2001.92.3c.1167.

[3] Grace Gaerlan M, Alpert PT, Cross C, Louis M, Kowalski S. Postural balance in young adults: the role of visual, vestibular and somatosensory systems. J Am Acad Nurse Pract. 2012; 24(6): 375-81. doi: 10.1111/j.1745-7599.2012.00699.x.

[4] Inukai Y, Otsuru N, Masaki M, Saito K, Miyaguchi S, Kojima S, *et al.* Effect of noisy galvanic vestibular stimulation on center of pressure sway of static standing posture. Brain Stimul. 2018; 11(1): 85-93. doi: 10.1016/j.brs.2017.10.007.

[5] Kuo AD. An optimal state estimation model of sensory integration in human postural balance. J Neural Eng. 2005; 2(3): S235-49. doi: 10.1088/1741-2560/2/3/S07.

[6] Angelaki DE, Cullen KE. Vestibular system: the many facets of a multimodal sense. Annu Rev Neurosci. 2008; 31: 125-50. doi: 10.1146/annurev.neuro.31.060407.125555.

[7] Welgampola MS, Ramsay E, Gleeson MJ, Day BL. Asymmetry of balance responses to monaural galvanic vestibular stimulation in subjects with vestibular schwannoma. Clin Neurophysiol. 2013; 124(9): 1835-9. doi: 10.1016/j.clinph.2013.03.015.

[8] Morita H, Kaji H, Ueta Y, Abe C. Understanding vestibular-related physiological functions could provide clues on adapting to a new gravitational environment. J Physiol Sci. 2020; 70(1): 17. doi: 10.1186/s12576-020-00744-3.

[9] Marchetti GF, Whitney SL, Redfern MS, Furman JM.

Factors associated with balance confidence in older adults with health conditions affecting the balance and vestibular system. Arch Phys med Rehabil. 2011; 92(11): 1884-91. doi: 10.1016/j.apmr.2011.06.015.

[10] Han BI, Song HS, Kim JS. Vestibular rehabilitation therapy: review of indications, mechanisms, and key exercises. J Clin Neurol. 2011; 7(4): 184-96. doi: 10.3988/jcn.2011.7.4.184.

[11] Sharma KG, Gupta AK. Efficacy and comparison of vestibular rehabilitation exercises on quality of life in patients with vestibular disorders. Indian J Otolaryngol Head Neck Surg. 2020; 72(4): 474-9. doi: 10.1007/s12070-020-01920-y.

[12] Kanekar N, Aruin AS. Improvement of anticipatory postural adjustments for balance control: effect of a single training session. J Electromyogr Kinesiol. 2015; 25(2): 400-5. doi: 10.1016/j.jelekin.2014.11.002.

[13] Sutter K, Oostwoud Wijdenes L, van Beers RJ, Medendorp WP. Even well-practiced movements benefit from repetition. J Neurophysiol. 2022; 127(5): 1407-16. doi: 10.1152/jn.00003.2022.

[14] Zemková E. Physiological mechanisms of exercise and its effects on postural sway: Does sport make a difference? Front Physiol. 2022; 13: 792875. doi: 10.3389/fphys.2022.792875.

[15] Kolimechkov S, Seijo M, Swaine I, Thirkell J, Colado JC, Naclerio F. Physiological effects of microcurrent and its application for maximising acute responses and chronic adaptations to exercise. Eur J Appl Physiol. 2023; 123(3): 451-65. doi: 10.1007/s00421-022-05097-w.

[16] Piras A, Zini L, Trofè A, Campa F, Raffi M. Effects of acute microcurrent electrical stimulation on muscle function and subsequent recovery strategy. Int J Environ Res Public Health. 2021; 18(9): 4597. doi: 10.3390/ijerph18094597.

[17] Williams NP, Kushwah N, Dhawan V, Zheng XS, Cui XT. Effects of central nervous system electrical stimulation on non-neuronal cells. Front Neurosci. 2022; 16: 967491. doi: 10.3389/fnins.2022.967491.

[18] Scinicariello AP, Eaton K, Inglis JT, Collins JJ. Enhancing human balance control with galvanic vestibular stimulation. Biol Cybern. 2001; 84(6): 475-80. doi: 10.1007/PL00007991.

[19] Lee A-R, Yu M, Kim J-H, Kim D-W, Kim J-JJJoBER. A response to postural response to sine curve vestibular electric stimulation during standing. J Biomed Eng Res. 2010; 31(3): 210-6. doi: 10.9718/JBER.2010.31.3.210

[20] Carvalho RL, Gomes MM, Franco LFdR, Abreu DCCd. Postural responses of galvanic vestibular stimulation: comparison between groups of older adults and young people. Revista Brasileira de Geriatria e Gerontologia. 2019; 22(05): e190091.

[21] Choi H-J, Kim S-SJJoKMR. Efficacy of microcurrent electrical neuromuscular stimulation with different types of stimulating electrodes. J Korean Med Rehabil. 2013; 23(3): 107-16.

[22] Mulavara AP, Fiedler MJ, Kofman IS, Wood SJ, Serrador JM, Peters B, *et al.* Improving balance function using vestibular stochastic resonance: optimizing stimulus characteristics. Exp Brain Res. 2011; 210(2): 303-12. doi: 10.1007/s00221-011-2633-z.

[23] Yi D, Lim H, Yim J. Effect of microcurrent stimulation on pain, shoulder function, and grip strength in early post-operative phase after rotator cuff repair. Medicina (Kaunas). 2021; 57(5): 491. doi: 10.3390/medicina57050491.

[24] Pires A, Silva TR, Torres MS, Diniz ML, Tavares MC, Gonçalves DU. Galvanic vestibular stimulation andits applications: a systematic review. Braz J Otorhinolaryngol. 2022; 88(Suppl 3): S202-S11. doi: 10.1016/j.bjorl.2022.05.010.

[25] Khowailed IA, Lee HJIJoSM. Neuromuscular control of ankle-stabilizing muscles-specific effects of sex and menstrual cycle. Int J Sports Med. 2021; 42(3):270-276. doi: 10.1055/a-1236-3654.

[26] Maharaj JN, Cresswell AG, Lichtwark GA. Tibialis anterior tendinous tissue plays a key role in energy absorption during human walking. J Exp Biol. 2019; 222(11): jeb191247. doi: 10.1242/jeb.191247.

[27] Juul-Kristensen B, Fallentin N, Hansson G-Å, Madeleine P, Andersen J, Ekdahl CJIJoIE. Physical workload during manual and mechanical deboning of poultry. Int J Ind Ergon. 2002; 29(2): 107-15. doi: 10.1016/S0169-8141(01)00051-8.

[28] Jonsson BJJohe. Measurement and evaluation of local muscular strain in the shoulder during constrained work. J Hum Ergol (Tokyo). 1982; 11(1): 73-88.

[29] Ko S-H, Yoon B-C, Kim J-S, Min K-O. Effects of microcurrent and high voltage pulsed galvanic current stimulation on fibular fracture healing of the rabbits. J Korea Contents Assoc. 2011; 11(10): 286-92. doi:10.5392/JKCA.2011.11.10.286

[30] Chen H, Hu Z, Chai Y, Tao E, Chen K, Asakawa T. Galvanic vestibular stimulation with low intensity improves dynamic balance. Transl Neurosci. 2021; 12(1): 512-21. doi: 10.1515/tnsci-2020-0197.

[31] Matsugi A, Oku K, Mori N. The effects of stochastic galvanic vestibular stimulation on body sway and muscle activity. Front Hum Neurosci. 2020; 14:591671. doi: 10.3389/fnhum.2020.591671.

[32] Brandt T, Daroff RB. Physical therapy for benign paroxysmal positional vertigo. Arch Otolaryngol. 1980; 106(8): 484-5. doi: 10.1001/archotol.1980.00790320036009.

[33] Hall CD, Herdman SJ, Whitney SL, Cass SP, Clendaniel RA, Fife TD, *et al.* Vestibular rehabilitation for peripheral vestibular hypofunction: an evidence-based clinical practice guideline: from the American Physical Therapy Association Neurology Section. J Neurol Phys Ther. 2016; 40(2): 124-55. doi: 10.1097/NPT.0000000000000120.

[34] Kolimechkov S, Seijo M, Swaine I, Thirkell J, Colado JC, Naclerio F. Physiological effects of microcurrent and its application for maximising acute responses and chronic adaptations to exercise. Eur J Appl Physiol. 2023;123(3):451-465. doi: 10.1007/s00421-022-05097-w.

Journal of Associated Medical Sciences 2025; 58(3): 137-146

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

Understanding occupation-based practice among Thai occupational therapy students: A mixed-methods study

Rapeepat Boonphirom[1,2], Tharathep Aoibumrung[1], Supaluck Phadsri[1]*

[1]Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Thanyarak Khonkaen Hospital, Khon Kaen, Thailand.

ARTICLE INFO

Article history:
Received 12 April 2025
Accepted as revised 16 June 2025
Available online 18 June 2025

Keywords:
Occupation-based practice, clinical experience, mixed-method study

ABSTRACT

Background: In Thailand, Occupation-based practice (OBP) has been emphasized in the curriculum of outcome-based education for Chiang Mai occupational therapy students towards the revised bachelor curriculum year 2021; however, it was essential in how they comprehensively understand and skills implementation in recent years.

Objective: The aim of this study was to explore an understanding of OBP among Thai occupational therapy students towards their clinical fieldwork and classroom experience.

Materials and methods: This study used a convergent mixed method by collecting data with a forty-item developed self-assessment questionnaire of the clinical fieldwork experience from third- and fourth-year students and employing focus group interviews with first- to fourth-year students between September and October 2022. Descriptive statistics were analyzed for seventy-five return questionnaires. Thirty-eight participants participated in a total of nine focus group interviews, and the qualitative data were analyzed by content and thematic analyses. Both sets of data were merged. An interpretation with six levels (remembering, understanding, applying, analyzing, evaluating, and creating) of cognitive domains of Bloom's revised taxonomy was used.

Results: The understanding of OBP in the clinical fieldwork experience years has ≥66.7% for all items within high self-assessment of understanding OBP in the six levels: 90.67%, 81.33%, 86.4%, 79.93%, 85.33%, and 83.47%, respectively. Two main themes, firstly, Occupation as the central focus of practice, and secondly, Importance of theoretical knowledge and experience, are presented. Analyzing, evaluating, and creating levels of a high cognitive domain were revealed in fourth-year students, while remembering, understanding, and applying levels were basic cognitive domains that support the students' understanding of OBP.

Conclusion: The results indicate an average comprehension of OBP in a high percentage, over 75% for overall levels with the supportive two themes, which is further useful in improving outcome-based learning and teaching of the occupational therapy curriculum.

Introduction

Occupation-based practice (OBP) is a concept and practice that emphasizes occupation, which is the core value of the occupational therapy profession.[1] Occupational therapists use occupation as a means and an end to promote clients' meaningful occupations.[1-3] Thus, OBP helps occupational therapists understand and focus on occupation through the lens of a client-centered and holistic approach by fostering occupational engagement,

* **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: supaluck.phad@cmu.ac.th
doi: 10.12982/JAMS.2025.082
E-ISSN: 2539-6056

such as activities of daily living (ADLs), work, education, leisure, and social participation.[3-5]

OBP challenges the power of occupation through the clinical implementation of both intervention and evaluation, such as stroke rehabilitation and marginalized youth programs,[5,6] highlighting OBP's transformative potential; however, its application continues to persist globally with barriers. The examples are the dominance of impairment-based practices, limited resources, systemic constraints, and time restrictions,[7-9] which are often found in the group of students and new graduates.[9-13] Regarding this, OBP and its implementation underlying the necessity of conceptual clarity and supportive systems in real contexts remain complex.

In Thailand, the study of OBP found working OBP to be consistent with integrated medical sciences,[13] although it is valued in improving services and achieving better occupational outcomes. This ambiguity suggests that the implementation of OBP from the perspective of Thai occupational therapists is still fragile. Interestingly, increasing evidence-based reasoning in OBP within the Thai context may help.

Thailand's occupational therapy curriculum is an important part of OBP knowledge improvement. The 2021 revised curriculum at Chiang Mai University introduces OBP early in alignment with the outcome-based education of the occupational therapy undergraduate program. The curriculum promotes pre-clinical education (Years 1 and 2), which builds a theoretical foundation through courses on anatomy, physiology, and core occupational therapy principles, while it supports the integration of OBP in advanced coursework and clinical internships during Years 3 and 4, ensuring graduates develop competencies required for professional practice. The clarification of the key occupational therapy terms,[2] including occupation-based (OB uses meaningful activities as both a means and goal of intervention), occupation-centered (OC centers occupation as the foundation of all reasoning and practice), and occupation-focused (OF focuses on resolving specific occupational performance issues) approaches, was initially discussed. Nevertheless, these concepts were disrupted by limited experience in clinical practice, especially during the COVID-19 pandemic, in fostering an understanding of clients' occupations in Thailand. Literature reviews[7-9,13] showed that occupational therapists struggle with their beliefs about the type of occupational therapy intervention as well, towards questioning on "what they do and how they do it".[2] Therefore, the challenges of unclear OBP concepts and complex skills implementation reveal an essential proposition for occupational therapists and students. To address these issues, embedding OBP principles into the curriculum is crucial to fostering students' understanding from pre-clinical education onwards.

This study used Bloom's Revised Taxonomy to develop a self-assessment questionnaire and interview guides of understanding OBP,[14,15] and to use as an analytical framework for data analysis following six levels of cognitive domains: remembering, understanding, applying, analyzing, evaluating, and creating. The aim of this research was to explore the understanding of OBP among Thai occupational therapy students towards their clinical fieldwork and classroom experience.

Materials and methods

Study design

This research used a concurrent mixed-methods design to mainly examine the understanding and application of OBP among Chiang Mai University occupational therapy students during the clinical fieldwork experience of the third- and fourth-year students through a quantitative study. A qualitative study was conducted through semi-structured focus group interviews for first- to fourth-year students.

Participants and recruitment

The population of third- and fourth-year occupational therapy students at Chiang Mai University was 100 (N=100). Inclusion criteria were enrolling in the first semester of the 2022 academic year, having clinical fieldwork experience for half of the clinical fieldwork courses (at least two groups of clients or more during the subject of clinical fieldwork practice 1 for Year 3 and clinical fieldwork practice 3 for Year 4), and volunteering for the research. The inclusion criteria were used for participants in both quantitative and qualitative studies. Participants were involved in completing a self-assessment questionnaire (10–15 minutes) and/or joining a focus group interview.

First- and second-year students, who had only theoretical experience and were enrolled under the OBE curriculum during the 2022 academic year, volunteered to participate in a focus group interview. Recruitment was carried out via announcements disseminated through student representatives, along with a distributed poster inviting voluntary participation.

Research instruments and data collection

The researchers developed a self-assessment questionnaire to examine the students' understanding of OBP towards clinical practice and classroom experience. The questionnaire uses yes/no questions. Scoring and interpretation of "yes" referred to "understanding the specific item", while "no" referred to "not understanding or unable to understand the specific item." The questionnaire ensured content validity with an Index of Item-Objective Congruence (IOC) assessment with three independent experts. The general criteria of the experts were obtaining a master's degree and having expertise in OBP, resulting from teaching or research. Additionally, one more specific expertise in OBP is needed, with at least 10 years of clinical experience or advanced experience in the measurement/assessment tool development in occupational therapy research. The first version of the questionnaire development consisted of 45 items. Four items scored IOC <0.50 were removed on the IOC process, and one was deleted due to redundancy. The revision questionnaire comprised 40 items, which underwent pilot testing with 16 volunteer students. The final version of the questionnaire had 40 items. The researcher distributed the questionnaire to

participants through class representatives, who had been assigned the appointment schedules in advance.

The researchers reviewed literature on OBP and the cognitive process of Bloom's revised taxonomy for developing the semi-structured interview guide, which was divided into the clinical experience and pre-clinical experience. The researchers interviewed the participants by using semi-structured and open-ended questions addressing the understanding and application of OBP during clinical fieldwork experience for third- and fourth-year students. The interview guide related to, for example, the ability to explain the concept of OBP from their perspective, to describe an understanding of OBP in application towards a situation's example, and the procedure in using assessment tools or designing a therapeutic intervention. The researcher used probing questions to deepen the exploration of their insight into the clinical implementation, client groups, and practice settings. In addition to pre-clinical experience, the researcher used the semi-structured interview guide to first- and second-year students to reflect an understanding of OBP's meaning and relevance to professional philosophy and the occupational therapy process. The interview guide related to, for example, their thoughts on OBP and its importance and value in occupational therapy, the applicable OBP in the occupational therapy process during class learning activities, and giving their reasons related to the principle.

The recruitment and data collection process occurred for both groups between September and October 2022. Focus group interviews were conducted at Chiang Mai University, with each session involving three to five participants per group and lasting 40-60 minutes. Each participant attended only one focus group session. The clinical and pre-clinical sessions were separately collected. Nine focus group interviews were conducted. The participants completed a consent form. They were briefed about the interview process and allowed to take notes and audio recordings before group interviews.

Data analysis

Quantitative data from the questionnaires were analyzed using descriptive statistics. The item interpretation followed the yes or no approach to understanding specific question items, while the percentage of mean score indicated understanding each level of the cognitive domains. This study applied Bloom's cut-off point to use in the data interpretation by categorizing into three categories with readjusted cut-off point consideration: the percentage of 76 and above shows a high self-assessment understanding of OBP at that level.[16,17] A medium self-assessment understanding of OBP is the percentage between 50 and 75, while a low self-assessment understanding of OBP is the percentage below 50.

Qualitative data from interviews were analyzed by content and thematic analyses.[18-20] The research team members, including a supervisor, double-checked the findings and interpretations to enhance the qualitative analytic rigor.

Results

Demographic data

The demographic characteristics of 75 respondents from the quantitative study are shown in Table 1. The results showed that the fourth-year occupational therapy students were 53.3%, and the third-year students were 46.7%. Most respondents were female, and almost all respondents were aged 20 years or older, with 98.7%.

Twenty-two participants with clinical experience participated in a total of five group interviews. All were the fourth-year students, aged 20 years or older, and most participants were female. They respond frequently to their clinical service experience for the client group with physical dysfunction and mental health. Likewise, their clinical fieldwork settings were in institutes or hospitals more than in community settings, except for a few who mentioned other settings.

Sixteen participants in the pre-clinical experience participated in a total of four focus group interviews. Four participants were first-year occupational therapy students, and the other twelve participants were second-year occupational therapy students, all aged 18 years or older.

Quantitative results

The quantitative results presented the frequency and percentage of six levels in understanding OBP based on Bloom's revised Taxonomy (Table 2).

Overall, six levels revealed a high self-assessment understanding of OBP: remembering (90.67%), understanding (81.33%), applying (86.40%), analyzing (78.94%), evaluating (85.33%), and creating (83.47%). The remembering level showed the highest percentage among the six levels, with the highest 96% of the "Participants were able to recall the meaning of OBP" item. Applying and evaluating levels were included in the top three levels, with "Participants were able to provide occupational therapy service with the belief that clients have the ability and potential to take action for their own health transformation" item was 98.7% the highest rate underlying the applying level, and "Participants were able to validate the organized information from the Occupational Profiles to understand the client's background, service needs and context" showed the highest percentage of the respondents. Afterwards, creating, understanding, and analyzing levels were exhibited from the fourth to the sixth. Questionnaire items with the highest percentage among such three levels were as follow: creating; "participants were able to design therapeutic activities aligned with the established goals" (94.70%), understanding; "participants were able to describe how occupational therapy views humans as active beings in creating health through occupational engagement" (82.70%) and "participant able to name an occupational therapy model and detail its key components" (82.70%), and analyzing; "participants were able to analyze how occupational therapy enables clients to learn through experiencing, thinking, and feeling by doing" (88%).

Table 1: Demographic data of quantitative study (N=75)

Demographic characteristic	Frequency (N)	Percentage (%)
Year of study		
3rd Year	35	46.70
4th Year	40	53.30
Gender		
Male	12	16.00
Female	63	84.00
Age		
Below 20 years	1	1.30
20 years and older	74	98.70
Clinical fieldwork experience in service groups (More than one response possible)		
Physical dysfunction	55	73.33
Mental health and psychiatry	54	72.00
Geriatrics	51	68.00
Pediatrics	49	65.33
Clinical fieldwork settings (More than one response possible)		
Institute/ hospital	71	94.67
Community settings	51	68.00
Other	7	9.33

Table 2: Frequency and percentage of OBP understanding levels based on Bloom's revised taxonomy.

Bloom's Taxonomy Level	Items *(Participants were able to …)*	3rd (N=35)	4th (N=40)	Total (N=75)	
		Yes N (%)	Yes N (%)	Yes N (%)	Yes N (%)
Remembering	- recall the meaning of occupation-based practice.	33 (94.30)	39 (97.50)	72 (96.00)	3 (4.00)
	- identify that occupational therapy services are based on the holistic approach.	31 (88.60)	33 (82.50)	64 (85.30)	11 (14.70)
	Average	32.00 (91.43)	36.00 (90.00)	68.00 (90.67)	7.00 (9.33)
Understanding	- explain how OBP helps in understanding clients' occupational demands.	27 (77.10)	33 (82.50)	60 (80.00)	15 (20.00)
	- describe how occupational therapy views humans as active beings in creating health through occupational engagement	31 (88.60)	31 (77.50)	62 (82.70)	13 (17.30)
	- provide examples to explain the term 'engagement in occupation.	28 (80.00)	32 (80.00)	60 (80.00)	15 (20.00)
	- name an occupational therapy model and detail its key components.	24 (68.60)	38 (95.00)	62 (82.70)	13 (17.30)
	Average	27.50 (78.57)	33.50 (83.75)	61.00 (81.33)	14.00 (18.67)
Applying	- provide occupational therapy services with the belief that clients have the ability and potential to take action for their own health transformation.	34 (97.10)	40 (100)	74 (98.70)	1 (1.30)
	- apply the client-centered approach in occupational therapy services.	32 (91.40)	39 (97.50)	71 (94.70)	4 (5.30)

Table 2: Frequency and percentage of OBP understanding levels based on Bloom's revised taxonomy. (Continue)

Bloom's Taxonomy Level	Items *(Participants were able to …)*	3rd (N=35)	4th (N=40)	Total (N=75)	
		Yes N (%)	Yes N (%)	Yes N (%)	Yes N (%)
	- summarize assessment results in alignment with theoretical concepts.	21 (60)	29 (72.50)	50 (66.70)	25 (33.30)
	- design occupational goals with client participation, focusing on meaningful occupations valued by the client.	28 (80.00)	31 (77.50)	59 (78.70)	16 (21.30)
	- review outcomes for further client-centered intervention by following up occupational therapy process and emphasizing occupational performance	26 (74.30)	39 (97.50)	65 (86.70)	10 (13.30)
	- recognize the use of occupation-based practice in evaluation, intervention, and outcome measurement during clinical training.	33 (94.30)	35 (87.50)	68 (90.70)	7 (9.30)
	- demonstrate an understanding OBP reflects the professional identity and roles.	32 (91.40)	35 (87.50)	67 (89.30)	8 (10.70)
	Average	29.40 (84.00)	35.40 (88.50)	64.80 (86.40)	10.20 (13.60)
Analyzing	- analyze how occupational therapy enables clients to learn through experiencing, thinking, and feeling by doing.	30 (85.70)	36 (90.00)	66 (88.00)	9 (12.00)
	- analyze the similarities, differences, and relationships of the models you use during clinical practice.	21 (60.00)	30 (75.00)	51 (68.00)	24 (32.00)
	- differentiate the tools, assessment methods, or approaches to determine whether they are based on measuring occupational performance as the target outcome.	21 (60.00)	29 (72.50)	50 (66.70)	25 (33.30)
	- set occupational goals that highlight occupational performance as a primary focus.	28 (80.00)	37 (92.50)	65 (86.70)	10 (13.30)
	- analyze the differences in types of occupational therapy intervention, such as differentiating between picking and shaping high-viscosity putty, molding clay, or preparing a sandwich for breakfast.	27 (77.10)	37 (92.50)	64 (85.30)	11 (14.70)
	Average	25.40 (72.57)	33.80 (84.50)	59.20 (78.93)	15.80 (21.07)
Evaluating	- determine the explanation of the concepts, theories, and/or rationale that support your decision in selecting an occupational therapy model.	24 (68.60)	33 (82.50)	57 (76.00)	18 (24.00)
	- validate the organized information from the Occupational Profile to understand the client's background, service needs, and context.	34 (97.10)	40 (100)	74 (98.70)	1 (1.30)
	- elect interviews, observations, and tests for assessments aimed at understanding the client's occupations and their impact on health.	33 (94.30)	39 (97.50)	72 (96.00)	3 (4.00)
	- accurately choose occupational performance assessment tools.	20 (57.10)	31 (77.50)	51 (68.00)	24 (32.00)

Table 2: Frequency and percentage of OBP understanding levels based on Bloom's revised taxonomy. (Continue)

Bloom's Taxonomy Level	Items *(Participants were able to …)*	3rd (N=35)	4th (N=40)	Total (N=75)	
		Yes N (%)	Yes N (%)	Yes N (%)	Yes N (%)
	- correctly conclude whether the selected assessments or methods align with evaluating the client's occupational performance.	22 (62.90)	36 (90.00)	58 (77.30)	17 (22.70)
	- recommend therapeutic activities in relation to the client's needs and context.	32 (91.40)	39 (97.50)	71 (94.70)	4 (5.30)
	- evaluate target outcomes as levels of occupational performance based on established goals.	29 (82.90)	36 (90.00)	65 (86.70)	10 (13.30)
	Average	27.70 (79.14)	36.30 (90.75)	64.00 (85.33)	11.00 (14.67)
Creating	- design assessment activities to identify the client's abilities, strengths, and assets required for performing their occupational needs.	19 (54.30)	38 (95.00)	57 (76.00)	18 (24.00)
	- design assessment activities that match the client's experiences and are close to their occupational context.	23 (65.70)	34 (85.00)	57 (76.00)	18 (24.00)
	- design assessment activities aligned with the occupation-based model you used as a reference.	22 (62.90)	37 (92.50)	59 (78.70)	16 (21.30)
	- design assessment activities that allow the client to express their abilities, such as in speaking, thinking, decision-making, behavioral expression, or other aspects.	24 (68.60)	38 (95.00)	62 (82.70)	13 (17.30)
	- design therapeutic activities that connect to the rehabilitation of client's daily living activities or work.	30 (85.70)	38 (95.00)	68 (90.70)	7 (9.30)
	- design therapeutic activities by integrating the foundation of medical knowledge and professional practice.	29 (82.90)	36 (90.00)	65 (86.70)	10 (13.30)
	- design therapeutic activities aligned with the established goals.	31 (88.60)	40 (100)	71 (94.70)	4 (3.30)
	- design therapeutic activities that are consistent with the theories being referenced.	25 (71.40)	36 (90.00)	61 (81.30)	14 (18.70)
	- create therapeutic activities that are interesting, challenging, or motivating for the client to engage in.	28 (80.00)	34 (85.00)	62 (82.70)	13 (17.30)
	- design therapeutic activities aimed at helping the clients divert their attention from distressing emotions to occupy them with the current task.	27 (77.10)	32 (80.00)	59 (78.70)	16 (21.30)
	- arrange the activity environment to be suitable for the therapeutic purposes (e.g., relaxation, stimulation, or focus enhancement).	28 (80.00)	34 (85.00)	62 (82.70)	13 (17.30)
	- modify or improve the design of therapeutic activities to facilitate achieving the therapeutic goals.	26 (74.30)	39 (97.50)	65 (86.70)	10 (13.30)

Table 2: Frequency and percentage of OBP understanding levels based on Bloom's revised taxonomy. (Continue)

Bloom's Taxonomy Level	Items *(Participants were able to …)*	3rd (N=35)	4th (N=40)	Total (N=75)	
		Yes N (%)	Yes N (%)	Yes N (%)	Yes N (%)
	- adapt activities to help the client return to meaningful and valued occupations or discover new occupations aligned with their interests.	29 (82.90)	39 (97.50)	68 (90.70)	7 (9.30)
	- improve clinical reasoning throughout clinical training by applying occupation-based practices.	20 (57.10)	40 (100)	60 (80.00)	15 (20.00)
	- integrate activity adaptations to make them both challenging and satisfying for the client, promoting personal health and well-being.	26 (74.30)	37 (92.50)	63 (84.00)	12 (16.00)
	Average	25.80 (73.71)	36.80 (92.00)	62.60 (83.47)	12.40 (16.53)

In each academic year, Year 3 students prioritized remembering (91.43%), applying (84%), evaluating (79.14%), understanding (78.57%), creating (73.71%), and analyzing (72.57%), respectively. They revealed a medium to high self-assessment understanding of OBP. While Year 4 students ranked all levels of understanding of OBP in high self-assessment, which consisted of creating (92%), evaluating (90.75%), remembering (90%), applying (88.5%), analyzing (84.5%), and understanding (83.75%) levels, respectively.

This study found an increase of 18.29% in the creating level in Year 4 compared to Year 3 students. Similarly, there was an 11.93% increase on the analyzing level and 11.61% on the evaluating level. These levels represented the results of their accumulating confidence of the self-assessment on the top tier Bloom's level, which demands the analysis performance: deconstruct information, identify underlying relationships, and differentiate between components; the evaluation performance: make informed judgments, critique arguments, or assess materials based on established criteria: and the creation performance: emphasizes the synthesis of information from multiple sources to formulate new ideas, designs, or products.

Qualitative results

The qualitative results presented an understanding of OBP in relation to the occupational therapy students' clinical and classroom experiences in two themes.

Theme 1

Occupation as the central focus of practice. The first theme emphasized the role of occupation at the core of both the conceptual framework and practical experiences in occupational therapy clinical experience. This theme included four subthemes. 1.1) Screening client data. The participants began the occupational therapy process by collecting clients' occupational performance history through observations and interviews. Screening data collection assists them in visualizing clients' occupational profiles and needs. 1.2) Selecting assessment tools focused on occupation. The participants learned to choose appropriate assessment tools in the measurement of the client's occupational performance based on the occupation-based model, particularly the Model of Human Occupation (MOHO) and the Canadian Occupational Performance Measure (COPM), to reinforce the importance of assessment tools focusing on occupation. 1.3) Determining occupational focus. The participants identified their integral knowledge in developing therapeutic activities by focusing on the client's occupation for their practice. Several clinical fieldwork settings are limited to hospitals, so an occupation-focused approach could be the closest to representing OBP. 1.4) Intrapersonal and interpersonal factors influencing OBP implementation. The participants acknowledged self-confidence, clinical instruction and supervision, and collaboration with peers and instructors in the classroom, influencing confidence in the implementation of OBP.

"We use screening data to better understand our clients before starting occupational therapy. By observing and conversing with clients, we develop a clear picture of their needs, which is essential as we begin our interventions."

"…Being occupation-based means working towards achieving occupational outcomes. …Ultimately, they must return to doing their own occupations".

"…I used MOHO with a client-centered approach-asking about their job, their future, and helping them explore occupations that matched their needs. Some might use COPM. … It is about focusing on occupation-based issues or assessing only what impacts their activities in their real context, clients' homes, and social situations, which helped us understand OBP better."

Theme 2

Importance of theoretical knowledge and experience. The second theme highlighted theoretical knowledge and preclinical experiences in bridging the transition from lower-order thinking (remembering, understanding, and applying) to higher-order thinking (analyzing, evaluating,

and creating) towards clinical scenarios. This theme consisted of five subthemes: 2.1) Client-centered practice and professional identity. The participants in both clinical and pre-clinical experience emphasized that an understanding of clients' meaningful occupations in OBP highlights client-centered practice and professional identity. 2.2) Understanding occupational profiles and assessment tools. This subtheme focuses on recognizing clients' occupational problems or needs to develop the occupational profiles. The participants learned that selecting appropriate assessment tools is crucial for evaluating occupational performance. However, those with limited clinical experience often struggled with data analysis and synthesizing information to create comprehensive occupational profiles. 2.3) Connecting theory to practice. The participants expressed that bridging classroom learning and real-world applications in fieldwork is essential to continuous clinical fieldwork. Instructor supervision and feedback are necessary to reinforce these connections and to engage them in understanding clients' occupations underlying the concept of OBP. 2.4) Distinguishing between OBP and occupation-focused approaches. The participants recognized the differences between OBP and occupation-focused interventions. Through discussions and simulations, they were able to explore how meaningful occupations evolve to enhance occupational performance alongside the occupational therapy process with professional reasoning, and further with clinical reasoning in fourth-year students, as well as improving confidence in the explanation and implementation of OBP. 2.5) Deepening understanding of OBP. This subtheme emphasized the role of critical reflection in enhancing participants' grasp of OBP. The participants know that clear theoretical concepts and technical terms can be learned through communication with instructors, which is essential. Collaborative lab activities and hands-on experience were valuable in promoting understanding of OBP concepts.

"During labs, we discuss and exchange ideas on case studies. This helps because our cases and approaches vary, providing multiple perspectives. I understood the meaning of technical terms more clearly when I combined theory from lectures with practical tasks."

"In a case scenario, I focus on what the client wants to do, like growing vegetables. I ask them what they want to plant and help them in the process. It highlights the client's willingness and occupation in a client-centered manner."

Discussion

The developmental progress of understanding OBP is explicitly reflected in the quantitative data, particularly fourth-year students' improvements in analyzing and creating levels of higher-order skills. Longer clinical fieldwork of practice results in advanced ability to analyze and create within OBP. The qualitative themes support quantitative results by suggesting clinical supervision, peer collaboration, and diverse practical settings as critical enablers of the cognitive growth.[21,22] The research results challenge the need to increase psychological confidence in the clinical implementation of OBP, while the need for this psychosocial development is similarly found in the literature,[22-24] which encourages skills improvement of OBP as occupational therapy students. Evidence-based practice (EBP) and writing the reflexive journal on the client's improvement in occupational goals and reflecting their own progression in OBP of both concept and clinical practice, are recommended.[22,25] Additionally, their comprehension would encourage work readiness amidst the barriers of OBP implementation in realistic contexts in the position of being occupational therapy practitioners soon or longer. Impairment-based medical service, work burden under related policies, and a dearth of assets and budgets are other examples of OBP confrontation in Thai occupational therapy service in Thailand;[13] however, the academic preparation would develop the work readiness of OBP application, like other countries.[6-12]

Among the clinical fieldwork experience group, the results reveal differences between the third-year and fourth-year students in their performance across Bloom's revised taxonomy levels, particularly in advanced cognitive processes. The fourth-year students showed markedly higher proficiency in the creating, analyzing, and evaluating levels, while third-year students confidently performed better only in the remembering level. Although Bloom's revised taxonomy is designed as a linear progression from basic to complex cognitive skills, the results of this study suggest that the learning process does not always occur sequentially; rather, it can progress non-linearly and may involve overlapping stages of cognitive development, as supported by qualitative results.

Additionally, the clinical fieldwork experience group demonstrated a theoretical understanding of applying occupation-based models, such as the Model of Human Occupation (MOHO) and the Canadian Model of Occupational Performance (CMOP). Their explanations logically linked core concepts to practical examples, addressing occupational problems derived from applying knowledge of occupational performance skills, performance patterns, and personal and environmental factors.[4] This ability reflects an integration of theoretical principles with practical reasoning, indicating a strong basic cognitive process. However, qualitative findings reveal implicit gaps in their ability to differentiate between nuanced occupational principles, particularly the distinctions among occupation-based (OB), occupation-centered (OC), and occupation-focused (OF) approaches.[2] These subtle differences remain challenging for the participants to fully understand.

Despite these challenges, the participants consistently highlighted occupation as a core focus of their clinical practice, demonstrating alignment with occupation-centered and occupation-focused principles.[2] Although the third- and fourth-year students were not part of the OBE curriculum, the structured clinical fieldwork course and the clinical staffs had successfully fostered the students' understanding of client-centered practice, which is an occupation as a pillar of practice. Students have exhibited enthusiasm, active engagement, and a strong

willingness to delve deeper into understanding OBP. In progressing further curriculum development, the primary academic goal of OBP should be to support the use of occupation and to nurture confidence in using occupation OBP throughout the occupational therapy process within supportive environments, ensuring alignment with the genuine OBP as articulated by Fisher.[2]

The preclinical fieldwork experience group gave the perspective of analyzing and creating skills as their weakest cognitive domains, likely due to their limited clinical experience and reliance on classroom-based learning. Their understanding reflects the critical role of clinical fieldwork in fostering higher levels of cognitive skills, which are essential for advanced problem-solving and clinical reasoning amidst future trends of occupational therapy education.[23,24] To discuss this, it is a necessity for the curriculum to integrate more case studies in the pre-clinical year, allowing students to practice solving clients' occupational problems or concerns through structured and real-world scenarios. Early entrance to clinical fieldwork experience is equally crucial, as it builds confidence, supports learning progression during clinical placements, and deepens understanding of the occupational therapy process. Providing students with greater autonomy to think and practice independently would further enhance their ability to develop the competencies required of occupational therapy clinicians.

Furthermore, gaining real-world experience in workplace settings-through opportunities for occupational therapy clinician observation, collaboration, and active practice-would moderate the stress of occupational therapy students in the professional transition process and work readiness. Embedding opportunities for service delivery development and fostering innovation in OBP within the curriculum are necessary steps for developing the professional identity of Thai occupational therapy amidst the digital technology era in health care promotion.

Limitation

The limitations found in participants from a single institution may limit the generalizability of the research results. The self-assessment questionnaire might limit the results from awareness of self-estimation. Future research should include other institutes to gain an understanding of OBP through various evaluations.

To improve understanding and application of OBP, occupational therapy curricula should integrate more practical and experiential learning activities, such as simulations, role-playing, and case studies. Furthermore, the incorporation of metacognitive reflection activities, such as guided journal reflection, peer discussions, and educator feedback, can enhance critical thinking and adaptability. We recommend a longitudinal study of the first- and second-year students to track the progression and to identify whether mentorship or specialized instruction is needed. Lastly, an evaluation of the graduate outcomes is essential for informing the development of continuing professional education programs, including the postgraduate education programs.

Conclusion

This study explored the understanding of OBP among occupational therapy students at Chiang Mai University by utilizing Bloom's revised taxonomy as an analytical framework. Quantitative and qualitative research results revealed that the occupational therapy students understood OBP with the basic lower-order cognitive domains. At the same time, higher-order skills were shown in third- and fourth-year students through clinical fieldwork experiences. The results, moreover, emphasized the need for tailored instructional strategies, particularly for second- and third-year students, to provide sheltered learning opportunities in the development of higher-level cognitive skills prior to clinical fieldwork as preparation. In conclusion, this research emphasizes the importance of aligning curriculum design and suggesting instructional methods to improve the understanding of OBP among Thai occupational therapy students, underlying the supportive principles of outcome-based education.

Ethical approval

Ethical approval for this study was obtained from the Research Ethics Committees of the Faculty of Associated Medical Sciences at Chiang Mai University (Approval Numbers 330/2565 and 356/2565). Participants provided written informed consent prior to data collection, and the research findings were reported anonymously.

Funding

This research received no external funding.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Rapeepat Boonphirom and Tharathep Aoibumrung; Methodology and data collection, Supaluck Phadsri; conceptualization, methodology, data collection, data curation, writing the original draft, review, and editing. All authors read and approved the final draft.

Acknowledgements

We extend our heartfelt gratitude to all participants for their invaluable contributions, time, and insights.

References

[1] Trombly CA. Occupation: Purposefulness and meaningfulness as therapeutic mechanism (Eleanor Clarke Slagle Lecture). Am J Occup Ther. 1995; 49(10): 690-72. doi: 10.5014/ajot.49.10.960.

[2] Fisher AG. Occupation-centered, occupation-based, occupation-focused: Same, same or different? Scand J Occup Ther. 2014; 20(3): 162-73. doi:10.3109/11038128.2012.754492.

[3] O'Brien JC, Hussey SM. Introduction to occupational therapy. 5th Ed. Missouri: Elsevier; 2018.

[4] American Occupational Therapy Association (AOTA). Occupational therapy practice framework: Domain and process (4th Edition). Am J Occup Ther. 2020;

74(Suppl 2): 7412410010p1-p87. doi: 10.5014/ajot.2020.74S2001.

[5] Wolf TJ, Chuh A, Floyd T, McInnis K, Williams E. Effectiveness of occupation-based interventions to improve areas of occupation and social participation after stroke: An evidence-based review. Am J Occup Ther. 2015; 69(1): 6901180060p1-6901180060p11. doi: 10.5014/ajot.2015.012195.

[6] Shea CK, Jackson N. Client perception of a client-centered and occupation-based intervention for at-risk youth. Scand J Occup Ther. 2015; 22(3): 173-80. doi: 10.3109/11038128.2014.958873.

[7] Hess-April L, Dennis L, Ganas N, Phiri L, Phoshoko P. Occupation-based practice in a tertiary hospital setting: Occupational therapists' perceptions and experiences. S Afr J Occup Ther. 2017; 47(3): 25-31. doi: 10.17159/2310-3833/2017/v47n3a5.

[8] Aas MH, Bonsaksen T. Exploring occupation-based practice among occupational therapists in hospitals and rehabilitation institutions. Scand J Occup Ther. 2022; 1-11. doi: 10.1080/11038128.2022.2059564.

[9] Tommaso AD, Wicks A, Scarvell J, Isbel S. Experiences of occupation based practice: An Australian phenomenological study of recently graduated occupational therapists. Br J Occup Ther. 2019; 82(7): 412-21. doi: 10.1177/0308022618823656.

[10] Coleman K, Senger N. Collaboration to impact occupation-based practice during level II fieldwork [master's thesis]. Department of Occupational Therapy, University of North Dakota; 2014: 43. Available from https://commons.und.edu/ot-grad/43.

[11] Keener AS, Hayden CL, Howell DM. Impacting occupational therapy assistant student knowledge about occupation and occupation-based practice. Open J Occup Ther. 2021; 9(2): 1-14. doi: 10.15453/2168-6408.1741.

[12] Jackson L, Bye K, Drumm D, Elizondo M, Murphy J. Student perspective on using evidence and occupation-based practice level II fieldwork. Am J Occup Ther. 2019; 73(4_Suppl_1): 7311505104p1. doi: 10.5014/ajot.2019.73S1-PO2031.

[13] Kaunnil A, Khemthong S, Sriphetcharawut S, Thicanpiang P, Sansri V, Thongchoomsin S, Permpoonputtana K, Smith CR. Occupational therapists' experiences and perspectives towards occupation-based practice in Thailand: A mixed-methods study. Br J Occup Ther. 2020; 84(1): 54-64. doi:10.1177/0308022620910402.

[14] Anderson LW, Krathwohl DR, editors. A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives. New York: Longman; 2001. Available from: https://www.quincycollege.edu/wp-content/uploadsAnderson-and Krathwohl_Revised-Blooms-Taxonomy.pdf

[15] Mernar TJ, Herzberger L. Understanding accreditation requirement trends of teaching occupation in occupational therapy curricula. Am J Occup Ther. 2024 ; 78(1): 7801347020. doi: 10.5014/ajot.2024.050342.

[16] Akalu Y, Ayelign B, Molla MD. Knowledge, attitude and practice towards COVID-19 among chronic disease patients at Addis Zemen Hospital, Northwest Ethiopia. Infect Drug Resist. 2020; 13: 1949. doi: 10.2147/IDR.S258736.

[17] Feleke BT, Wale MZ, Yirsaw MT. Knowledge, attitude and preventive practice towards COVID-19 and associated factors among outpatient service visitors at Debre Markos compressive specialized hospital, North-West Ethiopia, 2020. PLoS ONE. 2021; 16: e0251708. doi: 10.1371/journal.pone.0251708.

[18] Neuendorf KA. Content analysis and thematic analysis. In: Brough P, Editor. Advanced research methods for applied psychology: Design, analysis, and reporting. New York: Routledge Taylor & Francis; 2019: 211-23.

[19] Vaismoradi M, Jones J, Turunen H, Snelgrove S. Theme development in qualitative content analysis and thematic analysis. J Nurs Educ Pract. 2016; 6(5): 100-10. doi: 10.5430/jnep.v6n5p100.

[20] Vaismoradi M, Snelgrove S. Theme in qualitative content analysis and thematic analysis. Forum: Quali Soc Res. 2019; 20(3): 1-14. doi: 10.17169/fqs-20.3.3376.

[21] Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006; 3(2): 77-101. doi: 10.1191/1478088706qp063oa.

[22] Raiz A. Service learning in an occupation-based curriculum: Student commentary. Occup Ther Health Care. 2009; 21(1-2): 61-9. doi: 10.1080/J003v21n01_05.

[23] McLaughlin M, Bracciano AG. An exploration into effective pedagogies in occupational therapy education for the safe and effective use of physical agents. J Occup Ther Educ. 2023; 7(3). doi:10.26681/jote.2023.070303.

[24] Tyminski QP, Nguyen TM, McFadden K, Edwards D, Le V. Proposing a metacurriculum for occupational therapy education in 2025 and beyond. J Occup Ther Educ. 2019; 3(4). doi:10.26681/jote.2019.030404.

[25] Knecht-Sabres LJ. Experiential learning in occupational therapy: Can it enhance readiness for clinical practice? J Exp Educ. 2013; 36(1): 22-36. doi: 10.1177/1053825913481584.

Journal of Associated Medical Sciences 2025; 58(3): 147-155

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

Correlations between plasma clusterin levels and liver fibrosis in people living with HIV

Nopparat Deesophon[1], Janya Khattiya[2], Warisara Sretapunya[3], Chareeporn Akekawatchai[4,5*]

[1]Graduate Program in Medical Technology, Faculty of Allied Health Sciences, Thammasat University, Pathumthani Province, Thailand.
[2]Graduate Program in Biomedical Sciences, Faculty of Allied Health Sciences, Thammasat University, Pathumthani Province, Thailand.
[3]Nakorn Nayok Hospital, Nakorn Nayok Province, Thailand.
[4]Department of Medical Technology, Faculty of Allied Health Sciences, Thammasat University, Pathumthani Province, Thailand.
[5]Thammasat University Research Unit in Diagnostic Molecular Biology of Chronic Diseases related to Cancer (DMB-CDC), Pathumthani Province, Thailand.

ARTICLE INFO

Article history:
Received 2 March 2025
Accepted as revised 12 June 2025
Available online 20 June 2025

Keywords:
People living with HIV, PLWH, clusterin, liver fibrosis.

ABSTRACT

Background: Chronic liver disease has become a major health concern in people living with human immunodeficiency virus (PLWH). Clusterin has been involved in various pathologic conditions and its contributory role in liver complications in PLWH is still unclear.

Objective: This study aimed to detect plasma clusterin levels and assess their correlation with liver fibrosis in Thai PLWH.

Materials and methods: The study was conducted on 112 subjects, HIV-infected groups with (N=43) and without liver fibrosis (N=41), and uninfected controls (N=28). The subjects were evaluated for liver fibrosis using fibrosis-4 (FIB-4) score and aspartate aminotransferase to platelet ratio index (APRI), together with extracellular matrix fibrosis markers, laminin (LN), procollagen type III N-terminal peptide (PIIINP), hyaluronic acid (HA) and type IV collagen (IVC) and measured for plasma clusterin levels using an enzyme-linked immunosorbent assay.

Results: This study reported that medians of clusterin levels in uninfected controls, and the HIV patient groups without and with liver fibrosis were 1,779.38 (373.07-5,578.69), 1,602.73 (445.28-3,738.33) and 1,487.17 (0-3,111.33) µg/mL, respectively. Kruskal-Wallis test demonstrated a downward trend of clusterin levels in the HIV-infected groups without liver fibrosis and a significantly decreased level in the group with liver fibrosis (p<0.05), compared to those in uninfected controls. Pearson's correlation analysis indicated negative correlations of clusterin levels with fibrosis markers, FIB-4 score (r=-0.224, p=0.018), APRI (r=-0.211, p=0.026), LN (r=-0.284, P=0.005), HA (r=-0.234, p=0.013) and IVC (r=-0.299, p=0.002).

Conclusion: The study reported the lower levels of plasma clusterin in PLWH who develop liver fibrosis and their correlations with liver fibrosis assessed by the non-invasive markers.

Introduction

Human immunodeficiency virus (HIV) infection remains a major global health issue, with 39.9 million people living with HIV (PLWH) and 30.7 million people accessing anti-retroviral therapy at the end of 2023.[1] Effective ART has led to a decrease in acquired immunodeficiency syndrome (AIDS)-related mortality and morbidity. Life expectancy of PLWH has been closer to that of the general population. However, prolonged survival leads to increased evidence of non-AIDS-defining illnesses associated with age and premature ageing has been reported in PLWH.[2] Persistent viral replication in

∗ **Corresponding contributor.**
Author's Address: *Department of Medical Technology, Faculty of Allied Health Sciences, Thammasat University, Pathumthani Province, Thailand.*
E-mail address: ejareepo@tu.ac.th
doi: 10.12982/JAMS.2025.083
E-ISSN: 2539-6056

PLWH causes chronic immune activation and inflammation leading to an elevation of proinflammatory cytokines and premature thymic atrophy.[3] These increase the risk of developing non-AIDS-related comorbidities associated with immunosenescence in PLWH including chronic liver disease (CLD).[3,4] Multiple risk factors mainly hepatitis B and C coinfection, anti-retroviral regimens, and HIV replication in the liver have been suggested to influence chronic liver damage and inflammation, potentially leading to CLD and hepatocellular carcinoma.[4,5] Presently, molecular biomarkers associated with the progression of CLD in PLWH have been increasingly identified and require more investigation.[2]

Liver fibrosis is a common event caused by various etiologies, and its progression leads to advanced chronic liver disease, cirrhosis and eventually end-stage liver disease or hepatocellular carcinoma (HCC).[6,7] Therefore, determination of the degree of liver fibrosis is crucial for evaluating liver disease severity and progression.[6,7] Because the gold standard liver biopsy is an invasive procedure associated with the risk of complications, various non-invasive markers have been developed and suggested to facilitate clinical management of chronic liver diseases, including combinatorial serum marker, fibrosis-4 (FIB-4) score and aspartate aminotransferase (AST)-to-platelet ratio index (APRI), and direct liver fibrosis markers, hyaluronic acid (HA), type IV collagen (IVC), N-terminal polypeptide of type III procollagen (PIIINP) and laminin (LN).[6,7] Particularly, the direct biomarkers, HA, IVC, PIIINP and LN, which reflect the deposition and removal of extracellular matrix (ECM) during the disease progression, have been evaluated in terms of their correlations with hepatic function indices in liver cirrhosis and hepatoma caused by different etiologies.[7,8] The four direct biomarkers, together with tumor marker alpha fetoprotein (AFP) have also been utilized in predicting significant liver inflammation and fibrosis in chronic hepatitis B infection.[9,10] As accumulating studies support the importance of chronic liver disease in PLWH, monitoring liver fibrosis in PLWH has also been suggested.[4,5]

Clusterin (CLU) is synthesized as a highly glycosylated glycoprotein of 80 kDa, consisting of two polypeptide chains connected by disulfide bonds commonly found in the body fluid including serum and plasma. The secretory form of CLU, a heterodimeric complex of two 40-45 kDa subunits located in the extracellular space, exerts a chaperone-like activity by clearing cellular debris and misfolded stressed proteins.[11-13] Functionally, CLU is demonstrated to be cytoprotective and cytotoxic, resulting in conflicting consequences. Whereas cytoplasmic CLU expression was correlated with poor prognosis of hepatocellular carcinoma (HCC),[14] many studies supported the influence of CLU expression and function in the progression of infectious and non-infectious diseases.[15-19] The expression was associated with good prognosis in pancreatic adenocarcinoma and no correlation with breast carcinoma.[16,19] Moreover, the elevation of serum clusterin levels was reported in early rheumatoid arthritis,[17] while markedly lower serum clusterin levels were observed in hepatitis B-mediated acute-on-chronic liver failure.[18] Importantly, clusterin is considered one of the potential molecular markers of ageing with HIV[2] and a few studies suggest its association with disease progression and complications in PLWH.[20,21] However, there is no direct evidence indicating the contribution of clusterin to chronic liver disease in PLWH. Therefore, this study aimed to examine the expression level of plasma clusterin and its correlation with liver fibrosis, assessed noninvasively by fibrotic markers, in PLWH. The study provides more understanding of disease progression mechanisms and potential biomarkers for liver complications in PLWH.

Materials and methods

Study population

A total of 84 patients infected with HIV attending the Antiretroviral Therapy Clinic in Nakorn Nayok Hospital, Thailand, from September 2017 to October 2020 (aged 20-80 years, 56.0% male and 44.0% female) were recruited into the present study. The inclusion criteria were as follows: 1) the patients aged 18 years or older, 2) the patients with documented HIV infection, and 3) the patients with available blood samples and clinical data. The exclusion criteria were 1) the patients consuming alcohol, herbal medicine and steroid drugs, and 2) the patients having opportunistic infections. The patients were divided into 41 patients without liver fibrosis (36.6%) and 43 patients with significant liver fibrosis (38.4%). Twenty-eight uninfected controls, aged 25-72 years, 57.1% male and 42.9% female, were recruited from the Healthcare Clinic at Police General Hospital, Bangkok, Thailand, in November 2022. Inclusion criteria for controls were individuals aged >18 years with available blood samples and clinical data, and seronegative for HIV, hepatitis B and C virus infection. Individuals with significant liver fibrosis, assessed by fibrosis-4 (FIB-4) score and aspartate aminotransferase (AST)-to-platelet ratio index (APRI) were excluded.[22-24] The study protocol was reviewed and approved by the Human Ethics Committees No. 3, Thammasat University, Pathumthani Province, Thailand (Approve No. 121/2565) and the Certificated Biological Safety Committee of Thammasat University, Pathumthani, Thailand (Approve No. 079/2565). Figure 1 showed the recruitment process of participants in this study.

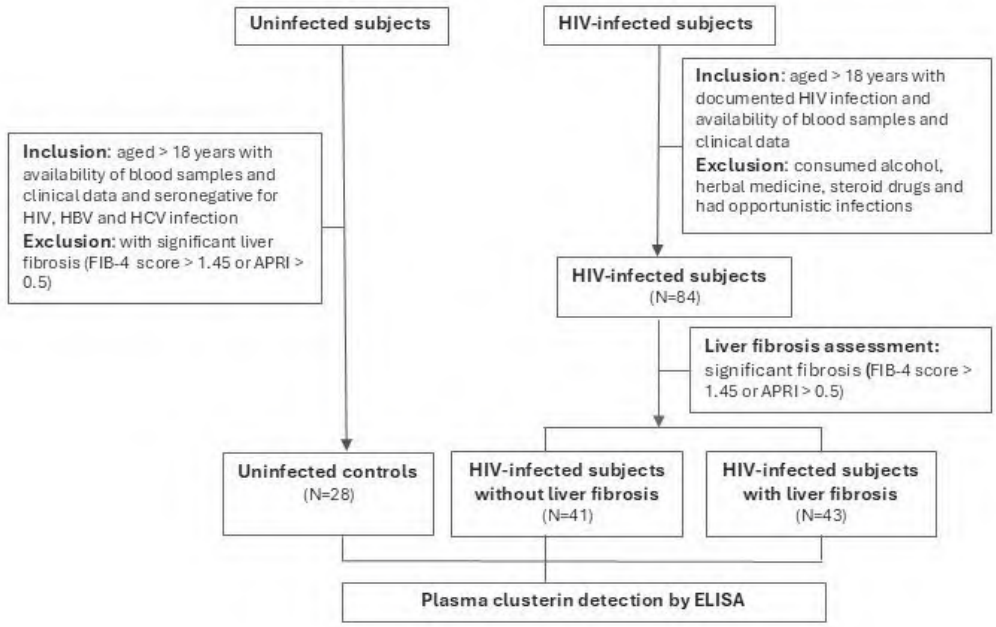


***Figure 1**. The recruitment process for participants in this study was based on a non-invasive assessment of liver fibrosis. FIB-4: fibrosis-4 score, APRI: AST to platelet ratio index, ELISA: enzyme-linked immunosorbent assay.*

Clinical data and laboratory investigation

Clinical and laboratory data were obtained as described in the previous studies.[25-27] Ethylene-diamine-tetra-acetic acid (EDTA) blood samples remaining after routine testing were subjected to plasma separation within 8 hours after blood collection and stored at -80°C until use. In this study, liver fibrosis in the subjects was determined by FIB-4 score, classified into class 1 (≤1.45), class 2 (1.46-3.25) and class 3 (>3.25) and by APRI, classified into class 1 (≤0.5), class 2 (0.51-1.5) and class 3 (>1.5). A significant liver fibrosis was defined as FIB-4 score >1.45 or APRI >0.5.[22-24] FIB-4 score was calculated using the formula: FIB-4 score = age [years] x aspartate aminotransferase (AST) level [U/L]/platelet count [10^9/L] x alanine aminotransferase (ALT)$^{1/2}$ [U/L].[22] APRI was calculated by the formula: APRI = (AST level [U/L]/upper limit of normal AST) x 100/ platelet count [10^9/L].[24] In addition, levels of extracellular matrix (ECM) fibrotic markers, laminin (LN), procollagen type III N-terminal peptide (PIIINP), hyaluronic acid (HA) and type IV collagen (IVC), in addition to tumor marker α-fetoprotein (AFP) in plasma samples were measured using electrochemiluminescence immunoassay (Mindray CL-900i and LN, PIIINP, HA, IVC and AFP test kits, all Mindray Medical International Co., Ltd.)

Detection of clusterin levels by sandwich enzyme-linked immunosorbent assay

Clusterin levels in plasma samples were measured using a commercial ELISA kit (Human clusterin Duoset ELISA kit, R&D Systems) according to the manufacturer's instructions. Clusterin concentrations were reported in µg/mL. The intra- and inter-assay coefficients of variability were less than 10% (4.85-9.22) and 15% (11.86), respectively.

Data analysis

Descriptive statistics, median and percentage were used to describe the characteristics of the study population. Chi-square and Fisher's exact tests analyzed differences of categorical variables between the study groups, whereas those with continuous variables were tested using one-way ANOVA, Mann-Whitney U, Kruskal-Wallis tests and Dunn's pos hoc tests. The $p<0.05$ was considered to indicate a statistically significant difference. The PASW Statistic 18 software (SPSS Inc.) and GraphPad Prism 9.1.1 (Dotmatics) were used for statistical analysis.

Results

Characteristics of the study population.

Table 1 demonstrated the clinical characteristics of the three study groups, HIV-infected subjects with (N=43) and without liver fibrosis (N=41) and uninfected controls (N=28). The data indicated no difference in the sex distribution but significant age differences between the study groups. Patients without liver fibrosis had lower median ages than those with liver fibrosis and uninfected controls ($p=0.004$). Median levels of non-invasive fibrotic markers, FIB-4 score, APRI, LN, PIIINP, HA and IVC, and tumor marker AFP significantly differed among the three groups (all $p<0.05$). Notably, analysis of the HIV-infected groups indicated that there is no difference between the distribution of patients with HBV and HCV coinfection in the HIV-infected groups with and without liver fibrosis. However, the data indicated significant differences in $CD4^+$ cell count, ARV regimen and duration of ART between the two HIV-infected groups (Table 1).

Table 1. Demographic data and clinical characteristics of uninfected controls and PLWH with or without liver fibrosis recruited in this study.

Characteristics (N=112)	Uninfected control group (N=28)	HIV patient group without liver fibrosis (N=41)	HIV patient group with liver fibrosis (N=43)	*p* value
Subjects	28 (25.0)	41 (36.6)	43 (38.4)	
Age (years)[b]	49.9±15.6)	41.2±9.2)	49.8±11.9)	0.004*
Gender[a]				
Male	16 (57.1)	19 (46.3)	28 (65.1)	
Female	12 (42.9)	22 (53.7)	15 (34.9)	0.221
Liver fibrosis markers[c]				
FIB-4 score	0.86 (0.28-1.40)	0.77 (0.35-1.45)	2.84 (0.80-52.62)	<0.001*
APRI	0.18 (0.12-0.32)	0.22 (0.12-0.46)	0.88 (0.22-36.44)	<0.001*
Liver fibrosis and tumor markers (ng/mL)[c]				
Laminin (N=99)	32.20 (10.06-61.51)	35.24 (20.11-53.05)	41.73 (24.66-96.32)	0.004*
Procollagen III N-terminal peptides (N=112)	10.07 (4.85-85.13)	7.49 (4.12-23.95)	11.66 (4.33-118.30)	0.001*
Hyaluronic acid (N=112)	51.10 (14.01-455.90)	23.44 (8.89-491.42)	76.18 (7.09-2549.84)	<0.001*
Type IV collagen (N=108)	13.87 (2.30-75.24)	38.75 (23.38-123.83)	48.82 (19.26-458.64)	<0.001*
Alpha-fetoprotein (N=112)	2.54 (1.06-5.92)	1.85 (0.95-5.78)	2.38 (0.91-74.54)	0.041*
Clusterin levels (µg/mL)	1,779.38 (373.07-5,578.69)	1,602.73 (445.28-3,738.33)	1,487.17 (0-3,111.33)	0.038*
HBV and HCV coinfection[b] (N, %)				
HIV monoinfection	ND	29 (82.9)	24 (70.6)	
HIV/HBV coinfection	ND	3 (8.6)	6 (17.6)	
HIV/HCV coinfection	ND	3 (8.6)	4 (11.8)	0.422
CD4$^+$ cell count (cells/µL)[d] (N=81)	ND	487 (4-873)	281 (2-1,282)	0.303
CD4$^+$ cell count (cells/µL)[b] (N=81)				
<350	ND	12 (30.0)	22 (53.7)	
≥350	ND	28 (70.0)	19 (46.3)	0.031*
Duration of ARV treatment[b] (N=83)	ND	82.00 (40.50-123.50)	27.00 (0.00-97.00)	0.087
Duration of ARV treatment[a] (N=83)				
Naive to ARV treatment	ND	2 (4.9)	9 (21.4)	
≤6 month	ND	3 (7.3)	5 (11.9)	
>6 months	ND	36 (87.8)	28 (66.7)	0.045*
ART[a] (N=83)				
Naive to ART	ND	2 (4.9)	9 (21.4)	
ART	ND	39 (95.1)	33 (78.6)	0.026*
ARV regimens[a] (N=72)				
Lamivudine/zidovudine/nevirapine	ND	10 (25.6)	3 (9.1)	
Tenofovir, efavirenz, lamivudine or emtricitabine	ND	20 (51.3)	14 (42.4)	
Others	ND	9 (23.1)	16 (48.5)	0.042*

***Note**: [a]number (%), [b]mean±SD, [c]median (range). Some variables had missing data, and N is given in parentheses, *p<0.05, ART: antiretroviral therapy, ARV: antiretroviral, AST: aspartate aminotransferase, ALT: alanine aminotransferase, FIB-4: fibrosis-4 score, APRI: AST to platelet ratio index.*

Plasma levels of FIB-4, APRI, ECM fibrotic markers, LN, PIIINP, HA and IVC, and tumor marker AFP were analyzed (Figure 1). Figure 1A-C indicated that levels of FIB-4 score, APRI and LN in the HIV-infected group without liver fibrosis were like those in the control group (Figure 1A-C). However, the levels in the infected group with liver fibrosis were significantly higher than those in the others. Figure 1D indicated that the levels of IVC were also significantly higher in the two HIV-infected groups, compared to the control group. However, the IVC levels in the infected groups with and without liver fibrosis were similar. In Figure 1E, levels of HA were significantly lower in HIV-infected patients without liver fibrosis than in the control groups, whereas the levels in the infected patients with liver fibrosis were significantly higher than those without liver fibrosis. PIIINP levels in the three study groups were like those observed with HA levels (Figure 2F). Additionally, there is a significant difference only between AFP levels in HIV-infected patients without liver fibrosis and uninfected controls (Figure 2G).

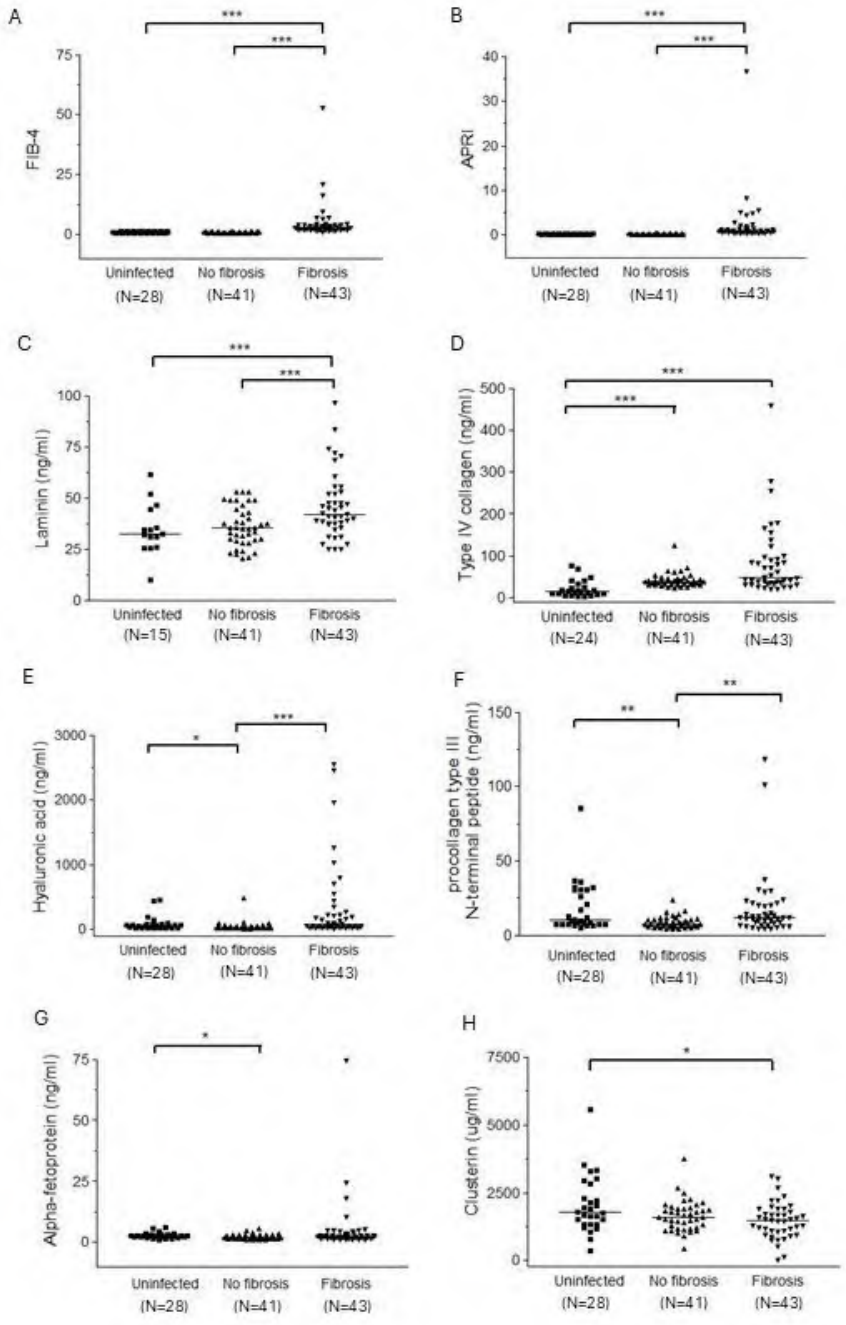


Figure 2. *Plasma levels of FIB-4 (A), APRI (B), and fibrosis markers, laminin (C), type IV collagen (D) hyaluronic acid (E), procollagen type III N-terminal peptide (F), alpha-fetoprotein (G) and clusterin (H) in HIV-infected patients with and without liver fibrosis compared to those in uninfected control groups. *p <0.05, **p <0.01 *** p <0.001, ****p<0.0001, and ns, p>0.05, FIB-4: fibrosis-4 score, APRI: AST to platelet ratio index*

Plasma clusterin levels in the uninfected controls, HIV-infected patient groups with and without liver fibrosis

Clusterin levels were detected in plasma samples from three study groups. Median levels in the uninfected control, HIV-infected groups with and without liver fibrosis were 1,779.38 µg/mL (373.07-5,578.69), 1,487.17 µg/mL (0-3,111.33) and 1,602.73 µg/mL (445.28-3,738.33), respectively. Statistical analysis indicated significant differences in clusterin levels between the three groups (Table 1). Figure 1H also showed a decreasing trend of clusterin levels in the HIV-infected groups, compared to uninfected controls. Importantly, a significant decrease in clusterin levels was observed in the infected group having liver fibrosis compared to the uninfected control group. In addition, Pearson's correlation analysis in the study groups, uninfected and HIV-infected patients, indicated reverse correlations of clusterin levels with AST, FIB-4 score, APRI, LN, HA and IVC (p =0.021, p=0.018, p=0.026, p=0.005, p=0.013 and p=0.002, respectively) (Figure3) and no correlation with ALT, PIIINP and AFP (Table 2). Interestingly, the analysis in the HIV-infected groups demonstrated no correlation between the clusterin levels and $CD4^+$ cell counts (Table 2).

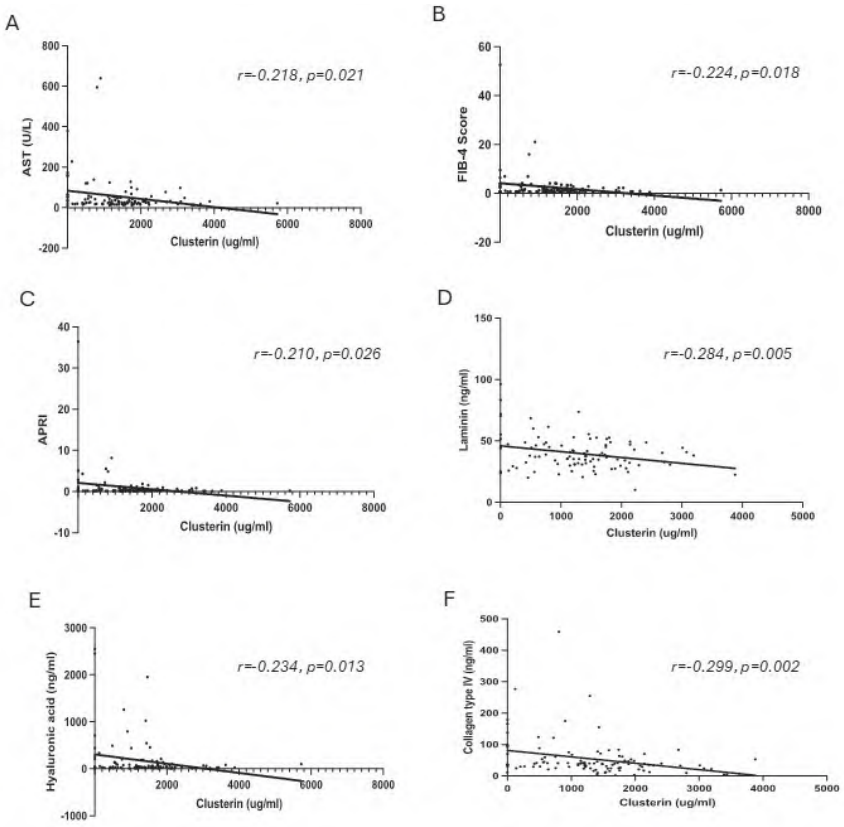


Figure 3. *Pearson correlation analysis of plasma clusterin levels with abnormal liver parameters in the uninfected and HIV-infected patients. The correlations of clusterin levels with AST (A), FIB-4 (B), APRI (C), laminin (D), hyaluronic acid (E) and type IV collagen (F) were demonstrated (p<0.05), AST: aspartate aminotransferase, FIB-4: fibrosis-4 score, APRI: AST to platelet ratio index*

Table 2. Pearson correlation of plasma clusterin levels with abnormal liver parameters and immune status in the uninfected and HIV-infected patients.

Abnormal liver parameters	Correlation coefficient *(r)*	*p* value
Aspartate aminotransferase	- 0.218	0.021*
Alanine aminotransferase	- 0.118	0.241
Fibrosis-4 score	- 0.224	0.018*
Aspartate aminotransferase to platelet ratio index	- 0.211	0.026*
Laminin	- 0.284	0.005*
Procollagen III (N-terminal peptides)	- 0.132	0.165
Hyaluronic acid	- 0.234	0.013*
Collagen type IV	- 0.299	0.002*
Alpha-fetoprotein	- 0.127	0.082
$CD4^+$ cell count (N=81)	0.146	0.192

***Note**: *data shown as p<0.05.*

Discussion

Accumulating studies indicate an increased burden of chronic liver disease in PLWH,[4,5] and monitoring for liver fibrosis in the PLWH has been suggested. Our previous studies have demonstrated the high prevalence of liver abnormalities and their risk factors in Thai PLWH.[25-27] In this study, the two HIV-infected and uninfected control groups were recruited and classified according to significant liver fibrosis evaluated using the non-invasive FIB-4 score (>1.45) and APRI (>0.5). The classification was supported by the additional analysis of median FIB-4, APRI and the established ECM fibrotic markers, LN, HA, IVC, and PIIINP, in the three study groups. Our data indicated the elevation of LN and IVC levels in the HIV-infected group with liver fibrosis compared to those without fibrosis and the controls, supporting the assessment of significant liver fibrosis using FIB-4 and APRI. Although there is no evidence showing the evaluation of the ECM biomarkers in HIV-mediated liver fibrosis, these biomarkers have been demonstrated previously to predict significant liver inflammation and fibrosis in chronic hepatitis B.[9,10] In this study, plasma clusterin levels and their correlation with significant liver fibrosis, as assessed by non-invasive fibrosis markers, have been reported in PLWH.

Hepatocyte apoptosis during liver injury is an initial event in all liver disease etiologies including chronic manageable HIV infection with ART.[28,29] HIV itself and the adverse effects of ART appear to be a major cause of hepatic injury, liver inflammation and fibrosis.[29] Clusterin is known to exert extracellular chaperone function by clearing cell debris and abnormal proteins.[11-13] It also protects cells from apoptosis induced by many stressors.[30,31] As hepatocytes are the major source of circulating clusterin,[32] hepatocyte damage may directly cause the reduction of plasma clusterin levels observed in the HIV-infected groups. In this study, the clusterin level was significantly reduced in the HIV-infected group with liver fibrosis, and negatively correlated with the well-established liver fibrosis markers, FIB-4 and APRL, and ECM markers, LN, HA, IVC, and PIIINP. This finding was consistent with the previous study indicating the lower clusterin levels as a predictor of the severity and prognosis of hepatitis B virus-related acute on chronic liver failure.[18] However, a few studies indicate the upregulation of clusterin in the brains of HIV-mediated acquired immunodeficiency syndrome and the increased serum concentrations of clusterin in HIV positive pregnancies.[20,21] This inconsistency may be due to specific mechanisms of complications in HIV disease. Notably, our analysis suggested no significant correlation of clusterin levels with the immune status of these HIV-infected patients assessed by $CD4^+$ cell counts, possibly indicating no involvement of circulating clustrin in the severity of underlying HIV disease in this study group. The findings generated from this study suggest the potential contribution of clusterin in the progression of liver inflammation and fibrosis in PLWHIV.

Limitations

This study had limitations due to the characteristics of the study groups. Firstly, the study groups have differences in age, $CD4^+$ cell count, ART, ARV regimens and ART duration. Our previous study in the HIV-infected groups suggested that age and $CD4^+$ cell count, but not ART, were predictive factors for liver fibrosis.[27] Therefore, these risk factors may partially influence the impact of clusterin on CLD progression in the PLWH. Secondly, there was no data on the status and treatment of HBV and HCV infection, especially HBV and HCV viral loads in these study groups. Although there was no significant difference between the distribution of HBV and/or HCV coinfection in PLWH with and without liver fibrosis in this study, the effect of viral hepatitis on liver fibrosis in this study group could not be ruled out. Thirdly, the stability of plasma analytes during storage is a concern in bioanalytical methods. In this study, to avoid the effect of freeze/thaw cycles and temperatures, the uninfected- and HIV-infected plasma samples were aliquoted and stored at -80 °C before examination at the same time. Even though the stability in the samples with long-term storage, which may affect the clusterin levels reported in this study, has not been clear, the previous study reported that an ELISA could detect variants of clusterin in the serum samples stored at

-80°C up to 24 months.[33] Lastly, the study was conducted in a relatively small number of subjects, which may limit statistical significance for variables tested, and further studies in a larger HIV-infected group are warranted.

Conclusion

The present study demonstrated plasma clusterin levels in PLWH who developed liver fibrosis and their correlations with non-invasive markers, FIB-4 and APRI, and ECM markers, LN, HA, IVC, and PIIINP, suggesting the contributory role of clusterin in the progression of liver fibrosis. This finding may provide more understanding in the development of liver inflammation and fibrosis in PLWH.

Conflict of interest

The authors declare no conflict of interest.

Funding

The study was supported by the Thailand Science Research and Innovation Fundamental Fund fiscal year 2023 (Project no.2493097) and Thammasat University Research Unit in Diagnostic Molecular Biology of Chronic Diseases related to Cancer (DMB-CDC).

Ethical approval

The study protocol has been approved by the human Ethics Committees No. 3, Thammasat University, Thailand (Project no. 121/2565).

Credit authorship contribution statement

Chareeporn Akekawatchai contributed to funding acquisition resources, supervision, study design, manuscript preparation, review and editing. Nopparat Deesophon performed data collection and analysis, and manuscript preparation. Janya khattiya contributed to data collection and analysis. Warisara Sretapunya participated in collecting blood samples and clinical data. All authors read and approved the final version of the manuscript.

Acknowledgements

Special thanks to all patients and staff in Nakorn Nayok Hospital for their contribution to this study.

References

[1] UNAIDS. Global HIV & AIDS statistics-Fact sheet. 2023. Available from: https://www.unaids.org/en/resources/fact-sheet.

[2] Rodes B, Cadinanos J, Esteban-Cantos A, Rodriguez-Centeno J, Arribas JR. Ageing with HIV: Challenges and biomarkers. EBioMedicine. 2022; 77: 103896 doi: 10.1016/j.ebiom.2022.103896.

[3] Zicari S, Sessa L, Cotugno N, Ruggiero A, Morrocchi E, Concato C, et al. Immune Activation, Inflammation, and Non-AIDS Co-Morbidities in HIV-Infected Patients under Long-Term ART. Viruses. 2019; 11(3): 200. doi: 10.3390/v11030200.

[4] Sherman KE, Thomas DL. HIV and liver disease: a comprehensive update. Top Antivir Med. 2022; 30(4): 547-58. Available from: https://www.ncbi.nlm.nih.gov/pubmed/36375129.

[5] Chamroonkul N, Bansal MB. HIV and the liver. Nat Rev Gastroenterol Hepatol. 2019; 16(1): 1-2. doi: 10.1038/s41575-018-0085-7.

[6] Nallagangula KS, Nagaraj SK, Venkataswamy L, Chandrappa M. Liver fibrosis: a compilation on the biomarkers status and their significance during disease progression. Future science OA. 2018; 4(1): FSO250. doi: 10.4155/fsoa-2017-0083.

[7] Lai JC, Liang LY, Wong GL. Noninvasive tests for liver fibrosis in 2024: are there different scales for different diseases? Gastroenterol Rep (Oxf). 2024; 12: goae024. doi: 10.1093/gastro/goae024.

[8] Ma HY, Dong L, Quan SZ, Li RY, Wang XR. Comparison of four markers of hepatic fibrosis and hepatic function indices in patients with liver cirrhosis and hepatoma. Ann Palliat Med. 2021; 10(4): 4108-21. doi:10.21037/apm-20-1623.

[9] Dong H, Xu C, Zhou W, Liao Y, Cao J, Li Z, et al. The combination of 5 serum markers compared to FibroScan to predict significant liver fibrosis in patients with chronic hepatitis B virus. Clin Chim Acta. 2018; 483: 145-50. doi:10.1016/j.cca.2018.04.036.

[10] Liao Y, Gong J, Zhou W, Dong H, Liang J, Luo M, et al. Serum liver fibrosis markers discriminate significant liver inflammation in chronic hepatitis B patients with normal or near-normal alanine aminotransferase. J Med Virol. 2019; 91(4): 642-9. doi:10.1002/jmv.25364.

[11] Aigelsreiter A, Janig E, Sostaric J, Pichler M, Unterthor D, Halasz J, et al. Clusterin expression in cholestasis, hepatocellular carcinoma and liver fibrosis. Histopathology. 2009; 54(5): 561-70. doi: 10.1111/j.1365-2559.2009.03258.x.

[12] Materia S, Cater MA, Klomp LW, Mercer JF, La Fontaine S. Clusterin (apolipoprotein J), a molecular chaperone that facilitates degradation of the copper-ATPases ATP7A and ATP7B. J Biol Chem. 2011; 286(12): 10073-83. doi: 10.1074/jbc.M110.190546.

[13] Rohne P, Prochnow H, Koch-Brandt C. The CLU-files: disentanglement of a mystery. Biomol Concepts. 2016; 7(1): 1-15. doi: 10.1515/bmc-2015-0026.

[14] Kang YK, Hong SW, Lee H, Kim WH. Overexpression of clusterin in human hepatocellular carcinoma. Hum Pathol. 2004; 35(11): 1340-6. doi: 10.1016/j.humpath.2004.07.021.

[15] Satapathy S, Wilson MR. The Dual Roles of Clusterin in Extracellular and Intracellular Proteostasis. Trends Biochem Sci. 2021; 46(8) :652-60. doi: 10.1016/j.tibs.2021.01.005.

[16] Redondo M, Villar E, Torres-Munoz J, Tellez T, Morell M, Petito CK. Overexpression of clusterin in human breast carcinoma. Am J Pathol. 2000; 157(2): 393-9. doi: 10.1016/S0002-9440(10)64552-X.

[17] Kropackova T, Mann H, Ruzickova O, Sleglova O, Vernerova L, Horvathova V, et al. Clusterin serum levels are elevated in patients with early rheumatoid arthritis and predict disease activity and treatment

response. Sci Rep. 2021; 11(1): 11525. doi: 10.1038/s41598-021-90973-2.

[18] Liu H, Li Y, Gao F, Meng P, Yu H, Wu T, et al. Serum Clusterin: A Potential Marker for Assessing the Clinical Severity and Short-Term Prognosis of Hepatitis B Virus-Related Acute-on-Chronic Liver Failure. Dis Markers. 2020; 2020: 8814841. doi: 10.1155/2020/8814841.

[19] So A, Sinnemann S, Huntsman D, Fazli L, Gleave M. Knockdown of the cytoprotective chaperone, clusterin, chemosensitizes human breast cancer cells both in vitro and in vivo. Mol Cancer Ther. 2005; 4(12): 1837-49. doi: 10.1158/1535-7163.MCT-05-0178.

[20] Mlambo ZP, Varaden D, Moodley J, Naicker T. Are concentrations of clusterin and beta-2-glycoprotein I dysregulated in HIV associated preeclampsia? Eur J Obstet Gynecol Reprod Biol. 2020; 251: 1-7. doi: 10.1016/j.ejogrb.2020.03.036.

[21] Torres-Munoz JE, Redondo M, Czeisler C, Roberts B, Tacoronte N, Petito CK. Upregulation of glial clusterin in brains of patients with AIDs. Brain Res. 2001; 888(2): 297-301. doi:10.1016/s0006-8993(00)03052-3.

[22] Foca E, Fabbiani M, Prosperi M, Quiros Roldan E, Castelli F, Maggiolo F, et al. Liver fibrosis progression and clinical outcomes are intertwined: role of CD4+ T-cell count and NRTI exposure from a large cohort of HIV/HCV-coinfected patients with detectable HCV-RNA: A MASTER cohort study. Medicine (Baltimore). 2016; 95(29): e4091. doi: 10.1097/MD.0000000000004091.

[23] Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology. 2006; 43(6): 1317-25. doi: 10.1002/hep.21178.

[24] Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology. 2003; 38(2): 518-26. doi: 10.1053/jhep.2003.50346.

[25] Akekawatchai C, Changsri K, Tunkor A, Phuegsilp C, Soimanee T, Fungkraja M, et al. Lack of Association between IFN-gamma, CXCL10 and TGF-beta1 Gene Polymorphisms and Liver Complication in HIV-infected Thais. Asian Pac J Cancer Prev. 2022; 23(4): 1279-84. doi: 10.31557/APJCP.2022.23.4.1279.

[26] Akekawatchai C, Sretapunya W, Pipatsatitpong D, Chuenchit T. Hepatitis B or C virus coinfection in and risks for transaminitis in human immunodeficiency virus - infected Thais on combined antiretroviral therapy. Asian Biomedicine. 2015; 9(3): 353-61.

[27] Chiraunyanann T, Changsri K, Sretapunya W, Yuenyongchaiwat K, Akekawatchai C. CXCL12 G801A polymorphism is associated with significant liver fibrosis in HIV-infected Thais: a cross-sectional study. Asian Pac J Allergy Immunol. 2019; 37(3): 162-70. doi: 10.12932/AP-160917-0162.

[28] Roehlen N, Crouchet E, Baumert TF. Liver Fibrosis: Mechanistic Concepts and Therapeutic Perspectives. Cells. 2020; 9(4): 875. doi:10.3390/cells9040875.

[29] Gruevska A, Moragrega AB, Cossarizza A, Esplugues JV, Blas-Garcia A, Apostolova N. Apoptosis of Hepatocytes: Relevance for HIV-Infected Patients under Treatment. Cells. 2021; 10(2): 410. doi: 10.3390/cells10020410.

[30] Cunin P, Beauvillain C, Miot C, Augusto JF, Preisser L, Blanchard S, et al. Clusterin facilitates apoptotic cell clearance and prevents apoptotic cell-induced autoimmune responses. Cell death & disease. 2016; 7(5): e2215. doi: 10.1038/cddis.2016.113.

[31] Sintich SM, Steinberg J, Kozlowski JM, Lee C, Pruden S, Sayeed S, et al. Cytotoxic sensitivity to tumor necrosis factor-alpha in PC3 and LNCaP prostatic cancer cells is regulated by extracellular levels of SGP-2 (clusterin). Prostate. 1999; 39(2): 87-93. doi:10.1002/(sici)1097-0045(19990501)39:2<87::aid-pros2>3.0.co;2-z.

[32] Burkey BF, Stuart WD, Harmony JA. Hepatic apolipoprotein J is secreted as a lipoprotein. J Lipid Res. 1992; 33(10): 1517-26. Available from: https://www.ncbi.nlm.nih.gov/pubmed/1431576.

[33] Fernandez Encinas L, Lluch N, Wu AHB, Kaski JC, Badimon L, Cubedo J. A Novel ELISA for the Quantification of Serum Levels of 2 Glycosylated Variants of Apolipoprotein J: Biomarkers for Myocardial Ischemia. J Appl Lab Med. 2023; 8(5): 917-30. doi: 10.1093/jalm/jfad034.

Journal of Associated Medical Sciences 2025; 58(3): 156-162

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Impact of self-reflection training using Kawa model on quality of life and academic performance among undergraduate occupational therapy students

Punitha* and Sofia Williams

Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha Nagar, Chennai, Tamilnadu, India.

ARTICLE INFO

Article history:
Received 4 April 2025
Accepted as revised 16 June 2025
Available online 23 June 2025

Keywords:
Self-reflection, KAWA model, academic performance, quality of life .

ABSTRACT

Background: Self-appraisal and reflection are crucial for occupational therapy students (OT) and practitioners, fostering learning and professional growth. The Kawa model, typically used clinically, may aid self-reflection by helping students recognize their strengths and challenges. However, while the model has been extensively studied in clinical settings, research on its application as a self-reflection tool for OT students remains limited.

Objectives: This study aimed to evaluate the impact of self-reflection training using the Kawa model on the quality of life and academic performance of undergraduate occupational therapy students.

Materials and methods: A randomized controlled trial was conducted. A total of 183 occupational therapy students were initially screened using the Study Skills Assessment Questionnaire. Based on poor and moderate study skill levels, along with teacher reports, 77 students met the inclusion criteria. Through simple randomization using the lottery method, 50 students were selected and assigned to either the experimental group (N=25) or the control group (N=25).The quality of life and academic performance were assessed using the WHOQOL-BREF and an academic performance scale. The experimental group received self-reflection training based on the Kawa model, while the control group underwent conventional occupational therapy for six months. Data analysis was performed using SPSS version 23.0, with paired and independent t-tests applied.

Results: Post-test scores showed statistically significant improvements in the experimental group in both quality of life (p<0.001) and academic performance (p<0.01). The results of this study showed that self-reflection training using the Kawa model was effective in enhancing both quality of life and academic performance among occupational therapy students.

Conclusion: The results suggested that self-reflection training with the Kawa model was beneficial for the experimental group, it improved quality of life and better academic performance of occupational therapy students after the intervention.

Introduction

Medical studies expose students to stress, causing mental health issues and professional burnout, making academic and personal success challenging during pre-clinical phase.[1] The World Health Organization (WHO) defines quality of life as, "an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns".[2] Medical and health-profession careers, including occupational therapy programs, are challenging due to high-intensity academic

* **Corresponding contributor.**
Author's Address: *Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha Nagar, Chennai, Tamilnadu, India.*
E-mail address: punithaarul5@gmail.com
doi: 10.12982/JAMS.2025.084
E-ISSN: 2539-6056

courses, leading to stress, burnout, and poor good point average (GPA). High academic achievements impact professional competence, causing pressure and lengthy revision periods.[3]

In response to these challenges, self-reflection is crucial for professional growth and learning in health profession education. During clinical rotations, occupational therapy students benefit from techniques like video recording, journaling, group discussions, and feedback which aid their personal and professional development. Research showed that both internal and interactive self-reflection help medical and allied health students transition into practice.[4] Reflective practice in occupational therapy is vital for learning from experience and skill development, supporting client-centered and goal-directed care. Self-reflection, a key aspect of emotional intelligence, is fundamental to professionalism and underpins adult education and lifelong learning.[5]

To support reflective practice, innovative tools like the KAWA model have gained prominence. Metaphors enhance self-reflection by adding meaning in which the Kawa Model, using a river as a life journey metaphor, serves as one such framework.[4] The Kawa model, developed by the Japanese Canadian occupational therapist Dr. Michael Iwama, uses the metaphor of a river (Kawa in Japanese) to represent life's journey. It illustrates how environmental factors influence the flow and balance of life, from birth to death or the achievement of a goal. Along the way, interactions with various elements shape experiences, challenges, and progress.[6]

The Kawa model illustrates how life's flow is influenced by surroundings and interactions. It consists of five components: water represents life's flow, riverbanks symbolize environmental factors, rocks signify obstacles, driftwood represents personal abilities, and space denotes opportunities for growth. As individuals encounter challenges, their flow may be disrupted, leading to stress and hindrances in achieving life goals. The model promotes self-awareness and integration of one's environment, skills, and aspirations. Development plans represented by space in the model helps to overcome obstacles, enhance well-being, and restore self-efficacy.[6]

The Kawa model has demonstrated practical value in occupational therapy education by enhancing students' cultural awareness, professional development, and self-reflection. It aids in identifying learning barriers, personal attributes, and contextual influences, and supports adaptive learning during clinical placements, particularly in remote settings. However, most existing studies use qualitative designs focused on self-reported reflections, often during fieldwork. These studies commonly lack objective measures, control groups, and adequate sample sizes, and do not assess outcomes like academic performance or quality of life—highlighting a gap in understanding the model's broader educational impact.[4,7] This highlights a critical gap in understanding the model's broader educational impact.

This gap underscores the need for empirical evidence to validate the model's effectiveness in educational contexts. This study aims to provide empirical evidence supporting the use of kawa model as self-reflection training for occupational therapy students, promoting well-being, academic success, and long-term professional sustainability. Therefore, this study aims to examine the impact of self-reflection training using the Kawa model on the quality of life and academic performance among occupational therapy students. It is hypothesized that the intervention group will demonstrate significantly greater improvements compared to the control group.

Materials and methods

Study design

The study employed a randomized controlled trial to evaluate the effect of Kawa model-based self-reflection training on occupational therapy students' quality of life and academic performance. Participants were randomly assigned to experimental or control groups using the lottery method. A single blinding method was used, in which the students were unaware of their group allocation.

Participants

This study received ethical approval from the Institutional scientific review board of Saveetha college of occupational therapy, with REF.no of SCOT/ISRB/085/2024 on 02.08.2024. Samples were collected from Saveetha college of occupational therapy college in Chennai. A total of 183 undergraduate occupational therapy students, aged 18 to 24 years, were initially screened using the Study Skills Assessment Questionnaire. Based on the inclusion criteria, students with poor and moderate study skills levels along with teacher reports, 77 students were identified as eligible. Using simple randomization through the lottery method, 50 students were selected and randomly assigned to either the experimental group (N=25) or the control group (N=25). Single blinding was implemented, wherein the students were unaware of their group allocation. Students with extensive prior experience with the Kawa model or participation in similar therapeutic interventions were excluded. All participants provided written informed consent prior to the study.

Instruments

WHOQOL-BREF

The WHOQOL-BREF was used to assess the quality of life among participants in this study. It is a 26-item self-report instrument divided into four domains: physical health, psychological health, social relationships, and environment. Each item is rated on a 5-point Likert scale, with higher scores indicating better quality of life. The WHOQOL-BREF demonstrates strong internal consistency, with a Cronbach's alpha coefficient of 0.896 for the entire scale. Test-retest reliability was also satisfactory for all domains, with correlation coefficients significant at the $p<0.01$ level, indicating good stability over time. However, the "Social Relationships" domain showed lower internal consistency ($\alpha=0.533$), which should be taken into consideration when interpreting results, as it may affect the reliability of findings within that domain.

Principal component analysis identified four components accounting for 49.5% of the variance.[8] For interpretive clarity, quality of life scores were categorized as follows: 0-20 = poor, 21-40 = moderate, 41-60 = good, and 61-80 = very good.[9]

Academic performance scale (APS)

The Academic Performance Scale (APS), developed by Christopher McGregory and colleagues at Saginaw Valley State University, is a self-report questionnaire assessing academic behaviors and actions. It consists of eight items rated on a 5-point scale, with higher scores indicating better academic performance. The scale has strong reliability, with an internal consistency of 0.89 and a test-retest reliability of 0.85. It also demonstrates satisfactory concurrent validity. Score interpretation is as follows: 33-40 (excellent), 25-32 (good), 17-24 (moderate), 9-16 (poor), and 0-8 (failing).[10] A recent study has successfully employed the APS to explore academic outcomes in college students, reinforcing its applicability.[10] In our study, the APS was selected due to its ability to capture self-reported academic behaviors aligned with the constructs targeted by the intervention.

Procedure

The experimental group (25 participants) received self-reflection training with the kawa model, while the control group (25 participants) underwent conventional occupational therapy. In self-reflection training using the Kawa model, the therapist introduces the model's key components-water, rocks, riverbanks, and riverbed. The participant then creates a river drawing that represents their life flow, allowing for free expression without time constraints. The participant also constructs a river cross-section, which serves as both a diagnostic and therapeutic tool for intervention planning and evaluation. Once the drawing is completed, the therapist and participant discuss its meaning and analyze the elements from the participant's perspective. Topics are then prioritized for intervention, focal points are identified, and interventions are implemented in a client-centered manner, tailored to individual needs and goals. The final stage involves assessing the effectiveness of the intervention. The control group received conventional occupational therapy, incorporating time management, relaxation techniques and academic performance guidance. The study duration was 6 months, one day per week, 24 sessions and 45 minutes per session. Following that, the post test was conducted with WHOQOL-BREF and Academic performance scale (Figure 1).

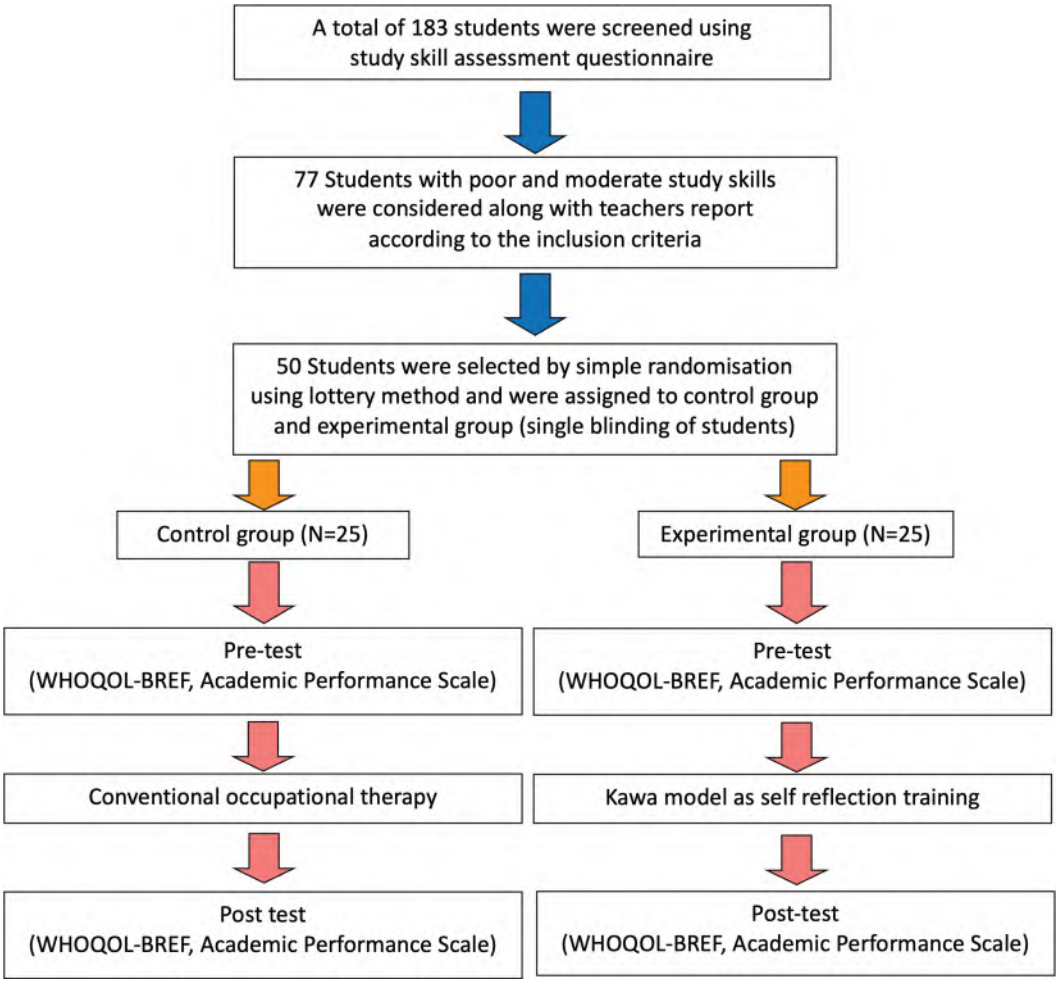


Figure 1. *Consort diagram of the study.*

Data analysis

Descriptive statistics (means, standard deviations, frequencies, and percentages) were used to summarize demographic and baseline data. Normality of the data was assessed using the Kolmogorov-Smirnov test. As the data were normally distributed, inferential statistics appropriate for randomised control trial design were applied. Participants were randomly allocated to experimental and control groups using the lottery method. Paired T-tests were used to compare pre- and post-intervention scores within groups, and independent t-tests were used to assess differences between groups. A 95% confidence interval was used in this study to estimate the precision of the observed effects. All statistical analyses were conducted using a significance level of $p<0.05$; results with $p<0.05$ were considered statistically significant, and those with $p<0.01$ were considered highly significant. The statistical analysis was done with the help of IBMSPSS version 23.0.

Results

Table 1 shows the baseline comparison between the control and experimental groups on demographic variables and outcome measures. All participants were undergraduate occupational therapy students. The mean age and gender distribution were comparable between the groups, with no statistically significant differences ($p>0.05$). Similarly, there were no significant differences in the baseline WHOQOL-BREF scores ($p=0.46$) and Academic performance scale scores ($p=0.26$), indicating that both groups were homogeneous at the start of the study.

Table 2 displays the statistical difference between the pre-test and post-test WHOQOL-BREF scores within each group. In the control group, the mean score increased slightly from 46.87 to 47.42 with a statistically significant p-value of 0.027, but the effect size (Cohen's d=0.06) was small. In contrast, the experimental group showed a substantial improvement in WHOQOL-BREF scores from 48.67 to 53.13, with a highly significant ($p<0.001$) and a moderate effect size (Cohen's d=0.50). This suggests that the intervention in the experimental group had a stronger impact on quality of life.

Table 3 presents the pre-test and post-test comparison of Academic performance scale scores within groups. The control group showed a small but statistically significant improvement from 18.24 to 19.08 ($p=0.032$, Cohen's d=0.32). The experimental group showed a larger increase in Academic performance scale scores from 18.92 to 21.68, with a highly significant ($p<0.001$) and a large effect size (Cohen's d=1.24), indicating a strong impact of the intervention on academic performance.

Table 4 compares the post-test scores between the control and experimental groups. The experimental group had significantly higher WHOQOL-BREF scores (53.13 vs 47.42; $p=0.025$, Cohen's d=0.65) and Academic performance scale scores (21.68 vs 19.08; $p=0.001$, Cohen's d=0.95) than the control group. These findings indicate that the intervention provided to the experimental group resulted in significantly better outcomes in both quality of life and academic performance.

Table 1. Baseline comparison of groups on demographics and outcome measures.

Variable	Control group (N=25)	Experimental group (N=25)	t/χ^2	*p* value
Age (years), Mean±SD	19.16±0.95	19.40±1.16	0.80	0.43
Gender	Female: 12 (48%) Male: 13 (52%)	Female: 13 (52%) Male: 12 (48%)	0.08	0.78
WHOQOL-BREF (pre-test)	46.87±8.62	48.67±8.91	0.74	0.46
Academic performance scale score (pre-test)	18.24±2.03	18.92±2.18	1.13	0.26

Table 2. Pre-post comparison of WHOQOL-BREF scores (Paired T-test).

Group	Test	Mean±SD	Mean Diff	t (df)	*p* value	Cohen's d
Control	Pre-test	46.87±8.62	0.55	-2.35 (24)	0.027*	0.06
	Post-test	47.42±8.69				
Experimental	Pre-test	48.67±8.91	4.46	-10.08 (24)	<0.001**	0.50
	Post-test	53.13±8.78				

Note: **significant at 5% level.*

Table 3. Pre-post comparison of academic performance scale scores (Paired T-test).

Group	Test	Mean±SD	Mean Diff	t (df)	*p* value	Cohen's d
Control	Pre-test	18.24±2.03	0.84	-2.28 (24)	0.032*	0.32
	Post-test	19.08±3.01				
Experimental	Pre-test	18.92±2.18	2.76	-20.80 (24)	<0.001**	1.24
	Post-test	21.68±2.25				

Note: **significant at 5% level.*

Table 4. Post-Test Comparison between groups (Independent T-test).

Outcome	Group	Mean±SD	Mean Diff	t (df)	*p* value	Cohen's d
WHOQOL-BREF	Control	47.42±8.69				
	Experimental	53.13±8.78	5.71	-2.31 (48)	0.025*	0.65
Academic performance scale score	Control	19.08±3.01				
	Experimental	21.68±2.25	2.60	-3.46 (48)	0.001**	0.95

Note: **significant at 5% level.*

Discussion

This study found that self-reflection training using the Kawa Model significantly improved both quality of life and academic performance among occupational therapy students, with greater gains than conventional interventions. These findings highlight the model's role in enhancing self-understanding, which supports better learning and overall well-being.

Table 2 presents the statistical analysis of pre-test and post-test quality of life (QoL) scores in both the control and experimental groups. In the control group, the mean QoL score increased slightly from 46.87 to 47.42 (T=-2.350, *p*=0.027*), indicating a statistically significant improvement following conventional occupational therapy. These findings are supported by the American Occupational Therapy Association's Vision 2025, which emphasizes the role of occupational therapy in enhancing health, well-being, and overall quality of life (AOTA, 2017).[11] Similarly, Espiritu reported that an occupation-based intervention positively impacted well-being among OT students in a mixed-methods study, further highlighting the potential of occupational therapy to support student mental health and quality of life.[12]

In contrast, the experimental group, which received self-reflection training using the Kawa Model alongside conventional OT, demonstrated a more pronounced improvement in QoL scores-from 48.67 to 53.13 (T=-10.075, *p*=0.000)-indicating a highly significant change. This aligns with previous findings on the model's adaptability and effectiveness across diverse populations. For instance, Galof reported improved QoL in individuals with visual impairments due to enhanced therapist-client interactions and greater independence,[13] while Hsiao HY found it effective in supporting recovery among individuals with ketamine use by fostering motivation and readiness for change.[14] Although these studies involved different populations, they emphasize the Kawa model's capacity to foster meaningful reflection and personal growth. In the present study, such reflective practices may have enabled students to better understand their life experiences, identify barriers and supports, and set realistic goals-ultimately contributing to enhanced confidence, resilience, and overall quality of life. However, the low reliability of the social relationship domain in the WHOQOL-BREF may limit the accuracy of related findings and be interpreted with caution.

Table 3 presents the statistical analysis of pre-test and post-test academic performance in both the control and experimental groups. In the control group, mean scores improved significantly from 18.24 to 19.08 (T=-2.281, *p*=0.032*), demonstrating the effectiveness of conventional occupational therapy (OT) interventions in enhancing academic performance. These results align with findings by Zolotnitsky *et al.*, who reported that OT-led coaching significantly improves self-regulation and structured learning among college students with disabilities,[15] and Nagata *et al.*, who highlighted the value of evidence-based OT in promoting academic engagement and well-being in students with learning differences.[16]

These complementary findings support the notion that OT interventions, whether conventional or coaching-based, contribute positively to academic outcomes, reinforcing the improvements observed in our control group. In the experimental group, which received self-reflection training using the Kawa Model, the mean score increased markedly from 18.92 to 21.68 (T=-20.804, *p*=0.000**), indicating a highly significant improvement. This finding supports Naidoo, who emphasized that the Kawa model fosters self-awareness and reflection, helping students better understand themselves in both personal and professional contexts.[7] In our study, the metaphor-based, structured Kawa model enabled deeper reflection, clearer goal setting, and more effective problem-solving, likely contributing to academic gains. These findings suggest that while conventional OT is beneficial, the Kawa model-based intervention has a greater impact on academic performance.

Tables 4 present the statistical analysis of post-test QoL and academic performance in both the experimental and control groups. A statistically significant difference was observed for both QoL (T=-2.312, *p*=0.025*) and academic performance (T=-3.458, p=0.001**), with the experimental group performing better. This may be due to the Kawa model's client-centered design, which focuses on self-reflection, goal setting, and understanding challenges using the river metaphor. Unlike standard OT, the Kawa model helps students visualize their obstacles (rocks), strengths (driftwood), and environment (riverbanks), helping them become more self-aware and better at coping. These strategies may have improved their mental health, self-confidence, and ability to bounce back from challenges, leading to better QoL and academic outcomes. This is supported by prior studies, including those by Blakeley *et al.* and Foo *et al*.[4,6] This study highlights the Kawa model as a culturally meaningful, structured approach that supported both mental and academic improvement in OT students. The findings suggest its potential to enhance personal insight, motivation, and resilience in OT education. Future research could explore its long-term impact, adaptability across settings, and use with other student groups. Clinically, it may complement standard OT by addressing emotional well-being and learning needs.

Conclusion

This study demonstrated that integrating the kawa model into occupational therapy education enhanced students' personal and academic development. The positive outcomes in quality of life, academic performance in the experimental group highlight the model's effectiveness, reinforcing its educational value. These results also suggest the model's potential for broader application beyond occupational therapy, encouraging further exploration in diverse educational settings and among healthcare professionals to enhance professional competence and support student success. Given the study's limited sample size and age range, future research should evaluate its effectiveness in larger, more diverse student groups and across various educational contexts to inform broader curriculum development and evidence-based training practices.

Conflict of interest

The authors declare no conflict of interest.

Ethical approval

This study has been approved by the Institution Scientific Review Board (ISRB) of Saveetha College of Occupational Therapy with Reference No. of SCOT/ISRB/085/2024.

Funding

No funding was received for this study

Credit authorship contribution statement

The main author designed the study, analyzed data, and prepared the manuscript. The second author conducted the study under the main author's supervision. All authors approved the final manuscript.

Acknowledgements

I sincerely thank Saveetha Institute of Medical and Technical Sciences, Saveetha college of Occupational Therapy, Dr. M. Arun Kumar, MOT, PhD, for his continuous encouragement and Dr. Benielraja, for their exemplary guidance, support throughout this research.

References

[1] Mishra B, Pathak R, Dwivedi A, Mishra HB, Gupta SS. Study of correlation between academic performance and quality of life among occupational therapy students. Int J Adv Res Ideas Innov Technol. 2018; 4(6): 497-501. Available from: https://www.ijariit.com/manuscripts/v4i6/V4I6-1322.pdf.

[2] Shareef MA, AlAmodi AA, Al-Khateeb AA, et al. The interplay between academic performance and quality of life among preclinical students. BMC Med Educ. 2015; 15(1): 193. doi: 10.1186/s12909-015-0476-1.

[3] Chattu VK, Sahu PK, Seedial N, et al. An exploratory study of quality of life and its relationship with academic performance among students in medical and other health professions. Med Sci (Basel). 2020; 8(2): 23. doi: 10.3390/medsci8020023.

[4] Blakely T, Potvin MC, Iwama M. The Kawa model's value for level II occupational therapy fieldwork students. J Occup Ther Educ. 2021; 5(4): 14. doi: 10.26681/jote.2021.050414.

[5] Iliff S, Tool G, Bowyer P, Parham L, Fletcher T, Freysteinson W. Self-reflection and its relationship to occupational competence and clinical performance in level II fieldwork. Internet J Allied Health Sci Pract. 2021; 19(3): Article 8. doi: 10.46743/1540-580x/2021.1988

[6] Foo KS, Kueh YC, Leong KJ, et al. Kawa model on mental health, sports and physical performance: A mini review. Asian J Sport Exerc Psychol. 2023; 3(2): 89-97. doi: 10.1016/j.ajsep.2023.08.002.

[7] Naidoo O, Christopher C, Lingah T, Moran M. The Kawa Model: A self-reflection tool for occupational therapy student development in practice placements in Australia. Occup Ther Int. 2023; 2023: 2768898. doi: 10.1155/2023/2768898.

[8] Ilić I, Šipetić S, Grujičić J, Mačužić IŽ, Kocić S, Ilić M. Psychometric properties of the World Health Organization's quality of life (WHOQOL-BREF) questionnaire in medical students. Medicina (Kaunas). 2019; 55(12): 772. doi: 10.3390/medicina55120772.

[9] Hadning I, Ainii NQ. An analysis of health workers' quality of life in Indonesia during COVID-19 pandemic. In: Proceedings of the 4th International Conference on Sustainable Innovation 2020–Health Science and Nursing (ICoSIHSN 2020). Atlantis Press; 2021. doi: 10.2991/ahsr.k.210115.085.

[10] Keerthana DMM. Exploring the relationship between hope, academic performance among college students.

Int J Multidiscip Res. 2024; 6(4): 1-7. doi: 10.36948/ijfmr.2024.v06i04.24592

[11] Pizzi MA, Richards LG. Promoting health, well-being, and quality of life in occupational therapy: A commitment to a paradigm shift for the next 100 years. Am J Occup Ther. 2017; 71(4): 7104170010p1-5. doi: 10.5014/ajot.2017.028456.

[12] Espiritu EW. Assessing the effectiveness of an occupation-based intervention in promoting OT student well-being. Am J Occup Ther. 2023; 77 (Suppl 2): 7711510288p1. doi: 10.5014/ajot.2023.77S2-PO288.

[13] Galof K. Improve quality of life client with visual impairment. Fam Med Prim Care Open Access. 2022; 6(2): 188. doi: 10.29011/2688-7460.100088.

[14] Hsiao HY. The effects of Kawa model therapeutic group for ketamine users [master's thesis]. Tainan (TW): National Cheng Kung University; 2016. Available from: https://ndltd.ncl.edu.tw/cgi-bin/gs32/gsweb.cgi?o=dnclcdr&s=id=%22104NCKU5738007%22.&searchmode=basic.

[15] Zolotnitsky LK, Beach C, Potvin MC, Ryan MA. Outcomes of OT-led coaching to support the success of students with disabilities in higher education. Am J Occup Ther. 2022; 76(Suppl 1): 7610510177p1. doi: 10.5014/ajot.2022.76s1-po177.

[16] Nagata R, Forry S, Lannigan EG. Occupational therapy interventions for college students with learning differences. Am J Occup Ther. 2022; 76(6): 7606390010. doi: 10.5014/ajot.2022.050057.

Journal of Associated Medical Sciences 2025; 58(3): 163-172

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Effects of distal robot-assisted therapy combining with task-oriented training on paretic upper extremity functions in clients with sub-acute stroke: A preliminary study

Pakpoom Jitrungruangchai, Sopida Apichai, Jananya P Dhippayom*

Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiangmai Province, Thailand.

ARTICLE INFO

Article history:
Received 13 March 2025
Accepted as revised 16 June 2025
Available online 26 June 2025

Keywords:
Stroke, robot-assisted therapy, task-oriented training, upper extremity rehabilitation, occupational therapy

ABSTRACT

Background: Robot-assisted therapy (RT) is revolutionizing stroke rehabilitation, offering new hope for improving upper extremity function. When integrated with task-oriented training (TOT), RT has the potential to enhance recovery by promoting more meaningful, functional outcomes for stroke survivors. Despite its promise, evidence on the effectiveness of distal RT combined with TOT remains limited, leaving clinicians and researchers searching for clearer evidence on the best rehabilitation strategies.

Objective: This study aimed to investigate the feasibility of distal RT combined with TOT on the function of the paretic upper extremity in sub-acute stroke patients.

Materials and methods: This preliminary randomized controlled, assessor-blinded study recruited 15 sub-acute stroke patients. Participants were randomly assigned to either the experimental group (EG; enrolled, N=8; completed, N=7) or the control group (CG; enrolled, N=7; completed, N=7). Both groups received conventional occupational therapy. The EG additionally underwent one hour of distal robot-assisted task-oriented training twice a week for six weeks. Paretic upper extremity function was assessed at baseline and post-rehabilitation using the Fugl-Meyer Assessment for Upper Extremity (FMA-UE) and the Functional Test for Hemiplegic Upper Extremity Thai Version (FTHUE-Thai version).

Results: Fourteen patients completed the program. No significant differences in post-intervention mean scores were found between groups. However, change scores were significantly higher in the EG for distal FMA-UE (p=0.002), total FMA-UE (p=0.004), and FTHUE-Thai Version (p=0.026). Additionally, effect size analysis also indicated high effects in the EG for proximal FMA-UE (d=0.94), distal FMA-UE (d=2.28), total FMA-UE (d=1.66), and FTHUE-Thai version scores (d=1.89), suggesting clinically meaningful improvements.

Conclusion: Combining distal RT with TOT has the potential to enhance upper extremity function in sub-acute stroke patients. Although further research is needed to confirm these findings, the results are promising for developing more effective rehabilitation strategies.

Introduction

Stroke is the second leading cause of death and a major cause of disability worldwide.[1] In Thailand, approximately 200,000 new stroke cases are diagnosed annually, imposing significant burdens on families, society, and healthcare systems.[2] Around 70% of stroke survivors experience upper limb function impairment, leading to difficulties in performing activities of daily living (ADLs) and a reduced quality of life (QoL).[3-6] Previous research emphasizes the importance of intensive upper limb rehabilitation within

* **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiangmai Province, Thailand.*
E-mail address: jananya.p@cmu.ac.th
doi: 10.12982/JAMS.2025.085
E-ISSN: 2539-6056

the first six months post-stroke particularly during the sub-acute phase (3-6 months), when brain recovery remains robust, and patients are in a more stable medical condition for intensive training.[7-10] Therefore, developing effective rehabilitation approaches for sub-acute stroke survivors is crucial.[11] Occupational therapy incorporates a variety of evidence-based interventions to address motor dysfunction, focusing on enhancing movement quality, task performance, and overall participation in daily activities.[12,13] To achieve these goals, advancements in the field have introduced technology-driven tools, such as robotic devices, alongside task-specific training to improve motor deficits more efficiently and effectively.

Robot-assisted therapy (RT) is widely used in stroke rehabilitation, as it delivers highly repetitive, consistent movements with greater training intensity, enhancing patient engagement and functional recovery.[14] Emerging studies suggest that RT may support somatosensory and motor cortex reorganization, making it a valuable tool for comprehensive stroke rehabilitation.[15-20] A systematic review and meta-analysis by Bertani et al. highlighted the potential benefits of RT for upper extremity rehabilitation in chronic stroke patients.[15] However, its effectiveness for sub-acute stroke patients remains inconclusive.[13] To date, most studies have predominantly employed proximal robotic devices that target the shoulder and elbow joints, while distal robotic rehabilitation, which targets wrist and finger function, has been relatively underexplored despite its importance for functional hand use.[13,19] Distal robotic devices offer a promising solution by facilitating isolated and coordinated hand movements including grasp-and-release activities while integrating multisensory feedback (visual, auditory, and proprioceptive) to enhance engagement and cortical activation.[16,21,22] The hand contains numerous motor units and sensory receptors responsible for fine and precise movements. When, the nervous system controlling hand function is damaged, prolonged and repeated stimulation is to achieve functional recovery.[23] Therefore, the benefits of distal robot-assisted therapy, such as the Gloreha Sinfonia, need to be further validated for distal upper extremity recovery.

Task-oriented training (TOT) aims to improve performance and function through goal-directed practice and repetition.[13,24,25] It requires an intensity exceeding 70% of total training time and is grounded in motor control, motor learning, and motor behavior principles. Previous study indicates that incorporating daily life activities and using real-life objects during training can boost motivation and improve movement efficiency in stroke patients.[26] Moreover, TOT promotes motor learning by enabling patients to transfer learned movements to their daily routines.[27,28] It also reduces long-term reliance on compensatory strategies for performing everyday tasks.[28,29] Therefore, combining distal RT with TOT could offer a more effective approach to improving function in stroke patients by engaging them in task-specific activities supported by robotics.

This study investigates the potential effects of integrating the distal RT with TOT to enhance upper limb function in sub-acute stroke patients. It hypothesizes that distal RT combining with TOT has the potential to enhance recovery of upper limb function. To the best of our knowledge, this is the first study to examine the integration of RT specifically designed for finger movements with TOT incorporating real-life objects and activities. This combined approach holds the potential to provide a more comprehensive solution for functional recovery after stroke.

Materials and methods

Design

This preliminary randomized controlled, assessor-blind study was reviewed and approved by the Thai Clinical Trials Registry Committee on July 2022 (TCTR identification number: TCTR20220719006). Sub-acute stroke patients were recruited from two hospitals and one health center, Chiangmai, Thailand, between August 2022 and December 2024. The study was conducted with ethics committee approval, and all participants signed informed consent forms prior to participation.

Participants

Between August 2022 and December 2024, 27 patients were assessed for eligibility. However, due to several recruitment challenges, including a limited number of eligible patients, participation inconvenience, travel constraints, and time limitations, 15 participants were initially enrolled in the study. Patients were eligible if they were over 18 years old, had experienced a first-time cerebrovascular disease resulting in hemiplegia for 3-6 months, were in Brunnstrom stages II-V of motor recovery for the arm and hand, and had a Modified Ashworth Scale score of less than 3. Exclusion criteria included cognitive impairment (MSET-10 Thai version score: ≤17 for primary education or ≤22 for higher education), peripheral nerve injury or musculoskeletal disorders (e.g., arthritis, muscle or bone injury), severe unilateral neglect, proprioceptive impairment, communication or vision difficulties, and prior treatments for spasticity such as botulinum toxin type A or acupuncture.

Participants were randomly assigned to either the experimental group (EG) or the control group (CG) using simple random sampling via the Research Randomizer tool. However, one participant from the EG withdrew due to an inability to complete at least 80% of the training sessions, resulting in a final sample of 14 participants. A flowchart outlining the recruitment and retention process was presented in Figure 1.

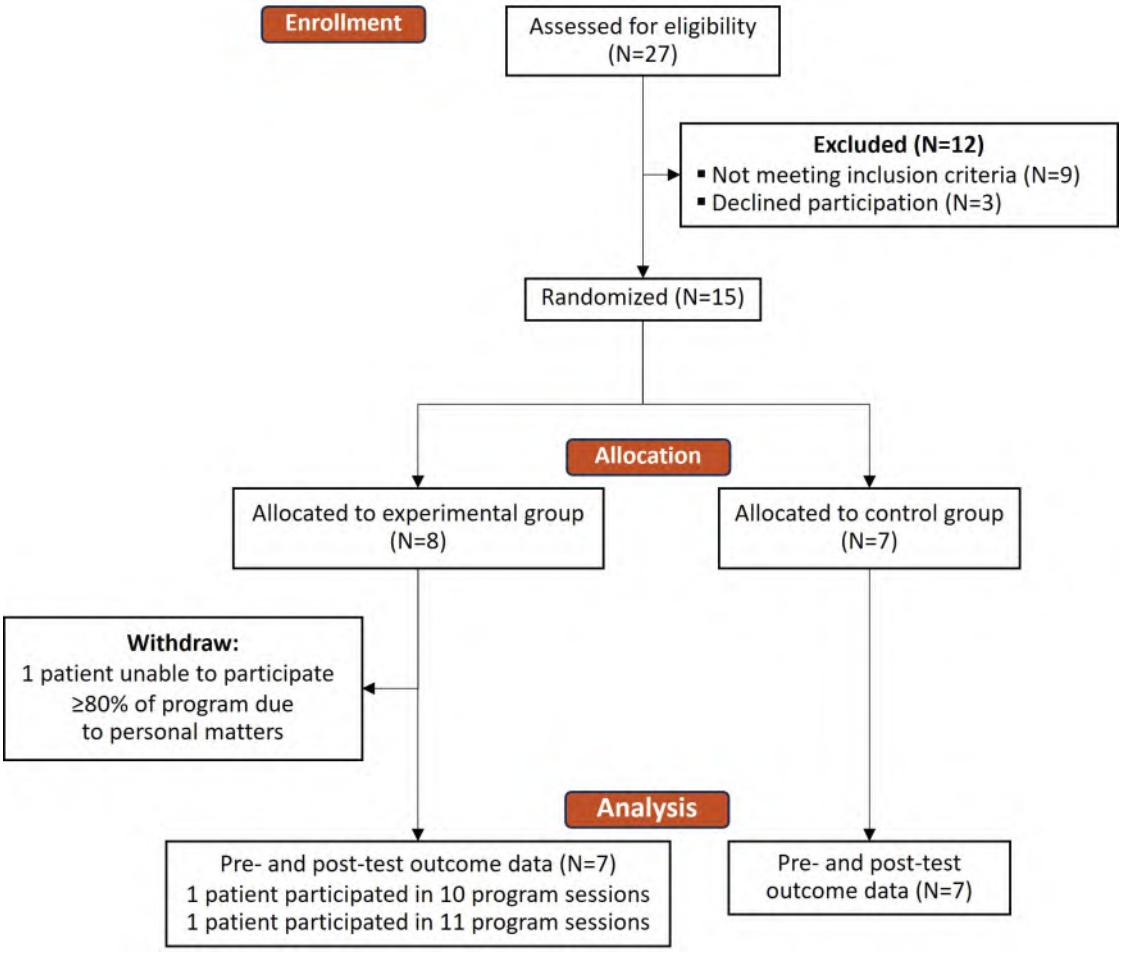


Figure 1. *Flow chart of the study*

Instruments

The Gloreha Sinfonia was employed as the primary intervention tool in this study. This distal robotic device is specifically designed to target the distal upper extremity, providing five degrees of freedom. Its dynamic support system stabilizes the proximal upper extremity against gravity during training. The device features a glove that provides partial or complete support for finger flexion and extension across six movement patterns. Additionally, it includes four training modes passive mobilization, bilateral training, active-assisted mobilization, and interactive games allowing therapists to customize rehabilitation programs to meet individual patient needs.[21,30]

This study utilized the Gloreha Sinfonia in its active-assisted mobilization mode, focusing on two key functions, including free grasping and tri-digital free pinch. The choice of function was based on the specific training activity and the participant's motor abilities. In this mode, the therapist could adjust glove sensitivity, finger flexion/extension support levels, and range of motion to match each participant's needs. The sensor glove detected individual finger movements and provided movement assistance based on the user's residual motor skills, ensuring a personalized and adaptive rehabilitation experience.

Interventions

All participants received conventional occupational therapy, which included 1-hour/sessions, at least 1 session/week for six weeks. The conventional program included spasticity inhibition, bilateral hand activities, grasp-and-release tasks, and pinch activities, customized by the therapist for each patient.

In addition to conventional therapy, the experimental group (EG) received robot-assisted task-oriented training using the Gloreha Sinfonia device for 1-hour sessions, twice a week, over six weeks. Each session consisted of two phases:

1. *Preparatory phase* (15 minutes): This phase included weight-bearing or prolonged stretching exercises aimed to normalized muscle tone.

2. *Robot-assisted task-oriented training phase* (45 minutes): During the first session, patients selected three out of five training activities, including drinking water from a bottle, folding fabric, using a fork and spoon for eating, using scissors, and opening/closing a book). The details of each activity were outlined in Table 1. These activities were chosen from the top 100 ranked activities, organized from the most to the least beneficial for stroke patients.[31] Only activities related to daily living that were analyzed and deemed suitable for

use with the Gloreha device were selected. The selection process began with the top-ranked activity and continued sequentially until a total of five activities were chosen. Each training session lasted 45 minutes, with 15 minutes per activity using the Gloreha device in active-assisted mode. Task difficulty and robotic assistance were adjusted based on individual progress. If participants mastered any task from the initial selection, they had the option to select new activities to maintain challenge and engagement. Training activities were performed in a randomized order to enhance motor learning and adaptability.

Table 1. Training activities.

Activity	Description	Chosen frequency (times)
Drinking water from a bottle	Patients reach out with the affected hand to grasp the bottle, bring it to their mouth as if drinking, then return it to the table and release it.	7
Folding fabric	Patients fold the fabric using both hands.	2
Using a fork and spoon for eating	Patients use both hands to hold a spoon and fork according to their preference, scoop beads from a plate, bring them to their mouth as if eating, then return them to the plate.	7
Using scissors	Patients use their affected hand to hold the scissors and their unaffected hand to hold the paper, then cut along specified lines in various shapes, such as straight lines, triangles, squares, and circles.	1
Opening/closing a book	Patients use their affected hand to open the book.	4

Note: Details and materials of each activity were adjusted according to the patient's abilities.

Outcome measures

Motor function was assessed at baseline and post-intervention using the following tools:

1. The Fugl-Meyer Assessment-Upper Extremity (FMA-UE) is a widely used tool for assessing upper extremity motor function. It includes 33 items divided into four categories: upper extremity (36 points), wrist (10 points), hand (14 points), and coordination (6 points). The total possible score is 66 points. The FMA-UE has demonstrated excellent inter-rater reliability (ICC=0.99), test-retest reliability (ICC=0.96), and strong construct validity (r=0.92).[32]

2. The Functional Test for Hemiplegic Upper Extremity-Thai Version (FTHUE-Thai Version) is a performance-based test adapted from the Hong Kong version to align with Thai cultural contexts.[33] It measures functional limitations due to upper extremity impairments. The FTHUE-Thai version comprises 14 tasks across seven functional levels, arranged in order of increasing difficulty. The tasks are related to activities of daily living, such as holding a bag, using a spoon to eat, and drinking water from a glass. Tasks are scored on a pass/fail basis: a 'pass (+)' is awarded if the task is completed within three minutes, while a 'fail (-)' is given if the task is incomplete or exceeds the time limit. Functional level scores are awarded upon the completion of all tasks within each level. Scores range from 1 to 7, with higher scores indicating greater functional ability. From level 5 onward, the time taken to complete each task is also recorded. The FTHUE-Thai version has demonstrated good content validity and strong inter-rater reliability.[34]

Statistical analysis

Descriptive statistics were used to summarize demographic and baseline characteristics. Continuous variables are presented as means ± standard deviations, while categorical variables are presented as frequencies and percentages (N, %). Between-group differences in baseline characteristics were analyzed using Fisher's exact test for categorical variables and multivariate analysis of variance (MANOVA) for continuous variables. The normality of the data was tested using the Shapiro-Wilk test, and due to the lack of normality, non-parametric statistics were used. Within-group comparisons were performed using the Wilcoxon signed-rank test. Between-group comparisons were assessed using the Mann-Whitney U test. A significance level of $p<0.05$ was used for all tests. Effect sizes were calculated using Cohen's *d* coefficient, where an effect size greater than 0.8 was considered large, between 0.2 and 0.8 moderate, and less than 0.2 small.[35,36]

Results

Baseline characteristics

The demographic and clinical characteristics of the participants at baseline were summarized in Table 2 and 3. There were no statistically significant differences between the EG and the CG across any measured baseline characteristics, confirming the comparability of the two groups at the start of the study.

Table 2. Baseline characteristics for categorical variables.

Characteristic	N (%)		
	EG (N=7)	CG (N=7)	*p*
Gender			1.000
Male	5 (71.43)	4 (57.14)	
Female	2 (28.57)	3 (42.86)	
Education level			0.266
High school	1 (14.29)	4 (57.14)	
Above high school	6 (85.71)	3 (42.86)	
Affected side			1.000
Left	4 (57.14)	3 (42.86)	
Right	3 (42.86)	4 (57.14)	
Etiology			0.592
Infraction	5 (71.43)	3 (42.86)	
Hemorrhage	2 (28.57)	4 (57.14)	

*Note: The p values were calculated according to the Fisher's Exact Test, EG: experimental group, CG: control group, *p<0.05.*

Table 3. Baseline characteristics for continuous variables.

Characteristic	Mean (SD)		
	EG (N=7)	CG (N=7)	*p*
Age (year)	53.29 (14.96)	64.29 (10.45)	0.137
Post-stroke duration (days)	127.14 (34.30)	139.71 (28.92)	0.473
Brunnstrom stage of arm	3.14 (0.69)	3.43 (0.79)	0.484
Brunnstrom stage of hand	3.57 (0.98)	3.71 (1.38)	0.827
MAS score of arms	1.21 (0.39)	1.07 (0.19)	0.403
MAS score of hands	0.71 (0.76)	1.00 (0.50)	0.421
MSET-10 score	26.43 (3.56)	25.57 (2.99)	0.634
Total duration of conventional therapy (hrs)	6.86 (2.27)	7.71 (2.93)	0.552

Note: The p values were calculated according to the multivariate analysis of variance (MANOVA), EG: experimental group, CG: control group, MAS: Modified Ashworth Scale, MSET-10: Mental State Examination T10.

Effects of robot-assisted therapy combining with task-oriented training on paretic upper extremity functions

Changes in upper extremity function from pre- to post-intervention were presented in Table 4. The results of the study found that there were significant improvements between pre- and post-intervention of proximal FMA-UE scores (p=0.018), distal FMA-UE scores (p=0.017), coordination FMA-UE scores (p=0.041), total FMA-UE scores (p = 0.018), and FTHUE-Thai Version scores (p=0.014) within the EG whereas there was only significant improvement between pre- and post-intervention of the proximal FMA-UE scores (p=0.026) and the total FMA-UE score (p=0.017) in the CG.

When compared between groups, no significant differences between group were found in mean scores of FMA-UE or FTHUE-Thai Versions after the intervention. However, significant change scores were observed in the distal FMA-UE (p=0.002), total FMA-UE (p=0.004), and FTHUE-Thai Version (p=0.026), indicating greater improvement in the EG (Table 5). These findings are further supported by Cohen's d effect sizes, which indicate a high effect for proximal FMA-UE scores (d=0.94), distal FMA-UE scores (d=2.28), total FMA-UE score (d=1.66), and FTHUE-Thai Version scores (d=1.89) (Table 5).

Table 4. Within-group comparison of proximal, distal, coordination, and total FMA-UE scores and FTHUE-Thai version scores between pre-test and post-test.

Outcome measure/group	Pre-test Mean (SD)	Post-test Mean (SD)	Within group *p value*
FMA-UE			
Proximal			
EG (N=7)	17.00 (6.03)	24.71 (8.44)	0.018*
CG (N=7)	18.00 (8.81)	21.00 (9.97)	0.026*
Distal			
EG (N=7)	8.86 (7.58)	15.43 (5.97)	0.017*
CG(N=7)	11.71 (8.75)	12.71 (8.18)	0.197
Coordination			
EG (N=7)	1.29 (1.89)	3.00 (2.16)	0.041*
CG (N=7)	1.86 (1.86)	2.71 (1.98)	0.102
Total			
EG (N=7)	27.14 (13.85)	43.14 (15.85)	0.018*
CG (N=7)	31.57 (17.83)	36.43 (18.50)	0.017*
FTHUE-Thai Version			
EG (N=7)	3.43 (0.53)	4.29 (0.49)	0.014*
CG (N=7)	3.71 (0.76)	3.86 (1.07)	0.317

*Note: The p values were calculated using Wilcoxon signed-rank test, FMA-UE: Fugl-Meyer Assessment-Upper Extremity, FTHUE-Thai Version: Functional Test for Hemiplegic Upper Extremity-Thai Version, EG: experimental group, CG: control group, *p<0.05*

Table 5. Between-group comparison of proximal, distal, coordination, and total FMA-UE scores and FTHUE-Thai version scores pre-test and post-test and the variance between CG and EG.

Outcome measure/group	Pre-test Mean (SD)	Post-test Mean (SD)	Change scores Mean (SD)	Cohen's *d*
FMA-UE				
Proximal				
EG (N=7)	17.00 (6.03)	24.71 (8.44)	7.71 (6.13)	0.94
CG (N=7)	18.00 (8.81)	21.00 (9.97)	3.00 (3.51)	
Between group *p* value	1.000	0.535	0.053	
Distal				
EG (N=7)	8.86 (7.58)	15.43 (5.97)	6.57 (2.82)	2.28
CG(N=7)	11.71 (8.75)	12.71 (8.18)	1.00 (2.00)	
Between group *p* value	0.805	0.710	0.002*	
Coordination				
EG (N=7)	1.29 (1.89)	3.00 (2.16)	1.71 (1.60)	0.55
CG (N=7)	1.86 (1.86)	2.71 (1.98)	0.86 (1.46)	
Between group *p* value	0.710	0.710	0.318	
Total				
EG (N=7)	27.14 (13.85)	43.14 (15.85)	15.86 (8.73)	1.66
CG (N=7)	31.57 (17.83)	36.43 (18.50)	4.86 (3.34)	
Between group *p* value	0.805	0.710	0.004*	

Table 5. Between-group comparison of proximal, distal, coordination, and total FMA-UE scores and FTHUE-Thai version scores pre-test and post-test and the variance between CG and EG. (Continue)

Outcome measure/group	Pre-test Mean (SD)	Post-test Mean (SD)	Change scores Mean (SD)	Cohen's *d*
FTHUE-Thai Version				
EG (N=7)	3.43 (0.53)	4.29 (0.49)	0.86 (0.38)	1.89
CG (N=7)	3.71 (0.76)	3.86 (1.07)	0.14 (0.38)	
Between group *p* value	0.535	0.209	0.026*	

*Note: The p values were calculated using Mann-Whitney U, FMA-UE: Fugl-Meyer Assessment-Upper Extremity, FTHUE-Thai Version: Functional Test for Hemiplegic Upper Extremity-Thai Version, EG: experimental group, CG: control group, Cohen's d=effect size between change scores,*p<0.05.*

Discussion

This study aimed to evaluate the effects of combining distal robot-assisted therapy (distal RT) with task-oriented training (TOT) on upper extremity function in sub-acute stroke patients. The findings demonstrated significant within-group improvements in the EG across proximal, distal, total FMA-UE scores, and FTHUE-Thai version scores after six weeks. In contrast, the CG showed significant improvement only in proximal and total FMA-UE scores. Although no statistically significant differences in the mean post-intervention score of both outcome measures were observed between groups, significant change scores were found for distal and total FMA-UE and FTHUE-Thai Version. These changes, confirmed by high effect sizes indicated clinically meaningful improvements in the EG, emphasizing the potential benefits of integrating TOT with distal RT.

Regarding the FMA-UE results, a significant improvement in total and distal FMA-UE change scores was observed in the EG compared to the CG. This may be attributed to the structured training program provided by the Gloreha Sinfonia device. The Gloreha Sinfonia is a robotic device designed to facilitate repetitive hand movements, such as grasp-and-release activities, while providing dynamic support to the proximal upper extremity against gravity. Its glove-based system offers adjustable support for finger flexion and extension, allowing therapists to customize rehabilitation programs for each patient. Interestingly, the effects of distal RT and TOT were supported not only by the moderate to high effect sizes observed through Cohen's *d* values but also by the significant differences in distal and total FMA-UE change scores. The EG exhibited an improvement of 7 points (8.29 vs 1.29) for distal FMA-UE and 11 points (15.86 vs 4.86) for total FMA-UE. These results exceed the minimal clinically important differences (MCID) for distal FMA-UE (1 to 3 points) and total FMA-UE (9 to 10 points). Although no significant difference was found in the proximal FMA-UE change score, the change in the EG was 7.71 points, which is within the MCID range for proximal FMA-UE (5 to 7 points), whereas the CG showed a change score below the MCID (3 points).[37] These findings suggest that the combined effects of distal RT and TOT may provide clinically meaningful improvements in upper extremity function for sub-acute stroke patients, as evidenced by the greater distal and total FMA-UE change scores in the EG. Given that the observed difference exceeds the MCID threshold for FMA-UE, this approach may hold significant potential for enhancing upper extremity recovery in stroke rehabilitation.

The significant improvement in FTHUE-Thai version change scores (p=0.026) further supports the effectiveness of distal RT combined with TOT in enhancing functional movement and facilitating the transfer of motor improvements to activities of daily living (ADLs). TOT emphasizes goal-directed practice and repetition, which are essential for motor learning and functional recovery.[24,38] Additionally, the active-assisted mode of distal RT allows patients to perform hand movements independently enhancing motor learning, motivation, and a sense of accomplishment. These findings align with studies such as Wu *et al*.[26] which demonstrated that real-life task integration promotes motor recovery and functional independence. Similarly, a study by Lee and Howe[39] highlighted the effectiveness of using real-life activities to improve motor function and ADLs in stroke patients.

The greater improvements in upper extremity function seen in the EG (Table 5) may be attributed to the combination of highly repetitive, intensive movements provided by distal RT and the goal-directed, task-specific nature of the TOT. Repetition and intensity are crucial for promoting neuroplasticity and motor learning, which are key components of post-stroke recovery.[7] The Gloreha Sinfonia facilitates repetitive hand and finger movements and provides dynamic support that enables initiation of proximal upper extremity motion against gravity-enhancing the functional execution of training tasks. Additionally, engaging patients in real-life activities, such as drinking water from a bottle, folding fabric, using a fork and spoon for eating, using scissors, and opening/closing a book, likely enhanced patient engagement and facilitated the application of motor gains to everyday tasks.[38] These findings align with those of Shi *et al.,*[40] which demonstrate that distal RT provides practical assistance for stroke patients, enabling them to perform training tasks more independently and intensively. This leads to improvements in upper extremity function, with effects that may persist for at least three months post-training. Although the selected activities were not fully client-centered due to technical limitations of the robotic device, including finger-only assistance, limited hand prehension types,

and a fixed glove-to-unit connection length-allowing participants to choose three out of five training activities provided a sense of autonomy and motivation, fostering sustained engagement and motor learning. Furthermore, the use of real materials in most activities, along with continuous feedback, helped reinforce functional movements and replicate real-world demands.[39,41] This approach aligns motor learning principles by creating practical, behaviorally relevant experiences essential for skill acquisition, ultimately leading to improved functional recovery in this study.

The results of this preliminary study suggest that combining distal RT with TOT is a feasible and potentially effective adjunct to conventional occupational therapy for improving upper extremity function in sub-acute stroke patients. These findings emphasize the importance of incorporating real-life tasks and distal RT into rehabilitation programs to enhance motor recovery and functional independence. However, further research with larger sample sizes is needed to confirm these results.

Limitation

The small sample size (N=14) was a limitation of this study, potentially contributing to the lack of statistically significant between-group differences. Larger studies are needed to confirm these findings. Moreover, the lack of long-term follow-up data may limit our understanding of the sustained effects of the intervention. Additionally, variations in stroke severity and patient-specific factors such as motivation and cognitive function may have influenced the outcomes.

Future research should focus on full-size randomized controlled trials to further investigate the effects of combining distal RT and TOT. Also, their long-term effects should be investigated. Additionally, exploring the impact of such interventions on different stroke populations (e.g., chronic stroke patients) could provide more comprehensive insights. Moreover, future studies should control for the total dosage of therapy across comparison groups to enable a more accurate assessment of intervention effects. Investigating the specific contribution of distal RT and TOT compared to conventional task-oriented training (TOT) alone would provide further insight into the added value of RT.

Conclusion

Although no statistically significant differences in mean scores were observed between the EG and CG, the significant score changes and medium to high effect sizes highlight the potential of combining distal RT with TOT for improving upper extremity function in sub-acute stroke patients. Further research with larger sample sizes, longer follow-up periods, is needed to confirm these finding and establish the long-term benefits of robot-assisted task-oriented training in stroke rehabilitation. The evidence can help therapists develop more effective training programs to improve upper limb function in stroke patients.

Ethical approval

This study received ethics approval from the Research Ethics Committee of Faculty of Associated Medical Sciences, Chiang Mai University (No. 267/2565), Maharaj Nakorn Chiang Mai Hospital (No. 355/2565) and Chiangmai Neurological Hospital (No. EC 021-67).

Funding

This research project was sponsored by the Faculty of Associated Medical Sciences and the Graduate School, Chiang Mai University, Chiang Mai, Thailand.

Conflict of interest

The authors declare no conflict of interests.

CRedit authorship contribution statement

Pakpoom Jitrungruangchai: conceptualization, methodology, investigation, formal analysis, writing-original draft, and visualization. **Jananya P Dhippayom:** conceptualization, methodology, writing-review and editing, and supervision. **Sopida Apichai:** validation, writing-review and editing, and supervision. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

The authors would like to give a special thanks to all stroke clients who volunteered themselves to the study as well as occupational therapists for their coordination and collaboration throughout the study.

References

[1] Feigin VL, Stark BA, Johnson CO, Roth GA, Bisignano C, Abady GG, et al. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol. 2021; 20(10): 795-820. doi: 10.1016/s1474-4422(21)00252-0.

[2] Tiamkao S, Ienghong K, Cheung LW, Celebi I, Suzuki T, Apiratwarakul K. Stroke incidence, rate of thrombolytic therapy, mortality in Thailand from 2009 to 2021. OAMJMS. 2022; 10(E): 110-5. doi: 10.3889/oamjms.2022.8051.

[3] Borschmann KN, Hayward KS. Recovery of upper limb function is greatest early after stroke but does continue to improve during the chronic phase: A two-year, observational study. Physiotherapy. 2020; 107: 216-23. doi: 10.1016/j.physio.2019.10.001.

[4] Nakayama H, Stig Jørgensen H, Otto Raaschou H, Skyhøj Olsen T. Recovery of upper extremity function in stroke patients: The Copenhagen stroke study. ACRM. 1994; 75(4): 394-8. doi: 10.1016/0003-9993(94)90161-9.

[5] Kainz A, Meisinger C, Linseisen J, Kirchberger I, Zickler P, Naumann M, et al. Changes of health-related quality of life within the 1st year after stroke–Results from a prospective sroke chort study. Front Neurol. 2021; 12. doi: 10.3389/fneur.2021.71531.

[6] Tiwari S, Joshi A, Rai N, Satpathy P. Impact of stroke on

quality of life of stroke survivors and their caregivers: A qualitative study from India. J Neurosci Rural Pract. 2021; 12(4): 680-8. doi: 10.1055/s-0041-1735323.

[7] Huang HC, Chung KC, Lai DC, Sung SF. The impact of timing and dose of rehabilitation delivery on functional recovery of stroke patients. J Chin Med Assoc. 2009; 72(5): 257-64. doi: 10.1016/s1726-4901(09)70066-8.

[8] Bath PM, Lees KR. Acute stroke. West J Med. 2000; 173(3): 209-12. doi: 10.1136/ewjm.173.3.209.

[9] Stinear C, Ackerley S, Byblow W. Rehabilitation is initiated early after stroke, but most motor rehabilitation trials are not: A systematic review. Stroke. 2013; 44(7): 2039-45. doi: 10.1161/strokeaha.113.000968.

[10] Teasell R, Bayona N, Bitensky J. Background concepts in stroke rehabilitation. EBRSR. 2008; 1. Available from: http://www.ebrsr.com/evidence-review/3-background-concepts-stroke-rehabilitation

[11] Pollock A, St George B, Fenton M, Firkins L. Top ten research priorities relating to life after stroke. Lancet Neurol. 2012; 11(3): 209. doi: 10.1016/s1474-4422(12)70029-7.

[12] Iruthayarajah J, Mirkowski M, Reg M, Iliescu A, Caughlin S, Harris J, et al. Upper extremity motor rehabilitation interventions. EBRSR. 2016. Available from: http://www.ebrsr.com/sites/default/files/Ch.%2010%20Upper%20Extremity%20Motor%20Interventions_v20.pdf.

[13] Hatem SM, Saussez G, Della Faille M, Prist V, Zhang X, Dispa D, et al. Rehabilitation of motor function after stroke: A multiple systematic review focused on techniques to stimulate upper extremity recovery. Front Hum Neurosci. 2016; 10: 442. doi: 10.3389/fnhum.2016.00442.

[14] Weber LM, Stein J. The use of robots in stroke rehabilitation: A narrative review. Neurorehab. 2018; 43(1): 99-110. doi: 10.3233/nre-172408.

[15] Bertani R, Melegari C, De Cola MC, Bramanti A, Bramanti P, Calabrò RS. Effects of robot-assisted upper limb rehabilitation in stroke patients: A systematic review with meta-analysis. Neurol Sci. 2017; 38(9): 1561-9. doi: 10.1007/s10072-017-2995-5.

[16] Lee H-C, Kuo F-L, Lin Y-N, Liou T-H, Lin J-C, Huang S-W. Effects of robot-assisted rehabilitation on hand function of people with stroke: A randomized, crossover-controlled, assessor-blinded study. AJOT. 2020; 75(1): 7501205020p1-p11. doi: 10.5014/ajot.2021.038232.

[17] Milia P, Peccini M, De Salvo F, Sfaldaroli A, Grelli C, Lucchesi G, et al. Rehabilitation with robotic glove (Gloreha) in poststroke patients. Digit Med. 2019; 5(2): 62-7. doi: 10.4103/digm.digm_3_19.

[18] Vanoglio F, Bernocchi P, Mulè C, Garofali F, Mora C, Taveggia G, et al. Feasibility and efficacy of a robotic device for hand rehabilitation in hemiplegic stroke patients: A randomized pilot controlled study. Clin Rehabil. 2017; 31(3): 351-60. doi: 10.1177/0269215516642606.

[19] Veerbeek JM, Langbroek-Amersfoort AC, van Wegen EEH, Meskers CGM, Kwakkel G. Effects of robot-assisted therapy for the upper limb after stroke: A systematic review and meta-analysis. Neurorehabil Neural Repair. 2017; 31(2): 107-21. doi: 10.1177/1545968316666957.

[20] Villafañe JH, Taveggia G, Galeri S, Bissolotti L, Mullè C, Imperio G, et al. Efficacy of short-term robot-assisted rehabilitation in patients with hand paralysis after stroke: A randomized clinical trial. Hand. 2018; 13(1): 95-102. doi: 10.1177/1558944717692096.

[21] Idrogenet. Gloreha Sinfonia Italy: Idrogenet coporation; 2021. Available from: https://www.gloreha.com/sinfonia/.

[22] Chang PH, Lee S-H, Koo K-M, Lee S-H, Jin S-H, Yeo SS, et al. The cortical activation pattern by a rehabilitation robotic hand: A functional NIRS study. Front Hum Neurosci. 2014; 8: article 49. doi: 10.3389/fnhum.2014.00049.

[23] Sobinov AR, Bensmaia SJ. The neural mechanisms of manual dexterity. Nat Rev Neurosci. 2021; 22(12): 741-57. doi: 10.1038/s41583-021-00528-7.

[24] Teasell R, Hussein N, Mirkowski M, Vanderlaan D, Saikaley M, Iruthayarajah MLJ. Hemiplegic upper extremity rehabilitation. In: Teasell R, Hussein N, Iruthayarajah J, Saikaley M, Longval M, Viana R, editors. The Stroke Rehabilitation Clinician's Handbook. London, Ontario, Canada: Heart & Stroke Foundation Canadian Partnership for Stroke Recovery; 2020: pp 1-60.

[25] Winstein CJ, Stein J, Arena R, Bates B, Cherney LR, Cramer SC, et al. Guidelines for adult stroke rehabilitation and recovery. Stroke. 2016; 47(6): e98-e169. doi: 10.1161/STR.0000000000000098.

[26] Wu C-y, Wong M-k, Lin K-c, Chen H-c. Effects of task goal and personal preference on seated reaching kinematics after stroke. Stroke. 2001; 32(1): 70-6. doi: 10.1161/01.STR.32.1.70.

[27] Pongtham P, Chingchit W, Dhippayom JP. Improving upper extremity function in chronic stroke using occupational therapy task-oriented approach. J Assoc Med Sci. 2021; 55(1): 52-9. Available from: https://he01.tci-thaijo.org/index.php/bulletinAMS/article/view/252539.

[28] Thant AA, Wanpen S, Nualnetr N, Puntumetakul R, Chatchawan U, Hla KM, et al. Effects of task-oriented training on upper extremity functional performance in patients with sub-acute stroke: A randomized controlled trial. J Phys Ther Sci. 2019; 31(1): 82-7. doi: 10.1589/jpts.31.82 .

[29] Winstein CJ, Rose DK, Tan SM, Lewthwaite R, Chui HC, Azen SP. A randomized controlled comparison of upper-extremity rehabilitation strategies in acute stroke: A pilot study of immediate and long-term outcomes. Arch Phys Med Rehabil. 2004; 85(4): 620-8. doi: 10.1016/j.apmr.2003.06.027.

[30] Borboni A, Mor M, Faglia R. Gloreha—hand robotic rehabilitation: Design, mechanical model, and experiments. J Dyn Sys Meas Control. 2016; 138(11).

doi: 10.1115/1.4033831.

[31] Thichanpiang P, Kaunnil A, Sansri V, Thongchoomsin S, Permpoonputtana K. Reliability and internal consistency of Thai activity card sort for stroke survivors in occupational therapy units. J Assoc Med Sci. 2020; 53(3). Available from: https://he01.tci-thaijo.org/index.php/bulletinAMS/article/view/241386.

[32] Platz T, Pinkowski C, van Wijck F, Kim I-H, di Bella P, Johnson G. Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: A multicentre study. Clin Rehabil. 2005; 19(4): 404-11. doi: 10.1191/0269215505cr832oa.

[33] Fong K, Ng B, Chan D, Chan E, Ma D, Au B, et al. Development of the Hong Kong version of the functional test for the hemiplegic upper extremity (FTHUE-HK). Hong Kong J Occup Ther. 2004; 14(1): 21-9. doi: 10.1016/S1569-1861(09)70025-7.

[34] Pingmuang P, Chinchai P, Dhippayom JP. Internal consistency and inter-rater reliability of the functional test for hemiplegic upper extremity in persons with hemiplegia-Thai version. ASEAN J Rehabil Med. 2016; 26(2): 36-46. doi: 10.14456/jtrm.2016.9.

[35] Cohen J. The Effect Size Index: *d*. Statistical power analysis for the behavioral sciences. 2nd Ed. New York: Lawrence Erlbaum Associates; 1988: pp 20-7

[36] Harikrishnan H, Abd halim N, Arjunan S. Exploring the digital game-based elements in mathematics education: A meta-analysis review. Univ J Educ. 2019; 7: 106-16. doi: 10.13189/ujer.2019.071613.

[37] Arya KN, Verma R, Garg RK. Estimating the minimal clinically important difference of an upper extremity recovery measure in subacute stroke patients. TSR. 2011; 18(Suppl 1): 599-610. doi: 10.1310/tsr18s01-599.

[38] Timmermans AA, Spooren AI, Kingma H, Seelen HA. Influence of task-oriented training content on skilled arm-hand performance in stroke: A systematic review. NNR. 2010; 24(9): 858-70. doi: 10.1177/1545968310368963.

[39] Lee CY, Howe TH. Effectiveness of activity-based task-oriented training on upper extremity recovery for adults with stroke: A systematic review. AJOT. 2024; 78(2): 7802180070. doi: 10.5014/ajot.2024.050391.

[40] Shi X, Yang C, Lee PC, Xie D, Ye Z, Li Z, et al. An interactive soft robotic hand-task training system with wireless task boards and daily objects on post-stroke rehabilitation. Wearable Technologies. 2025; 6: e4. doi: 10.1017/wtc.2024.10.

[41] Zhao M, Wang G, Wang A, Cheng LJ, Lau Y. Robot-assisted distal training improves upper limb dexterity and function after stroke: A systematic review and meta-regression. Neurol Sci. 2022; 43(3): 1641-57. doi: 10.1007/s10072-022-05913-3.

Journal of Associated Medical Sciences 2025; 58(3): 173-183

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

The preparation of Au decorated on ZnO nanorods by comparative DCMS/HIPIMS techniques for antibacterial activity

Sukon Kalasung[1], Saksorn Limwichean[2], Pitak Eiamchai[2], Mati Horprathum[2], Noppadon Nuntawong[2], Viyapol Patthanasettakul[2], Artitaya Yatsomboon[1*]

[1]Faculty of Medicine, Bangkokthonburi University, Thawi Watthana, Bangkok, Thailand.
[2]National Electronics and Computer Technology Center, National Science and Development Agency, Pathum Thani, Thailand.

ARTICLE INFO

Article history:
Received 7 February 2025
Accepted as revised 29 May 2025
Available online 27 June 2025

Keywords:
Au-decorated ZnO NRs, hybrid nanomaterials, novel antimicrobial agents, Escherichia coli, Staphylococcus aureus.

ABSTRACT

Background: The mortality rate of antimicrobial-resistant infections has increased dramatically worldwide due to the increased use of antibiotics. The rise of antibiotic-resistant bacteria has highlighted the need to develop novel materials with antimicrobial properties, and nanotechnology offers promising prospects for the development of new therapeutic approaches. Currently, hybrid nanomaterials are interesting alternatives that enhance the physical and antibacterial properties of nanomaterials with a large surface area, making them efficient and biocompatible.

Objectives: This study evaluated the antibacterial activity of Au nanoparticle-decorated ZnO Nanorods (NRs) with different characteristics of Au nanoparticles (Au NPs) on the ZnO surface.

Materials and methods: ZnO NRs were grown on a silicon wafer using the hydrothermal method, and Au NPs were decorated on the ZnO NRs surface by DC magnetron sputtering and high-power impulse magnetron sputtering (HiPIMS) techniques for comparison. The physical morphologies and crystallinity of the ZnO NRs and Au-nanoparticle-decorated ZnO NRs were investigated by field-emission electron microscopy (FE-SEM), transmission electron microscopy (TEM), and X-ray diffraction (XRD).

Results: FE-SEM results indicated changes in the physical morphologies of the Au NPs on the ZnO NRs. The antibacterial efficacy of the ZnO NRs and Au-decorated ZnO NRs against *Escherichia coli* and *Staphylococcus aureus* was evaluated under UV light irradiation with bacterial concentrations ranging from 10^0 to 10^8 CFU/mL to assess their inhibitory effects using the plate count technique.

Conclusion: The results demonstrated that the proposed Au-ZnO NRs exhibited a significant inhibitory effect on the growth of *Escherichia coli* indicating the potential of Au NPs decorated ZnO NRs as a novel antimicrobial material. Importantly, the results highlight the influence of bacterial concentration on the effectiveness of Au-ZnO NRs, offering insights for future applications in combating antibiotic-resistant bacteria.

Introduction

Antimicrobial resistance (AMR) is an escalating global health issue, characterized by the diminishing efficacy of antibiotics against bacterial pathogens. This crisis, driven by the overuse and misuse of antimicrobial agents, poses severe threats to healthcare systems, agriculture, and global economy. The World Health Organization (WHO) has projected that, by 2050, AMR-related infections could result in 10 million deaths annually, surpassing cancer and cardiovascular diseases as the leading causes of mortality.[1,2] The rapid emergence of resistant strains

∗ **Corresponding contributor.**
Author's Address: *Faculty of Medicine, Bangkokthonburi University, Thawi Watthana, Bangkok, Thailand.*
E-mail address: artitaya.yat@bkkthon.ac.th
doi: 10.12982/JAMS.2025.086
E-ISSN: 2539-6056

of bacteria, such as *Escherichia coli* and *Staphylococcus aureus*, underscores the urgent need for innovative strategies to manage bacterial infections. Nanotechnology offers promising solutions for AMR by leveraging the unique properties of nanoscale materials. Zinc oxide nanorods (ZnO NRs) are among the most studied nanomaterials for their intrinsic antibacterial activity, which is attributed to the generation of reactive oxygen species (ROS) and the release of zinc ions that disrupt bacterial membranes and DNA.[3] Hybrid nanostructures, such as gold nanoparticle (AuNP)-decorated ZnO NRs, have shown synergistic effects that enhance antimicrobial efficacy.[4] The addition of AuNPs amplifies ROS production through plasmonic effects and provides additional antibacterial mechanisms, such as ion release and physical damage to bacterial cell walls.[5] Recent studies have explored various synthesis techniques for Au-ZnO hybrid nanomaterials. Among these, sputtering methods, including direct current magnetron sputtering (DCMS) and high-power impulse magnetron sputtering (HIPIMS), have been widely used because of their ability to produce uniform coatings and control particle size.[6] HIPIMS, in particular, offers high-energy pulsed deposition, resulting in superior nanoparticle distribution and enhanced antibacterial properties compared to traditional methods. Despite these advancements, the comparative efficacy of Au-ZnO NRs fabricated by DCMS and HIPIMS remains underexplored, particularly in terms of their performance against clinically relevant pathogens, such as *E. coli* and *S. aureus.*[7] While significant progress has been made in developing ZnO-based nanostructures, several gaps remain.[8] The influence of fabrication techniques on the physical, chemical, and antibacterial properties of Au-ZnO NRs is not fully understood. Furthermore, there is limited knowledge of the optimal conditions for enhancing the performance of these materials under UVA light exposure, which is a critical factor in their potential applications in healthcare and environmental settings. Addressing these gaps is essential for advancing the practical use of hybrid nanomaterials to combat AMR.

This study aimed to investigate the preparation and antibacterial efficacy of Au-decorated ZnO NRs fabricated using the DCMS and HIPIMS techniques. The specific objectives were 1) to compare the physical and structural properties of Au-ZnO NRs synthesized using DCMS and HIPIMS, 2) to evaluate their antibacterial performance against *E. coli* and *S. aureus* under UVA exposure, and 3) to identify the production conditions that maximize antibacterial efficacy. The significance of this study lies in its potential to provide a scientific basis for optimizing the design and production of hybrid nanomaterials. By addressing the challenges associated with AMR, this study contributes to the development of effective antimicrobial agents that can be integrated into medical devices, water treatment systems, and other critical applications.

Materials and methods

Materials

The bacterial strains used in this study were *Escherichia coli* (ATCC 25922) and *Staphylococcus aureus* (ATCC 29213). The culture medium used for bacterial growth and experimentation consisted of Nutrient Broth (NB; Difco, USA) and Nutrient Agar (NA; Oxoid, UK). Reagents and solutions included Normal Saline Solution (0.9% NaCl; Sigma-Aldrich, USA) and the McFarland Standard 0.5 (0.5; Remel, USA) for bacterial standardization. The equipment used included a UV-A light source (220 W; Philips, Netherlands), an incubator set at 37 °C (Thermo Fisher Scientific, USA), a densitometer DEN-2 model (Biosan, Latvia) for turbidity measurement, and a plate counter for bacterial colony enumeration. Zinc nitrate hexahydrate ($Zn(NO_3)_2 \cdot 6H_2O$, 98%) was purchased from Laboratory Reagent & Fine Chemical (LOBA Chemie). HMTA ($C_6H_{12}N_4$, ≥99.0%) was purchased from Sigma-Aldrich (St. Louis, MO, USA). A Zn sputtering target (2-inch diameter and 99.99 % purity) was purchased from Kurt J. Lesker (Pennsylvania, USA). A gold (Au) sputtering target (2-inch diameter and 96.5% purity) was obtained from Siam Gold Gallery Co., Ltd. (Thailand).

Preparation of the Au nanoparticles-decorated ZnO NRs

Au-decorated ZnO nanorods were fabricated using a multistep process, as illustrated in Figure 1. First, 100 nm ZnO films on 2×4 cm^2 n-type silicon (100) wafers were used as base substrates. These films were prepared via DC magnetron sputtering using a Zn target operated at 100 W with argon and oxygen flow rates of 20 and 40 sccm, respectively. Next, the base substrates were immersed in a mixed solution of zinc nitrate hexahydrate and HMTA at a 1:1 ratio in 200 ml of deionized (DI) water at a concentration of 10 mM. For the hydrothermal synthesis, the solution-containing substrates were transferred to an autoclave set at 90 °C for 4 hrs. The hydrothermally grown ZnO NRs arrays were carefully removed, rinsed, and dried under ambient air. Finally, the prepared ZnO NR templates were decorated with Au nanoparticles using pulsed DC magnetron sputtering (DCMS) at 100 W, 5 mTorr operating pressure, and a 20 sccm argon flow rate for deposition times of 15 sec and 60 sec. High-power impulse magnetron sputtering (HiPIMS) was employed, with a pulse width of 100 µs, frequency of 700 Hz, 5 mTorr operating pressure, and 20 sccm argon flow rate for deposition times of 15 sec and 60 sec (AJA International, Inc.; ATC 2000-F). In this study, Au NPs were decorated on ZnO NR surfaces using a comparative DCMS and HiPIMS approach to investigate their antibacterial activity.

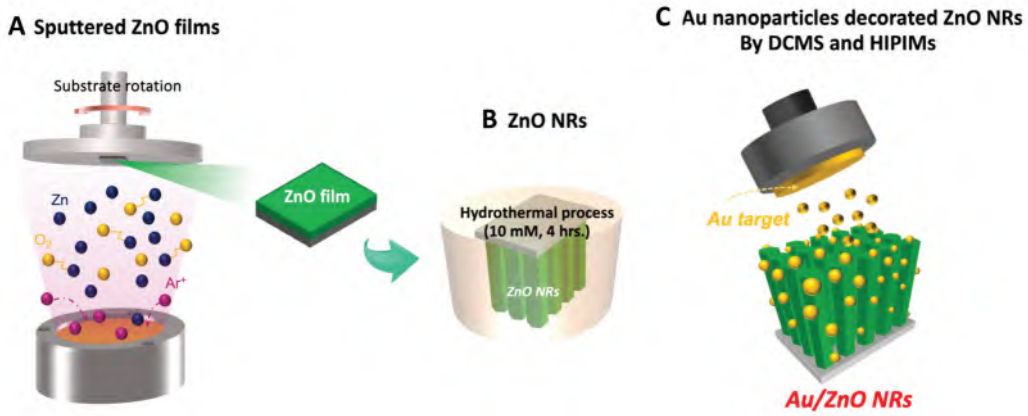


Figure 1. *Preparation procedure for design of Au-nanoparticles decorated ZnO nanorods.*

Characterization of the Au nanoparticles -decorated ZnO nanorods

The physical morphologies of the fabricated samples were analyzed using field-emission scanning electron microscopy (FE-SEM; Hitachi High Tech; SU8030) at an accelerating voltage of 10 kV. Top-view and cross-sectional FE-SEM images were captured for both the ZnO NRs and Au-nanoparticle decorated ZnO NRs, with measurements such as diameter and height taken at up to 50 positions for each sample. The crystallinity of the ZnO NRs and Au-decorated ZnO NRs was investigated using X-ray diffraction (XRD). Further analyses of the physical morphologies and elemental/compound compositions were performed on a single Au-decorated ZnO NRs sample. These nanorods were carefully scraped onto copper grids and examined using a transmission electron microscope (TEM; JEOL, JEM-2100 Plus) operated at an accelerating voltage of 200 kV. Additionally, Energy-dispersive X-ray spectroscopy (EDS) was performed to map the elemental distribution of Au, Zn, and O. Finally, high-resolution TEM (HR-TEM) imaging and selected area electron diffraction (SAED) were carried out specifically on the topmost and central regions of the selected Au-decorated ZnO nanorods.

Antibacterial Test

The antibacterial activity of the Au-decorated ZnO nanorods (Au-ZnO NRs) was evaluated against *Escherichia coli* and *Staphylococcus aureus* under UV-A exposure. Both bacterial strains were first cultured in 3mL of Nutrient Broth (NB) at 37 °C with shaking at 150 rpm for 18-24 hrs to reach the exponential growth phase. The bacterial suspensions were then standardized to the McFarland 0.5 turbidity standard, corresponding to approximately 1.5×10^8 CFU/mL, using a densitometer. Serial dilutions were subsequently performed using sterile saline or NB to achieve bacterial concentrations ranging from 10 CFU/mL to 10 CFU/mL. For antibacterial testing, sterilized Au-ZnO NR samples were placed in sterile Petri dishes (five samples per dish), and 100µL of bacterial suspension was applied to each sample. The dishes were exposed to UV-A light (220 W) for 1 hr at a fixed distance to ensure consistent irradiation. Control groups containing bacterial suspensions without nanorods were subjected to the same UV-A exposure conditions. After treatment, all samples were incubated at 37 °C for 18–24 hrs to allow the surviving bacteria to form visible colonies. The number of colony-forming units (CFUs) was counted using the plate count method, and antibacterial efficacy was determined by calculating the Killing Rate using the following formula:

$$\text{Killing rate (\%)} = [(\text{CFU\_control} - \text{CFU\_treated}) / \text{CFU\_control}] \times 100.$$

Statistical analysis

The experimental results are expressed as the mean ± standard deviation. Statistical significance was analyzed using one-way ANOVA, followed by post-hoc tests to compare multiple groups. Statistical significance was set at $p<0.05$. All experiments were performed in triplicate to ensure reproducibility and minimize variability.

Results

Figure 2A presents the top-view and cross-sectional FE-SEM images of the ZnO NRs and Au-decorated ZnO NRs, along with histograms illustrating the distribution of the NR diameters and lengths measured from a sample size of 2 mm^2. Figure 2B displays the Gaussian-fitted distributions of (i) the diameters and (ii) lengths of the ZnO NRs and Au-decorated ZnO NRs. The average diameter was approximately 120 nm, with a standard deviation (SD) of 30 nm, while the average length was approximately 2,221 nm, with an SD of 46 nm. Figure 3A clearly shows the morphology of the Au nanoparticles deposited on the surface of the ZnO NRs. Figure 3B Additionally, after Au deposition, the samples were examined using X-ray diffraction (XRD) to confirm the presence of Au nanoparticles on the ZnO NR surfaces. The XRD patterns of the ZnO NRs exhibited diffraction peaks corresponding to the wurtzite structure, with prominent peaks observed at 34.8°, 45.7°, and 63.2°, which were indexed to the (002), (102), and (103) planes of ZnO, respectively. (JCPDS card No. 01-070-8072).

After Au decoration, an additional diffraction peak appeared at 38.272°, corresponding to the (111) crystallographic plane of face-centered cubic (FCC) Au (JCPDS card No. 04-004-8456), confirming the successful deposition of Au nanoparticles. This result indicates that the Au nanoparticles exhibit a strong preferential (111) orientation,[4,9-11] suggesting a highly ordered and textured

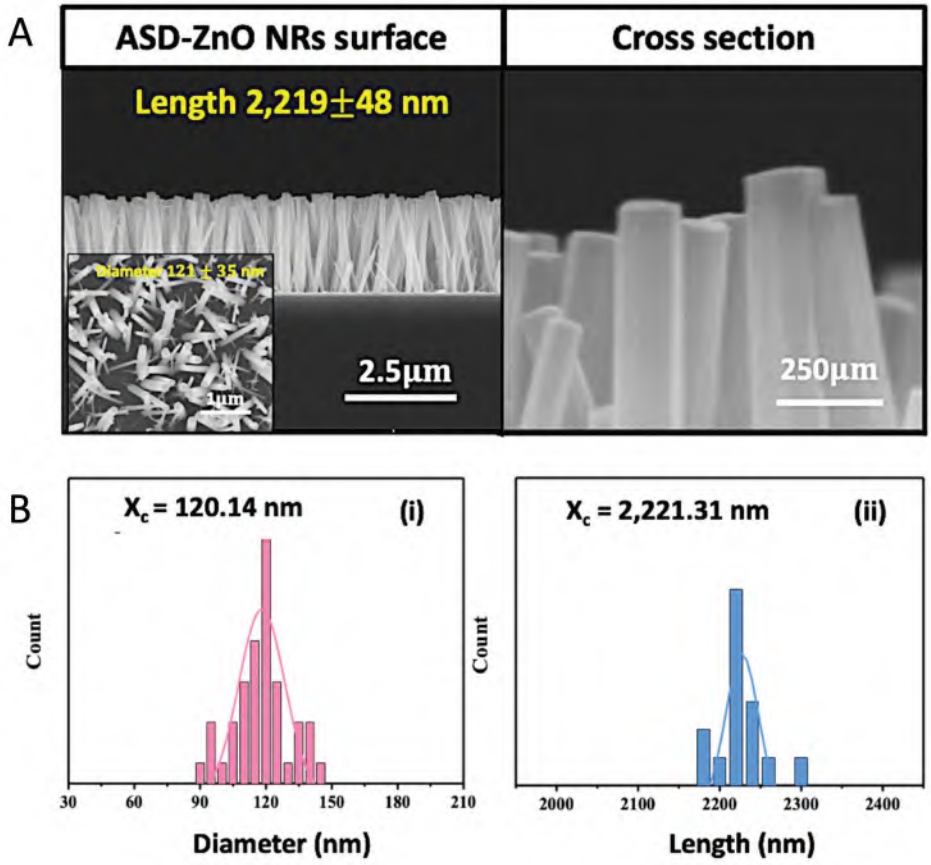


Figure 2. *Morphological and dimensional characterization of ASD-ZnO nanorods (NRs). A: FE-SEM image showing the physical morphologies of ASD-ZnO NRs, B: corresponding distribution plots of the diameter (pink) and length (blue) of the ZnO NRs templates.*

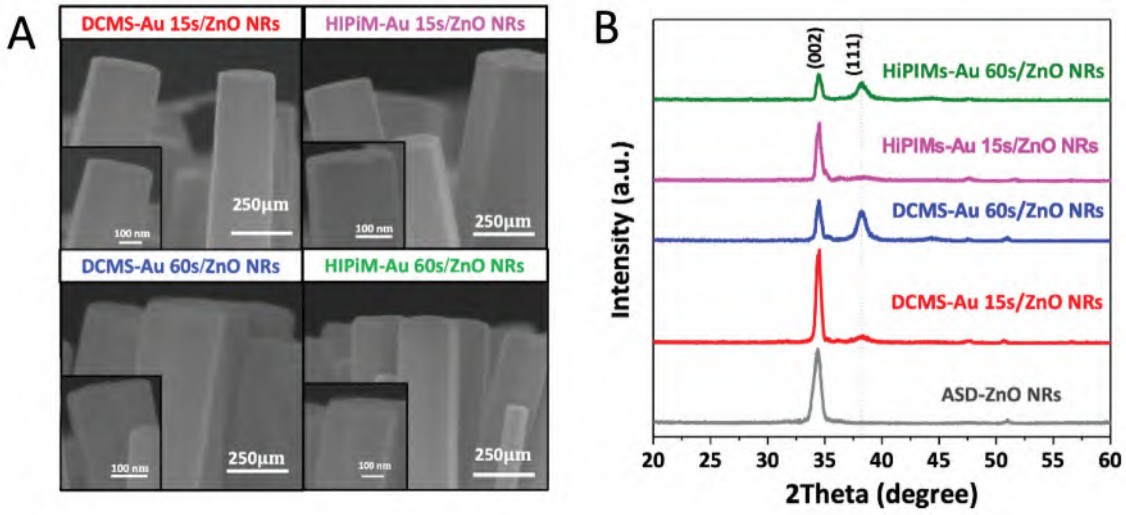


Figure 3. *Surface morphology and crystal structure of Au-decorated ZnO nanorods (NRs) prepared by DCMS and HiPIMS techniques. A: FE-SEM images showing the surface morphologies of Au-decorated ZnO NRs fabricated using DCMS and HiPIMS techniques at different Au deposition times (15 and 60 sec). Distinct differences in the distribution of Au nanoparticles were clearly observed. Red, blue, pink, and green text annotations indicate DCMS-Au 15 sec/ZnO NRs, DCMS-Au 60 sec/ZnO NRs, HiPIMS-Au 15 sec/ZnO NRs, and HiPIMS-Au 60 sec/ZnO NRs, respectively, B: XRD patterns of ZnO NRs and Au-decorated ZnO NRs prepared using the DCMS and HiPIMS techniques are shown. The gray line corresponds to the ZnO NRs, with diffraction peaks at 34.8°, 45.7°, and 63.2°. After Au deposition, an additional diffraction peak was observed at 38.272°. Red, blue, pink, and green lines represent the DCMS-Au 15 sec/ZnO NRs, DCMS-Au 60 sec/ZnO NRs, HIPIMS-Au 15 sec/ZnO NRs, and HIPIMS-Au 60 sec/ZnO NRs, respectively.*

structure. The preferential (111) growth of Au is expected to enhance the surface reactivity of Au-ZnO NRs.

To provide clearer evidence of how the gold nanoparticle morphology on the ZnO NRs surface varies with different deposition times (15 and 60 sec) and to support a more robust structural interpretation, we have included additional analysis beyond the FE-SEM images shown in Figure 3a), which has limitations and does not clearly reveal the presence or distribution of gold on the ZnO surface. Additional TEM images of the DCMS-Au 15 sec/ZnO NRs and DCMS-Au 60 sec/ZnO NRs samples are shown in Figure 4. Figure 4A illustrates the morphology of the Au-decorated ZnO nanorods after 15 sec of gold deposition, in which discrete Au nanoparticles were uniformly distributed over the nanorod surface. In contrast, Figure 4B (60 sec deposition) clearly shows a significant coalescence of Au nanoparticles, forming a continuous on the nanorod surfaces.

Furthermore, high-magnification TEM images at 1,000,000×, shown in Figure 4A and 4B for DCMS-Au 15 sec/ZnO NRs and DCMS-Au 60 sec/ZnO NRs, respectively, clearly reveal the crystalline structure of the AuNPs deposited on the ZnO NRs. In particular, the inset (i) displays distinct lattice fringes corresponding to the gold crystallites. These observations were further supported by elemental analysis via EDX mapping. As shown in Figure 5A, for the 15 sec Au-deposited sample, EDX elemental maps clearly illustrate the spatial distribution of gold (red; l), oxygen (blue; m), and zinc (green; n). Gold nanoparticles were observed as discrete, well-dispersed spots on the ZnO NR surfaces, consistent with the particle-like morphology observed in the corresponding TEM inset (i). In contrast, Figure 5B shows a significant morphological transformation for the 60 sec Au-deposited sample. The EDX map reveals a continuous gold (red; l) film covering the ZnO nanorod surfaces, as corroborated by inset (i) of the TEM image. This indicates substantial nanoparticle coalescence and film formation as a function of increased deposition time. These results confirm the morphological transition of Au as a function of deposition time.

The antibacterial efficacies of ZnO NRs and Au-ZnO NRs were evaluated. From figure 6A the antibacterial performance of ZnO NRs and Au-ZnO NRs was assessed against *Escherichia coli* at an initial concentration of 10^6 CFU/mL under UV-A exposure for 1 hr. The results revealed significant variations in bacterial inhibition depending on sputtering technique and deposition duration. ASD-ZnO NRs exhibited the lowest antibacterial efficiency (Figure 6B), with a killing rate of 26.64%, indicating limited antibacterial activity. In contrast, Au-ZnO NRs demonstrated enhanced antibacterial efficacy. Among the samples, the HiPIMS-Au 15 sec condition achieved the highest bacterial reduction of 99.62%, highlighting the superior antibacterial effect of the HiPIMS sputtering. Other Au-decorated samples, such as HiPIMS-Au 60 sec (89.62%), DCMS-Au 15 sec (85.69%), and DCMS-Au 60 sec (84.93%), also exhibited significant bacterial reduction compared to the control. These findings indicate that both the sputtering technique and the deposition duration play crucial roles in determining the antibacterial efficacy of Au-ZnO NRs.

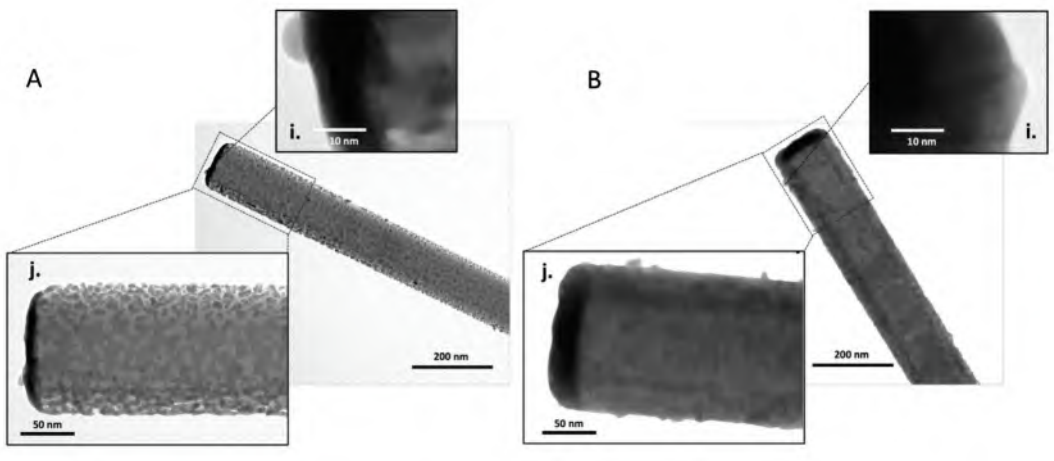


Figure 4. *High-resolution TEM images of a single Au-decorated ZnO nanorod fabricated using DCMS at two deposition times. A: 15 sec, B: 60 sec, i: insets clearly show lattice fringes corresponding to crystalline Au, confirming the presence of AuNPs on ZnO NRs.*

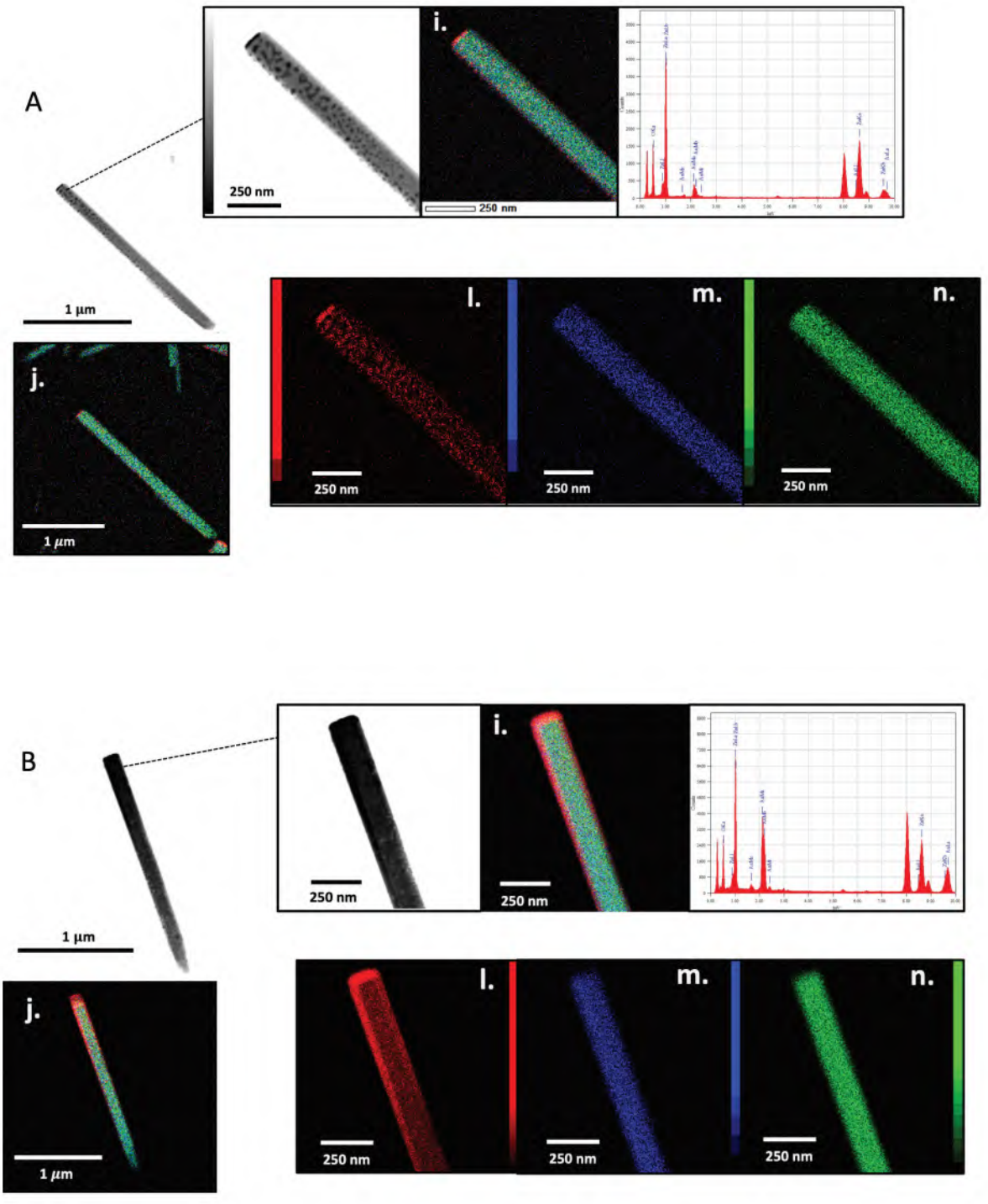


Figure 5. *EDX elemental mapping of a single Au-decorated ZnO nanorod. A: DCMS-Au 15 sec; B: DCMS-Au 60 sec. The elemental distributions of Au (red, l), O (blue, m), and Zn (green, n) are presented. For the 15 sec sample, Au appears as dispersed nanoparticles on the ZnO nanorods, as shown in Figure 5A (i, j), whereas for the 60 sec sample, Au forms a continuous film coating the nanorods, as shown as Figure 5B (I, j).*

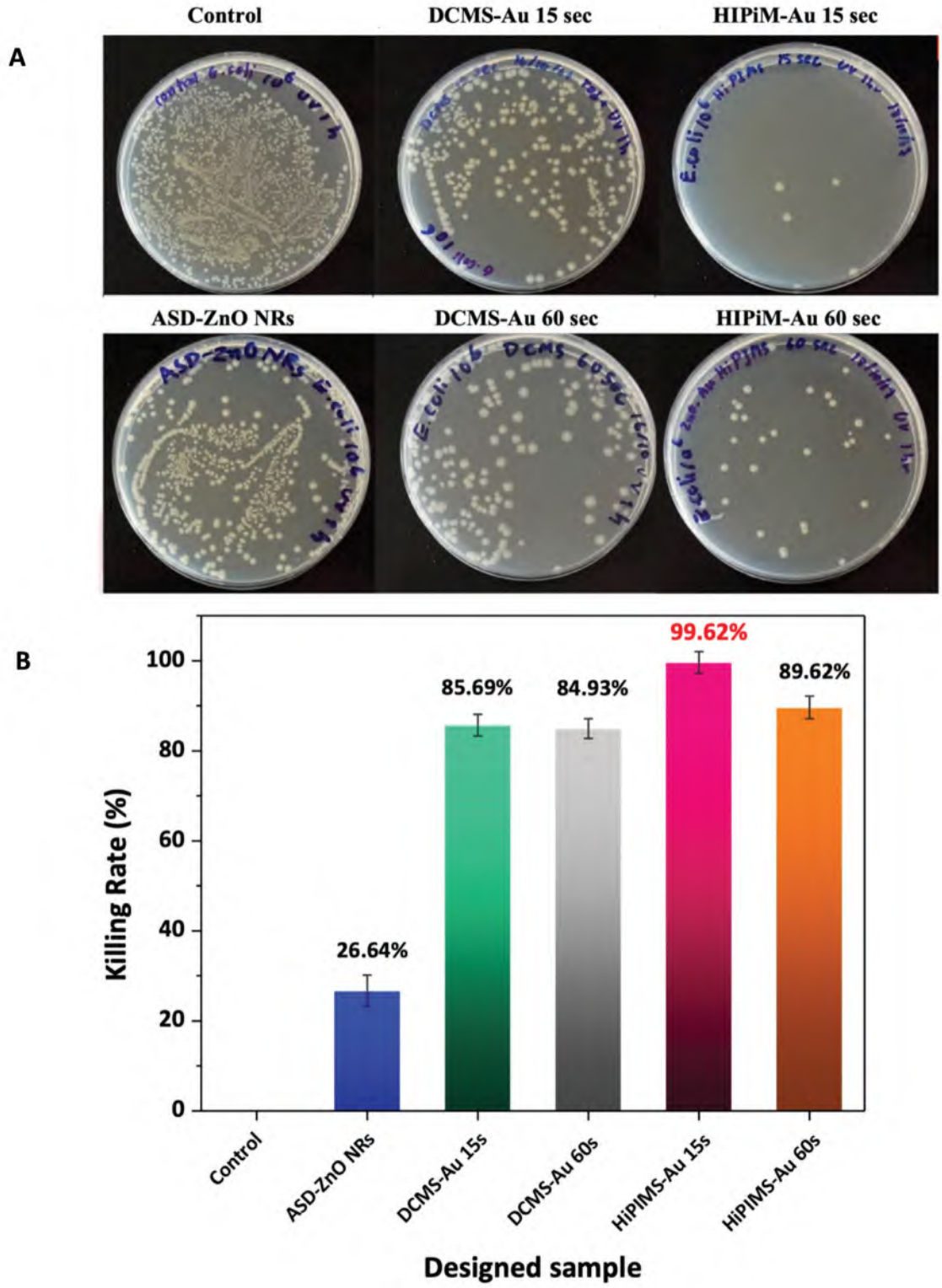


Figure 6. *Comparative antibacterial efficacy of ZnO nanorods and Au-decorated ZnO nanorods fabricated by DCMS and HiPIMS techniques under UV-A exposure. A: antibacterial efficacy of ZnO and Au-ZnO NRs was evaluated against Escherichia coli at a concentration of 10^6 CFU/mL under UV-A exposure for 1 hr. The results demonstrate significant variations in antibacterial activity based on the sputtering technique and duration. B: antibacterial killing rate of ZnO NRs and Au-ZnO NRs against Escherichia coli under UV-A exposure.*

Discussion

Figure 2A shows significant changes in the physical morphology of the NRs before and after the Au decoration process. The ZnO NRs exhibited a hexagonal structure with uniform distribution across the surface of the sample. As shown in Figure 3A, at a deposition time of 15 sec, the Au nanoparticles appeared as small, isolated particles. When the deposition time was increased to 60 sec, the Au nanoparticles formed an island-like structure that covered the entire surface of the NRs. The FE-SEM and TEM images clearly show the presence of Au nanoparticles coated on the ZnO NRs, compared to the FE-SEM images before Au deposition, as shown in Figure 2A and Figure 4, respectively. Figure 3B clearly shows a strong peak (002), which corresponds to the preferred ZnO growth direction along the c-axis. The other diffraction peak at 38.2°, corresponding to the (111) plane, was indexed as Au. From the antibacterial efficacy testing of the ZnO NRs and Au-ZnO NRs, the untreated control samples exhibited dense bacterial colonies, indicating no reduction in bacterial growth (Figure 6a, top left). This confirmed that UV-A exposure alone did not significantly inhibit bacterial survival. ASD-ZnO NRs (Pure ZnO): The ZnO NRs without Au decoration showed limited antibacterial activity, achieving a killing rate of 26.64% (Figure 6B).

This reflects the inherent but insufficient antibacterial properties of ZnO, which can generate reactive oxygen species (ROS), but not at levels sufficient to eradicate high bacterial loads. DCMS-decorated Au-ZnO NRs: Au-ZnO NRs prepared via DCMS sputtering for 15 sec demonstrated a killing rate of 85.69% (Figure 6A), top middle, and Figure 6B, representing a significant enhancement compared with pure ZnO NRs. Extending the sputtering time to 60 sec yielded a slightly lower killing rate of 84.93%, possibly because of nanoparticle aggregation, which reduced the effective surface area for bacterial interaction and ROS generation. This hypothesis is strongly supported by the high-resolution TEM and EDX mapping analyses presented in Figures 4 Figure 5 illustrates the morphology of Au-decorated ZnO NRs after 15 sec of gold deposition, where discrete Au nanoparticles are uniformly distributed across the NR surface. In contrast, Figure 4B (60 sec deposition) clearly reveals significant nanoparticle coalescence, resulting in a continuous film-like coverage of Au on the ZnO surface. This morphological transformation can be explained by the classical nucleation and growth theory. At shorter deposition times, the limited atomic flux and surface diffusion favor the formation of discrete nanoclusters. As the deposition time increased, the surface accumulated Au atoms, enhancing surface diffusion and promoting coalescence through grain-boundary migration mechanisms. This results in larger aggregates or film-like structures that reduce the specific surface area and limit effective bacterium-nanomaterial interactions. Furthermore, EDX elemental mapping supported this observation. For the 15 sec, sample gold was clearly seen as isolated (Figure 5A), well-dispersed nanoparticles, while the 60 sec sample (Figure 5B) revealed a near-continuous gold layer enveloping the ZnO nanorods. These findings provide quantitative and visual confirmation that extended deposition time promotes nanoparticle aggregation, which, in turn, diminishes antibacterial efficacy.

Taken together, the TEM and EDX mapping data corroborate the observed decline in the bacterial killing rate at 60 sec and validate the mechanistic hypothesis that nanoparticle aggregation negatively impacts antibacterial performance.

HIPIMS-decorated Au-ZnO NRs: HiPIMS sputtering for 15 sec achieved the highest antibacterial efficacy, with a killing rate of 99.62% (Figure 6A), top right, and Figure 6B). This indicated nearly complete bacterial eradication. HiPIMS sputtering for 60 sec resulted in a killing rate of 89.62% (Figure 6B), which, although lower than the 15 sec HiPIMS treatment, still outperformed both DCMS samples. These results highlight the critical role of sputtering technique and duration in determining the antibacterial efficacy of Au-ZnO NRs. HiPIMS-decorated samples consistently outperformed DCMS samples, likely due to the pulsed plasma's ability to deposit smaller, well-distributed Au nanoparticles, and the high energy from the HIPIMS power source enabled the Au nanoparticles to sputter onto the surface and adhere well to the ZnO NRs surface.[8]

To further elaborate on the advantages of HiPIMS over DCMS techniques, HiPIMS generates high peak power densities, resulting in the creation of plasma with a significantly higher ionization fraction compared to conventional DCMS. In HiPIMS, the plasma contains a larger proportion of ionized species (30-70%) than neutral atoms, enabling better energy transfer to the target material. Consequently, the ejected particles possessed higher kinetic energy, leading to the deposition of smaller, more uniformly distributed nanoparticles with strong adhesion to the ZnO nanorod surfaces. In contrast, DCMS produces a lower plasma density and a predominance of neutral atoms, resulting in the formation of larger nanoparticles with a less uniform distribution and weaker adhesion to the substrate. The smaller particle size and superior dispersion achieved by HiPIMS increased the surface area available for bacterial interactions and facilitated greater generation of reactive oxygen species (ROS) under UVA light exposure, thereby enhancing antibacterial efficacy. Furthermore, the stronger adhesion provided by the HiPIMS deposition ensures greater durability of the antibacterial coating during practical applications. These mechanisms collectively explain the superior antibacterial performance observed for the HiPIMS-Au 15sec/ZnO NRs compared to that of the DCMS-decorated samples. This observation is consistent with previous reports, which demonstrated that HiPIMS techniques yield denser, smoother, and more functional nanostructured films than DCMS approaches.[9,10,12,]

This significantly improves the performance, as the nanoparticles work synergistically with ZnO to enhance ROS generation under UV-A light and disrupt bacterial cell membranes through additional mechanisms, such as ion release. The superior performance of HiPIMS Au-ZnO NRs, particularly with a sputtering duration of 15 sec, underscores

their potential as effective antibacterial agents.[13] These findings suggest that optimizing sputtering conditions can n not only improve antibacterial efficacy, but also reduce material and production costs, making them suitable for applications in healthcare, water purification, and environmental remediation.

In this study, the antibacterial activity of samples exposed to UV-A was evaluated. Future studies should investigate the efficacy of these materials under other irradiation conditions. However, the experimental data demonstrated that the use of ZnO NRs, including those decorated with Au nanoparticles, via both DCMS and HiPIMS methods did not result in a significant reduction in *S. aureus* colonies (Supplementary data). Despite UV-A exposure, the nanorods did not effectively inhibit the growth of *S. aureus*. This suggests that the antibacterial properties observed in *E. coli* may not extend to *S. aureus* under the tested conditions, indicating a potential limitation in the effectiveness of Au-decorated ZnO NRs against this specific bacterial strain.[14]

These results indicated that additional modifications or alternative approaches may be necessary to achieve effective antibacterial activity against *S. aureus*. Future studies should explore the long-term stability and reusability of the HiPIMS-Au-ZnO NRs in real-world applications. Further investigation is warranted to examine the effects of Au nanoparticle size, shape, and distribution on antibacterial performance. Additionally, the development of scalable and cost-effective fabrication techniques for HiPIMS-decorated nanomaterials is recommended.

Limitation

This study was limited to *E. coli*. Expanding the scope to include other bacterial species, especially multi-drug-resistant strains, is essential.

Conclusion

The antibacterial efficacy of ZnO NRs decorated with Au nanoparticles was tested against *Escherichia coli have successfully achieved* at 10^6 CFU/mL under UVA exposure for 1 hr. Untreated controls showed dense bacterial growth. ASD-ZnO NRs displayed limited activity, with a killing rate of 26.64%. Au decoration via DCMS for 15 sec and 60 sec achieved moderate antibacterial effects (85.69% and 84.93% killing rates, respectively). HiPIMS for 15 sec showed the highest efficacy with a 99.62% killing rate, followed by HiPIMS for 60 sec at 89.62%.

These results highlight the enhanced antibacterial activity of Au-decorated ZnO NRs, especially with HiPIMS for 15 sec. Au decoration significantly enhanced the antibacterial efficacy of the ZnO NRs, with HiPIMS sputtering at 15 sec providing the highest performance. These results highlight the potential of Au-ZnO NRs as innovative solutions for combating bacterial infections, particularly in the context of antimicrobial resistance. Further studies should explore scalability, long-term stability, and efficacy against a broader spectrum of pathogens to facilitate their practical application.

Funding

This research received financial support from Bangkokthonburi University.

Conflict of interest

The authors declare that they have no conflicts of interest.

Credit authorship contribution statement

Sukon Kalasung: conceptualization, methodology, formal analysis, investigation, writing original draft, writing review and editing, visualization, validation; **Saksorn Limwichean:** conceptualization, methodology, investigation, validation, resources; **Pitak Eiamchai:** methodology, investigation, resources; **Mati Horprathum:** methodology, investigation, resources; **Noppadon Nuntawong:** methodology, investigation, resources; **Viyapol Patthanasettakul:** methodology, investigation, resources; **Artitaya Yatsomboon:** conceptualization, methodology, form alanalysis, investigation, writing review and editing, visualization, validation.

Acknowledgements

The authors would like to thank Bangkokthonburi University, National Electronics and Computer Technology Center (NECTEC), and National Science and Technology Development Agency (NSTDA) for their financial support through research.

References

[1] O'Neill J. Tackling drug-resistant infections globally: final report and recommendations. London: Wellcome Trust; 2016.

[2] Naghavi M, Vollset SE, Ikuta KS, Swetschinski LR, Gray AP, Wool EE, et al. Global burden of bacterial antimicrobial resistance 1990-2021: a systematic analysis with forecasts to 2050. The Lancet. 2024; 404(10459): 1199-226. doi: 10.1016/S0140-6736(24) 01867-1.

[3] Xie J, Li H, Zhang T, Song B, Wang X, Gu Z. Recent advances in ZnO nanomaterial-mediated biological applications and action mechanisms. Nanomaterials. 2023;13(9):1500. doi: 10.3390/nano13091500.

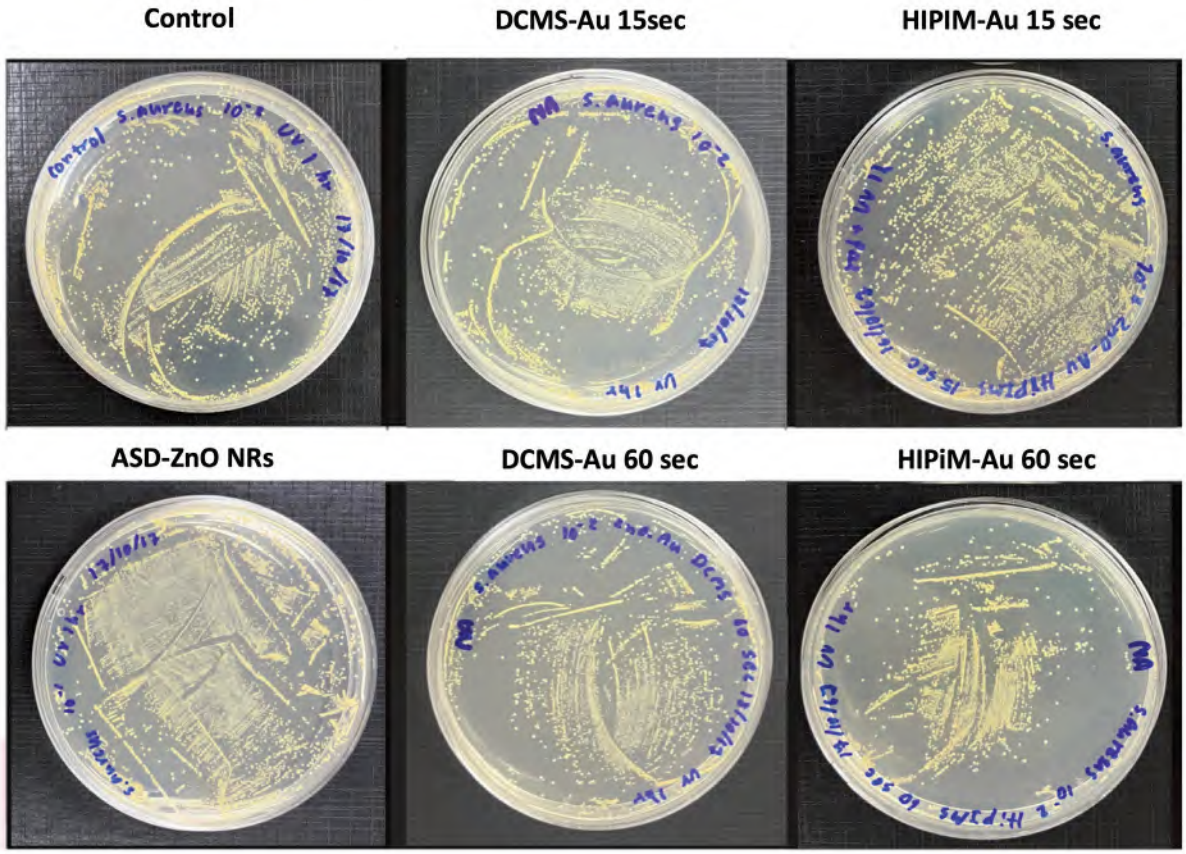
[4] Kalasung S, Aiempanakit K, Chatnuntawech I, Limsuwan N, Lertborworn K, Patthanasettakul V, et al. Trace-level detection and classifications of pentaerythritol tetranitrate via geometrically optimized film-based Au/ZnO SERS sensors. Sens Actuators B Chem. 2022; 366: 131986. doi: 10.1016/j. snb.2022.131986.

[5] Khan SS, Ullah I, Ullah S, An R, Xu H, Nie K, et al. Recent advances in the surface functionalization of nanomaterials for antimicrobial applications. Materials. 2021;14(22): 6932. doi: 10.3390/ma14226932.

[6] Nuchuay P, Laongwan C, Promcham W, Somboonsaksri P, Kalasung S, Chananonnawathorn C, et al. A study of the electrical and optical properties of AZO thin film by controlling pulse frequency of HiPIMS. J.Met. Mater.Miner.2023; 33(2): 103-7. doi: 10.55713/jmmm.

v33i2.1696.
[7] Sornsanit K, Horprathum M, Eiamchai P, Chananonnawathorn C, Kalasung S, Kaewkhao J. Enhanced antibacterial activity by Au nanoparticle decorated ZnO nanorods. Key Eng Mater. 2016;113–6.
[8] Dediu V, Busila M, Tucureanu V, Bucur FI, Iliescu FS, Brincoveanu O, et al. Synthesis of ZnO/Au Nanocomposite for Antibacterial Applications. Nanomaterials. 2022; 12(21): 3832 doi: 10.3390/nano12213832.
[9] Fageria P, Gangopadhyay S, Pande S. Synthesis of ZnO/Au and ZnO/Ag nanoparticles and their photocatalytic application using UV and visible light. RSC Adv. 2014; 4(48): 24962-72. doi: 10.1039/C4RA03158J.
[10] Kalasung S, Kopwitthaya A, Horprathum M, Kaewkhao J, Tuscharoen S, Eiamchai P, et al. Functionalization of Au nanoparticles on ZnO nanorods through low-temperature synthesis. Key Eng Mater. 2016; 675-6: 45-8. doi: 10.4028/www.scientific.net/KEM.675-676.45.
[11] Kalasung S, Chatnuntawech I, Patthanasettakul V, Limwichean S, Lertborworn K, Horprathum M, et al. Au-decorated ZnO nanorod arrays for SERS-active substrates towards trace detection and classification of pentaerythritol tetranitrate. Mater Today Proc. 2022; 56: 2245–2251. doi: 10.1016/j.matpr.2021.03.511
[12] Busila M, Musat V, Alexandru P, Romanitan C, Brincoveanu O, Tucureanu V, et al. Antibacterial and Photocatalytic Activity of ZnO/Au and ZnO/Ag Nanocomposites. Int J Mol Sci. 2023; 24(23): 16939 doi: 10.3390/ijms242316939
[13] Abebe B, Zereffa EA, Tadesse A, Murthy HCA. A Review on Enhancing the Antibacterial Activity of ZnO: Mechanisms and Microscopic Investigation, Nanoscale Res. Lett. 2020; 15: 190 doi: 10.1186/s11671-020-03418-6.
[14] Mlynarczyk-Bonikowska B, Kowalewski C, Krolak-Ulinska A, Marusza W. Molecular Mechanisms of Drug Resistance in *Staphylococcus aureus*. Int J Mol Sci. 2022 Jul 22; 23(15): 8088. doi: 10.3390/ijms23158088.

Supplementary Data



The antibacterial efficacy of ZnO nanorods (ZnO NRs) and gold-decorated ZnO nanorods (Au-ZnO NRs) was evaluated against *Staphylococcus aureus* at a concentration of 10^6 CFU/mL under UV-A exposure for 1 hr. The results showed no significant differences in antibacterial activity against *Staphylococcus aureus* among the control group, ZnO NRs, and Au-decorated ZnO NRs, regardless of the sputtering technique or duration.

Journal of Associated Medical Sciences 2025; 58(3): 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

Active ageing level and interest in activities among older people in community group membership, Chiang Mai, Thailand

Donyaporn Srijomthong, Supawadee Putthinoi, Autchariya Punyakaew*

Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 11 March 2025
Accepted as revised 29 June 2025
Available online 1 July 2025

Keywords:
Active ageing, interest, activities, occupational therapy, older people

ABSTRACT

Background: An efficient strategy for an ageing population is to promote a healthy lifestyle that emphasizes improving and preserving functional capabilities to support the well-being of older people. From an occupational therapy perspective, one must consider the interests and needs of older people when analyzing their activities, as these factors influence improving their personal skills.

Objective: This study aimed to examine the levels of active ageing and the specific areas of interest among older people participating in community groups within Nongpakang Village, Chiang Mai, Thailand.

Materials and methods: This research study involved thirty older people aged 60 years and above, all determined to have no cognitive impairments based on assessments conducted using the Mental State Examination T10 (MSET10). The study employed a cross-sectional survey to examine participants' levels of active ageing and their interest in various activities. Data was collected through structured interview-based questionnaires, ensuring a comprehensive assessment of these factors.

Results: The findings indicate a high level of active ageing among older people, as reflected in the Active Ageing Index (AAI) score of 0.94. Participants' interest in activities was categorized into five domains: manual skills, physical sports, activities of daily living (ADLs), educational and cultural activities, and social recreation. Among these categories, most participants demonstrated a strong preference for ADL and social recreation. Conversely, engagement in handicraft-related activities and manual skills was relatively low, suggesting limited interest in these domains.

Conclusion: The study revealed that the mean score within this AAI group was notably high, indicating a strong level of active ageing among participants. Furthermore, the findings suggest that older people exhibit diverse interests in activities, which vary according to their backgrounds. A comprehensive understanding of these interests can enhance therapeutic interventions, fostering meaningful participation in everyday activities and promoting overall well-being.

Introduction

The ageing population represents a global phenomenon. Nearly all developed countries are currently undergoing demographic ageing, while many developing countries are observing a rapid rise in their ageing populations. In 2005, Thailand moved into an ageing society, as individuals aged 60 or older established 10% of the nation's population. Thailand became a completely aged society in 2023.[1] As people age, the probability of dealing with health problems and degenerative conditions increases.[2] Consequently, older people could burden their families, communities,

* **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: autchariya.punyakaew@cmu.ac.th
doi: 10.12982/JAMS.2025.087
E-ISSN: 2539-6056

and society. This will affect the quality of life for older people and result in a strengthened burden on Thai society. In 2002, the World Health Organization (WHO) introduced the idea of active ageing to address the increasing population of older people. The objective is to maintain the active engagement of older people in society, enabling ongoing access to healthcare, social participation, and security.[3] The United Nations Development Programme (UNDP) 2005 categorized the active ageing levels into high, moderate, and low classifications.[4]

The Thai government has acknowledged the importance of addressing the ageing population by launching numerous significant projects to enhance engagement across all societal strata. The execution of policies that promote and support active ageing in the elderly depends on the cooperation of all sectors. In 2017, the National Statistical Office (NSO) of Thailand introduced the active ageing concept, integrating a new element known as the enabling environment for active ageing to evaluate the active ageing status of Thai elders in the Thai context. The findings indicated a moderate level of active ageing, with older individuals in Northern Thailand exhibiting the highest levels compared to other regions.[5] The different levels of active ageing reflect the health status and quality of life of older people based on the four parts of the AAI. The concept of active ageing is complex, encompassing health indicators and social, environmental, and economic dimensions, which differ among contextual and cultural perspectives, requiring careful consideration. Consequently, it is essential to examine active ageing in every component. Promoting proactive planning and strategic preparation for ageing among older people profoundly impacts their overall quality of life and level of independence. By fostering awareness and encouraging engagement in anticipatory strategies, individuals are better equipped to navigate the physical, psychological, and social changes associated with ageing. This approach supports sustained autonomy and enhances well-being, enabling older people to participate actively in daily activities and social interactions.

Occupational therapy (OT) highlights its significance in promoting physical and mental well-being in older people. OT is crucial in helping seniors maintain their independence and improve their daily living skills. The independence of older people influences the value and quality of life. In the geriatric field, occupational therapy focuses on promoting health and well-being and promoting participation in meaningful activities for older people.[6] Moreover, occupational therapy emphasizes a client-centered approach that prioritizes everyone's desires and needs, particularly their interests in performing activities. Occupational therapists analyze older people's interests and values in their occupation and help them discover abilities, interests, and skills based on their existing experiences.[7] However, when providing rehabilitation and health promotion intervention for older people, the context must also be considered because it will affect motivation to perform activities. Occupation is unique and tailored to everyone's older people, which will reflect the ability and the uniqueness of that person.[8] The ability of older people to perform activities depends on their interest in each one. If older people can perform activities they are interested in, it will help them fulfill their life roles. The interests of older people are different for each age. Therefore, the survey of older people's interests reflects their lifestyle preferences. This will lead to promoting and advising older people about maintaining or adjusting their lifestyle appropriately according to their age.

Ageing relates directly to health problems in terms of physical health, mental well-being, and functional ability.[9] The importance of promoting active ageing in older people relates to performing various activities efficiently. Supporting older people at different levels of age should suit everyone's abilities. Occupational therapists must focus on helping older people engage in activities that align with their interests, which will contribute to developing their skills. This study aimed to investigate the levels of active ageing and the activity interests of older people who are members of the Senior Activity and Service Center of Nongpakrang Subdistrict Municipality, Mueang District, Chiang Mai, Thailand. By focusing on this specific population, the research aims to provide valuable insights into their ageing experiences and preferences, thereby contributing to the development of tailored programs and interventions that support their well-being and active engagement.

Materials and methods

Study Design

This study was a cross-sectional survey of older people that provided data on the active ageing level and interest in activities in a study area.

Study Setting and Participants

The study setting was a Nongpakang village in Chiang Mai, Thailand. The participants were 30 older people who attended community groups in Chiang Mai. The inclusion criteria were being 60 years or older, being voluntary participants in the study, and being cooperative in understanding the questionnaire. The exclusion criterion was cognitive deficit found by using the MSET10 (scored according to the participant's education).[10] All of the participants gave their informed consent to take part in this study.

Data Collection

The data were collected between August and November 2023. Researchers visited the community to explain the study's objectives and obtained informed consent from all participants. Socio-demographic and cognitive data were gathered using the MSET10 screening tool. Active ageing levels were assessed with the AAI,[11] which comprises four equally weighted components: health, social participation, security, and an enabling environment for active ageing. Interview-based questionnaires were administered to explore participants' interests in daily and meaningful activities.[12] The questionnaire contained 36 items across five domains, including manual skills, educational and cultural activities, physical sports, activities of daily living

(ADL), and social recreation. The tool had undergone content validity testing and demonstrated high internal consistency (Cronbach's alpha=0.91). Interviews were conducted individually in a private setting at the community center.

Data Analysis

Socio-demographic characteristics were analyzed using frequencies and percentages. AAI scores and activity interest scores were examined using means and standard deviations. The AAI was calculated as a weighted average of four components: health, participation, security, and an enabling environment for active ageing. Scores ranged from 0 to 1 and were categorized as low (0.000-0.499), moderate (0.500-0.799), and high (0.800-1.000).

Results

A total of 30 older people agreed to participate in this study. The results were divided into three parts: socio-demographic characteristics of participants, AAI, and interests in activities as follows.

Socio-demographic characteristics of participants

The socio-demographic characteristics of the study participants are outlined in Table 1, providing a detailed overview of their profiles. Most participants identified as Buddhists (96.67%), with a smaller proportion identifying as Christians (3.33%). Regarding gender distribution, females constituted a significantly larger percentage than males (83.33% and 16.67%, respectively). Participants' ages ranged from 60 to over 80 years, with the largest age group being 70-74 (33.33%). Marital status data revealed equal proportions of married and widowed individuals (46.67% each), while single participants represented a smaller group (6.67%). Educational attainment varied among participants, with 36.67% having completed less than high school, 26.67% being high school graduates, 33.34% having obtained a bachelor's degree, and 3.33% reporting some college education. Employment status showed that 53.33% of participants were actively working, while 46.67% were not employed during the study. Additionally, most participants (76.67%) lived in single-parent families, whereas 23.33% resided in extended-family households.

Table 1. Socio-demographic characteristics of participants (N=30).

Characteristics	N	%
Gender		
Male	5	16.67
Female	25	83.33
Age (years)		
60-64	3	10.00
65-69	8	26.67
70-74	10	33.33
75-79	7	23.33
>80	2	6.67
Marital status		
Single	2	6.67
Married	14	46.67
Widowed	14	46.67
Education		
Less than high school	11	36.67
High school graduate	8	26.67
Some college	1	3.33
College graduate (bachelor's degree)	10	33.34
Religion		
Buddhism	29	96.67
Christian	1	3.33
Current working status		
Not working	14	46.67
Working	16	53.33
Family type		
Single parent family	23	76.67
Extended family	7	23.33
Chronic health conditions		
No	5	16.66
Yes*	25	83.34

*Note: *older people with chronic health conditions include hyperlipidemia, heart disease, hypertension, arthritis, and diabetes.*

Active Ageing Levels

The characteristics of active ageing among community group members are presented in Table 2. The active ageing levels were classified based on the AAI. All thirty participants demonstrated a high level of active ageing, encompassing health, social participation, security, and an enabling environment for active ageing. The average AAI score within this group was 0.94, reflecting a consistently high level of active ageing across participants.

Interest in Activities

The activities can be divided into five categories: manual skills, physical sport, ADL, educational and cultural activities, and social recreation. Most of the participants were interested in ADL and social recreation. The activities are shown in Table 3.

Table 2. Active ageing Index by dimensions (N=30).

Components	Active ageing levels			(Mean±SD)	Interpretation
	High N (%)	Moderate N (%)	Low N (%)		
Health	26 (86.67%)	4 (13.33%)	0 (0.00%)	0.884±0.074	High
Social participation	29 (96.67%)	1 (3.33%)	0 (0.00%)	0.956±0.072	High
Security	30 (100.00%)	0 (0.00%)	0 (0.00%)	0.937±0.070	High
Enabling environment for active ageing	30 (100.00%)	0 (0.00%)	0 (0.00%)	1.000±0.000	High
Total AAI score				**0.944±0.040**	**High**

Table 3. Interest in activity categories (N=30).

Activity category	Interested N (%)	Not interested N (%)
Manual skill	22 (73.33%)	8 (26.67%)
Educational and cultural activities	27 (90.00%)	3 (10.00%)
Physical sport	29 (96.67%)	1 (3.33%)
ADL	30 (100.00%)	0 (0.00%)
Social recreation	30 (100.00%)	0 (0.00%)

This study categorizes activities into five distinct domains. The manual skills category (5 activities) includes tasks requiring technical proficiency and craftsmanship, such as sewing, car repair, handicrafts, woodcarving, and artistic endeavors like painting and drawing. The educational and cultural activities category (2 activities) emphasizes intellectual engagement through writing and reading, fostering cognitive stimulation and cultural appreciation. The physical sports category (3 activities) consists of sports participation, cycling, and general physical exercise, all contributing to physical well-being and mobility. The ADL category (8 activities) encompasses essential daily tasks that support independence and functionality, including mending clothes, shopping, dressing, laundry, house cleaning, cooking, driving, and ironing. Lastly, the social recreation category (18 activities) covers a diverse range of leisure and communal activities, such as gardening, poetry writing, social gatherings, club participation, card games, listening to the radio, watching television or movies, dancing, visiting others, chess and checkers, dominoes and puzzle games, festival-related activities, collecting items, singing, baking, religious practices, volunteer work, and musical pursuits. These categories provide a structured framework for assessing participants' engagement in meaningful activities.

Discussion

The result of AAI is calculated based on four components: health, social participation, security, and an enabling environment for active ageing. This study utilized a small, community-based sample, which may limit the generalizability of the findings. The results indicate that older individuals attending community groups in Nongpakang exhibited a high level of active ageing, with an overall mean AAI score of 0.944. These findings are consistent with those of Muengmoon[13], who reported that older adults engaged in community groups demonstrate a high level of active ageing. Such results suggest that structured community participation plays a significant role in promoting well-being and facilitating active ageing.

The high AAI scores observed in this study may reflect the municipality's ongoing support for older adults through

the development of community centers, health promotion initiatives, and skill development programs. Wongsala *et al*. discussed the three pillars of active ageing, emphasizing the importance of involving older adults in programs and activities specifically designed for them.[9] This research highlights the significance of access to culturally relevant resources and engagement in local daily-life activities as essential factors in promoting active ageing. Furthermore, previous studies indicate that access to these resources empowers older adults to maintain autonomy and social inclusion, reinforcing the importance of a supportive community environment. [14] These findings suggest that policymakers should continue developing health and social policies that align with older adults' perspectives and support their autonomy, participation, and security-ultimately contributing to an improved quality of life.

Regarding activity preferences, older adults demonstrated the highest interest levels in ADLs and social recreation. ADLs, which encompass fundamental self-care tasks, are essential for maintaining independence and preventing institutionalization. Limitations in performing ADLs can result in adverse outcomes, such as increased dependency and a diminished quality of life.[15,16] Therefore, ensuring that older adults retain the ability to perform ADLs should be a primary focus of age-related programs. Social recreation was also universally valued, likely due to its comprehensive benefits for both physical and mental health.[17] These activities foster social interaction and emotional well-being and help mitigate feelings of loneliness and isolation, which are prevalent among older adults.

Among the activities examined, manual skills received the lowest interest among older adults, with woodcarving being the least preferred. Manual skill-based tasks require high attention to detail and involve various physical abilities, such as manual dexterity, agility, and bodily strength. These activities are often perceived as complex and time-consuming.[18] In contrast, woodcarving demands fine motor skills, hand strength, and endurance, which tend to decline with age.[19,20] Age-related deteriorations in motor function, especially in tasks requiring precision, may contribute to the lack of interest in such activities. Modifying these tasks to accommodate physical limitations or providing assistive tools may enhance engagement among older participants.

Finally, as highlighted by Punyakaew *et al*. this study examines active ageing levels and time-use patterns among elderly individuals in a suburban Thai community.[21] The findings indicate that individuals with higher levels of active ageing engage more frequently in leisure and social participation. Understanding time allocation among healthy elderly individuals is essential for developing effective active ageing strategies, as increased leisure and social interaction can enhance the quality of life for those with moderate and low levels of active ageing. Consequently, participation in meaningful daily activities fosters a sense of purpose and personal fulfillment among older adults. Promoting healthy ageing requires a comprehensive approach that integrates medical and physical health interventions with robust social and psychological support systems. Programs designed to improve lifestyle quality through personalized activities-tailored to individuals' physical capabilities and personal interests-may mitigate the effects of age-related diseases and support ageing with dignity.

Limitations

This study has several limitations that should be acknowledged. First, the sample was limited to older adults attending a community group in Nongpakang village, Chiang Mai, Thailand. Consequently, the study did not encompass diverse subgroups, such as home-bound or bed-bound elders, nor did it account for variations across ethnic, racial, or cultural backgrounds. As such, the findings cannot be generalized to all older adults in Chiang Mai or Thailand. Second, the relatively small sample size further limits the generalizability of the results. Future research should include a larger and more diverse sample, incorporating older adults from various settings, including those who are home-bound or bed-bound, and individuals from different geographic regions.

Another limitation is that while this study explored older adults' interests in activities, it did not examine the underlying factors influencing these choices in depth. Preliminary findings suggest potential gender differences in activity interests. Therefore, future studies should investigate demographic and psychosocial factors, such as gender, health status, and cultural influences, that may shape older adults' engagement in meaningful activities.

These findings underscore the importance of considering older adults' preferences and functional capacities when designing and implementing activity programs. Occupational therapists and healthcare providers can utilize this evidence to develop tailored interventions that align with older adults' needs and abilities, enhancing health, well-being, and active ageing across diverse populations.

Conclusion

This study examined the active ageing levels of older adults by assessing the AAI, which ranges from 0 to 1, and exploring their interests in activities. The findings indicate that older adults participating in community groups in Nongpakang village, Chiang Mai, Thailand, exhibited a high level of active ageing, with a mean AAI score of 0.94. All four AAI components-health, social participation, security, and enabling environment-were also at high levels.

Regarding activity preferences, most participants expressed strong interest in ADLs and social recreation, whereas manual skills, particularly woodcarving, were the least favored. These results highlight the diversity of activity preferences among older adults and underscore the importance of aligning activity programs with their capabilities and interests.

These findings suggest that occupational therapists and other health professionals should consider older adults' preferences and functional capacities when designing and implementing interventions to promote active ageing.

Ethical approval

This study was approved by the Ethics Committee of the Faculty of Associate Medical Sciences, Chiang Mai University (number: AMSEC-66EX060).

Funding

This research received no external funding.

Conflict of interest

The authors declare no conflicts of interest concerning the research, authorship, or publication of this article.

CReDIT authorship contribution statement

Donyaporn Srijomthong: data analysis; **Supawadee Putthinoi**: data collection and analysis; **Autchariya Punyakaew**: research design and manuscript publication.

Acknowledgements

The authors extend their gratitude to all participants who were members of the Community Group in Nongpakang Village, Chiang Mai, Thailand. This study was financially supported by the Office of Research Administration and the Faculty of Associated Medical Sciences, Chiang Mai University, Thailand.

References

[1] Thailand Development Research Institute. Promotion of active ageing and quality of life in old age and preparation for a complete aged society in Thailand. 2023 [cited 2024 October 25]. Available from: https://tdri.or.th/wp-content/uploads/2023/11/Volume-38-Number-3-September-2023. pdf.

[2] Michael AP, Lorie GR. Promoting health, well-being, and quality of life in occupational therapy: A commitment to a paradigm shifts for the next 100 years. Am J Occup Ther. 2017; 71(4): 71041700-10p1-5. doi: 10.5014/ajot.2017.028456.

[3] World Health Organization. Active ageing: A policy framework. Geneva: WHO; 2002.

[4] United Nations Development Programme. Human Development Report 2005: International Cooperation at a Crossroads - Aid, Trade and Security in an Unequal World. New York: UNDP; 2005.

[5] National Statistical Office of Thailand. Active ageing index of Thai elderly 2017. 2017 [cited 2024 Aug 15]. Available from: http://www.nso.go.th/sites/2014en/active_ageing.

[6] American Occupational Therapy Association. Productive ageing. 2010 [cited 2024 Mar 20]. Available from: https://www.aota.arg/-/media/Corporate/Files/Practice/Ageing/Distinct-Value-Productive-Ageing.pdf.

[7] Kielhofner G. A Model of Human Occupation: Theory and Application. 4th Ed. Baltimore: Lippincott Williams & Wilkins; 2008.

[8] Ha Bowyer P, Bélanger R, Briand C, Heras C, Kinébanian A, Launiainen H, et al. International efforts to disseminate and develop the Model of Human Occupation. Occup Ther Health Care. 2008;22: 1-24. doi:10.1080/07380570801989291.

[9] Wongsala M, Anbäcken E, Rosendahl S. Active ageing - perspectives on health, participation, and security among older adults in northeastern Thailand - a qualitative study. BMC Geriatr. 2021; 21(41): 1-10. doi.org/10.1186/s12877-020-01981-2.

[10] Dementia Association of Thailand. The Mental State Examination T10 (MSET10). 2018 [cited 2024 Aug 15]. Available from: http://www.thaidementia.org/core/File/839.pdf.

[11] Punyakaew A, Hsu HY, Lersilp S, Putthinoi S. Development and psychometric properties of a questionnaire to measure the active ageing index for older people. J Assoc Med Sci. 2022;56(1):150-8. doi:10.12982/JAMS.2023.030.

[12] Muengudon S. Active ageing level and interest in older people. [Term paper]. Chiang Mai: Faculty of Associated Medical Sciences, Chiang Mai University; 2019. [in Thai].

[13] Lakkatham T. Active ageing level and functional ability in performing activities and participation in social activities in elderly. [Term paper]. Chiang Mai: Faculty of Associated Medical Sciences, Chiang Mai University; 2022. [in Thai].

[14] Muengmoon P. Active ageing level, basic activities of daily living, and instrumental activities of daily living. [Term paper]. Chiang Mai: Faculty of Associated Medical Sciences, Chiang Mai University; 2022. [in Thai].

[15] Jeenmuang K, Kaewsawas S, Thanapop C, Thanapop S. Social support, active ageing perception and practices among educational staff in the primary educational service, Nakhon-Si-Thammarat province, Southern Thailand. Soc Sci. 2023; 12(9): 486. https://doi.org/10.3390/socsci12090486.

[16] Gao J, Gao Q, Huo L, Yang J. Impaired activity of daily living status of older adults and its influencing factors: A cross-sectional study. Int J Environ Res Public Health. 2022; 19(23): 15607. doi:10.3390/ijerph192315607.

[17] Jin G, Qing G, Liting H, Jianchuang Y. Impaired activity of daily living status of older adults and its influencing factors: A cross-sectional study. Int J Environ Res Public Health. 2022; 19(23): 1-10. doi: 10.3390/ijerph192315607.

[18] Tohan MM, Ahmed F, Juie IJ, et al. Outdoor recreational activities and mental well-being of geriatric people in Bangladesh: Structural equation modeling. Discov Psychol. 2024; 4(33): 1-16 https://doi.org/10.1007/s44202-024-00131-8.

[19] European Commission, Cedefop. Physical and manual skills and competences. 2021 [cited 2024 Mar 20]. Available from: https://www.cedefop.europa.eu/en/tools/vet-glossary/glossary/fysieke-vaardigheden-en-handvaardigheden.

[20] Emilio C, Sonia T, Francesca F, et al. You are as old as the connectivity you keep: Distinct neurophysiological mechanisms underlying age-related changes in hand dexterity and strength. Arch Med Res. 2025; 56(1): 1-16. https://doi.org/10.1016/j.arcmed.2024.103031.

[21] Punyakaew A, Lersilp S, Puttinoi S. Active ageing level and time use of elderly persons in a Thai suburban community. Occup Ther Int. 2019; 2019: 1-8. doi: 10.1155/2019/7092695.

Journal of Associated Medical Sciences 2025; 58(3): 191-203

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

Extended verbal fluency in older adults: Results of a 2-minute test across animal, object, and food categories with frequently listed words

Nicha Kripanan*, Somjit Ruamsuk, Thanwarat Artayakul, Isara Suttichujit

Sirindhorn National Medical Rehabilitation Institute, Nonthaburi Province, Thailand.

ARTICLE INFO

Article history:
Received 25 February 2025
Accepted as revised 26 June 2025
Available online 7 July 2025

Keywords:
Two-minute verbal fluency, Thai older, Nonthaburi, Thai language, frequently listed words

ABSTRACT

Background: Verbal fluency tests are widely used to assess cognitive function in dementia and evaluate word retrieval in stroke, typically within one minute. Although extending the test duration improves sensitivity to cognitive decline, 2-minute data in older Thai adults remain limited.

Objectives: This study aimed to evaluate the performance of older adults in Nonthaburi on a 2-minute verbal fluency test across animal, object, and food categories, reporting frequently listed words and exploring influential factors such as age, gender, and education.

Materials and Methods: This cross-sectional study recruited 147 healthy adults aged 60-89, categorized into three age groups: 60-69, 70-79, and 80-89. All participants were Central Thai speakers, had no history of neurological disorders, scored above 23 on the Thai Mental State Examination (TMSE), and underwent an oral reading of the Noo Jaew Passage and an oral motor examination by speech-language pathologists. Participants completed a 2-minute verbal fluency task in three categories: animal, object, and food. Responses were transcribed and analyzed using one-way ANOVA, independent t-tests, and regression analysis to examine the relationships between verbal fluency performance and relevant variables. Inter-rater reliability was assessed using the Intraclass Correlation Coefficient (ICC).

Results: Participants had an average age of 70.6 (SD=7.3) years, and 75% were females. The average TMSE score was 28.3 (SD=1.5), and the average years of education was 13.4 (SD=4.6). The 2-minute test yielded an average of 26.3 (SD=7.0) animals, 32.0 (SD=10.2) objects, and 24.2 (SD=7.3) foods. Significant differences were found across age groups ($p \leq 0.001$), with the 60-69 group outperforming older groups in the animal category. Gender influenced performance only in the food category. ICC values ranged from 0.982 to 0.997, indicating excellent inter-rater reliability.

Conclusion: This study reported the performance of older Thai adults on a 2-minute verbal fluency test, highlighting the effects of age, education, gender, and language-specific scoring. Frequently listed words may inform culturally relevant assessments and training materials. Future research should investigate alternative measures beyond word count to enhance cognitive assessments in clinical settings.

Introduction

Verbal fluency assessments are valuable tools for evaluating cognitive impairments associated with stroke and dementia. These tests measure the capacity for word retrieval in aphasia and cognitive functions, typically within one minute. Verbal fluency has two types: semantic fluency, which involves listing words that belong to a specific category, and phonemic fluency, which requires generating

* **Corresponding contributor.**
Author's Address: *Sirindhorn National Medical Rehabilitation Institute, Nonthaburi Province, Thailand.*
E-mail address: slp.nicha@gmail.com
doi: 10.12982/JAMS.2025.088
E-ISSN: 2539-6056

words that begin with a specified letter.[1] Phonemic fluency tests require literacy skills, whereas semantic fluency tests are typically less demanding but can be influenced by educational levels.[2] Previous studies on the 1-minute verbal fluency of Thai adults have reported an average of 17.3-19.4 words in the animal category,[2-5] 20.2 words in the object category, and 15.0 words in the food category.[4] Expanding beyond the animal category can enhance assessment options.[4]

Thailand has transitioned into an "Aging Society," with 19% of its 66 million population, or 13 million individuals, aged 60 years or older.[6] In Nonthaburi, 20.5% of the residents are seniors.[6] This aging population is at an increased risk for dementia, with a prevalence rate ranging from 2.4% in adults over 45[7] to 3.4-9.9% in those 60 years and older in Thailand.[8,9] These situations contribute to the need for more nuanced cognitive tests. Extending the administration time of the verbal fluency test to two minutes increases sensitivity to detect cognitive changes associated with aging. A longer duration allows for observing declines in word retrieval capabilities, offering a more comprehensive assessment.[10] Individuals with neurodegenerative diseases, such as those with amyotrophic lateral sclerosis (ALS), produced a significantly lower number of responses (p=0.008).[11] Cognitive decline is a common issue among individuals with various neurological conditions, and having a baseline for healthy older adults can facilitate early detection and intervention. There is no established data on 2-minute verbal fluency performance among older adults in Thailand.

This study aimed to address this gap by assessing performance on a 2-minute verbal fluency test involving animal, object, and food categories among older Nonthaburi individuals aged 60-69, 70-79, and 80-89. This study also aimed to identify factors such as age, gender, and education influencing word count. In line with previous research, this study hypothesized that verbal fluency performance might decline slightly with age, and that gender, education, and cognitive status could also be possible contributing factors to the number of words generated. These insights will assist in interpreting the results and planning targeted training. Moreover, the study reported frequently listed words by older Thai individuals. The collection of these words will help create assessment tools and treatment programs for Thai patients.

Materials and methods

This study employed a cross-sectional design and was reported following the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines. The research involved a group of healthy elderly individuals from Nonthaburi, Thailand. The sample size was calculated using the finite sample proportion method through the n4Studies application,[12] based on 2019 data from the Department of Older Persons, which reported a senior population of 239,410 in Nonthaburi.[13] The participants were organized into age groups: 60-69, 70-79, and 80-89, reflecting the demographic distribution.[14] The calculated sample size was 140, with an additional 5% added for potential data loss, resulting in 147 participants. The study included 147 participants, with 84 individuals aged 60-69, 45 aged 70-79, and 18 aged 80-89. Purposive sampling was utilized to select participants who were either healthy clients or caregivers at the Sirindhorn National Medical Rehabilitation Institute and elderly individuals engaged in activities at the Center for Older People's Quality of Life in Nonthaburi. Data were collected in person at these locations from October 2020 to July 2021. No participants withdrew from the study during data collection.

Inclusion and exclusion criteria: the study targeted healthy Thai individuals aged between 60 and 89 who spoke the Central Thai dialect. Eligibility criteria required that participants have no history of cerebral or neurological diseases, no severe visual or auditory impairments, and the ability to perform daily tasks independently. Participants were excluded if their TMSE (Thai Mental State Examination) score was ≤23,[15] if they demonstrated an inability to read or repeat the passage intelligibly, or if they failed the oral motor examination.

Participants were initially screened for dementia using the TMSE.[15] Visual and auditory functions were assessed by inquiring about any existing impairments and through practical tests to confirm their ability to see images and read text, which was particularly important during the TMSE and while reading the Noo Jaew Passage.[16] Auditory capabilities were evaluated using a finger-rubbing test[17] to ensure auditory clarity. The oral reading of the Noo Jaew Passage helped assess speech intelligibility; participants who were unable to read were asked to repeat the text after hearing it from an examiner. Oral motor examinations were conducted by speech-language pathologists (SLPs) to assess the functionality of speech-related organs.

All participants provided written informed consent before participating in the study. Screenings and assessments were carried out by researchers and speech-language pathologists (SLPs), ensuring that only participants with clear and intelligible speech were included in the study.

Participants were given two minutes per category (animal, object, and food) to generate as many words as possible without any cues. Categories were randomly assigned to avoid bias, and no examples were provided. The researchers used neutral nonverbal expressions, such as slight smiling or nodding, to encourage appropriate responses, without giving any corrective or negative feedback. Participants who paused or expressed difficulty were gently encouraged to continue. The researchers defined "food" as any edible, typically complete dish. Each valid and intelligible word within its category was scored once. Repeated words, words not in the target language, intrusions (words outside the intended category), and non-specific terms (e.g., "cooked food," "fried food," or "fish" without specifying how it was prepared) received no points. Variations of words (beginning or ending with the same word) were scored up to two times. For instance, "noodles" and "fish noodles" would result in a score only for "fish noodles." Similarly, "fried chicken," "fried fish," and "fried meat" would collectively score only twice. Synonyms such as "dog" and "canine" (or "สุนัข" and "หมา"

in Thai) or "TV" and "television" (or "ทีวี" and "โทรทัศน์" in Thai) were counted once.

Data analysis

The researchers transcribed the recordings and tallied the words produced. Inter-rater reliability was assessed on a randomly selected 10% of responses (N=14) using alphanumer cal codes. Sample size for estimating the intraclass correlation coefficient (ICC) was calculated using the n4Studies Plus application.[18,19] Based on a two-tailed test (expected agreement=0.95, minimum acceptable value=0.75, two raters, α=0.05, power=80%), the required sample size was 12; thus, the selected 14 participants provided sufficient power for reliability analysis. A third speech therapist, with 18 years of clinical experience and blinded to the initial scores, independently re-scored these recordings. ICCs were calculated with 95% confidence interval using a two-way random-effects model with an absolute agreement definition. Values less than 0.5 indicated poor reliability; values between 0.5 and 0.75 indicated moderate reliability; values between 0.75 and 0.9 indicated good reliability, and values greater than 0.90 indicated excellent reliability.[20]

Demographic data were presented as percentages, means, standard deviations, and 95% confidence intervals. A one-way ANOVA was conducted to compare means across the three age groups; Independent T-tests were used to compare means between genders; and multiple regression analysis was employed to investigate how age, TMSE score, gender, and education influenced word counts in each category. All statistical analyses were performed using SPSS 29.0.

Results

The study included 147 healthy older adults in Nonthaburi, 110 women and 37 men, 70 participants from the Center for Older People's Quality of Life, and 77 from Sirindhorn National Medical Rehabilitation Institute. The average TMSE score across the sample was 28.3 (SD=1.5). Statistical analysis using one-way ANOVA demonstrated significant differences in TMSE scores among the three age groups (p=0.017, η^2=0.055, medium effect), as detailed in Table 1. A Chi-square test revealed no significant gender distribution differences among age groups (p=0.096, Cramér's V=0.179, small effect). Bonferroni-adjusted post hoc comparisons following one-way ANOVA revealed that individuals aged 60-69 exhibited significantly higher TMSE scores than those aged 80-89 (p=0.024).

Table 1 Demographic data and average words within 2 minutes of each age group.

Variables	60-69 (N=84)	70-79 (N=45)	80-89 (N=18)	Total (N=147)	*p* value	Effect size
Age	65.1±2.7 (64.5, 65.7)	75.5±1.5 (75.0, 75.9)	84.1±1.9 (83.2, 85.1)	70.6 ± 7.3 (69.4, 71.8)	<0.001**	0.902
Female (%)	67 (80%)	33 (73%)	10 (56%)	110 (75%)	0.096	0.179
TMSE	28.5±1.4 (28.2, 28.8)	28.1±1.5 (27.6, 28.5)	27.5±1.8 (26.6, 28.4)	28.3±1.5 (28.0, 28.5)	0.017*	0.055
Education (years)	13.5±4.5 (12.6, 14.5)	13.6±4.4 (12.2, 14.9)	12.3±5.7 (9.5, 15.2)	13.4±4.6 (12.6, 14.1)	0.586	0.007
Animal	28.0±6.1 (26.6, 29.3)	24.8±6.7 (22.7, 26.8)	22.2±9.3 (17.5, 26.8)	26.3±7.0 (25.1, 27.4)	0.001**	0.090
Object	34.6±9.1 (32.6, 36.6)	31.0±10.4 (27.9, 34.1)	22.4±8.3 (18.3, 26.5)	32.0±10.2 (30.4, 33.7)	< 0.001**	0.152
Food	26.1±7.2 (24.5, 27.6)	23.1±5.4 (21.4, 24.7)	18.3±8.5 (14.1, 22.6)	24.2±7.3 (23.0, 25.4)	< 0.001**	0.124

*Note: Values are reported as mean±SD (95% CI lower, 95% CI upper), p values from one-way ANOVA and Chi-square test (Female), *p<0.05, **p<0.01, considered statistically significant. Effect sizes are reported as Eta squared (η^2; 0.01=small, 0.06=medium, 0.14=large), and Cramér's V for Female (0.10=small, 0.30=medium, 0.50=large).*[21]

Regarding education, the average length of education across the sample was 13.4 (SD=4.6), showing no significant differences between the age groups (p=0.586, η^2=0.007, small effect). Education levels among participants were as follows: one individual was illiterate with non-formal education; 17 individuals had completed primary education (4-6 years); 38 had completed secondary education (7-12 years); and 91 had completed tertiary education (13-21 years). Within the tertiary education category, 12 held diplomas, 50 held bachelor's degrees, 27 held master's degrees, and two had earned doctoral degrees.

The 2-minute verbal fluency test yielded an average of 26.3 (SD=7.0) words for animals, 32.0 (SD=10.2) words for objects, and 24.2 (SD=7.3) words for food. There were marked differences in performance across these categories among different age groups (p≤0.001, η^2=0.090-0.152, medium to large effects). Bonferroni-adjusted post hoc comparisons indicated that participants aged 60-69 produced significantly more animal words than those aged 70-79 and 80-89 (p=0.034, p=0.004). The 80-89 age group produced significantly fewer words for objects compared to the 60–69 and 70-79 groups (p<0.001, p=0.004), and for

food compared to the 60-69 and 70-79 groups (p<0.001, p=0.045), as presented in Table 2.

A multiple linear regression analysis exploring the factors influencing verbal fluency revealed that advancing age significantly predicted lower word counts across all categories (β=-0.213 to -0.277, p≤0.005), with adjusted R^2 values of 0.241 for animals, 0.402 for objects, and 0.312 for food. In contrast, higher TMSE scores were associated with increased verbal fluency in all categories (β=0.251 to 0.313, p≤0.002), and extended period of education was also significantly correlated with greater word counts in the animal, object, and food categories (β=0.215 to 0.323, p≤0.008), as shown in Table 3.

Table 2 Comparative analysis of age, TMSE, education, and verbal fluency between different age groups.

Variables	60-69 vs 70-79	60-69 vs 80-89	70-79 vs 80-89
Age	-10.4 (-11.4, -9.4) p<0.001**	-19.0 (-20.5, -17.6) p<0.001**	-8.6 (-10.2, -7.1) p<0.001**
TMSE	0.5 (-0.2, 1.1) p=0.269	1.0 (0.1, 2.0) p=0.024*	0.6 (-0.4, 1.6) p=0.521
Education (years)	0.0 (-2.1, 2.0) p=1.000	1.2 (-1.7, 4.1) p=0.965	1.2 (-1.9, 4.3) p=1.000
Animal	3.2 (0.2, 6.2) p=0.034*	5.8 (1.5, 10.1) p=0.004**	2.6 (-2.0, 7.2) p=0.516
Object	3.6 (-0.6, 7.8) p=0.119	12.2 (6.3, 18.1) p<0.001**	8.6 (2.3, 15.0) p=0.004**
Food	3.0 (-0.1, 6.1) p=0.062	7.7 (3.4, 12.1) p<0.001**	4.7 (0.1, 9.4) p=0.045*

*Note: Values are reported as mean difference (95% CI lower, 95% CI upper), p values were obtained from Bonferroni-adjusted post hoc comparisons following one-way ANOVA, *p<0.05, and **p<0.01, considered statistically significant.*

Table 3 Regression coefficients and significance levels of each verbal fluency (N=147).

Predictor	Animal Adjusted R^2=0.241	Object Adjusted R^2=0.402	Food Adjusted R^2=0.312
Age	B=-0.207 β=-0.213 p=0.005**	B=-0.374 β=-0.268 p<0.001**	B=-0.279 β=-0.277 p<0.001**
TMSE	B=1.310 β=0.283 p=0.001**	B=2.089 β=0.313 p<0.001**	B=1.206 β=0.251 p=0.002**
Education	B=0.329 β=0.215 p=0.008**	B=0.714 β=0.323 p<0.001**	B=0.383 β=0.241 p=0.002**
Gender	B=1.451 β=0.090 p=0.225	B=2.881 β=0.124 p=0.061	B=3.339 β=0.199 p=0.005**

*Note: Analyses were conducted using data from 147 participants. B: unstandardized regression coefficients β: standardized coefficients and corresponding p-values, *p<0.05 and **p<0.01, considered statistically significant, adjusted R^2: percentage of variance in the target variable explained by the predictors, a value of 1 reflects perfect prediction, whereas a value less than or equal to 0 indicates that the model has no predictive value.*

Gender did not significantly affect word count in the animal and object categories (p=0.225 and 0.061, respectively), but it did significantly impact the food category according to regression analysis (β=0.199, p=0.005). This finding was consistent with the results from the independent t-test, where gender differences were statistically significant in average age (p=0.043, d=0.387) and word count in the food category (p=0.008, d=- 0.515), as presented in Table 4.

For reliability, Table 5 shows that ICC ranged from 0.982 to 0.997, indicating excellent inter-rater reliability. Scores obtained from the examiner with less clinical experience strongly agreed with the senior assessors.

Table 4 Comparison of characteristics by gender.

Variables	Female (N=110)	Male (N=37)	*p* value	Cohen's *d*
Age	69.9±6.8 (68.7, 71.1)	72.7±8.3 (70.1, 75.3)	0.043*	0.387
TMSE	28.3±1.5 (28.0, 28.6)	28.2±1.5 (27.7, 28.7)	0.919	-0.019
Education	13.1±4.8 (12.2, 14.0)	14.3±3.9 (13.0, 15.5)	0.165	0.265
Animal	26.7±6.9 (25.4, 28.0)	25.0±7.3 (22.8, 27.4)	0.215	-0.237
Object	32.8±10.2 (30.8, 34.6)	29.7±9.8 (26.5, 33.2)	0.106	-0.309
Food	25.1±6.8 (23.9, 26.3)	21.4±8.2 (18.7, 24.2)	0.008**	-0.515

*Note: Values are reported as mean±SD (95% CI lower, 95% CI upper), p values were obtained from an independent T-test; *p<0.05 and **p<0.01, considered statistically significant, effect sizes are reported as d (Cohen's d; 0.2=small, 0.5=medium, 0.8=large).*[21]

Table 5 Inter-rater reliability.

(N=14)	Intraclass correlation[b]	95% Confidence interval	
		Lower bound	Upper bound
Animal	0.989	0.966	0.996
Object	0.997	0.992	0.999
Food	0.982	0.946	0.994

Note: [b]Type A intraclass correlation coefficients using a two-way random-effects model with absolute agreement definition.

Discussion

This study is the first known investigation of 2-minute verbal fluency among older Thai adults, focusing on animal, object, and food vocabulary. A literature review revealed no prior studies on word counts in the object or food categories in 2-minute fluency, with existing research limited to the animal category fluency.[2-5] Therefore, our comparisons are confined to 1-minute fluency within the same demographic[4] and 2-minute animal fluency across languages,[11,22,23] as shown in Tables 6-8.

Table 6 Comparison of 1-minute and 2-minute verbal fluency test results by age group and category.

Category	Age group	1-minute test	2-minute test	Difference
Animal	60-69 (N=84)	20.6±5.0 (19.5, 21.7)	28.0±6.1 (26.6, 29.3)	7.3±7.9 (5.6, 9.0)
	70-79 (N=45)	18.1±5.1 (16.6, 19.6)	24.8±6.7 (22.7, 26.8)	6.7±8.4 (4.2, 9.1)
	80-89 (N=18)	16.6±5.2 (14.2, 18.9)	22.2±9.3 (17.5, 26.8)	5.6±10.7 (0.7, 10.5)
	60-89 (N=147)	19.4±5.3 (18.5, 20.2)	26.3±7.0 (25.1, 27.4)	6.9±8.8 (5.5, 8.3)
Object	60-69 (N=84)	21.9±6.2 (20.5, 23.2)	34.6±9.1 (32.6, 36.6)	12.8±11.0 (10.4, 15.1)
	70-79 (N=45)	19.4±6.9 (17.4, 21.5)	31.0±10.4 (27.9, 34.1)	11.6±12.5 (7.9, 15.2)
	80-89 (N=18)	14.3±4.9 (12.0, 16.6)	22.4±8.3 (18.3, 26.5)	8.1±9.6 (3.7, 12.6)
	60-89 (N=147)	20.2±6.7 (19.1, 21.3)	32.0±10.2 (30.4, 33.7)	11.8±12.2 (9.9, 13.8)

Table 6 Comparison of 1-minute and 2-minute verbal fluency test results by age group and category. (Continue)

Category	Age group	1-minute test	2-minute test	Difference
Food	60-69 (N=84)	16.4±4.7 (15.3, 17.4)	26.1±7.2 (24.9, 27.6)	9.7±8.6 (7.9, 11.5)
	70-79 (N=45)	13.7±3.2 (12.8, 14.7)	23.1±5.4 (21.5, 24.7)	9.4±6.3 (7.5, 11.2)
	80-89 (N=18)	12.1±4.6 (10.0, 14.2)	18.3±8.5 (14.1, 22.6)	6.2±9.7 (1.8, 10.7)
	60-89 (N=147)	15.0±4.6 (14.3, 15.8)	24.2±7.3 (23.0, 25.4)	9.2±8.6 (7.8, 10.6)

Note: Values are reported as mean±SD (95% CI lower, 95% CI upper).

Table 7 Comparison of 1-minute and 2-minute verbal fluency test results by gender and category.

Category	Gender	1-Minute Test	2-Minute test	Difference
Animal	Female (N=110)	19.7±5.4 (18.7, 20.7)	26.7±6.9 (25.4, 28.0)	7.0±8.8 (5.4, 8.6)
	Male (N=37)	18.4±4.8 (16.8, 19.9)	25.0±7.3 (22.8, 27.4)	6.7±8.7 (3.9, 9.5)
	All (N=147)	19.4±5.3 (18.5, 20.2)	26.3±7.0 (25.1, 27.4)	6.9±8.8 (5.5, 8.3)
Object	Female (N=110)	20.7±7.0 (19.4, 21.9)	32.8±10.2 (30.8, 34.6)	9.5±8.2 (8.0, 11.1)
	Male (N=37)	18.8±5.8 (17.0, 20.7)	29.7±9.8 (26.5, 33.2)	8.1±9.2 (5.2, 11.1)
	All (N=147)	20.2±6.7 (19.1, 21.3)	32.0±10.2 (30.4, 33.7)	9.2±8.6 (7.8, 10.6)
Food	Female (N=110)	15.6±4.6 (14.7, 16.5)	25.1±6.8 (23.9, 26.3)	12.1±12.3 (9.8, 14.5)
	Male (N=37)	13.3±4.0 (12.0, 14.6)	21.4±8.2 (18.7, 24.2)	10.9±11.4 (7.2, 14.5)
	All (N=147)	15.0±4.6 (14.3, 15.8)	24.2±7.3 (23.0, 25.4)	11.8±12.2 (9.9, 13.8)

Note: Values are reported as mean±SD (95% CI lower, 95% CI upper).

Table 8 Comparison of 2-minute verbal fluency test results across studies and populations.

Study	Population	N	Age	Semantic verbal fluency (animals)	Language
This Study	Healthy	147	70.6±7.3 (69.4, 71.8)	26.3±7.0 (25.1, 27.4)	Thai
Barois *et al.*[22]	MS patients	68	52.3±12.1	24.6±8.4	French
	Healthy	33	49.4±9.6	34.9±7.5	French
Scholtissen *et al.*[23]	PD patients	25	66.4±10.6	29.7±8.1	Dutch
	Healthy	15	66.6±13.1	34.3±11.9	Dutch
Perez *et al.*[11]	ALS patients	42	Median 62 (IQR=15)	Median 25 (IQR=13)	Spanish
	Healthy	42	Median 62 (IQR=16)	Median 30.5 (IQR=13)	Spanish

Note: Values are reported as mean±SD or median, IQR: interquartile range, 75th percentile-25th percentile).

The analysis revealed excellent inter-rater reliability, with ICC values ranging from 0.982 to 0.997 across all categories and time intervals, as values above 0.90 indicate excellent agreement.[20] These findings align with those reported by Carnero-Pardo et al.,[24] who found an ICC of 0.96 for the semantic verbal fluency test. Although Woods *et al*.[25] used a computerized scoring system and reported test–retest rather than inter-rater reliability, the ICCs were lower: 0.77 for semantic and 0.91 for phonemic verbal fluency. The high inter-rater reliability observed in the present study may be attributed to scoring procedures and audio-recorded responses. This level of consistency supports the robustness and objectivity of the data used for further analysis.

Compared to the 1-minute results,[4] Table 6 shows that the 60-69 age group produced the highest word increase with the extra minute across all categories, with gains of 7.3 words for animal, 12.8 for object, and 9.7 for food. The 70-79 group also outperformed the 80–89 group in each category. When comparing by gender in Table 7, females consistently produced more words than males across all categories, with increases of 7.0 words for animal, 9.5 for object, and 12.1 for food.

Verbal fluency scores vary across languages and are influenced by age, education, cultural context, and scoring criteria.[4,26,27] In 1-minute verbal fluency Thai studies, findings from Charernboon and Suttichujit *et al*. are consistent, with both studies reporting an average of 19.4 words, with standard deviations of 5.0 and 5.3, respectively.[2,4] During the 1-minute tests, the 60-69 age group outperformed the 70-79 and 80-89 groups in the animal and food categories.[4] When extended to two minutes, only their animal category scores remained significantly higher. In contrast, the 80s group showed lower performance than the 60s and 70s groups in the object category during the 1-minute test, which further declined in the food category during the 2-minute test.

Extending the test duration to 2 minutes provides a more nuanced view of verbal fluency, capturing delays and discrepancies that may signal early cognitive decline in older adults. This additional time helps reveal variations in word production strategies across different categories. Regression analysis showed that while age is associated with reduced word output across all categories, higher TMSE scores and more education correlate with increased word production. Our study observed significant gender differences only in the food category, possibly due to cultural factors, as Thai women traditionally play a central role in food preparation.[28] Cultural influences, including gender roles and educational access, can impact verbal fluency across languages.

The rationale for using a 2-minute test to uncover cognitive challenges aligns with research on verbal fluency differences between patients and healthy controls, as shown in Table 8. Studies by Barois *et al*., Scholtissen *et al*., and Perez *et al*. have investigated how neurological conditions like multiple sclerosis (MS), Parkinson's disease (PD), and amyotrophic lateral sclerosis (ALS) impact verbal fluency.[11,22,23] Compared to healthy controls, ALS and MS patients produced significantly fewer responses (p=0.008; p<0.001).[11,22] MS patients also demonstrated significantly longer delays before the first word, slower production speeds, and extended inflection times (p<0.001), suggesting substantial executive and linguistic processing difficulties.[22] In contrast, PD patients showed no significant differences in total word count, switching, or clustering abilities (p=0.16, p=0.48, p=0.99, respectively), indicating that PD primarily affects motor switching and concept-shifting rather than the processes of cognitive switching needed in the fluency task.[23]

Compared with 2-minute results from other studies, Barois *et al*.[22] found a significant difference in semantic verbal fluency scores between French-speaking MS patients and healthy controls (p<0.001), and Perez *et al*.[11] also found a significant difference between Spanish-speaking ALS patients and healthy controls (p=0.008). In contrast, Scholtissen *et al*.[23] reported no significant difference between Dutch-speaking PD patients and healthy controls (p=0.16), as detailed in Table 8.

In this study, the lower word count compared to other languages likely stems from distinct scoring criteria and the structure of the Thai language, which encourages repetition. In the animal category, Thai vocabulary includes many subgroups, such as words beginning with ปลา (pla:; fish) and นก (nók; bird) or set phrases like หมูหมากาไก่ (mŭ: mă: ka: kàj; pig, dog, crow, chicken). Food names also follow similar patterns based on protein type, as seen with ก๋วยเตี๋ยว (kŭaj tĭ:aw; noodles) and ข้าวผัด (kʰâ:w pʰàt; fried rice) variations, making it easier to recall familiar patterns rather than unique items. Additionally, longer, complex dish names slow down recall. For instance, ข้าวหน้าปลาซาบะทอดซีอิ๊ว (kʰâ:w nâ:p la: sa: bàʔ tʰɔ̂:t si: ʔíw; grilled saba with soy sauce over rice, eight syllables) and ต้มจืดผักกาดขาวใส่เต้าหู้หลอด (tôm tɕù:t pʰàk kà:t kʰă:w sàj tâw hû: lɔ̀:t; clear soup with cabbage and tofu, nine syllables) require more time to articulate, limiting opportunities for additional responses.

The scoring rule of this study, which limits credit to two words per subgroup, further challenges participants by restricting patterned responses. This approach mainly affects categories like animal and food, where similar prefixes and subgroup patterns are common, whereas the object category shows fewer subgroup repetitions, allowing more unique responses. This scoring approach aligns with the criteria outlined in Olabarrieta-Landa et al.'s study.[29] Superordinate words (e.g., "fish") were allowed if specific examples from that category (e.g., "shark," "sardine") were not also included. Proper names, unrelated words, repeated words, and variations due to number, diminutives, or augmentations were excluded. In the animal category, words showing gender variations (e.g., "cow," "bull") and developmental stages of the same animal (e.g., "calf," "cow") were accepted. Additionally, extinct animals and mythical or magical creatures were allowed.[29]

Perez's and Scholtissen's studies[11,23] used Troyer *et al*.'s method,[30] which involves counting clusters and switching in verbal fluency tasks. In this approach, clusters are divided into subgroups, such as farm animals, pets,

aquatic animals, and insects, and switching is measured by tracking transitions between clusters to reflect participants' cognitive flexibility. In contrast, Barois's study did not measure clusters or switching but instead evaluated factors like first-word delay and inflection time to capture processing speed and executive function.[22]

Based on these findings, the following suggestions are proposed to support clinical decision-making regarding verbal fluency assessment. In patients with brain injuries, such as traumatic brain injuries or cerebrovascular accidents, a 1-minute verbal fluency test is generally sufficient. Extending the test to 2 minutes may cause stress for participants who experience word-finding difficulties. However, for adults without brain lesions, a 1-minute test may be insufficient to detect subtle cognitive impairments related to language function. In such cases, a 2-minute test serves as a more suitable tool. It is easy to administer, does not require specialized equipment or motor function, and is not time-consuming. Regarding category selection, the food category appears less influenced by participants' educational level, making it particularly suitable for individuals with low or no formal education.[4] However, men who do not typically engage in cooking activities may be disadvantaged by this category and might perform better with object-related tasks. Moreover, repeated administration of the animal category may result in task familiarity or learning effects. Therefore, incorporating alternative categories can help reduce potential bias and support a more accurate assessment result. The collected word lists may help develop culturally appropriate tools for assessment and intervention in Thai adults.

Limitation

This study has several limitations. The sample included relatively few participants aged 80 and above, and most participants were from Nonthaburi province, which is like Bangkok, but may not fully represent the broader Thai older adult population. These factors may limit the generalizability of the findings to other sociocultural contexts. Therefore, variables such as age, education, and cultural background should be considered when interpreting verbal fluency performance.

To enhance representativeness and external validity, future research should recruit participants from diverse regions, age groups, and socioeconomic backgrounds. There is also a need to develop a comprehensive Thai verbal fluency assessment that incorporates both semantic and phonemic components. Although the F-A-S format is widely adopted for phonemic verbal fluency testing in English,[31,32] Thai-language studies have thus far been limited to one-minute tasks using consonants such as /k/ ("ก")[33] and /s/ ("ส")[3] and evidence on 2-minute phonemic fluency still lacking. Moreover, the current study focused solely on category-based semantic fluency; phonemic fluency and more nuanced performance features such as clustering and switching were not assessed. Furthermore, additional performance features, including clustering (grouping words into subcategories), switching (shifting between clusters), intrusions, and perseverations, should be explored to capture the full complexity of verbal fluency and improve its diagnostic utility in both clinical and research contexts.

Conclusion

This study provides valuable insights into 2-minute verbal fluency among older Thai adults, particularly in the unique linguistic and cultural contexts of animal, object, and food vocabulary. Our findings highlight the effects of age, education, and gender on verbal fluency performance and the role of language-specific scoring criteria that influence response diversity, especially in categories like animal and food. Extending the fluency test to 2 minutes proved effective in identifying variations in cognitive processing and early signs of cognitive decline, suggesting its potential as an early detection tool. Comparisons with previous research also emphasize the need for flexible, culturally relevant scoring systems, given how linguistic structure impacts fluency performance.

Ethical approval

This study received ethical approval from the Sirindhorn National Medical Rehabilitation Institute Human Ethics Committee in Nonthaburi Province, Thailand (Approval No. 63017).

Funding

This research was supported by the Pay for Performance (P4P) of Sirindhorn National Medical Rehabiitation Institute in 2022.

Conflict of interest

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

CReDIT authorship contribution statement

Nicha Kripanan: conceptualization, methodology, validation, formal analysis, investigation, data curation, writing - original draft, visualization, project administration; **Somjit Ruamsuk:** investigation, resources, writing - review and editing, supervision, funding acquisition; **Thanwarat Artayakul:** validation, data curation, writing - review and editing, visualization; **Isara Suttichujit:** investigation, data curation, funding acquisition.

Acknowledgements

The authors would like to express their sincere gratitude to the institute's staff for their support and assistance throughout the research process. We also appreciate the cooperation and participation of all study participants and their families. The contents of this study are solely the responsibility of the authors and do not necessarily represent the official views of Sirindhorn National Medical Rehabilitation Institute.

References

[1] Whiteside DM, Kealey T, Semla M, Luu H, Rice L, Basso MR, et al. Verbal Fluency: Language or Executive Function Measure? Appl Neuropsychol Adult. 2016; 23(1): 29-34. doi: 10.1080/23279095.2015.1004574.

[2] Charernboon T. Verbal fluency in the Thai elderly with mild cognitive impairment and elderly with dementia. J Ment Health Thai. 2018; 26(2): 91-102.
[3] Muangpaisan W, Intalapaporn S, Assantachai P. Digit span and verbal fluency tests in patients with mild cognitive impairment and normal subjects in Thai-community. J Med Assoc Thai. 2010; 93(2): 224-30. PMID: 20302005.
[4] Suttichujit I, Kripanan N, Ruamsuk S. The Study on Verbal Fluency in Older Adults in Nonthaburi Province. J Assoc Med Sci. 2024; 57(3): 230-6. doi: 10.12982JAMS.2024.065.
[5] Teerapong W. The comparison of language abilities of Thai aphasic patients and Thai normal subjects by using Thai Adaptation of Western Aphasia Battery [Thesis]. Nakhon Pathom: Mahidol University; 2000.
[6] Foundation of Thai Gerontology Research and Development Institute (TGRI). Situation of the Thai Older Persons 2022. Nakhon Pathom: Institute for Population and Social Research, Mahidol University; 2022.
[7] Wangtongkum S, Sucharitkul P, Silprasert N, Inthrachak R. Prevalence of dementia among population age over 45 years in Chiang Mai, Thailand. J Med Assoc Thai. 2008; 91(11): 1685-90. PMID: 19127790.
[8] Jitapunkul S, Kunanusont C, Phoolcharoen W, Suriyawongpaisal P. Prevalence estimation of dementia among Thai elderly: a national survey. J Med Assoc Thai. 2001; 84(4): 461-7. PMID: 11460954.
[9] Senanarong V, Poungvarin N, Sukhatunga K, Prayoonwiwat N, Chaisewikul R, Petchurai R, et al. Cognitive status in the community dwelling Thai elderly. J Med Assoc Thai. 2001; 84(3): 408-16. PMID: 11460944.
[10] Holtzer R, Goldin Y, Donovick PJ. Extending the administration time of the letter fluency testincreases sensitivity to cognitive status in aging. Exp Aging Res. 2009; 35(3): 317-26. doi: 10.1080/03610730902922119. PMID: 19449244; PMCID: PMC2832286.
[11] Perez M, Amayra I, Lazaro E, García M, Martínez O, Caballero P, Berrocoso S, López-Paz JF, Al-Rashaida M, Rodríguez AA, Luna P, Varona L. Intrusion errors during verbal fluency task in amyotrophic lateral sclerosis. PLoS One. 2020; 15(5): e0233349. doi: 10.1371/journal.pone.0233349.
[12] Ngamjarus C. n4Studies: Sample size calculation for an epidemiological study on a smart device. Siriraj Med J. 2016; 68(3): 160-70.
[13] Statistics of Elderly Person in Thailand 2019 [Internet]. 2019. Available from: www.dop.go.th/download/knowledge/th1580099938-275_1.pdf.
[14] Department of Older Persons. Situation of The Thai Elderly 2019. Nakhon Pathom: Institute for Population and Social Research, Mahidol University, and Foundation of Thai Gerontology Research and Development Institute (TGRI); 2019. Available from: https://ipsr.mahidol.ac.th/wp-content/uploads/2022/03/Report-File-613.pdf.
[15] Train the Brain Forum Committee. Thai mental state examination (TMSE). Siriraj Hosp Gaz. 1993; 45: 661-74.
[16] Sindermsuk D. The survey of speech defects among Prathom 4 students in Mitsampan school group [Thesis]. Nakhon Pathom: Mahidol University; 1986.
[17] Torres-Russotto D, Landau WM, Harding GW, Bohne BA, Sun K, Sinatra PM. Calibrated finger rubs auditory screening test (CALFRAST). Neurology. 2009; 72(18): 1595-600. doi: 10.1212/WNL.0b013e3181a41280.
[18] Ngamjarus C, Pattanittum P. n4Studies Plus: application for sample size calculation in health science research. App Store; 2024. Available from: https://he02.tci-thaijo.org/index.php/sirirajmedj/article/view/58342.
[19] Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. Stat Med. 2012; 31(29): 3972-81. doi: 10.1002/sim.5466. Epub 2012 Jul 4. PMID: 22764084.
[20] 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: 10.1016/j.jcm.2017.10.001.
[21] Cohen J. Statistical Power Analysis for the Behavioral Science. 2nd Ed. New York: Routledge; 1988.
[22[Barois E, Sagawa Y, Yilmaz S, Magnin E, Decavel P. What (more) can verbal fluency tell us about multiple sclerosis? Ann Phys Rehabil Med. 2021; 64(2): 101394. doi: 10.1016/j.rehab.2020.05.002.
[23] Scholtissen B, Dijkstra J, Reithler J, Leentjens AF. Verbal fluency in Parkinson's disease: results of a 2-min fluency test. Acta Neuropsychiatr. 2006; 18(1): 38-41. doi: 10.1111/j.0924-2708.2006.00122.x.
[24] Carnero-Pardo C, Gurpegui M, Sánchez-Cantalejo E, Camello AG, Santervás PG, Ríos MT. [Reliability assessment of the EUROTEST]. Neurologia. 2007; 22(3): 153-8. doi: 10.1016/j.nrl.2010.09.021.
[25] Woods DL, Wyma JM, Herron TJ, Yund EW. Computerized Analysis of Verbal Fluency: Normative Data and the Effects of Repeated Testing, Simulated Malingering, and Traumatic Brain Injury. PLoS One. 2016; 11(12): e0166439. doi: 10.1371/journal.pone.0166439.
[26] Ardila A. A cross-linguistic comparison of category verbal fluency test (ANIMALS): a systematic review. Arch Clin Neuropsychol. 2020;35(2):213-25. doi: 10.1093/arclin/acz060.
[27] Villalobos D, Torres-Simón L, Pacios J, Paúl N, Del Río D. A Systematic Review of Normative Data for Verbal Fluency Test in Different Languages. Neuropsychol Rev. 2023; 33(4): 733-64. doi: 10.1007/s11065-022-09549-0.
[28] Chandrangam K. The relationship between gender and women's self-value in Bangkok metropolis [Thesis]. Bangkok: Chulalongkorn University; 2005.
[29] Olabarrieta-Landa L, Torre EL, López-Mugartza JC, Bialystok E, Arango-Lasprilla JC. Verbal fluency tests: developing a new model of administration and scoring for Spanish language. NeuroRehabilitation. 2017; 41(2): 539-565. doi: 10.3233/NRE-162102.
[30] Troyer AK, Moscovitch M, Winocur G, Alexander MP, Stuss D. Clustering and switching on verbal fluency:

the effects of focal frontal- and temporal-lobe lesions. Neuropsychologia. 1998; 36(6): 499-504. doi: 10.1016/s0028-3932(97)00152-8.

[31] Tombaugh TN, Kozak J, Rees L. Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. Arch Clin Neuropsychol. 1999; 14(2): 167-77. PMID: 14590600.

[32] Machado TH, Fichman HC, Santos EL, Carvalho VA, Fialho PP, Koenig AM, et al. Normative data for healthy elderly on the phonemic verbal fluency task - FAS. Dement Neuropsychol. 2009; 3(1): 55-60. doi: 10.1590/S1980-57642009DN30100011.

[33] Metarugcheep S, Punyabukkana P, Wanvarie D, Hemrungrojn S, Chunharas C, Pratanwanich PN. Selecting the most important features for predicting mild cognitive impairment from Thai Verbal Fluency Assessments. Sensors. 2022; 22(15): 5813. doi: 10.3390/s22155813.

Appendix

Frequently listed words from the 2-minute verbal fluency test results from all participants

This table presents the 30 most frequently listed words from a verbal fluency task, categorized into three semantic groups: Object, Animal, and Food. Each entry includes the Thai word, its International Phonetic Alphabet (IPA) transcription, the English translation, and the frequency count based on participant responses. In cases where multiple words shared the same frequency at the final rank (rank 30), those words are presented together under the same ranking.

	Object		**Animal**		**Food**	
1	โต๊ะ /tó?/ Table	147	ช้าง /tɕʰáːŋ/ Elephant	137	ก๋วยเตี๋ยว /kŭaj tĭːaw/ Noodles	141
2	รถยนต์ /rót jon/ Car	134	ควาย /kʰwaːj/ Buffalo	130	แกงจืด/ต้มจืด /kɛːŋ tɕɯ̀ːt/, /tôm tɕɯ̀ːt/ Clear soup	130
3	เก้าอี้ /kâw ʔîː/ Chair	125	วัว /wuːa/ Cow	130	ข้าวผัด /kʰâːw pʰàt/ Fried rice	106
4	รองเท้า /rɔːŋ tʰáːw/ Shoes	111	หมา /mă:/ Dog	129	ผัดกะเพรา /pʰàt kàʔ pʰraw/ Stir fried with basil	101
5	ช้อน /tɕʰɔ́ːn/ Spoon	107	แมว /mɛːw/ Cat	123	ต้มยำ /tôm jam/ Tom yum soup	97
6	เตียง /tiːaŋ/ Bed	101	ไก่ /kàj/ Chicken	115	แกงส้ม /kɛːŋ sôm/ Sour curry	95
7	พัดลม /pʰát lom/ Fan	94	ม้า /máː/ Horse	112	ยำ /jam/ Spicy salad	94
8	กางเกง /kaːŋ keːŋ/ Pants	93	ปลา /plaː/ Fish	110	แกงเขียวหวาน /kɛːŋ kʰĭːaw wăːn/ Green curry	89
9	ปากกา /pàːk kaː/ Pen	91	นก /nók/ Bird	107	น้ำพริก /nám pʰrík/ Chili paste	82
10	กระทะ /kràʔ tʰáʔ/ Pan	88	เสือ /sŭːa/ Tiger	105	ปลาทอด /plaː tʰɔ̂ːt/ Fried fish	69
11	ดินสอ /din sɔ̌ː/ Pencil	85	หมู /mŭː/ Pig	101	แกงเผ็ด /kɛːŋ pʰèt/ Spicy curry	63
12	โทรศัพท์ /tʰoː rá sàp/ Phone	82	งู /ŋuː/ Snake	97	ผัดผัก /pʰàt pʰàk/ Stir-fried vegetables	63
13	หม้อ /mɔ̂ː/ Pot	81	ลิง /liŋ/ Monkey	96	ราดหน้า /râːt nâː/ Stir-fried noodles in gravy sauce	62
14	นาฬิกา /naː líʔ kaː/ Clock / Watch	80	สิงโต /sĭŋ toː/ Lion	95	ผัดเผ็ด /pʰàt pʰèt/ Spicy stir-fried curry	57

	Object		Animal		Food	
15	เสื้อ /sɯ̂ːa/ Shirt	80	ยีราฟ /jiːráːp/ Giraffe	88	ลาบ /lâːp/ Spicy minced meat salad	57
16	กระเป๋า /kràʔ păw/ Bag	79	จระเข้ /tɕɔː ráʔ kʰêː/ Crocodile	86	ไข่เจียว /kʰàj tɕiːaw/ Omelet	54
17	ทีวี /tʰiː wiː/ TV	78	เป็ด /pèt/ Duck	76	ข้าวเหนียว /kʰâːw nǐːaw/ Sticky rice	54
18	ชาม /tɕʰaːm/ Bowl	74	กระต่าย /kràʔ tàːj/ Rabbit	67	ส้มตำ /sôm tam/ Papaya salad	52
19	ตู้เย็น /tûː jen/ Refrigerator	74	หนู /nǔː/ Rat	64	ขนมจีน /kʰà nɔ̌m tɕiːn/ Fermented rice noodles	50
20	จาน /tɕaːn/ Plate	72	กระรอก /kràʔ rɔ̂ːk/ Squirrel	61	ผัดไทย /pʰàt tʰaj/ Pad Thai	48
21	แก้ว /kêːw/ Glass	70	จิ้งจก /tɕîŋ tɕòk/ Lizard	60	แกงมัสมั่น /kɛːŋ mát sà màn/ Massaman curry	42
22	ตู้ /tûː/ Cabinet	70	แรด /rɛ̂ːt/ Rhinoceros	57	ผัดซีอิ๊ว /pʰàt siː ʔíw/ Stir-fried noodles in soy sauce	42
23	เสื้อผ้า /sɯ̂ːa pʰâː/ Clothes	63	เต่า /tàw/ Turtle	55	แกงไก่ /kɛːŋ kàj/ Chicken curry	39
24	ถ้วย /tʰûaj/ Cup	62	ปลาวาฬ /plaː waːn/ Whale	53	แกงเลียง /kɛːŋ liːaŋ/ Herbal mixed vegetable soup	39
25	ถุงเท้า /tʰŭŋ tʰáːw/ Socks	60	ม้าลาย /máː laːj/ Zebra	53	สลัด /sà làt/ Salad	38
26	แว่นตา /wɛ̂ːn taː/ Glasses	59	กวาง /kwaːŋ/ Deer	49	ไก่ย่าง /kàj jâːŋ/ Grilled chicken	37
27	หมอน /mɔ̌ːn/ Pillow	59	ชะนี /tɕʰáʔ niː/ Gibbon	46	สเต็ก /sà tèk/ Steak	37
28	ขวด /kʰùːat/ Bottle	54	อิปโปฯ /híp poː/ Hippopotamus	46	บะหมี่ /bàʔ mìː/ Egg noodles	36
29	ส้อม /sɔ̂ːm/ Fork	53	ตุ๊กแก /túk kɛː/ Gecko	45	ข้าวต้ม /kʰâːw tôm/ Boiled rice	34

	Object		Animal		Food	
30	กะละมัง /kàʔ láʔ maŋ/ Basin	50	หมี /mǐː/ Bear	42	ทองหยอด /tʰɔːŋ jɔ̀ːt/ Golden egg-yolk drops	33
	สบู่ /sà bùː/ Soap	50				
	แอร์ /ʔɛː/ Air conditioner	50				
	ยาสีฟัน /jaː sǐː fan/ Toothpaste	50				

Note: IPA transcriptions were from thai-language.com; English dish names were cross-referenced and adapted from thaifoodmaster.com.

Journal of Associated Medical Sciences 2025; 58(3): 204-215

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

Translation, cross-cultural adaptation and psychometric properties of the Thai version of the Trunk Control Measurement Scale for children with cerebral palsy

Arisa Parameyong*, Jitapa Chawawisuttikool, Sirinun Boripuntakul, Sauwaluk Dacha

Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 9 January 2025
Accepted as revised 1 July 2025
Available online 7 July 2025

Keywords:
Clinical assessment, trunk control, motor disabilities, trunk impairment, pediatric health.

ABSTRACT

Background: Trunk impairment is an important clinical feature in children with cerebral palsy (CP), therefore assessment of trunk control is a crucial part of functional assessment in children with CP.

Objectives: The study aimed to translate the Trunk Control Measurement Scale (TCMS) from the original English into a Thai version and demonstrate the reliability and validity of the TCMS-TH among children with spastic CP.

Materials and methods: The TCMS was translated using standard cross-cultural adaptation guidelines. The final TCMS-TH was tested on 29 children with spastic CP aged 5-15 years. For inter-rater reliability, two assessors independently scored the children from the video recording using the TCMS-TH. One month after the first assessment, a second assessment was performed to determine intra-rater reliability. The convergent validity of the TCMS-TH was assessed by comparison to the Gross Motor Function Measure (GMFM-88). The intraclass correlation coefficient (ICC) and Pearson's correlation coefficient were used to analyze reliability and validity.

Results: The TCMS-TH had good internal consistency, with Cronbach's alpha = 0.916 for the total TCMS-TH. Inter-rater reliability was good to excellent, ICC = 0.837-0.955, whereas intra-rater reliability was excellent, ICC =0.918-0.995. The TCMS-TH and the GMFM has moderate to high correlation (r=0.631-0.810).

Conclusion: The TCMS-TH is a reliable and valid tool for assessing trunk control in children with spastic CP. However, its outcomes are primarily generalizable to observations obtained from video recordings rather than direct in-person assessments. Despite this limitation, the TCMS-TH remains a valuable tool for both clinical practice and research settings.

Introduction

Cerebral palsy (CP) is defined as a group of disorders caused by damage to the immature brain after birth or in the developing fetus. This is a non-progressive pathological condition characterized by abnormal posture, loss of selective motor control, uncoordinated movement and postural control dysfunction that significantly interferes with movement development and daily living activities.[1-3] The main problem associated with motor disabilities in children with CP is thought to be a lack of postural control.[4,5] The trunk is an important body segment associated with postural control and has been recognized as a primary factor for daily activities such as upper extremity function, gross motor skills, and self-care.[6,7] At around 9 months old, typically developing children can

* **Corresponding contributor.**
Author's Address: *Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: arisa.p@cmu.ac.th
doi: 10.12982/JAMS.2025.089
E-ISSN: 2539-6056

maintain head and trunk posture by resisting gravity. However, previous studies reported that children with CP have delayed sitting and gross motor function due to impaired trunk control.[8] Moreover, children with CP have increased trunk sway during sitting[9] and standing[10] when compared to typically developing children, reflecting poor trunk control during upright positions. Heyrman *et al*.[11] found that trunk control in sitting has a relationship with lower limb movement during gait in children with spastic diplegia CP. Additionally, Mendoza *et al*.[12] demonstrated that there was a statistically significant relationship between sitting ability and the capacity of gross motor function. Therefore, assessment of trunk performance in children with CP is necessary for clinical evaluation.

To date, clinical tools used to evaluate clinical features related to trunk control impairment among children with CP in Thailand are limited. The Gross Motor Function Measure (GMFM) is a standardized assessment tool commonly used in the clinical evaluation of children with CP.[13,14] This four-point scale consists of 88 items divided into five dimensions of gross motor function, but it is not designed to measure trunk control. The most common assessment tool specifically for measuring trunk control in children with CP is the Segmental Assessment of Trunk Control (SATCo).[15,16] Despite its good reliability and validity in children with CP, the SATCo has limitations in that it only measures static trunk control and has a ceiling effect in the independent sitting child. The Trunk Control Measurement Scale (TCMS) was developed by Heyrman *et al*. in 2011[17] and was adapted from the Trunk Impairment Scale (TIS).[18] The TIS was primarily used in stroke patients to assess trunk control in static and dynamic measurements and showed good psychometric properties in clinical and research settings. Appropriate modification and additions to the TIS were made according to the clinical features of impaired trunk control in children with CP. Heyrman *et al*.[17] indicated that the TCMS has good reliability in children with spastic CP, with intraclass correlation coefficient (ICC) values of 0.91-0.99 for inter-rater and test-retest reliability. In addition to the convergent validity found in Dimensions B-E of the GMFM, the Spearman rank correlation coefficient for convergent validity showed acceptable values of 0.6-0.87 between the TCMS and the GMFM. TCMS is a widely recognized clinical tool designed to assess trunk control in CP. Since its initial development, the TCMS has been translated into multiple languages, including Korean,[19] German,[20] and Spanish,[21] to facilitate its use across diverse clinical settings. Currently, few clinical tools are available for assessing trunk impairment among children with CP in Thailand, and no prior studies have translated a Thai version of the TCMS.

This study aimed to translate a Thai version of the TCMS (TCMS-TH) and provide a beneficial tool to assess trunk control in clinical and research settings among Thai children with CP. The TCMS was translated into the Thai language as a cross-cultural adaptation. In addition, the psychometric properties of the TCMS-TH were investigated in Thai children with CP.

Materials and methods

Trunk Control Measurement Scale (TCMS)

The TCMS is a 15-item clinical assessment tool designed to measure trunk control during functional activities in children with CP. The TCMS consists of two main components for trunk control: static and dynamic sitting balance. Static sitting balance measures trunk control during movement of the upper and lower limbs. However, dynamic sitting balance consists of two subsets: selective movement control and dynamic reaching. Selective movement control measures selective movement of the trunk in three planes (frontal, sagittal and transverse), whereas dynamic reaching measures active trunk movement beyond the base of support during reaching activities. All items are scored on a two-, three- or four-point ordinal scale and bilaterally evaluated in relation to clinical activities. The total score of the TCMS is 0-58, with a higher score indicating better performance.

Translation and cross-cultural adaptation process

Permission to translate the original TCMS was obtained from the test developer[17], ensuring compliance with authorized use and adherence to copyright requirements. The translation process of the TCMS from the original English to the Thai language was performed according to the recommended guidelines.[22] First, two native Thai speakers independently translated the original English into Thai: one had 10 years of experience as a pediatric physical therapist and the other was an English teacher with no medical background. Then, the two drafts of the TCMS-TH were merged into one after discussion between the translators and the research teams. This consensus TCMS-TH version was then translated back into English by two translators who were bilingual in English and Thai but had no medical background. Finally, all the translators and research teams met to discuss and finalize the TCMS-TH. Next, the content validity of the TCMS-TH was approved by analyzing the item-objective congruence (IOC) from the three expert reviewers. The items that had IOC equal to or higher than 0.5 were acceptable.[23] The expert reviewers included two physical therapists who had experience as lecturers in the pediatric field at leading universities for 20 years and one pediatric physical therapist who had experience in evaluating and providing treatment to children with CP for more than 10 years.

Participants

The sample size was calculated using G*Power with the following parameters: effect size=0.5, α=0.05, and power = 0.8. Based on this calculation, a minimum of 29 participants was required. Participants were children aged 5-15 years diagnosed with spastic CP from a children's treatment center and special education school. All participants can sit independently for at least 30 minutes and comprehend the test instructions. A total of 40 children were initially recruited but 11 were excluded due to uncorrected visual problems and other diagnoses such as autism. Therefore, eligible participants were 29 children (14 males, 15 females) with a mean age of 11.3 years: 20

with spastic bilateral CP and 9 with spastic unilateral CP. The Gross Motor Function Classification System (GMFCS) ranges from Level I to Level IV. Two children were classified as Level I, seven as Level II, thirteen as Level III and seven as Level IV. The characteristics of the participants are presented in Table 1.

Table 1. Characteristics of the participants.

Characteristics	Mean (SD)	Range
Age (years)	11.3 (2.9)	5-15
Height (cm)	130.5 (12.3)	112-157
Weight (kg)	31.4 (10.2)	15-51.4
Gender*		
Male	14 (48.3)	
Female	15 (51.7)	
GMFCS level*		
I	2 (6.9)	
II	7 (24.1)	
III	13 (44.8)	
IV	7 (24.2)	
CP distribution*		
Bilateral	20 (69)	
Unilateral	9 (31)	

*Note: GMFCS: Gross Motor Function Classification System, CP: cerebral palsy, *data presented as number (%).*

This study was approved by the Human Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University, Thailand (AMSEC-65FB-001). A parent or guardian of each child provided written informed consent, and all children provided verbal assent before participation in this study.

Reliability study

Assessment of the TCMS-TH in all participants was administered by a pediatric physical therapist in a quiet room. For the testing protocol, a physical therapist verbally explained and demonstrated the movement in each item with respect to the original version. After that, participants were allowed to perform the test with manual guidance and then were requested to execute the test independently. Testing lasted 20-30 minutes, during which a resting period was allowed to prevent fatigue. The testing of each item was recorded on video in the sagittal and coronal planes. The videos were kept secret and deleted after completion of the study.

For inter-rater reliability, two pediatric physical therapists with at least 10 years of experience were trained to score the TCMS-TH by discussing and watching an instructional video with the research team before assessment. Then, two assessors independently used the video recording to score the TCMS-TH for each participant. To minimize the recall memory of the scoring data, a second assessment was performed 1 month after the first assessment to assess intra-rater reliability.

Convergent validity study

The TCMS-TH was validated with the GMFM-88 version, a standardized functional measurement tool used by physical therapists or rehabilitation specialists to measure the quantified change in gross motor function of children with CP aged 5 months to 16 years.[24] The GMFM-88 is composed of five dimensions: lying and rolling (17 items; 51 points); sitting (20 items; 60 points); crawling and kneeling (14 items; 42 points); standing (13 items; 39 points); and walking, running and jumping (24 items; 72 points). The GMFM is a criterion-referenced tool, with items scored on a four-point ordinal scale (0 = cannot initiate; 1 = initiates; 2 = partially completes item; 3 = completes item independently). The total score from the five dimensions is 264 points. Higher scores indicated greater capacity of gross motor function.[25] The estimated time required to complete the GMFM-88 test was approximately 45 minutes per participant. In the present study, the assessor also received training for use of the GMFM before scoring. Assessment of the GMFM for each participant was performed on the same day after a 1-hour rest from the TCMS video recorded in the reliability study.

Statistical analysis

The data were normally distributed, as assessed by the Shapiro-Wilk test. Cronbach's alpha was used to assess the internal consistency of the total and subscales. Inter-rater and intra-rater reliability for total and subscales were analyzed by ICC. The ICC(2,k) model was used for inter-rater reliability and the ICC(3,k) model for intra-rater reliability.

ICC values of >0.9 indicate excellent reliability, 0.70-0.89 as good, 0.5-0.69 as moderate and <0.5 indicates poor reliability.[26] The standard error of measurement (SEM) and minimal detectable change (MDC) were calculated as $SEM = SD \times \sqrt{1 - ICC}$ and $MDC = SEM \times 1.96 \times \sqrt{2}$. Pearson's correlation coefficient was used to determine the convergent validity between the TCMS-TH and percent score of the GMFM. Interpreted correlation values of > 0.9 indicate a very high correlation, 0.7–0.89 as high, 0.5-0.69 as moderate and <0.5 indicates low correlation.[27] All data were analyzed using SPSS version 17 and the significance level was set at $p<0.05$.

Results

Translation process

During the translation process, the expert committees suggested that some sentences of the TCMS-TH version quite complicated to understand, especially for people who have no medical background. Therefore, a revision was made before the psychometric property studies. The average IOC value was 0.84. For each item, the IOC was between 0.67 and 1.00, which indicated good content validity. The IOC of the TCMS-TH version is presented in Table 2.

Reliability

Internal consistency

Cronbach's alpha was 0.916 for the total TCMS-TH score, which indicated excellent internal consistency. For Cronbach's alpha of subscales, the static sitting balance was 0.702, selective movement control was 0.883 and the dynamic reaching subscale was 0.804, suggesting acceptable and good internal consistency (Table 3).

Table 2. Item-objective congruence (IOC) of each item of the Thai version of the Trunk Control Measurement Scale (TCMS-TH).

TCMS-TH items	IOC	TCMS-TH items	IOC
Item 1	1.00	Item 9a	1.00
Item 2	1.00	Item 9b	0.67
Item 3	1.00	Item 9c	1.00
Item 4	1.00	Item 10a	0.67
Item 5	1.00	Item 10b	0.67
Item 6a	1.00	Item 11a	0.67
Item 6b	1.00	Item 11b	0.67
Item 7a	1.00	Item 12a	0.67
Item 7b	1.00	Item 12b	0.67
Item 8a	0.67	Item 13	0.67
Item 8b	0.67	Item 14	0.67
Item 8c	1.00	Item 15	0.67
Total TCMS	0.84		

Table 3. Score and internal consistency of the Thai version of the Trunk Control Measurement Scale (TCMS-TH).

TCMS-TH	Mean (SD)	Minimum	Maximum	Cronbach's alpha
Static sitting balance (/20)	13.33 (4.97)	3	20	0.702
Selective movement control (/28)	12.18 (6.74)	2	22	0.883
Dynamic reaching (/10)	3.96 (3.24)	0	5	0.804
Total TCMS (/58)	27.48 (14.03)	6	46	0.916

Intra-rater and inter-rater reliability

For the intra-rater reliability, the ICC values showed excellent reliability, with ICC=0.995 (95% CI=0.989-0.998; $p<0.05$) for the total scores of the TCMS-TH. The ICC of the subscales ranged from 0.918 to 0.990. The SEM ranged from 0.49 to 0.82 and the MDC ranged from 0.70 to 1.16 for total scores of the TCMS-TH, static sitting balances, selective movement control and dynamic reaching.

For inter-rater reliability, the ICC values showed excellent reliability except for the dynamic reaching subscale, which showed good reliability. The ICC value was 0.955 (95% CI=0.901-0.980; $p<0.05$) for the total scores of the TCMS-TH. The ICC of the subscales ranged from 0.837 to 0.940. The SEM ranged from 0.66 to 1.96 and the MDC ranged from 0.94 to 2.77 for the total scores of the TCMS-TH, static sitting balances, selective movement control and dynamic reaching. The intra-rater, inter-rater reliability, SEM and MDC are presented in Table 4.

Table 4. Inter- and intra-rater reliability of the Thai version of the Trunk Control Measurement Scale (TCMS-TH).

TCMS-TH	Inter-rater reliability (95% CI)	SEM	MDC	Intra-rater reliability (95% CI)	SEM	MDC
Static sitting balance (/20)	0.929 (0.843-0.968)	0.66	0.94	0.989 (0.976-0.995)	0.49	0.70
Selective movement control (/28)	0.940 (0.865-0.973)	1.28	1.82	0.990 (0.977-0.995)	0.56	0.79
Dynamic reaching (/10)	0.837 (0.637-0.927)	0.85	1.20	0.918 (0.818-0.963)	0.57	0.81
Total TCMS (/58)	0.955 (0.901-0.980)	1.96	2.77	0.995 (0.989–0.998)	0.82	1.16

Note: SEM: standard error of measurement, MDC: minimal detectable change.

Convergent validity

Pearson's correlation coefficient indicated that the total scores of the TCMS-TH and three subscales were significantly correlated with all dimensions and the total percent score of the GMFM. However, the static sitting balance subscales were found to have a low to moderate correlation with the GMFM, with the coefficient varying between 0.482 and 0.631 (p<0.05). The dynamic sitting balance of the TCMS-TH, including selective movement control and dynamic reaching subset, demonstrated moderate to high correlations with the GMFM. The correlation coefficients ranged from 0.567 to 0.810 (p<0.01). Notably, high correlations were observed with GMFM Dimension B, Dimension E, and the total percent score of GMFM, with coefficients ranging from 0.712 to 0.810 (p<0.01). Pearson's correlation coefficients between the TCMS and the GMFM are summarized in Table 5.

Table 5. Pearson's correlation coefficients for the Trunk Control Measurement Scale (TCMS) and the Gross Motor Function Measure (GMFM).

TCMS	GMFM					
	Dimension A	Dimension B	Dimension C	Dimension D	Dimension E	Total
Static sitting balance	0.482*	0.566**	0.573**	0.499**	0.583**	0.631**
Selective movement control	0.567**	0.722**	0.640**	0.601**	0.712**	0.760**
Dynamic reaching	0.623**	0.712**	0.671**	0.687**	0.742**	0.810**
Total TCMS	0.589**	0.714**	0.669**	0.627**	0.723**	0.779**

*Note: *Correlation is significant at the 0.05 level,** Correlation is significant at the 0.01 level.*

Discussion

Trunk control plays a crucial role in facilitating functional activities in children with CP. A study by Choi *et al*.[28] revealed that children with spastic CP exhibit reduced trunk control, and the severity and type of motor disability influence the degree of impairment. Furthermore, there is a notable association between trunk control impairment and the GMFCS levels, with the trunk control scores decrease significantly as the GMFCS levels rise.[8] These findings underscore the essentiality of adequate trunk control for executing various movements and engaging in functional activities. Consequently, there is a growing need for reliable assessment tools to evaluate trunk control in clinical settings. In this study, the decision was made to translate the Trunk Control Measurement Scale (TCMS) into the Thai version due to its comprehensive assessment of trunk control, encompassing both static and dynamic aspects during functional activities. Moreover, the TCMS score provides valuable insights into changes in trunk ability over time, making it a valuable tool for evaluating the quality of movement in both static and dynamic actions among children with CP. In the present study, the TCMS-TH demonstrated excellent inter-item correlation in measuring trunk control in children with CP, (Cronbach's alpha=0.916). In addition to measuring internal consistency, we also evaluated inter- and intra-rater reliability since tester administration directly influences the results. Our findings revealed that the TCMS-TH total and subscale scores exhibited excellent intra- and inter-rater reliability, with all intraclass correlation coefficients (ICCs) exceeding 0.9. These results align with those reported in the original English[17] and Korean-translated version.[19] However, the exception was inter-rater reliability in the dynamic reaching subscales, which showed good reliability (ICC=0.837). It should be note that the inter-rater reliability of the dynamic reaching subscales exhibited good reliability (ICC=0.837), which may be attributed from the subjective nature of the phrase "difficulties in performance" encompassing factors such as slowness and increased effort. The subjective nature of this assessment criterion can pose challenges for raters in making accurate scoring decisions. Interestingly, this study observed that children with spastic unilateral CP encountered challenges when performing dynamic reaching tasks on the affected side due to spasticity and limited range of motion in the upper extremities, in addition to trunk impairment.

The measurement errors for the TCMS-TH were calculated in this study, and the SEM for the TCMS-TH total score was found to be slightly higher compared to the original version. This may be attributed to the wider age range of children included in this study compared to the original version.

In terms of convergent validity, the results indicated strong correlation coefficients between the total score of the TCMS-TH and the total percent score of the GMFM. However, when analyzing the subscales of the TCMS-TH, the data were inconsistent with the original. Specifically, a significant but low to moderate correlation was observed between Dimension A of the GMFM and the TCMS-TH total score (Table 5). Furthermore, all dimensions of the GMFM displayed low to moderate correlation with the static sitting balance subscale of the TCMS-TH. This inconsistency could be explained by the fact that the items in all dimensions of the GMFM primarily assess functional activities that involve greater dynamics. Conversely, the dynamic sitting balance exhibited moderate to high correlations with the GMFM, particularly with Dimension B (Sitting), Dimension E (Walking, Running, and Jumping), and the total percent Score of the GMFM. The highest correlation was observed between the total score of the GMFM and the dynamic reaching subscale of the TCMS-TH (r=0.810). This finding highlights the importance of trunk control as a fundamental component for maintaining sitting stability, facilitating reaching movements, ensuring postural balance, and enhancing overall gross motor function.

Children with neuromotor deficits commonly face challenges in achieving selective control of movement and often exhibit excessive co-contraction, characterized by the simultaneous activation of agonist and antagonist muscles.[4] The core region of the body, encompassing both the trunk and pelvis, plays a pivotal role in maintaining sitting postural control. Notably, research has demonstrated that achieving independent sitting by the age of 2 years is a robust predictor of ambulatory capabilities in children with CP.[29] The excessive coupling of the pelvis to the trunk is a common compensatory mechanism employed to enhance stability in children with CP. However, in the selective movement subscale of the TCMS, to achieve the mobility needed to perform the task, children must decrease the level of antagonistic co-activation to attain the necessary mobility for task performance. Consequently, children exhibiting higher levels of co-activation tend to receive lower scores in the selective movement subscale. The reduction in antagonistic co-activation directly contributes to improved gross motor performance.

The study conducted by Heyrman *et al*.[17] focused exclusively on children aged 8 to 15 years. In contrast, our study aimed to broaden the age range by including children as young as 5 years old. However, with younger children, it is important to ensure that they fully understand the test instructions. Our findings support the robust psychometric properties of the TCMS-TH as a valuable clinical tool for assessing trunk control in children with CP. Based on the results of this study, we recommend that pediatric therapists, well-versed in the intricacies of CP, employ the TCMS-TH as an assessment instrument. It is imperative that the assessors undergo proper training before administering the test. Certain items within the scale require particular attention, such as the careful observation of compensatory movements, including trunk muscle shortening and elongation, during each task.

Additionally, variations in study methodology, including test-retest agreement using Kappa statistics, exclusion criteria, and the distribution of participants across the GMFCS levels, may influence the reliability and validity outcomes reported. Differences in these aspects between studies should be carefully considered when comparing results, as they may impact the observed correlations and generalizability of findings. We firmly believe that this measurement tool will prove beneficial to clinicians and researchers when assessing and evaluating the efficacy of interventions in children with CP.

Limitations

This study has several limitations. First, the use of the same assessor for both the GMFM and TCMS-TH may have introduced rater bias. To minimize this, assessments were conducted under standardized instructions, and another researcher was responsible for summarizing the raw data provided by the assessor. Second, its outcomes are primarily generalizable to observations obtained from video recordings rather than direct in-person assessments. For future studies, the responsiveness and discriminate validity of the TCMS-TH should be determined. Furthermore, expanding the evaluation of the psychometric properties of the TCMS-TH to include other types of cerebral palsy (CP) would provide a more comprehensive understanding of its applicability and validity across different CP subtypes.

Conclusion

The measurement tool for assessing trunk control, the TCMS-TH, has good to excellent observer reliability for both inter-rater and intra-rater reliability. In addition, the total score of the TCMS-TH showed a high correlation with the percent score of the GMFM. This study reported good psychometric properties of the TCMS-TH for assessment of trunk control in children with spastic CP aged 5-15 years. This measurement would be very helpful to pediatric therapists in clinical practice and research settings.

Ethical approval

The study was approved by the Human Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University, Thailand (AMSEC-65FB-001).

Funding

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

Conflict of interests

The authors declare no conflict of interests.

CReDIT authorship contribution statement

Conceptualizations and methodology: Arisa Parameyong, Jitapa Chawawisuttikool, Sirinun Boripuntakul, Sauwaluk Dacha; **Data collection, data acquisition and data analysis**: Arisa Parameyong, Jitapa Chawawisuttikool, Sirinun Boripuntakul, Sauwaluk Dacha; **drafting the manuscript**: Arisa Parameyong. All authors have read and agreed to the final version of the manuscript.

Acknowledgements

This work was supported by a research grant from the Faculty of Associated Medical Sciences, Chiang Mai University. We would like to thank Ariya N., Changprajak S, and Thumvivornkul P., for their contributions to video processing. In addition, we are thankful expert committee, the parents and children who participated in this study as well as the research teams from Rajanagarindra Institute of child development, for assistance with participant recruitment.

References

[1] Armstrong RW. Definition and classification of cerebral palsy. Dev Med Child Neurol. 2007; 49(3): 166. doi: 10.1111/j.1469-8749.2007.00166.x.

[2] Bax M, Goldstein M, Rosenbaum P, Leviton A, Paneth N, Dan B, et al. Proposed definition and classification of cerebral palsy, April 2005. Dev Med Child Neurol. 2005; 47(8): 571-6. doi: 10.1017/s001216220500112x.

[3] Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, et al. A report: the definition and classification of cerebral palsy April 2006. Dev Med Child Neurol Suppl. 2007; 109: 8-14. doi:10.1111/j14699-8749.2007.tb12610.x.

[4] Carlberg EB, Hadders-Algra M. Postural dysfunction in children with cerebral palsy: some implications for therapeutic guidance. Neural Plast. 2005;12:221-8.

[5] Rose J, Wolff DR, Jones VK, Bloch DA, Oehlert JW, Gamble JG. Postural balance in children with cerebral palsy. Dev Med Child Neurol. 2002; 44(1): 58-63. doi: 10.1017/s0012162201001669.

[6] Reid DT. The effects of the saddle seat on seated postural control and upper-extremity movement in children with cerebral palsy. Dev Med Child Neurol. 1996; 38(9): 805-15. doi: 10.1111/j.1469-8749.1996.tb15115.x.

[7] Rachwani J, Santamaria V, Saavedra SL, Woollacott MH. The development of trunk control and its relation to reaching in infancy: a longitudinal study. Front Hum Neurosci. 2015; 9:94. doi: 10.3389/fnhum.2015.00094.

[8] Heyrman L, Desloovere K, Molenaers G, Verheyden G, Klingels K, Monbaliu E, et al. Clinical characteristics of impaired trunk control in children with spastic cerebral palsy. Res Dev Disabil. 2013; 34: 327-34. doi: 10.1016/j.ridd.2012.08.015.

[9] Liao SF, Yang TF, Hsu TC, Chan RC, Wei TS. Differences in seated postural control in children with spastic cerebral palsy and children who are typically developing. Am J Phys Med Rehabil. 2003; 82(8): 622-6. doi: 10.1097/01.PHM.0000073817.51377.51.

[10] Donker SF, Ledebt A, Roerdink M, Savelsbergh GJ, Beek PJ. Children with cerebral palsy exhibit greater and more regular postural sway than typically developing children. Exp Brain Res. 2008; 184(3): 363-70. doi: 10.1007/s00221-007-1105-y.

[11] Heyrman L, Feys H, Molenaers G, Jaspers E, Monari D, Nieuwenhuys A, et al. Altered trunk movements during gait in children with spastic diplegia: compensatory or underlying trunk control deficit? Res Dev Disabil. 2014; 35(9): 2044-52. doi: 10.1016/j.ridd.2014.04.031.

[12] Montero Mendoza S, Gomez-Conesa A, Hidalgo Montesinos MD. Association between gross motor function and postural control in sitting in children with Cerebral Palsy: a correlational study in Spain. BMC pediatr. 2015; 15: 124. doi: 10.1186/s12887-015-0442-4.

[13] Nordmark E, Hägglund G, Jarnlo GB. Reliability of the gross motor function measure in cerebral palsy. Scand J Rehabil Med. 1997;29(1):25-8.

[14] Lundkvist Josenby A, Jarnlo GB, Gummesson C, Nordmark E. Longitudinal construct validity of the GMFM-88 total score and goal total score and the GMFM-66 score in a 5-year follow-up study. Phys Ther. 2009; 89(4): 342-50. doi: 10.2522/ptj.20080037.

[15] Butler PB, Saavedra S, Sofranac M, Jarvis SE, Woollacott MH. Refinement, reliability, and validity of the segmental assessment of trunk control. Pediatr Phys Ther. 2010; 22(3): 246-57. doi: 10.1097/PEP.0b013e3181e69490.

[16] Hansen L, Erhardsen KT, Bencke J, Magnusson SP, Curtis DJ. The Reliability of the Segmental Assessment of Trunk Control (SATCo) in Children with Cerebral Palsy. Phys Occup Ther Pediatr. 2018; 38(3): 291-304. doi: 10.1080/01942638.2017.1337662.

[17] Heyrman L, Molenaers G, Desloovere K, Verheyden G, De Cat J, Monbaliu E, et al. A clinical tool to measure trunk control in children with cerebral palsy: the Trunk Control Measurement Scale. Res Dev Disabil. 2011; 32(6): 2624-35. doi: 10.1016/j.ridd.2011.06.012.

[18] Verheyden G, Nieuwboer A, Mertin J, Preger R, Kiekens C, De Weerdt W. The Trunk Impairment Scale: a new tool to measure motor impairment of the trunk after stroke. Clin Rehabil. 2004; 18(3): 326-34. doi: 10.1191/0269215504cr733oa.

[19] Jeon JY, Shin WS. Reliability and validity of the Korean version of the trunk control measurement scale (TCMS-K) for children with cerebral palsy. Res Dev Disabil. 2014; 35(3): 581-90. doi: 10.1016/j.ridd.2014.01.009.

[20] Mitteregger E, Marsico P, Balzer J, van Hedel HJA. Translation and construct validity of the Trunk Control Measurement Scale in children and youths with brain lesions. Res Dev Disabil. 2015; 45-46: 343-52. doi: 10.1016/j.ridd.2015.08.007.

[21] Lopez J, Estrada C, Gomez C, Egea-Gámez R, Valera Calero J, Casas P, et al. Trunk Control Measurement

Scale (TCMS): Psychometric Properties of Cross-Cultural Adaptation and Validation of the Spanish Version. Int J Environ Res Public Health. 2023; 20: 5144. doi: 10.3390/ijerph20065144.

[22] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000; 25(24): 3186-91. doi: 10.1097/00007632-200012150-00014.

[23] Turner R, Carlson L. Indexes of Item-Objective Congruence for Multidimensional Items. Int J Test. 2003; 3: 163-71. doi: 10.1207/S15327574IJT0302_5.

[24] Ko J, Kim M. Reliability and responsiveness of the gross motor function measure-88 in children with cerebral palsy. Physical therapy. 2013; 93(3): 393-400. doi: 10.2522/ptj.20110374.

[25] Weis R. Gross Motor Function Measure (GMFM-66 and GMFM-88) User's Manual. J Paediatr Neurol. 2004; 8: 111-2. doi: 10.1016/j.ejpn.2003.11.003.

[26] Domholdt E. Physical Therapy Research : Principles and Applications. Philadelphia [Pennsylvania]: W.B. Saunders Company; 2000.

[27] Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med. J. 2012; 24(3): 69-71.

[28] Choi Y-E, Jung H-R, Kim J-H. Comparison of Trunk Control on Gross Motor Function and Topography in Children with Spastic Cerebral Palsy. J Korean Soc of Phys Med. 2019; 14: 45-53. doi: 10.13066/kspm.2019.14.4.45.

[29] Keeratisiroj O, Thawinchai N, Siritaratiwat W, Buntragulpoontawee M. Prognostic Predictors for Ambulation in Thai Children With Cerebral Palsy Aged 2 to 18 Years. J. Child Neurol. 2015; 30: 1812-8. doi: 10.1177/0883073815582267.

Appendix
Trunk control measurement scale -Thai version (TCMS-TH)

คำชี้แจงในการทดสอบ

ควรถอดอุปกรณ์เสริม รองเท้า และ/หรืออุปกรณ์ช่วยพยุงลำตัวออก ท่าเริ่มต้นเป็นท่าเดียวกันในแต่ละข้อ โดยผู้ป่วยนั่งบนขอบเตียงที่ใช้รักษาโดยไม่มีการช่วยพยุงบริเวณหลัง แขน และเท้า ต้นขาสัมผัสชิดกับเตียง มือวางบนขาใกล้กับลำตัว

ให้ผู้ป่วยนั่งตัวตรงทุกครั้งเมื่อเริ่มการทดสอบในแต่ละข้อและคอยกระตุ้นให้ผู้ป่วยพยายามนั่งตัวตรงในระหว่างการทดสอบ

คำว่าตัวตรง หมายถึง การนั่งตัวตรงมากที่สุดเท่าที่เด็กจะสามารถทำได้ ท่าเริ่มต้นอาจจะแตกต่างกันได้ในเด็กแต่ละคน ท่าเริ่มต้น คือ ท่าที่ใช้เป็นจุดอ้างอิงเพื่อระบุการเคลื่อนไหวที่เบี่ยงเบน และ/หรือมีการเคลื่อนไหวชดเชย แต่ละข้อทำการทดสอบ 3 ครั้ง ใช้คะแนนในครั้งที่ทำการทดสอบได้ดีที่สุด

ถ้าเด็กทำการทดสอบในข้อย่อย “การทรงตัวในท่านั่งขณะอยู่นิ่ง” มีการใช้แขนข้างหนึ่งช่วยพยุง การใช้มือ 1 ข้างใช้แค่เพียงวางราบบนเตียงโดยไม่เกาะถือว่ายอมรับได้

ข้อที่			สองข้าง/ ซ้าย	ขวา
	การทรงตัวในท่านั่งขณะอยู่นิ่ง ขั้นตอนการทดสอบ: ผู้ทดสอบอธิบายกับผู้ป่วยในแต่ละข้อและสาธิตได้ถ้าต้องการ			
1	**ท่าเริ่มต้น** (นั่งโดยไม่ช่วยพยุง มือวางบนขา) ให้ผู้ป่วยนั่งตัวตรงและอยู่ในท่านี้เป็นเวลา 10 วินาที	ผู้ป่วยล้ม หรือสามารถนั่งตัวตรงได้ต่อเมื่อใช้แขนสองข้างช่วยพยุง	☐ 0	
		ผู้ป่วยสามารถนั่งตัวตรงได้ต่อเมื่อใช้แขนหนึ่งข้างช่วยพยุงเป็นเวลา 10 วินาที	☐ 1	
		ผู้ป่วยสามารถนั่งตัวตรงอยู่ในท่านี้โดยไม่ใช้แขนช่วยพยุงเป็นเวลา 10 วินาที *ถ้าคะแนน = 0 ดังนั้นคะแนนรวม = 0*	☐ 2	
2	**ท่าเริ่มต้น** ผู้ป่วยยกแขนทั้งสองข้างขึ้นสูงระดับสายตาเป็นเวลา 1 วินาทีและกลับสู่ท่าเริ่มต้น	ผู้ป่วยล้ม หรือไม่สามารถยกแขนขึ้นได้	☐ 0	
		ผู้ป่วยสามารถยกแขนขึ้นโดยไม่ล้มแต่มีการเคลื่อนไหวชดเชย การเคลื่อนไหวชดเชยที่อาจเกิดขึ้น ได้แก่ (1) เอนไปด้านหลัง (2) ลำตัวงอมากขึ้น (3) เอียงไปด้านข้าง (4) อื่น ๆ	☐ 1	
		ผู้ป่วยยกแขนขึ้นโดยไม่มีการเคลื่อนไหวชดเชย	☐ 2	
3	**ท่าเริ่มต้น** ผู้ทดสอบยกขาผู้ป่วยข้างหนึ่งไขว้บนขาอีกข้างหนึ่ง	ผู้ป่วยล้มหรือไม่สามารถนั่งไขว้ขาได้ หรือสามารถนั่งได้ต่อเมื่อใช้แขนสองข้างช่วยพยุง	☐ 0	☐ 0
		ผู้ป่วยสามารถนั่งได้ต่อเมื่อใช้แขนหนึ่งข้างช่วยพยุงเป็นเวลา 10 วินาที	☐ 1	☐ 1
		ผู้ป่วยสามารถนั่งได้โดยไม่ใช้แขนช่วยพยุงเป็นเวลา 10 วินาที	☐ 2	☐ 2
4	**ท่าเริ่มต้น** ผู้ป่วยยกขาข้างหนึ่งไขว้บนขาอีกข้างหนึ่ง (อนุญาตให้ใช้มือข้างหนึ่งช่วยได้) ‘เล็กน้อย’= มีการเคลื่อนไหวของลำตัวเล็กน้อยโดยไม่มีทีท่าของการเสียสมดุลของลำตัวขณะที่มีการเคลื่อนไหวของขา ‘ชัดเจน’= มีท่าทีที่ชัดเจนของการเสียสมดุล เช่น เอียงไปด้านข้าง หรืองอตัว	ผู้ป่วยล้ม หรือไม่สามารถไขว้ขาได้ หรือสามารถไขว้ขาได้ต่อเมื่อใช้แขนสองข้างช่วยพยุง	☐ 0	☐ 0
		ผู้ป่วยสามารถไขว้ขาได้ต่อเมื่อใช้แขนข้างหนึ่งช่วยพยุง	☐ 1	☐ 1
		ผู้ป่วยสามารถไขว้ขาได้โดยไม่ใช้แขนช่วยพยุงแต่มีการเคลื่อนของลำตัวจากตำแหน่งเดิมอย่างชัดเจน	☐ 2	☐ 2
		ผู้ป่วยสามารถไขว้ขาได้โดยมีการเคลื่อนของลำตัวจากตำแหน่งเดิมเล็กน้อย	☐ 3	☐ 3
5	**ท่าเริ่มต้น** ผู้ป่วยกางขาข้างหนึ่งออกมากกว่า 10 เซนติเมตร และกลับสู่ท่าเริ่มต้น (ความกว้าง 10 เซนติเมตร = ความกว้างของเข่า) ‘เล็กน้อย’= มีการเคลื่อนไหวของลำตัวเล็กน้อยโดยไม่มีทีท่าของการเสียสมดุลของลำตัวขณะที่มีการเคลื่อนไหวของขา ‘ชัดเจน’= มีท่าทีที่ชัดเจนของการเสียสมดุล เช่น เอียงไปด้านข้างหรืองอลำตัว	ผู้ป่วยล้มหรือไม่สามารถกางขาได้ หรือสามารถกางขาได้ต่อเมื่อใช้แขนสองข้างช่วยพยุง	☐ 0	☐ 0
		ผู้ป่วยสามารถกางขาได้ต่อเมื่อใช้แขนหนึ่งข้างช่วยพยุง	☐ 1	☐ 1
		ผู้ป่วยสามารถกางขาได้โดยไม่ใช้แขนช่วยพยุงแต่มีการเคลื่อนของลำตัวออกจากตำแหน่งเดิมอย่างชัดเจน	☐ 2	☐ 2
		ผู้ป่วยสามารถกางขาได้โดยมีการเคลื่อนของลำตัวออกจากตำแหน่งเดิมเล็กน้อย	☐ 3	☐ 3
			คะแนนรวมการทรงตัวในท่านั่งขณะอยู่นิ่ง	**/20**

ข้อที่			สองข้าง/ ซ้าย	ขวา
	การทรงตัวในท่านั่งขณะเคลื่อนไหว การควบคุมการเคลื่อนไหวแบบแยกส่วน ขั้นตอนการทดสอบ: อันดับแรกผู้ทดสอบอธิบายและสาธิตแต่ละหัวข้อให้ผู้ป่วย อันดับสอง สาธิตกับผู้ป่วยโดยผู้ทดสอบทำให้เกิดการเคลื่อนไหว อันดับที่สาม ให้ผู้ป่วยเคลื่อนไหวโดยผู้ทดสอบช่วยให้เกิดการเคลื่อนไหว ตามด้วยผู้ป่วยพยายามทำการเคลื่อนไหวด้วยตัวเอง 3 ครั้ง			
6 ก	**ท่าเริ่มต้น** – กอดอก ให้ผู้ป่วยเอนไปด้านหน้าโดยลำตัวตรง ประมาณ 45 องศา และกลับสู่ท่าเริ่มต้น ปฏิกิริยาการตั้งตรงของศีรษะตามปกติ ได้แก่ การเงยของศีรษะเล็กน้อยไม่นับเป็นการเคลื่อนไหวชดเชย	ผู้ป่วยล้ม หรือไม่สามารถไปถึงตำแหน่งเป้าหมายได้	☐ 0	
		ผู้ป่วยสามารถเอนไปด้านหน้าได้ *ถ้าคะแนน = 0 ดังนั้นข้อ 6ข = 0*	☐ 1	
6 ข		ผู้ป่วยมีการเคลื่อนไหวชดเชย (1) เงยศีรษะมากขึ้น (2) งอลำตัวมากขึ้น (3) แอ่นกระดูกสันหลังมากขึ้น (4) งอเข่ามากขึ้น (5) อื่น ๆ	☐ 0	
		ผู้ป่วยสามารถเอนไปด้านหน้าได้โดยไม่มีการเคลื่อนไหวชดเชย	☐ 1	
7 ก	**ท่าเริ่มต้น** – กอดอก ให้ผู้ป่วย<u>เอนไปด้านหลัง</u>โดยลำตัวตรง ประมาณ 45 องศา และกลับสู่ท่าเริ่มต้น ปฏิกิริยาการตั้งตรงของศีรษะตามปกติ ได้แก่ การก้มของศีรษะเล็กน้อยไม่นับเป็นการเคลื่อนไหวชดเชย	ผู้ป่วยล้ม หรือไม่สามารถไปสู่ตำแหน่งเป้าหมายได้	☐ 0	
		ผู้ป่วยสามารถเอนไปด้านหลังได้ *ถ้าคะแนน = 0 ดังนั้นข้อ 7ข = 0*	☐ 1	
7 ข		ผู้ป่วยมีการเคลื่อนไหวชดเชย (1) ก้มศีรษะมากขึ้น (2) งอลำตัวมากขึ้น (3) เหยียดเข่ามากขึ้น (4) อื่น ๆ	☐ 0	
		ผู้ป่วยสามารถเอนไปด้านหลังได้โดยไม่มีการเคลื่อนไหวชดเชย	☐ 1	
8 ก	**ท่าเริ่มต้น** ให้ผู้ป่วยใช้ข้อศอกแตะที่เตียงระดับเดียวกับหัวกระดูกต้นขา (โดยมีการหดสั้นของกล้ามเนื้อลำตัวข้างเดียวกัน และมีการยืดยาวของกล้ามเนื้อลำตัวด้านตรงข้าม) และกลับสู่ท่าเริ่มต้น	ผู้ป่วยล้ม หรือข้อศอกไม่แตะที่เตียง	☐ 0	☐ 0
		ผู้ป่วยสามารถใช้ข้อศอกแตะที่เตียงได้ *ถ้าคะแนน = 0 ดังนั้นข้อ 8ข และ 8ค = 0*	☐ 1	☐ 1
8 ข		ผู้ป่วย (1) ไม่มีการหดสั้น/ยืดยาวของกล้ามเนื้อลำตัว หรือ (2) มีการหดสั้น/ยืดยาวของกล้ามเนื้อลำตัวในทิศทางตรงกันข้าม	☐ 0	☐ 0
		ผู้ป่วยมีการหดสั้น/ยืดยาวของกล้ามเนื้อลำตัวตามที่คาดหวัง *ถ้าคะแนน = 0 ดังนั้นข้อ 8ค = 0*	☐ 1	☐ 1
8 ค		ผู้ป่วยมีการเคลื่อนไหวชดเชย: (1) งอลำตัวมากขึ้น (2) เอนไปด้านหน้าหรือด้านหลัง (3) ยักสะโพกขึ้น (4) อื่น ๆ	☐ 0	☐ 0
		ผู้ป่วยสามารถใช้ข้อศอกแตะที่เตียงได้โดยไม่มีการเคลื่อนไหวชดเชย	☐ 1	☐ 1
9 ก	**ท่าเริ่มต้น** ให้ผู้ป่วยยักสะโพกขึ้นหนึ่งข้างและกลับสู่ท่าเริ่มต้น ไม่อนุญาตให้ยกต้นขาขึ้น	ผู้ป่วยล้ม หรือไม่สามารถยักสะโพกขึ้นได้	☐ 0	☐ 0
		ผู้ป่วยสามารถยักสะโพกขึ้นได้ *ถ้าคะแนน = 0 ดังนั้นข้อ 9ข และ 9ค = 0*	☐ 1	☐ 1
9 ข		ผู้ป่วยไม่มีการหดสั้น/ยืดยาวของกล้ามเนื้อลำตัว	☐ 0	☐ 0
		ผู้ป่วยมีการหดสั้น / ยืดยาวของกล้ามเนื้อลำตัว เพียงบางส่วน (บางส่วน= ช่วงการเคลื่อนไหวสั้น และ/หรือน้อย)	☐ 1	☐ 1
		ผู้ป่วยมีการหดสั้น/ยืดยาวของกล้ามเนื้อลำตัว ตามที่คาดหวัง *ถ้าคะแนน = 0 ดังนั้นข้อ 9ค = 0*	☐ 2	☐ 2

ข้อที่			สองข้าง/ ซ้าย	ขวา
9 ค		ผู้ป่วยมีการเคลื่อนไหวชดเชย: (1) มีการเอียงศีรษะไปด้านตรงข้าม (2) มีการเอียงของลำตัวไปด้านข้างจากตำแหน่งเดิมอย่างชัดเจน (3) อื่น ๆ	☐ 0	☐ 0
		ผู้ป่วยยักสะโพกขึ้นได้โดยไม่มีการเคลื่อนไหวชดเชย	☐ 1	☐ 1
10 ก	**ท่าเริ่มต้น** – กอดอก ให้ผู้ป่วยหมุนลำตัวส่วนบนโดยไม่หันศีรษะ การเคลื่อนไหวเริ่มจากส่วนไหล่	ผู้ป่วย (1) ล้ม (2) ไม่สามารถหมุนลำตัวได้ หรือ (3) ไม่สามารถแยกการเคลื่อนไหวในการหมุนลำตัวส่วนบน (เคลื่อนไหวแบบเป็นท่อน)	☐ 0	
		ผู้ป่วยแยกการเคลื่อนไหวในการหมุนลำตัวส่วนบนได้บางส่วน (บางส่วน = ไม่สมมาตร, ช่วงการเคลื่อนไหวน้อย, เคลื่อนไหวไหล่มากกว่าลำตัว)	☐ 1	
		ผู้ป่วยแยกการเคลื่อนไหวในการหมุนลำตัวส่วนบนได้ตามที่คาดหวัง *ถ้าคะแนน = 0 ดังนั้นข้อ 10ข = 0*	☐ 2	
10 ข		ผู้ป่วยหมุนลำตัวส่วนบนพร้อมกับการหมุนของศีรษะ	☐ 0	
		ผู้ป่วยหมุนลำตัวส่วนบนโดยไม่มีการหมุนของศีรษะ	☐ 1	
11 ก	**ท่าเริ่มต้น** – กอดอก ให้ผู้ป่วยหมุนลำตัวส่วนล่าง โดยไม่หันศีรษะ การเคลื่อนไหวเริ่มจากส่วนของเชิงกราน	ผู้ป่วย (1) ล้ม (2) ไม่สามารถหมุนลำตัวได้ หรือ (3) ไม่สามารถแยกการเคลื่อนไหวในการหมุนลำตัวส่วนล่าง (เคลื่อนไหวแบบเป็นท่อน)	☐ 0	
		ผู้ป่วยแยกการเคลื่อนไหวในการหมุนลำตัวส่วนล่างได้บางส่วน (บางส่วน = ไม่สมมาตร, ช่วงการเคลื่อนไหวน้อย, เคลื่อนไหวลำตัวส่วนบนเพิ่มเติม)	☐ 1	
		ผู้ป่วยแยกการเคลื่อนไหวในการหมุนลำตัวได้ตามที่คาดหวัง *ถ้าคะแนน = 0 ดังนั้นข้อ 11ข = 0*	☐ 2	
11 ข		ผู้ป่วยมีการเคลื่อนไหวชดเชยโดยมีการหมุนของเชิงกราน	☐ 0	
		ผู้ป่วยหมุนลำตัวส่วนล่างโดยไม่มีการเคลื่อนไหวชดเชย	☐ 1	
12 ก	**ท่าเริ่มต้น** – กอดอก ให้ผู้ป่วยเคลื่อนไหวเชิงกรานสลับหน้า-หลัง 3 ครั้ง การเคลื่อนไหวเชิงกรานสลับหน้า-หลัง = การเคลื่อนไหวของเชิงกรานโดยเอียงไปด้านข้างร่วมกับการหมุนสลับด้านซ้ายและขวา	ผู้ป่วยล้ม หรือไม่สามารถเคลื่อนไหวเชิงกรานสลับหน้า-หลัง ได้แก่ ไม่มีการเคลื่อนของร่างกายจากตำแหน่งเดิมไม่ว่าทิศทางใด	☐ 0	
		ผู้ป่วยสามารถเคลื่อนไหวเชิงกรานสลับหน้า-หลัง ได้เพียงบางส่วน (บางส่วน = ส่วนใหญ่เอียงไปด้านข้างและมีการหมุนเล็กน้อย ช่วงการเคลื่อนไหวน้อย ใช้ความพยายามมาก)	☐ 1	
		ผู้ป่วยสามารถเคลื่อนไหวเชิงกรานสลับหน้า-หลัง ได้โดยเอียงไปด้านข้างร่วมกับการหมุนด้านเดียว และทำได้บางส่วนในอีกด้าน	☐ 2	
		ผู้ป่วยสามารถเคลื่อนไหวเชิงกรานสลับหน้า-หลัง โดยเอียงไปด้านข้างร่วมกับการหมุนทั้งสองทิศทาง *ถ้าคะแนน = 0 ดังนั้นข้อ 12ข = 0*	☐ 3	
12 ข		ผู้ป่วยชดเชยการเคลื่อนไหวโดยมีการเคลื่อนของลำตัวจากตำแหน่งเดิมอย่างมาก	☐ 0	
		ผู้ป่วยสามารถเคลื่อนไหวเชิงกรานสลับหน้า-หลัง โดยไม่มีการเคลื่อนไหวชดเชย	☐ 1	
			คะแนนรวมการควบคุมการเคลื่อนไหวแบบแยกส่วน	**/28**

ข้อที่			สองข้าง/ ซ้าย	ขวา
	การประเมินขณะเคลื่อนไหวโดยการเอื้อม (ปฏิกิริยาการรักษาสมดุล) ขั้นตอนการทดสอบ: ผู้ทดสอบอธิบายแต่ละข้อให้ผู้ป่วยและตามด้วยผู้ป่วยทำการเคลื่อนไหว 3 ครั้ง			
13	**ท่าเริ่มต้น** – แขนเหยียดตรงไปข้างหน้า ให้ผู้ป่วยเอื้อมแขนทั้งสองข้างไปสู่เป้าหมายในระดับสายตา ระยะห่างเท่ากับความยาวแขนท่อนล่างและกลับสู่ท่าเริ่มต้น	ผู้ป่วยล้ม หรือไม่สามารถเอื้อมไปยังเป้าหมาย	☐ 0	
		ผู้ป่วยเอื้อมไปยังเป้าหมายอย่างยากลำบาก ความยากลำบาก คือ (1) ต้องใช้ความพยายามอย่างมาก ได้แก่ ช้า และยาก หรือ (2) ใช้มือช่วยพยุงบ้างเมื่อกลับเข้าสู่ท่าเริ่มต้น	☐ 1	
		ผู้ป่วยเอื้อมไปยังเป้าหมายและกลับสู่ท่าเริ่มต้นโดยไม่ยากลำบาก	☐ 2	
14	**ท่าเริ่มต้น** –กางแขนออกด้านข้างและมืออีกข้างวางบนขา ให้ผู้ป่วยเอื้อมแขนที่เหยียดออกไปยังเป้าหมายในระดับสายตา ระยะห่างเท่ากับความยาวแขนท่อนล่างและกลับสู่ท่าเริ่มต้น	ผู้ป่วยล้ม หรือไม่สามารถเอื้อมไปยังเป้าหมาย	☐ 0	☐ 0
		ผู้ป่วยเอื้อมไปยังเป้าหมายด้วยความยากลำบาก ความยากลำบาก คือ (1) ต้องใช้ความพยายามมาก ได้แก่ ช้า และยาก หรือ (2) ใช้มือช่วยพยุงบ้างเมื่อกลับเข้าสู่ท่าเริ่มต้น	☐ 1	☐ 1
		ผู้ป่วยเอื้อมไปยังเป้าหมายและกลับสู่ท่าเริ่มต้นโดยไม่ยากลำบาก	☐ 2	☐ 2
15	**ท่าเริ่มต้น** – กางแขนข้างหนึ่งออกด้านข้างและมืออีกข้างวางบนขา ให้ผู้ป่วยเอื้อมแขนที่กางออกข้ามแนวกลางลำตัว (เอื้อมไปยังด้านตรงกันข้าม) และกลับสู่ท่าเริ่มต้น เป้าหมายอยู่ในระดับสายตา ระยะห่างเท่ากับความยาวแขนท่อนล่าง	ผู้ป่วยล้ม หรือไม่สามารถเอื้อมไปยังเป้าหมาย	☐ 0	☐ 0
		ผู้ป่วยเอื้อมไปยังเป้าหมายอย่างยากลำบาก ความยากลำบาก คือ (1) ต้องใช้ความพยายามอย่างมาก ได้แก่ ช้า และยาก หรือ (2) ใช้มือช่วยพยุงบ้างเมื่อกลับสู่ท่าเริ่มต้น	☐ 1	☐ 1
		ผู้ป่วยเอื้อมไปยังเป้าหมายและกลับสู่ท่าเริ่มต้นโดยไม่ยากลำบาก	☐ 2	☐ 2
			คะแนนรวมการเอื้อมขณะมีการเคลื่อนไหว	**/10**
			คะแนนรวมทั้งหมด	**/58**

Journal of Associated Medical Sciences 2025; 58(3): 216-220

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Effect of a multisensory integration intervention on policewomen with migraine

Beniel Raja Gnanadurai and Sharon Magadalene JDR*

Department of Occupational Therapy, Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India.

ARTICLE INFO

Article history:
Received 13 September 2024
Accepted as revised 10 April 2025
Available online 8 July 2025

Keywords:
Multisensory integration, migraine, policewomen.

ABSTRACT

Background: Multisensory integration is used for people who are having migraine attacks. When multisensory functions such as visual and auditory stimulation are applied, the level of migraine triggers and attacks decreases gradually.

Objectives: The study aimed to determine the effect of multisensory integration intervention on policewomen with migraine.

Materials and methods: The study is based on a quasi-experimental study with 30 policewomen (15 in the experimental group and 15 in the control group) aged between 20 and 60. The MIDAS (Migraine Disability Assessment Scale) was used to measure the migraine level among policewomen.

Results: There is a significant reduction in migraine levels among female police officers. The data shows that the pre-test and post-test scores of both experimental and control groups were significantly different. Since the *p* value is lesser than 0.05, alternate hypothesis is accepted. Hence, there is a statistically significant difference in post-test scores between experimental and control group. This suggests that the intervention received by the experimental group had greater improvement when compared to the control group.

Conclusion: Multisensory integration has a significant effect on reducing migraines among female police officers. This study's benefit is to reduce the impact of migraines on individuals through multisensory integration-based activities.

Introduction

Migraine is derived from the Greek word 'Hemikrania', later converted into Latin as "hemigranea'. A migraine is a headache that causes severe throbbing pain or a pushing sensation, usually on one side of the head.[1] Migraine attacks are frequently accompanied by nausea, vomiting, and a heightened sensitivity to light and sound. They can persist for hours or even days, and the pain can be severe enough to disrupt your daily activities. It is a headache characterized by recurrent attacks of moderate to severe throbbing and pulsating pain on one side of the head. It occurs most frequently in the morning, some people have migraines at predictable times, such as before menstruation or on weekends following a stressful week throughout workspace. It is a brain and nervous system disorder.[2]

Multisensory integration is the process of merging inputs from multiple senses to create a perception that is distinct from, and cannot be easily broken down into, its individual sensory components.[3] This integration is operationally defined as a statistically significant difference between the response elicited by combining sensory

* **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Saveetha College of Occupational Therapy, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India.*
E-mail address: jdrmagdalenesharon7@gmail.com
doi: 10.12982/JAMS.2025.090
E-ISSN: 2539-6056

stimuli and the response elicited by the most effective single stimulus. In terms of single-neuron physiology, this comparison is made between the total number of neural impulses or firing rates generated by the combined stimuli versus those generated by each stimulus individually.[4,5]

In summary, multisensory integration involves various processes where information from multiple senses is combined to create an enhanced or diminished response. However, other multisensory computations, such as comparing stimulus features (like shape) across different modalities or detecting cross-modal timing and rhythm, require the individual sensory inputs to retain their distinct identities rather than merge into a single output.[4] As a result, these processes likely have distinct mechanisms and developmental trajectories, yet not fully understood. The issues the policewomen face is mostly workplace stress due to the work. Migraines occur three to four episodes for an individual in the group.[6,7] Therefore, this study is aimed to reduce the migraine levels by multisensory integration of policewomen.

Materials and methods

The MIDAS was created to assess headache-related disability and enhance communication between physicians and patients regarding the impact of migraines on daily life. The MIDAS questionnaire consists of five questions that measure the amount of lost time across three areas: work-related activities, household chores, and family, social, and leisure activities. It is designed for physicians, nurses, pharmacists, and other healthcare practitioners. MIDAS questionnaire through a standard method proved its validity and reliability to be used for evaluation of migraine.[8]

Interpretation

MIDAS scores closely align with physician assessments of illness severity and treatment needs. The scoring system is categorized as follows: little or no disability (5 to 10), moderate disability (11 to 20), and severe disability (more than 20). As a crucial component of a comprehensive approach that includes educational, diagnostic, and therapeutic strategies, the MIDAS questionnaire can significantly enhance the management of migraine and other headache types. A randomized, placebo-controlled trial has demonstrated that the MIDAS grade can guide the selection of initial treatment in a stratified care model.[9]

Participants

Policewomen with migraine were invited to participate in the study. The inclusion criteria were policewomen with moderate to severe migraine, aged from 20 to 60 years and scores above 5 of are included. Participants were excluded if they had visual impairment and conditions like pathological and neurological. The Ethical Committee (SCOTISRB059) authorized the trial, and eligible subjects provided written informed permission before participating in the study.

Protocol

Participants were interviewed and assessed for their demographic details (including name and age) and migraine features such as headache, trigger duration, causes, and severity using MIDAS (Migraine Disability Assessment Test). The MIDAS is a 5-item Questionnaire that assesses disability caused by migraine. The score ranges from 0 to 270. Higher scores indicate greater disability caused by migraine. Policewomen with scores of 21 or greater considered severe. Eligible participants were divided into 2 groups, intervention and control. The participants in the intervention group were exposed to a multisensory program when the control group received conventional occupational therapy. The participants in both groups were followed for 3 months to determine the effect of the migraine and its triggers (Table 1).

Table 1. The intervention and activities.

Sessions	Intervention	Activities
1	Introduction and assessment	- Establish rapport and trust - Discuss the client about migraine history and triggers
2, 25	Visual stimulation	- Introduce soothing visual aids like calming images and soft lighting
3	Sound therapy	- Introduce calming sound or music
4	Olfactory stimulation	- Explore different essential oils for stimulation - Combine olfactory stimulation with guided visual techniques
5, 6, 18, 19, 20, 26	Tactile stimulation	- Incorporate tactile stimulation through textures and object-based activities
7,8,28	Auditory stimulation	- Calming images with soothing music
10, 11, 21, 23, 30, 31, 32, 33, 34	Sensory-scenario-based training	- Integration for sensory experience in training - Immersive stimulation to enhance senses
12, 15, 29	Combined visual and sound therapy	- Calming images and lighting with soothing music - Flash lighting infused with sound effects
16, 17, 24, 27	Migrarelief device	- Migrarelief device was used for reduction of pain
35	Post-test	- Analyze post test score for migraine
36	Personalized integration and review	- Review of the effectiveness of each technique

Statistical analysis

The data was analyzed by the Wilcoxon signed rank test to test the statistical difference between pre-test and post-test of experimental and control group. Mann Whitney U test was applied for the comparison of post-test score, which identifies whether there exists statistically difference in consideration of the treatment given. The level of p=0.05 was measured to be statistically significant. The statistical analysis was performed using IBM SPSS version 24.0.

Results

At the beginning, 55 individual policewomen with migraines participated in this research. Nevertheless, 25 were excluded because they had mild migraine (N=12), visual and auditory impairments (N=10), and no interest in participating (N=3). Finally, 30 samples completed this research. Thirty participants were divided into two groups of 15: intervention and control groups.

There is a significant reduction in migraine level among female police officers. The data analyzed the pre-test and post-test scores of both control group (Figure 1) and experimental group (Figure 2), showing a significant difference between groups (Table 2). Since the p<0.05, we accept the alternative hypothesis. Hence, there is a statistically significant difference in post-test scores between experimental and control group of the MIDAS. This suggests that the intervention received by the experimental group had more improvement when compared to the control group. Since the p=0.00 is lower than 0.05, we accept the alternative hypothesis. Hence, there is a statistically significant difference between pre- test and post-test scores in the control group of the MIDAS. This suggests that the intervention received by the control group had significant improvement.

In the Experimental group, since the p=0.001 is less than 0.05, alternate hypothesis is accepted. Hence, there is a statistically significant difference in experimental group between pre-test and post-test scores of MIDAS (Table 3). This suggests that the intervention received by the experimental group had significant improvement. Since the p=0.037 is lesser than 0.05, alternate hypothesis is accepted. Hence, there is a statistically significant difference in post-test scores between experimental and control group of the MIDAS (Table 4). This suggests that the intervention received by the experimental group had more improvement when compared to the control group.

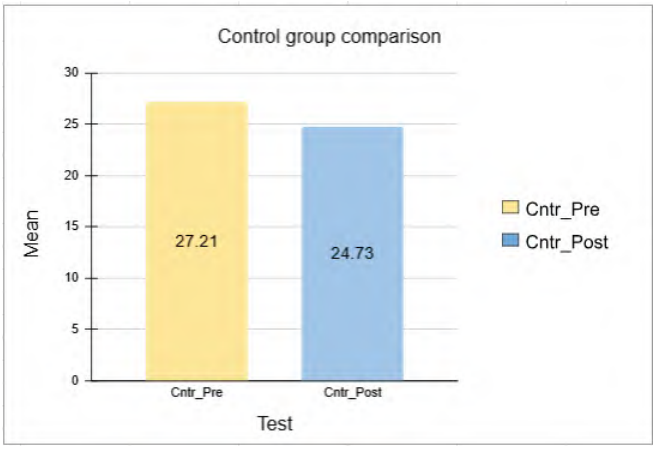


***Figure 1**. Statistical analysis of pre- test and post- test in control group*

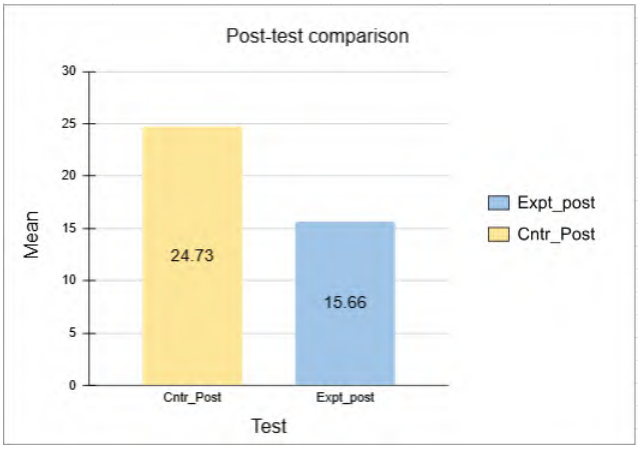


***Figure 2**. Statistical analysis between the post- test scores of the control and experimental group.*

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

Test	Mean	SD	N	Z value	*p* value
Cntr_Pre	27.2	12.22527	15	-3.496	0.00*
Cntr_Post	24.7333	11.73193	15		

***Note:** *significant at 5% alpha level, Cntr: control group.*

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

Test	Mean	SD	N	Z value	*p* value
Expt_Pre	26.4	9.24121	15	-3.417	0.001*
Expt_Post	15.6667	7.95224	15		

***Note:** *significant at 5% alpha level, Expt: experimental group.*

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

Group	Mean	SD	N	Z value	*p* value
Cntr_Post	24.7333	11.73193	15	2.094	0.037*
Expt_Post	15.6667	7.95224	15		

***Note:** *significant at 5% alpha level, Cntr: control group, Expt: experimental group.*

Discussion

The study aimed to find the effect of multisensory integration on policewomen with migraines. It was conducted at the St Thomas Mount and Guindy police station in Chennai.

A total 30 policewomen with moderate or severe migraine were selected by the selection criteria described in the methodology. They were allocated to experimental and control group using convenient sampling technique. The age of the chosen sample ranged between 20 to 60 years. It includes only policewomen officers.

Both experimental and control groups were measured with the MIDAS scale to identify the level of migraine. The experimental group underwent multisensory integration therapy intervention, whereas the control group underwent conventional occupational therapy. The multisensory integration therapy intervention session consists of 36 sessions in 3 months, each lasting 45 minutes. The effectiveness of the intervention was analyzed by comparing the pre-test and post–test values of the control and experimental groups.

In the control group, the conventional occupational therapy program included occupational based activities and relaxation techniques. Since the p=0.00 is less than 0.05, there is a statistically significant difference between pre- and post-test scores in the Control Group of MIDAS. This suggests that the intervention received by the control group had significant improvement. The result of this study is like another study reviewed by Philips EG, which explored occupational therapy's role in headache and migraine management for women in military aviation fields through a biopsychosocial approach. This study shows conventional occupational therapy intervention reduces migraine among individuals.[10]

Table 2 and Figure 1 illustrate the results of the experimental group's multisensory integration-based occupational therapy intervention. The p=0.001, below the significance level of 0.05, indicates that the difference between pre-test and post-test scores on the MIDAS scale is statistically significant. This finding suggests that the intervention led to a notable improvement in migraine symptoms among participants. Additionally, 9 out of 15 participants reported greater satisfaction with the outcomes of multisensory activities, such as visual and tactile stimulation, which further supports the effectiveness of the intervention in reducing migraines.

Table 4 and Figure 2 compare post-test MIDAS scores between the experimental and control groups. The statistical analysis reveals a significant difference, with the experimental group improving more than the control group. This highlights the superior effectiveness of the multisensory integration-based intervention compared to the control condition. These results are consistent with findings from Schwedt *et al*., which demonstrated that multisensory integration therapy effectively reduces migraines.[11] The study by Schwedt *et al* also found significant improvements in various sensory stimulations—visual, tactile, auditory, and olfactory—benefitting patients with migraine disabilities in their daily lives. However, the study has some limitations regarding its findings and clinical implications. The study was done with a limited duration of 3 months, a randomized controlled study was not used, the blinding method was not used, and convenient samplings could have influenced the results of the study.[11]

Conclusion

This study examined the effect of the intervention on migraines level among policewomen over 3 months. A total of 30 policewomen with moderate to severe migraines were selected. They were divided into two groups: the experimental group, consisting of 15 policewomen who received multi-sensory integration therapy, and the control group, with 15 policewomen, who received conventional occupational therapy. The results indicated a statistically significant difference between both groups' pre-test and post-test scores. Further analysis showed that the experimental group experienced a notable improvement compared to the control group, highlighting the effectiveness of multi-sensory integration therapy in alleviating migraines among policewomen.

Ethical approval

The study has been approved by the Institution Scientific Review Board (ISRB) of Saveetha of Occupational Therapy, with Reference No SCOT/ ISRB /059 /2023.

Funding

This research received no external funding.

Conflict of interest

There are no conflicts of interest.

CRedit authorship contribution statement

Benielraja Gnanadurai: study design, data analysis and manuscript preparation; Sharon Magadalene: study conduct under the primary author's supervision. All author's approved the manuscript.

Acknowledgements

I would like to express my sincere gratitude to everyone who supported me throughout this work. I am especially thankful to my advisor/supervisor for their valuable guidance, encouragement, and constructive feedback. I would also like to thank the police authority for conducting the course, which was instrumental to completing this research. Additionally, my sincere thanks go to my colleagues, friends, and family for their continued support and motivation during the study and writing process.

References

[1] Amiri P, Kazeminasab S, Nejadghaderi SA, Mohammadinasab R, Pourfathi H, Araj-Khodaei M, et al. Migraine: A review on its history, global epidemiology, risk factors, and comorbidities. Front Neurol. 2022; 12: 800605. doi: 10.3389/fneur.2021.800605.

[2] Haghdoost F, Togha M. Migraine management:

Non-pharmacological points for patients and health care professionals. Open Med. 2022; 17(1): 1869-82. doi: 10.1515/med-2022-0598.

[3] De Dhaem OB, Sakai F. Migraine in the workplace. Neurol Sci. 2022; 27: 100408. doi: 10.1016/j.ensci.2022.100408.

[4] Stein BE, Stanford TR, Rowland BA. Development of multisensory integration from the perspective of the individual neuron. Nat Rev Neurosci. 2014; 15(8): 520-35. doi: 10.1038/nrn3742.

[5] Buse DC, Rupnow MFT, Lipton RB. Assessing and managing all aspects of migraine: migraine attacks, migraine-related functional impairment, common comorbidities, and quality of life. Mayo Clin Proc. 2009; 84(5): 422-35. doi: 10.1016/s0025-6196(11)60561-2.

[6] Leso V, Gervetti P, Mauro S, Macrini MC, Ercolano ML, Iavicoli I. Shift work and migraine: A systematic review. Occup Health. 2020; 62(1): e12116. doi: 10.1136/bmjopen-2013-002791.

[7] Garbarino S, Cuomo G, Chiorri C, Magnavita N. Association of work-related stress with mental health problems in a special police force unit. BMJ Open. 2013; 3(7): e002791. doi: 10.1136/bmjopen-2013-002791.

[8] Stewart WF, Lipton RB, Dowson AJ, Sawyer J. Development and testing of the Migraine Disability Assessment (MIDAS) questionnaire to assess headache-related disability. Neurology. 2001; 56(Suppl 1): S20-8. doi: 10.1212/wnl.56.suppl_1.s20

[9] Peng KP, Wang SJ. Migraine diagnosis: Screening items, instruments, and scales. Acta Anaesthesiol Taiwan. 2012; 50(2): 69-73. doi: 10.1016/j.aat.2012.05.002.

[10] Philips EG. Exploring occupational therapy's role in headache & migraine management for women in military aviation fields through a biopsychosocial approach. Boston (MA): Boston University; 2023.

[11] Schwedt TJ. Chronic migraine. BMJ. 2014; 348: g1416. doi: 10.1136/bmj.g1416.

Journal of Associated Medical Sciences 2025; 58(3): 221-227

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

Psychometric properties of the Thai version of Allen Cognitive Level Screen-5 (Thai ACLS-5) for individuals with schizophrenia

Kaewalin Sriyuang and Pornpen Sirisatayawong*

Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 9 March 2025
Accepted as revised 5 July 2025
Available online 8 July 2025

Keywords:
Construct validity, Known-group method, Inter-rater reliability, Schizophrenia, Functional cognitive assessment.

ABSTRACT

Background: Cognitive impairment is a critical concern for individuals with schizophrenia, affecting their occupational performance, daily activities, and the potential for rehabilitation and skill development. The Allen Cognitive Level Screen-5 (ACLS-5) is a standardized, performance-based assessment designed to evaluate functional cognitive abilities and limitations, providing guidance for individualized interventions. However, there is currently no Thai translation for the ACLS-5.

Objectives: This study aimed to translate the ACLS-5 into Thai using a back-translation method and to evaluate its psychometric properties, focusing on construct validity and inter-rater reliability, to ensure its suitability for use in Thai clinical settings.

Materials and methods: The ACLS-5 was translated from English into Thai following Beaton's protocol, a five-step process for cross-cultural adaptation, and the pre-final version was tested on a sample of 30 individuals with schizophrenia to assess cultural adaptation and content validity. Construct validity was evaluated using a sample of 102 participants, comprising of 51 individuals with schizophrenia and 51 healthy individuals. Inter-rater reliability was examined by two occupational therapists who independently evaluated 38 participants from the schizophrenia group.

Results: In the translation and cultural adaptation phase, the names of the stitching tasks were modified to match Thai culture. Feedback from the Allen Cognitive Group suggested that the terms lacked direct Thai equivalents; consequently, they remained in their original English form. The pre-final Thai ACLS-5 was tested with 30 individuals with schizophrenia, in which it was reported that the instructions were clear and easy to understand. The Mann-Whitney U test indicated a significant difference in construct validity between individuals with schizophrenia and healthy controls ($Z=-6.511$, $p<0.000$), confirming the assessment's effectiveness in distinguishing cognitive impairment. The intraclass correlation coefficient (ICC) for inter-rater reliability was 0.932, demonstrating excellent reliability.

Conclusion: These findings confirm that the Thai ACLS-5 is a valid and reliable assessment for evaluating functional cognition in individuals with schizophrenia. It allows for a quick and effective assessment of functional cognition in clinical settings. Future research should examine usage in different populations to further establish its utilization.

* **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: pornpen.siri@cmu.ac.th
doi: 10.12982/JAMS.2025.091
E-ISSN: 2539-6056

Introduction

Schizophrenia is a mental disorder characterized by abnormalities in neurotransmitter systems, leading to significant cognitive impairments.[1] Symptoms usually appear in adolescence or early adulthood, affecting various cognitive domains, such as attention, verbal fluency, working memory, processing speed, and executive functions.[2] These

deficits are associated with functional skills affecting their daily life activities.[3] A study of the relationships between occupational engagement and cognitive functioning among persons with schizophrenia suggests that cognitive functioning influences occupational engagement and the capacity to perform daily activities.[4]

The evaluation of cognitive abilities is an important part in the occupational therapy process. Assessing cognitive abilities in daily activities allows for the identification of strengths and limitations, as well as providing how an individual acquires new skills and adjusts to environmental strategies that support daily living.[5,6] There are various cognitive assessment tools available, such as self-report measures, performance-based tests, and computer-based measures. However, performance-based assessments are most used for psychiatric patients.[7]

One of the performance-based assessments is the Allen Cognitive Level Screen-5 (ACLS-5), a standardized tool based on the Cognitive Disabilities Model (CDM). This model suggests that functional performance in daily activities is influenced by the interaction between cognitive abilities and environmental contexts. The model categorizes cognitive abilities into six levels, reflecting brain processing capacity, learning potential, problem-solving skills, and performance capacity. These levels are determined through the observation of performance on three leather stitching tasks with increasing complexity.[8] The ACLS-5 facilitates the classification of individuals based on cognitive performance and is useful in the development of specific treatment strategies that relate to their cognitive abilities.[9] Additionally, the application of the ACLS-5 in clinical settings enables professionals to promptly and effectively evaluate cognitive capacities, hence impacting the formulation of treatment plans and subsequent interventions after discharge.[10]

Studies on the psychometric properties of the ACLS-5 have demonstrated strong inter-rater reliability across all versions of the ACLS and LACLS (1978-2007), with correlation coefficients ranging from r=0.91 to 0.99. [11] A study involving 225 individuals in a mental health setting found a significant association between functional cognition, as measured by the LACLS-5, and the level of independence (r=0.55, p<.01).[11] The ACLS-5 has been translated and validated in multiple languages, for example, Japanese,[11] Chinese,[12,13] and Turkish,[14] demonstrating robust psychometric properties. Significant correlations have been reported with cognitive and functional assessments, including the Mini-Mental State Examination (r=0.90, p<.0001),[11] the Chinese Mini-Mental State Examination (Z=0.61, p<.01),[12] the Social Functioning Scale (Z=0.382-0.480, p<.05),[14] and the Wisconsin Card Sorting Test (Z=0.257-0.557, p<.01). [14] Overall, findings consistently support the construct validity, reliability, and discriminant validity of the ACLS-5, confirming its effectiveness in evaluating functional cognition across diverse populations.

The ACLS-5 is widely utilized for evaluating functional cognition and designing interventions for daily living activities. However, there have been no reports of the translation of the ACLS-5 into Thai.

This study aimed to translate the ACLS-5 from English to Thai and evaluate its psychometric properties, including content validity, construct validity, and inter-rater reliability.

Materials and methods

Study design

This study was conducted in two phases: 1) the translation and cultural adaptation of the ACLS-5, and 2) the evaluation of psychometric properties, which included construct validity and inter-rater reliability.

Phase 1: Translation and cultural adaptation

The researchers obtained permission from the Allen Cognitive Group for the translation of the administration and scoring manuals. The translation process followed the Beaton's guideline[15] and consisted of five steps:

Step I: Initial translation

ACLS-5 was independently translated from English to Thai by two bilingual translators. One translator had a medical background with expertise in psychiatric occupational therapy, while the other had no medical background. This step produced two initial Thai versions, labeled T1 and T2.

Step II: Synthesis of translations

Two translators and the researchers synthesized T1, T2, and the original ACLS-5 to develop a unified version, referred to T-12. Both translators reviewed and approved this synthesized version.

Step III: Back translation

Two bilingual translators who were blind to the original ACLS-5 independently back-translated the T-12 version into English. This step produced two English versions, BT1 and BT2. Both versions were thoroughly reviewed to confirm their consistency with the original ACLS-5 content. The outcome of this step was a combined English version (BT12).

Step IV: Expert panel discussion

An expert panel was focused on semantic equivalence, idiomatic equivalence, and conceptual equivalence to ensure the accuracy and appropriateness of the translation. The panel consisted of seven members: four translators involved in steps I-III, one English language expert, one research methodology expert, and one occupational therapist. Through consultations and discussions, the panel refined the assessment, resulting in a pre-final version. Additionally, the BT12 English version was approved and verified by the original developers (The Allen Cognitive Group) to confirm consistency with the original content.

Step V: Pre-testing

The pre-final Thai version of the ACLS-5 was administered to 30 individuals with schizophrenia. They

were recruited from a psychiatric hospital in Chiang Mai Province, Thailand, using purposive sampling. The inclusion criteria were as follows:

1. A diagnosis of schizophrenia based on the International Classification of Diseases, 10th Revision (ICD-10) criteria
2. Stable psychiatric symptoms, indicated by a score of the Positive and Negative Syndrome Scale Thai version (PANSS-T)
3. Aged between 20 and 65 years
4. No auditory or visual impairments
5. Ability to communicate in Thai
6. No history of brain injury or neurological disorders
7. No history of experience with stitching leather

After completing the assessment, interviews were conducted to assess the clarity and comprehensibility of the assessment. Feedback was analyzed, and necessary modifications were made to refine the instrument. Following this process, the final version of the Thai ACLS-5 was completed.

Phase 2: Psychometric properties study

Construct validity of the Thai ACLS-5 was examined using the known-group technique. The sample size for the study was calculated using the G*Power 3.1.9.4 software, with an effect size of 0.5, a power level of 0.8, and a significance level of α=0.05.[14] Based on this calculation, a total of 102 participants were included in the study, comprising 51 individuals with schizophrenia and 51 healthy individuals. Ethical approval for the study was obtained from the Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University (EC Number: AMSEC-67FB-001) and the Human Research Ethics Committee at Suanprung Hospital (EC Number: SPH. IRB002/2567SCs_Ful).

Individuals with schizophrenia were recruited from a psychiatric hospital in Chiang Mai Province, Thailand, using purposive sampling. The inclusion criteria for this group were identical to those used in Phase 1.

A control group of healthy individuals was recruited from the community in Chiang Mai Province, Thailand, through purposive sampling. The inclusion criteria of the control group were as follows:

1. Aged between 20 and 65 years
2. No auditory or visual impairments
3. Ability to communicate in Thai
4. No history of psychiatric disorder
5. No history of brain injury or neurological disorders
6. No history of experience with stitching leather

To ensure comparability, the healthy controls were demographically matched with the schizophrenia group based on gender, age, education level, and work experience.

Inter-rater reliability was assessed in a subgroup of 38 individuals with schizophrenia from the construct validity sample. Two occupational therapists with experience in psychiatric patients were trained by the researcher to administer the Thai ACLS-5. One therapist administered the assessment while the other observed, with both independently scoring the participants to minimize bias and ensure objective evaluation.

Instruments

The Allen Cognitive Level Screen-5 (ACLS-5)

The Allen Cognitive Level Screen-5 (ACLS-5) is a standardized assessment tool developed within the framework of the Cognitive Disabilities Model (CDM).[16] The ACLS-5 consists of three visuomotor tasks of increasing complexity: the Running Stitch, the Whipstitch, and the Single Cordovan Stitch. Each task requires participants to follow the examiner's verbal and demonstrative instructions, utilizing feedback from their motor actions to make necessary corrections. Scoring ranges from cognitive level 3.0 to 5.8, with higher scores indicating better cognitive functioning. The assessment typically takes 15 to 20 minutes to complete and reflects the capacity of the individual to learn and adapt while performing various activities.

The Positive and Negative Syndrome Scale Thai version (PANSS-T)

The Positive and Negative Syndrome Scale Thai version (PANSS-T)[17] is a tool designed to assess symptom severity in individuals with schizophrenia, based on clinical interviews and observations. The PANSS-T has been assessed for both reliability and criterion validity in Thai patients with schizophrenia. Significant correlations between the PANSS and PANSS-T scales were observed, including positive symptoms (r=0.89), negative symptoms (r=0.72), and general psychopathology (r=0.88), all with p<.0001. Additionally, the agreement for each item ranged from 70% to 100%, demonstrating strong reliability.[17]

In this study, a composite scale score was calculated by subtracting the negative scale score from the positive scale score. The composite scale score below 0, indicating the relative predominance of no positive symptoms.

Statistical analysis

All statistical analyses were conducted using SPSS software. Descriptive statistics were used to summarize demographic data. The Chi-square test was employed to compare demographic characteristics between the individuals with schizophrenia and healthy controls. The Mann-Whitney U test was used to assess construct validity by comparing the Thai ACLS-5 scores between individuals with schizophrenia and healthy individuals. Inter-rater reliability was evaluated using the intraclass correlation coefficient (ICC).

Results

Phase 1: Translation and cultural adaptation

The Thai ACLS-5 was reviewed by an expert committee for semantic and idiomatic equivalence. Seven experts discussed key occupational therapy terms (such as visuospatial-motor, cognitive activity demands, performance, and cues), stitching task names (such as

Running Stitch, Whipstitch, and Single Cordovan Stitch), and the phrasing of instructional directions. During this review, terms without direct Thai equivalents were identified. The Allen Cognitive Group commented on specific terms, such as "Allen Cognitive Level," "Cross-in-Back Error," and "Twisted Lace Errors," which lacked direct Thai translations. To address these issues, the stitching task names were adapted to align with Thai cultural practices, phrase patterns were clarified, and the original technical terms were retained when necessary.

The pilot testing was conducted on 30 individuals with schizophrenia, with their demographic characteristics presented in Table 1. The results of the pilot testing, as shown in Table 2, include the Thai ACLS-5 scores and the time taken to complete the three tasks. Participant feedback indicated that the instructions were clear and easy to understand.

Phase 2: Psychometric properties study

The analysis of demographic characteristics, as shown in Table 3, revealed no differences between schizophrenia and healthy groups. The construct validity of the Thai ACLS-5 was assessed by comparing scores between individuals with schizophrenia and healthy controls. The results indicated a significant difference between the two groups (Z=-6.511, p<.000), as presented in Table 4. Additionally, the comparison of time taken to complete the Thai ACLS-5 between the two groups, analyzed using the Independent Samples T-Test, is shown in Table 5.

The inter-rater reliability analysis demonstrated a high level of agreement for the Thai ACLS-5, with an intraclass correlation coefficient (ICC) of 0.932 (Table 6), indicating excellent consistency between raters.

Table 1. Characteristics of individuals with schizophrenia (N=30) in the pilot testing.

Characteristics		Schizophrenia	%
Gender	Male	17	56.7
	Female	13	43.3
Age (years)	20-40	11	36.7
	41-60	18	60.0
	61-65	1	3.3
Education	Primary	11	36.7
	Secondary	18	60.0
	University	1	3.3
Work experience	None	18	60.0
	Farmer	4	13.3
	Self-employed	5	16.7
	Employee	3	10.0

Table 2. Result of pilot testing for the Thai ACLS-5 in schizophrenia (N=30).

Result of Testing	Minutes	Max	Mean (SD)	Median
Thai ACLS-5 score	4.4	5.6	5.17 (0.31)	5.2
Time used (minute)	2.49	28.10	13.19 (6.12)	

Table 3. Comparison of demographic data between schizophrenia (N=51) and healthy controls (N=51) using the Chi-square test.

Characteristics		Schizophrenia	%	Healthy controls	%	*p* value
Gender	Male	39	76.5	19	37.3	0.748
	Female	12	23.5	32	62.7	
Age (years)	20-40	23	45.1	19	37.3	0.942
	41-60	24	47.1	22	43.1	
	61-65	4	7.8	10	19.6	
Education	Primary	15	29.4	18	35.3	0.156
	Secondary	29	56.9	28	54.9	
	University	7	13.7	5	9.8	
Work experience	None	16	31.4	5	9.8	0.229
	Farmer	5	9.8	3	5.9	
	Self-employed	24	47.1	36	70.6	
	Employee	6	11.7	7	13.7	

Table 4. Comparison of the Thai ACLS-5 scores between schizophrenia (N=51) and healthy controls (N=51) using the Mann-whiney U test.

	Schizophrenia			Healthy controls			Z	*p* value
	Min	Max	Mean (SD)	Min	Max	Mean (SD)		
Thai ACLS-5 score	3.4	5.8	4.8(0.65)	5.2	5.8	5.5(0.17)	-6.511	0.000

Table 5. Comparison of time used of the Thai ACLS-5 scores between schizophrenia (N=51) and healthy controls (N=51) using Independent Samples T-Test.

	Schizophrenia			Healthy controls			*p* value
	Min	Max	Mean (SD)	Min	Max	Mean (SD)	
Time used (minute)	1.36	24.34	10.24 (4.31)	1.00	23.10	9.50 (5.39)	0.106

Table 6 Inter-rater reliability analysis of the Thai ACLS-5 scores between two raters using the intraclass correlation coefficient (ICC) test.

	Rater 1			Rater 2			ICC	95% CI
	Min	Max	Mean (SD)	Min	Max	Mean (SD)		
Thai ACLS-5 score (N=38)	3.4	5.8	4.7 (0.65)	3.2	5.8	4.7 (0.64)	0.932	0.873-0.964

Discussion

The back-translation process, following Beaton's protocol, was utilized to ensure the concepts and content of the original assessment. A key component of this method is the expert panel review, which serves as a content validity. The panel improved the translation to ensure accuracy, clarity, and consistency with the original assessment.[18] In this study, the expert panel revised the names of the three stitching tasks to better reflect Thai cultural contexts for the purpose of defining the type of stitching. The adjustments improved the instrument's meaning and clarity within Thai context. Additionally, in pilot testing of individuals with schizophrenia, task instructions were simplified while keeping meaning. Participant feedback indicated that the instructions were clear and easy to understand.

The construct validity of the Thai ACLS-5 indicated that individuals with schizophrenia had significantly lower scores compared to healthy controls (Z=-6.511, p<0.000); supporting the hypothesis that the ACLS-5 is effective in distinguishing individuals with cognitive impairment. The results highlight that cognitive impairment plays a crucial role in determining functional cognition in schizophrenia, impacting more than 80% of individuals with the condition.[19] Previous studies conducted by Ozturk, Bumin, Ozturk, and Akyurek[14] examined the ACLS-5 in a sample of 61 individuals with schizophrenia compared to 61 healthy participants; reporting significantly lower scores in the schizophrenia group. Like this study, their findings confirmed that the Thai ACLS-5 effectively differentiates individuals with cognitive impairment.

The inter-rater reliability analysis of the Thai ACLS-5 demonstrated an Intraclass Correlation Coefficient (ICC) of 0.932, which indicated excellent reliability.[20] Various versions of the ACL have demonstrated excellent inter-rater reliability, with correlation coefficients ranging from 0.92 to 0.99.[11] For instance, the Cantonese version of the ACLS reported high inter-rater reliability (ICC=0.98),[13] as their raters received training in assessment. In this study, raters had prior experience with the cognitive disability model in schizophrenia and received introduction and training on the assessment instruction and scoring criteria. Therefore, it is recommended that training in the cognitive disability model and assessment instructions be provided before administering the Thai ACLS-5 to ensure a comprehensive understanding of the assessment process and scoring. This training enhances accuracy in identifying cognitive levels. Additionally, the ACLS-5 includes clear and standardized instructions, promoting consistency in administration and interpretation among examiners. Moreover, participants completed the three stitching tasks within thirty minutes, highlighting that the Thai ACLS-5 is a quick and practical assessment tool. In clinical settings, it allows professionals to efficiently evaluate the cognitive abilities of patients, thus facilitating informed treatment and post-discharge intervention recommendations.[10,21]

Limitations

This study has several limitations. The individuals with schizophrenia were recruited from a single setting, and construct validity used the known-groups method. To strengthen the evidence base, future research should include participants from multiple centers and conduct further validation with additional cognitive assessment tools. Additionally, examining psychometric properties such as test-retest reliability and intra-rater reliability is recommended.

Conclusion

The Thai ACLS-5 has been effectively translated and culturally adapted to align with Thai context. The findings confirm its reliability, validity, and practicality as an assessment for evaluating functional cognition in individuals with schizophrenia. The assessment demonstrates significant potential for occupational therapists in Thailand, providing important perspectives for intervention strategies.

Future research should examine its application across diverse populations and settings to further validate its utility in both clinical practice and research.

Ethical approval

The Research Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University (EC Number: AMSEC-67FB-001) and the Human Research Ethics Committee at Suanprung hospital (EC Number: SPH.IRB002/2567SCs_Ful) both gave their approval to this research study.

Funding

This research received funding from the Faculty of Associated Medical Sciences, Chiang Mai University.

Conflict of interest

The authors declare no conflicts of interest

CRedit authorship contribution statement

Kaewalin Sriyuang: data collection, trained occupational therapists training to serve as research assistants, and statistical analysis, manuscript preparation; **Pornpen Sirisatayawong**: critical comments providing and contributing to develop research, analysis, and writing.

Acknowledgements

The researchers thank all participants from the psychiatric hospital and subdistrict community in Chiang Mai Province, Thailand, for their participation in this study. We also sincerely appreciate the staff from both settings for their support in data collection.

References

[1] Medalia A, Revheim N. Dealing with cognitive dysfunction associated with psychiatric disabilities. Walter Boppert: The New York State Office of Mental Health Family; 2002.

[2] Bowie CR, Harvey PD. Cognitive deficits and functional outcome in schizophrenia. Neuropsychiatr Dis Treat. 2006; 2(4): 531-6. doi: 10.2147/nedt.2006.2.4.531.

[3] Shimada T, Ohori M, Inagaki Y, Shimooka Y, Sugimura N, Ishihara I, et al. A multicenter, randomized controlled trial of individualized occupational therapy for patients with schizophrenia in Japan. PLoS One. 2018; 13(10): e0205549. doi: 10.1371/journal.pone.0193869.

[4] Lexen A, Bejerholm U. Occupational engagement and cognitive functioning among persons with schizophrenia: an explorative study. Scand J Occup Ther. 2018; 25(3): 172-9. doi: 10.1080/11038128.2017.1290135.

[5] American Occupational Therapy Association. Role of OT in assessing functional cognition. 2021. Available from: https://www.aota.org/practice/practice-essentials/payment-policy/medicare1/medicare---role-of-ot-in-assessing-functional-cognition.

[6] Pitman T. Mini-Mental State Examination and Large Allen Cognitive Level Screen: predictive validity for discharge disposition among patients of a skilled nursing facility. Tacoma, WA: University of Puget Sound; 2010.

[7] Bakkour N, Samp J, Akhras K, El Hammi E, Soussi I, Zahra F, et al. Systematic review of appropriate cognitive assessment instruments used in clinical trials of schizophrenia, major depressive disorder and bipolar disorder. Psychiatry Res. 2014; 216(3): 291-302. doi: 10.1016/j.psychres.2014.02.014.

[8] Allen CK, Austin SL, David SK, Earhart CA, McCraith DB, Riska-Williams L. Manual for the Allen Cognitive Level Screen-5 (ACLS-5) and Large Allen Cognitive Level Screen-5 (LACLS-5). Camarillo, CA: ACLS and LACLS Committee; 2007.

[9] Rojo-Mota G, Pedrero-Perez EJ, Huertas-Hoyas E, Merritt B, MacKenzie D. Allen Cognitive Level Screen for the classification of subjects treated for addiction. Scand J Occup Ther. 2017; 24(4): 290-8. doi: 10.3109/11038128.2016.1161071.

[10] Scanlan JN, Still M. Functional profile of mental health consumers assessed by occupational therapists: level of independence and associations with functional cognition. Psychiatry Res. 2013; 208(1): 29-32. doi: 10.1016/j.psychres.2013.02.032.

[11] Williams LR, Earhart CA, David SK. ACLS-5 and LACLS-5 Test: psychometric properties and use of scores for evidence-based practice 2018. Available from: https://allencognitive.com/wp-content/uploads/CopyrightIVReportfPsychometricsACLS-5_6-12-2016.pdf.

[12] Leung SB, Man DWK. Validity of the Chinese Version of the Allen Cognitive Level Screen Assessment for individuals with schizophrenia. Occup Ther J Res. 2007; 27(1): 31-40. doi: 10.1177/153944920702700105.

[13] Chan CL, Ray S, Trudeau S. The development and reliability of the Cantonese version of the Allen Cognitive Level Screen. Occup Ther Int. 2001; 8(4): 287-300. doi: 10.1002/oti.152.

[14] Ozturk KL, Bumin G, Ozturk E, Akyurek G. Investigation of the validity and reliability of the Turkish adaptation of Allen Cognitive Level Screen-5 (ACLS-5) with individuals with schizophrenia. Occup Ther Ment Healt. 2022; 39(4): 1-17. doi: 10.1080/0164212X.2022.2135671.

[15] Beaton D, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000; 25(24): 3186-91. doi: 10.1097/00007632-200012150-00014.

[16] Katz N, Toglia J. Cognition, occupation, and participation across the life span: neuroscience, neurorehabilitation, and models of intervention in occupational therapy. 4th ed. North Bethesda, MA: AOTA Press; 2018.

[17] Nilchaikovit T., Uneanong S., Kessawai D., Thomyangkoon P. The Thai version of the Positive and Negative Syndrome Scale (PANSS) for schizophrenia: criterion validity and interrater reliability. J Med Assoc Thai. 2000; 83(6): 646-51. PMID: 10932491.

[18] Kitreerawutiwong N. Public health research: from principle to practice. Phitsanulok: Naresuan University Publishing House; 2017.

[19] Sachs G, Bannick G, Maihofer EIJ, Voracek M, Purdon

SE, Erfurth A. Dimensionality analysis of the German version of the Screen for Cognitive Impairment in Psychiatry (SCIP-G) Schizophr Res Cognition. 2022; 29: 1-6. doi: 10.1016/j.scog.2022.100259.

[20] 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: 10.1016/ j. jcm.2016.02.012.

[21] Stewart K, Hancock N, Stancliffe RJ. Factors related to hospital utilisation for people living with schizophrenia: examining Allen's Cognitive Level Scores, recommended supports and routinely collected variables. Aust Occup Ther J. 2019; 66(5): 591-602. doi: 10.1111/1440-1630. 12597.

Journal of Associated Medical Sciences 2025; 58(3): 228-236

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

Factor associated with dental caries in alcohol dependences, Thailand

Namsukh Kongkalai[1*], Prapornpan Utamachote[2], Supatra Sang-in[3]

[1]Suanprung Psychiatric Hospital, Chiang Mai Province, Thailand.
[2]Somdej Phranangchaosirikit Hospital, Chonburi Province, Thailand.
[3]Faculty of Dentistry, Chiangmai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 20 May 2025
Accepted as revised 6 July 2025
Available online 9 July 2025

Keywords:
Dental caries, DMFT, dental diseases, alcohol dependence, alcohol dependent patients.

ABSTRACT

Background: Alcohol dependences have higher risk to dental caries from having poor oral hygiene practice and exposure to acidic drink.

Objectives: To determine prevalence and associating factors associated to dental caries in alcohol dependences that are useful for designing targeted dental caries prevention and promotion program for this specific population.

Materials and methods: A cross-sectional survey was conducted involving 450 alcohol dependent patients admitted at male in-patient ward from November 2015 to September 2016 at Suanprung Psychiatric Hospital, Chiang Mai Province. Data in terms of number of decayed teeth, missing teeth due to decayed and filled teeth (DMFT), socio-demographic information and potential associating factors to dental caries experience were collected using structured interviews, questionnaires and oral examination. Descriptive statistics and multiple logistic regression were performed.

Results: Age of participants ranged from 18-69 years old. Mean DMFT was 7.86±6.37. Participants who were diagnosed as alcohol dependence and single had higher number of decayed teeth compared to others. Age, satisfaction to own oral health and self-perceived oral health were associating factors to DMFT.

Conclusion: Dental caries was a high prevalent problem among alcohol dependences. Dental caries prevention and promotion program targeting high prevalence sub-group of population with special focus on increasing self-satisfaction and self-perceived of own oral health should be considered.

Introduction

Thailand has the high alcohol consumption per capita in 15+ years old at 8.3 mL compared to 6.4 mL worldwide.[1] Among drinkers, the alcohol per capita consumption can be as high as 26.2 mL in males 15+ years old and 9.2 mL in females.

Alcohol is an addictive substance either stimulates or suppresses central nervous system which subsequently can induce mental illness. Over half (56%) of drinkers suffered from mental problem more than physical health problem.[2] The Thai National Mental Health Survey 2008 reported that mental health problem was self-reported by 11.7% of drinkers, 42.4% of binge drinkers and 58.8% of alcohol dependences.[2] Among alcohol dependences, 36.9% were diagnosed with anxiety, 27.9% depression and 1.9% bi-polar disorder. Schizophrenic was also found.[3]

Alcohol dependences are also in high risk of oral diseases. Alcohol is an acidic drink and can weaken enamel and increase caries susceptibility. Poor oral hygiene practice

* **Corresponding contributor.**
Author's Address: *Suanprung Psychiatric Hospital, Chiang Mai Province, Thailand.*
E-mail address: Dentkawila@hotmail.com
doi: 10.12982/JAMS.2025.092
E-ISSN: 2539-6056

is commonly found amongst alcohol dependences.[4,5] A study by Hornecker *et al*. found that 52% of alcohol dependences did not brush their teeth.[6]

Suanprung Psychiatric Hospital is a large tertiary psychiatric hospital that provides holistic services for patients with mental health and psychiatric problems in the North of Thailand. It provided mental health services for 4,307 alcohol dependences in 2014. To integrate oral health into general health, it is necessary to understand factors affecting dental caries in this specific population to plan and design dental caries prevention and promotion programs. While there is an increase of evidence highlighting negative impact of alcohol on oral health, study on dental caries determinants in Thai alcohol dependences is scarce.

Factors associated to dental caries in alcohol dependences are useful for designing targeted dental caries prevention and promotion program for this specific population. Moreover, it will be helpful for the occupational therapist to design a suitable program and intervention to promote the patients' quality of life.

Therefore, the objective of this study was to identify associating factors to dental caries in alcohol dependences.

Materials and methods

A cross-sectional survey was conducted at male in-patient ward at Suanprung Psychiatric Hospital between November 2015 and September 2016 using convenient sampling. Ethics approval was granted by the Human Ethics Committee, Suanprung hospital (study ID 21/2015). Inclusion criteria were:

- Being diagnosed with alcohol dependence (ICD F10.20) or mental and behavioural disorders due to use of alcohol: psychotic disorder (ICD F10.50)
- No or mild alcohol withdrawal (≤7 scores of the Clinical Institute Withdrawal Assessment for Alcohol revised version or CIWA-Ar)[7]
- Good co-operative (<2 scores of 1, 2, 6 dimensions of the Health of the National Outcome Scales or HoNOS)[8]
- Can communicate in Thai
- Consent to participate to the study

Potential participants were excluded if platelet count was less than 100,000/cu.mm^3, diagnosed liver cirrhosis or hyperthyroidism or acute hepatitis; have heart disease that requires premedication before oral examination and procedure, and have immunocompromised condition that is in high risk of infection.

Questionnaire survey

A predefined survey consisted of questions related to participant's socio-demographic, alcohol consumption, dental service experience and oral hygiene practice was collected by a trained registered nurse. Face-to-face validation was conducted with the research team.

Clinical examination

A dentist examined dental caries status on a dental chair and recorded number of decayed (D), missing due to decay (M) and filled due to decayed teeth (F) which made up to missing teeth due to decayed and filled teeth (DMFT) index. The oral examination adopted the WHO Oral Health Survey Basic Methods 5th Edition.[9] Intra-rater reliability was 0.88.

Data analysis

STATA version 14.0 was used to manage and analyse data. Descriptive statistics of data collected by questionnaires and oral examination was performed to generate frequencies, percentages, means and standard deviations of variables. Spearman rank, Wilcoxon Rank Sum test and Kruskal Wallis test were conducted to assess categorical data comparison. Multiple logistic regression with significant level of 0.05 were examined.

Results

Average age was 41.8±9.26 years old (ranged between 18-69 years). Average number of teeth in a mouth of participants was 28.12±3.73. Average DMFT was 7.86±6.37. Of these 84% did not received treatment for decayed teeth and 67% lost their teeth due to caries. Table 1 shows caries experience in specific age groups and characteristics of participants. Significant associating factors (p<0.05) to the DMFT were age, diagnostic group, occupation, had dental treatment in the past, having chewing problem, self-perceived oral health status, satisfaction of own oral health, and time of the last clean and scale.

Table 1. Study variables description by caries experience and index.

Variables	Mean (SD)			
	DMFT	DT	MT	FT
Age (years)				
≤30	5.57 (4.11)	4.24 (3.49)	0.67 (1.42)	0.65 (1.66)
31-40	7.56 (5.91)	4.88 (4.31)	2.16 (2.98)	0.53 (1.64)
41-50	7.88 (6.75)	3.97 (3.98)	3.66 (4.25)	0.25 (0.85)
>50	9.80 (7.11)	4.53 (4.49)	4.79 (4.86)	0.48 (1.72)
p value[a]	<0.001	0.203	<0.001	0.033
ICD				
F10.20 alcohol dependence	8.28 (6.47)	4.74 (4.30)	3.04 (3.86)	0.50 (1.51)
F10.50 mental and behavioural disorders due to use of alcohol: psychotic disorder	7.11 (6.14)	3.86 (3.82)	2.93 (4.08)	0.31 (1.29)
p value[b]	0.034	0.025	0.354	0.12
Hospital admission				
1-5	7.62 (6.24)	4.40 (4.10)	2.83 (3.78)	0.39 (1.34)
>5	9.67 (7.08)	4.62 (4.59)	4.31 (4.86)	0.75 (2.01)
p value[a]	0.069	0.504	0.077	0.330
Body mass index (BMI)				
<18.5	7.57 (5.52)	3.82 (3.33)	3.04 (3.98)	0.30 (1.07)
18.5-22.9	8.02 (6.58)	4.68 (4.33)	3.62 (4.45)	0.11 (0.53)
23.0-24.9	7.19 (5.70)	3.95 (3.68)	2.58 (3.71)	0.66 (1.74)
25.0-29.9	8.28 (6.72)	4.57 (4.05)	2.67 (3.57)	1.04 (2.40)
≥30	8.62 (8.94)	4.92 (6.93)	3.00 (3.19)	0.69 (2.21)
p value[c]	0.947	0.940	0.229	<0.001
Highest education				
≤Year 9th	7.45 (5.94)	4.29 (3.93)	2.90 (3.89)	0.25 (1.05)
High school, certificate, diploma	9.28 (7.68)	5.17 (5.12)	3.62 (4.39)	0.49 (1.04)
Bachelor and postgraduate	8.26 (6.46)	3.80 (3.36)	2.37 (3.05)	2.09 (3.28)
p value[c]	0.209	0.579	0.159	<0.001
Occupation				
Unemployed, labour, farmer, others	7.80 (6.34)	4.48 (4.07)	2.99 (4.02)	0.32 (1.27)
Public servant	6.63 (6.16)	2.83 (3.26)	2.47 (3.07)	1.33 (2.29)
Business owner/merchant	10.30 (7.13)	5.43 (5.83)	3.83 (3.60)	1.04 (1.97)
p value[c]	0.047	0.042	0.365	<0.001
Marital status				
Single	8.14 (6.20)	5.08 (4.32)	2.45 (3.58)	0.62 (1.82)
Married	7.86 (6.82)	4.17 (4.11)	3.34 (4.22)	0.36 (1.24)
Widow/divorce/separate	7.32 (5.77)	3.72 (3.76)	3.35 (3.92)	0.25 (0.82)
p value[c]	0.533	0.014	0.052	0.164
Number of family member				
1-2	8.01 (6.32)	4.33 (4.08)	3.19 (3.95)	0.49 (1.67)
3-4	7.87 (6.22)	4.42 (3.98)	3.01 (4.09)	0.44 (1.31)
5-6	7.78 (6.96)	5.02 (4.90)	2.48 (3.24)	0.27 (1.11)
≥7	7.00 (6.30)	2.58 (2.29)	3.68 (4.83)	0.74 (2.10)
p value[a]	0.640	0.830	0.295	0.915

Table 1. Study variables description by caries experience and index (continued).

Variables	Mean (SD)			
	DMFT	DT	MT	FT
Average monthly household income (Baht)				
≤3,000	8.39 (6.00)	4.16 (3.66)	3.76 (4.03)	0.46 (1.69)
3,000-25,000	7.98 (6.54)	4.70 (4.31)	2.98 (4.09)	0.30 (1.09)
>25,000	6.70 (5.83)	3.30 (3.66)	2.28 (2.80)	1.11 (2.28)
p value[a]	0.202	0.714	0.007	<0.001
Having medical condition				
No	7.50 (6.03)	4.27 (3.99)	2.80 (3.54)	0.43 (1.53)
Yes	8.22 (6.69)	4.58 (4.31)	3.20 (4.29)	0.44 (1.33)
p value[b]	0.312	0.526	0.659	0.426
Smoking history				
Never smoke	7.55 (6.70)	4.56 (4.31)	2.30 (3.73)	0.68 (1.88)
Smoked	7.99 (6.24)	4.37 (4.09)	3.28 (4.00)	0.34 (1.19)
p value[b]	0.236	0.705	<0.001	0.016
Quit smoking (years)				
≤5	6.06 (4.81)	3.17 (3.74)	2.06 (1.98)	0.83 (2.15)
>5	7.00 (5.70)	2.14 (2.77)	4.14 (4.15)	0.71 (2.13)
p value[a]	0.606	0.289	0.585	0.299
Cigarette consumption/day				
≤10	8.30 (6.42)	4.55 (4.15)	3.48 (4.19)	0.26 (0.96)
>10	7.47 (5.93)	4.53 (4.06)	2.55 (3.34)	0.39 (1.37)
p value[a]	0.638	0.545	0.111	0.181
Age started smoking (years)				
≤20	7.89 (6.00)	4.36 (4.09)	3.20(3.88)	0.32 (1.17)
>20	8.62 (7.68)	4.40 (4.12)	3.81 (4.68)	0.40 (1.33)
p value[a]	0.184	0.212	0.819	0.108
Age started drinking (years)				
≤20	7.94 (6.42)	4.52 (4.25)	2.97 (3.95)	0.45 (1.46)
>20	7.47 (6.13)	3.95 (3.62)	3.16 (3.88)	0.36 (1.29)
p value[a]	0.600	0.094	0.447	0.372
Alcohol dependence (years)				
≤10	7.65 (6.41)	4.43 (4.09)	2.75 (3.79)	0.46 (1.52)
11-20	8.59 (6.47)	4.73 (4.24)	3.37 (4.22)	0.49 (1.42)
>20	7.36 (5.79)	3.50 (4.36)	3.83 (4.14)	0.03 (0.17)
p value[a]	0.095	0.594	<0.001	0.858
Average daily standard drink				
≤20	8.39 (6.56)	4.32 (3.77)	3.40 (4.44)	0.66 (1.98)
21-40	7.78 (6.36)	4.65 (4.42)	2.78 (3.51)	0.34 (1.13)
>41	7.07 (6.01)	3.82 (3.86)	2.96 (4.29)	0.29 (0.95)
p value[a]	0.355	0.715	0.171	0.047
Highest CIWA scores				
1-7 (less severe)	7.63 (6.48)	4.19 (4.52)	3.11 (3.48)	0.33 (0.97)
8-14 (moderate)	7.95 (6.44)	4.34 (3.90)	3.10 (4.47)	0.51 (1.79)
15-19 (severe)	9.46 (7.41)	5.05 (4.43)	4.05 (5.29)	0.35 (1.37)
≥20 (very severe)	7.76 (5.89)	4.91 (4.16)	2.65 (3.26)	0.20 (0.63)
p value[a]	0.313	0.047	0.439	0.599

Table 1. Study variables description by caries experience and index (continued).

Variables	Mean (SD)			
	DMFT	DT	MT	FT
Number of drinkers in family				
none	7.86 (6.80)	3.93 (3.88)	3.31 (4.61)	0.62 (1.72)
drinker	6.86 (4.39)	3.68 (2.71)	2.27 (2.76)	0.91 (2.00)
Alcohol dependence	7.94 (6.25)	4.79 (4.38)	2.87 (3.55)	0.28 (1.14)
p value[c]	0.842	0.230	0.342	0.679
Received dental services				
No	7.29 (6.50)	4.54 (4.29)	2.75 (4.15)	0
Yes	8.26 (6.26)	4.34 (4.06)	3.17 (3.78)	0.74 (1.81)
p value[b]	0.034	0.833	0.008	NA
Received oral hygiene instruction				
No	7.50 (6.26)	4.43 (4.17)	2.83 (4.06)	0.24 (1.01)
Yes	8.64 (6.56)	4.41 (4.13)	3.37 (3.65)	0.86 (2.02)
p value[b]	0.056	0.830	0.029	<0.001
Number of tooth brushing per day				
1	8.38 (6.36)	4.93 (4.14)	2.71 (3.82)	0.44 (1.55)
2	5.44 (4.75)	4.30 (4.17)	2.97 (3.80)	0.45 (1.47)
≥3	4.33 (4.04)	5.20 (4.09)	3.05 (4.08)	0.36 (1.07)
Irregular	5.33 (2.51)	3.20 (3.42)	7.00 (9.87)	0
p value[c]	0.639	0.254	0.656	0.788
Brush teeth after getting up				
Never	9.50 (10.61)	6.50 (6.36)	2.00 (2.83)	1.00 (1.41)
Sometimes	13.00 (9.97)	5.00 (3.67)	7.80 (9.63)	0.20 (0.45)
Always	7.79 (6.31)	4.41 (4.16)	2.95 (3.82)	0.44 (1.44)
p value[c]	0.210	0.406	0.343	0.389
Brush teeth after having lunch				
Never	7.70 (6.31)	4.42 (4.23)	2.91 (3.78)	0.37 (1.28)
Sometimes	9.25 (6.86)	4.67 (3.83)	3.65 (4.75)	0.94 (2.31)
Always	7.21 (6.00)	3.79 (3.50)	3.10 (4.63)	0.32 (0.94)
p value[c]	0.236	0.624	0.763	0.089
Tooth brushing before bed				
Never	9.24 (6.96)	5.00 (4.59)	3.56 (4.57)	0.68 (1.95)
Sometimes	8.03 (6.13)	4.56 (3.32)	2.64 (4.18)	0.83 (2.34)
Always	7.76 (6.36)	4.38 (4.20)	3.00 (3.88)	0.38 (1.28)
p value[c]	0.446	0.551	0.843	0.614
Types of toothbrush brittle				
Extra soft	10.0 (10.34)	6.71 (8.16)	3.00 (2.94)	0.28 (0.76)
Soft	8.44 (6.74)	4.85 (4.25)	3.06 (4.14)	0.52 (1.43)
Medium	7.24 (6.01)	4.00 (4.00)	2.76 (3.49)	0.48 (1.60)
Hard	8.08 (6.28)	4.39 (3.89)	3.49 (4.70)	0.20 (1.08)
Not sure/do not know	7.62 (4.03)	4.88 (3.31)	2.62 (2.26)	0.12 (0.35)
p value[c]	0.550	0.191	0.967	0.172

Table 1. Study variables description by caries experience and index (continued).

Variables	Mean (SD)			
	DMFT	DT	MT	FT
Types of dentifrice				
Normal	8.05 (7.80)	3.95 (5.49)	3.05 (2.60)	1.05 (2.60)
Fluoride	7.68 (6.33)	4.42 (4.10)	2.92 (3.95)	0.34 (1.19)
Sensitive/gum care	8.14 (4.95)	3.57 (2.88)	2.86 (2.85)	1.71 (3.68)
Herbal	9.36 (6.74)	5.02 (4.10)	4.02 (4.88)	0.31 (0.89)
Do not use dentifrice	7.88 (5.22)	4.00 (4.33)	2.35 (2.76)	1.53 (2.79)
p value[c]	0.530	0.603	0.355	0.016
Have chewing problem				
No	6.54 (5.95)	3.98 (4.18)	2.08 (3.01)	0.47 (1.66)
A bit	8.76 (6.51)	4.72 (4.11)	3.63 (4.38)	0.41 (1.26)
A lot	9.67 (7.50)	5.67 (4.73)	3.33 (3.06)	0.67 (1.15)
p value[c]	<0.001	0.049	<0.001	0.476
Self-perceived oral health				
Very poor	13.78 (8.45)	6.50 (6.70)	6.28 (5.18)	1.00 (1.81)
Poor	10.07 (7.10)	5.98 (4.62)	3.85 (4.58)	0.24 (0.98)
Fair	6.88 (5.50)	3.77 (3.50)	2.64 (3.57)	0.46 (1.44)
Good/very good	5.74 (5.25)	3.36 (3.60)	1.83 (2.68)	0.54 (1.91)
p value[c]	<0.001	<0.001	<0.001	0.080
Satisfaction of own oral health				
Little	10.33 (7.44)	5.92 (5.00)	3.95 (4.61)	0.46 (1.38)
Moderate	6.91 (5.52)	3.80 (3.55)	2.70 (3.61)	0.40 (1.34)
A lot	5.29 (6.15)	3.65 (3.92)	0.82 (1.24)	0.82 (2.90)
p value[c]	<0.001	<0.001	<0.001	0.820
Frequency of having dental visit				
Only when having symptom	8.38 (6.36)	4.37 (4.09)	3.33 (3.90)	0.68 (1.80)
<once a year	5.44 (4.75)	3.56 (3.61)	1.11 (1.45)	0.78 (1.20)
>once a year	4.83 (3.06)	1.83 (1.47)	1.67 (1.86)	1.33 (1.21)
p value[c]	0.126	0.248	0.101	0.045
Did you have dental appointment last year?				
No	7.82 (6.05)	4.21 (3.89)	3.15 (3.96)	0.46 (1.41)
Cannot remember	10.25 (6.02)	3.00 (2.16)	6.75 (5.56)	0.50 (0.58)
Yes	8.76 (6.70)	4.45 (4.39)	3.16 (3.47)	1.14 (2.23)
p value[c]	0.436	0.923	0.160	0.003
Number of dental appointments				
1	7.98 (6.26)	4.46 (4.48)	2.77 (3.25)	0.75 (1.62)
2	11.08 (7.56)	4.62 (4.11)	4.46 (3.91)	2.00 (3.06)
≥3	11.33 (8.03)	4.11 (4.51)	4.33 (4.18)	2.89 (3.69)
p value[c]	0.218	0.893	0.136	0.003
When was the last scale and clean?				
< 6 months	6.93 (5.25)	2.86 (3.30)	2.93 (3.54)	1.14 (1.96)
6-12 months	9.36 (7.24)	4.54 (4.15)	3.41 (4.32)	1.41 (2.63)
Never	9.41 (5.91)	4.88 (4.06)	3.98 (4.02)	0.55 (1.53)
>12 months/cannot remember	4.92 (5.52)	3.06 (3.84)	1.39 (2.24)	0.46 (1.37)
p value[c]	<0.001	0.003	<0.001	0.012

***Note**: [a]Spearman Rank, [b]Wilcoxon Rank Sum Test, [c]Kruskal Wallis Test, NA: not available.*

Multiple logistic regression analysis found that age, diagnostic group, self-perceived oral health, satisfaction of own oral health are significant predicting factors of DMFT. While predicting factors of DT were diagnostic group, marital status, self-perceive oral health and satisfaction of own oral health. Predicting factors of MT were age, smoking, had dental treatment in the past and self-perceived oral health whereas those received oral health education and educational level, and BMI were significant factors of FT (Table 2).

Table 2. Multiple logistic regression analysis results with selected variables.

Dependent variable	Associating factors	Adjusted β (95%CI)	*p* value
DMFT	Age (years)	0.10 (0.04, 0.16)	0.001
(R^2 12.2)	Diagnosed with ICD F10.20	1.17 (0.01, 2.33)	0.049
	Self-perceived oral health	-1.66 (-2.56, -0.75)	<0.001
	Satisfaction of own oral health	-1.64 (-2.97, -0.31)	0.031
DT	Diagnosed with ICD F10.20	1.01 (0.22,1.79)	0.012
(R^2 9.3)	Marital status		
	- Single	1.00	
	- Married	-1.03 (-1.87, -0.19)	0.016
	- Widow/divorce/separate	-1.34 (-2.35, -0.34)	0.009
	Self-perceived oral health	-0.91 (-1.50, -0.31)	0.003
	Satisfaction of own oral health	-0.94 (-1.82, -0.06)	0.036
MT	Age (years)	0.14 (0.11, 0.18)	<0.001
(R^2 18.1)	Smoking	0.94 (0.20, 1.68)	0.013
	Received dental services	0.72 (0.04, 1.40)	0.036
	Self-perceived oral health	-0.97 (-1.41, -0.53)	<0.001
FT	Received oral hygiene instruction	0.39 (0.10, 0.67)	0.007
(R^2 18.1)	Highest education	0.56 (0.34, 0.77)	<0.001
	Body Mass Index	0.05 (0.01, 0.08)	0.013

***Note**: DMFT: decayed/missing/filled teeth, DT: decayed teeth, MT: missing teeth, FT: filled teeth.*

Discussion

Thailand had a high alcohol consumption per capita and was at the 3rd rank in Asia followed South Korea and Japan. Alcohol consumption per capita in South Korea, Japan and Thailand were 12.3 mL, 7.2 mL and 7.1 mL, respectively.[10] Based on the Thai national survey of drinking behaviour and psychiatric disorders in 2008, 5.3 million individuals (10.9%) had excess alcohol consumption with the highest rate in the North as 13.3%.[11] This statistic was consistent with the 2017 reports by the Center of Alcohol Studies and found that the proportion of current drinkers was highest in the North at 35.4% compared to the national average at 28.6%.[12,13]

This cross-sectional study is the first survey examining dental caries status of alcohol dependences with and without mental and behavioural disorders at a large psychiatric hospital in the North of Thailand. Study participants were recruited from male in-patient ward because over 90% of alcohol dependence patients at the study site were male. The large proportion of male was also reported in other Thai and international studies.[14-19] Our findings confirmed previous evidence that alcohol dependence related to socioeconomic disadvantages e.g. highest education was primary school, income was 5,000-30,000 Baht and labour.[10,19,20]

Most of participants (70.9%) were ex-smokers, 63.8% were current smokers. It is common that alcohol drinkers were also smokers.[21] Participants of our study started drinking at 17.93 years old and smoking at 17.06 year old, which confirmed the previous findings that found significant relationship between drinking and smoking behaviour.[22] The Thai National Mental Health Survey 2013 found that the younger an individual start drinking, the higher risk of bad alcohol consumption behaviour. Having family member(s) who was(were) drinker can increase the risk of having bad alcohol consumption behaviour up to 4.1 times.[23]

Study participants had significantly poorer dental caries status. Participants aged 35-44 years having decay tooth/teeth was DMFT 7.31±5.96 among participants compared to DMFT 6.0±5.14 national average and DMFT 5.8±4.85 regional average. The percentage of individual having decay tooth/teeth in participants was 2.4 times higher than the national average (84.7% vs 35.2%).

While the poor oral health was a prominent problem in this population, dental service utilisation was low as 41.3% never get dental services and 68.7% never received oral hygiene instruction. Untreated dental caries was found in 84.2% total decayed teeth. Previous study found that alcohol dependence had poor hygiene practice.[6] However, 80.7% participants in our study brushed their teeth twice a day.

Participants who were diagnosed as alcohol dependence and single had a greater number of decayed teeth compared

to others. Evidence showed that alcohol and medication for alcohol dependence can cause dry mouth and reduce salivary flow which increased risk of dental caries development and progress.[24] Our result found that DMFT significantly related to age, satisfaction to own oral health and self-perceived oral health.

Conclusion

The findings suggested that dental caries prevention program should focus on alcohol dependences who are single given their highest dental caries prevalence. Dental caries prevention and promotion program and oral hygiene instruction should be implemented as soon as a person being diagnosed as alcohol dependence with an emphasis on improving self-perception and satisfaction of own oral health.

Although the participants can represent Thai alcohol dependence population, the current study design allowed us to only explore association of dental caries and its determinants using collected data at a certain time. Longitudinal studies are required if researchers wish to learn more about causal mechanism and observe the pattern of changes.

Ethical approval

Ethics approval was granted by the Human Ethics Committee, Suanprung hospital (study ID 21/2015).

Funding

This study was funded by the Mental Health Department, Ministry of Public Health Research Grant.

Conflict of interest

Authors declared no conflict of interest to declare.

CRedit authorship contribution statement

Namsukh Kongkalai: proposal developing, funding application, research conduct, data analysis, manuscript preparing and approve; **Praporpan Utamachote**: research performing under supervision; **Supatra Sang-in**: data analysis and approve.

Acknowledgements

This study was funded by the Mental Health Department, Ministry of Public Health Research Grant. We would like to thank study participants, Suanprung Psychiatric Hospital staff who involved and supported this study. Special thanks to Dr. Paranee Chawalwut, Associate Professor Anon Jaruakkara and Dr. Chitchanok Ruengon for their guidance and support.

References

[1] World Health Organisation. Thailand: Alcohol consumption levels and patterns. In: Organisation WH, editor https://www.who.int/substance_abuse/publications/global_alcohol_report/profiles/tha.pdf?ua=12016.

[2] International Health Policy Program. Alcohol consumption and impacts in Thailand. Nonthaburi, Thailand: Ministry of Public Health; 2013.

[3] Wigun S. Experiences of recurrence among persons with alcohol induced psychotic disorder. Chaingmai: Chaingmai University; 2005.

[4] Sakki TK, Knuuttila ML, Vimpari SS, Hartikainen MS. Association of lifestyle with periodontal health. Community Dent Oral Epidemiol. 1995; 23(3): 155-8. doi: 10.1111/j.1600-0528.1995.tb00220.x.

[5] Novacek G, Plachetzky U, Potzi R, Lentner S, Slavicek R, Gangl A, et al. Dental and periodontal disease in patients with cirrhosis--role of etiology of liver disease. J Hepatol. 1995 ; 22(5): 576-82. doi: 10.1016/ 0168-8278(95)80453-6.

[6] Hornecker E, Muuss T, Ehrenreich H, Mausberg RF. A pilot study on the oral conditions of severely alcohol addicted persons. J Contemp Dent Pract. 2003; 15(4): 51-9.

[7] Sullivan JT, Sykora K, Schneiderman J, Naranjo CA, Sellers EM. Assessment of alcohol withdrawal: the revised clinical institute withdrawal assessment for alcohol scale (CIWA-Ar). Br J Addict. 1989; 84(11): 1353-7. doi: 10.1111/j.1360-0443.1989.tb00737.x.

[8] Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A. Health of the Nation Outcome Scales (HoNOS). Research and development. Br J Psychiatry. 1998 Jan:172:11-8. doi: 10.1192/bjp.172.1.11.World Health Organisation. Oral Health Surveys-Basic Methods. Geneva: World Health Organisation; 2013.

[9] Sookrung K, Prayoon P, San-In N. Rueng Lao Gao 10: outstanding studies in the past decade and policy change to reduce alcohol problem in Thailand. Bangkok: Center for Alcohol Studies; 2014.

[10] Kittiranttanapaiboon P, Chansirimongkol B, Kamwongpin M, Jutha W, Kongsook T, Leejongpermpoon J. Prevalence of behaviour disturbance related to alcohol consumption and psychiatric comorbidity in Thai population: Thai national epidemiology study 2008. The 6th National Alcohol Conference; Bangkok, 2010.

[11] Wichaidit W, McNeil E, Saingam D, Assanagkornchai S. Alcohol consumption in Thai society 2017. Bangkok: Center for Alcohol Studies; 2017.

[12] Vichitkunakorn P, Tanaree A. Provincial alcohol report 2017. Bangkok: Center for Alcohol Studies; 2017.

[13] Teeson M, Hall W, Salade T, Mills K, Grove R, Mewton L, et al. Prevalence and correlates of DSM-IV alcohol abuse and dependence in Australia: findings of the 2007 National Survey of Mental Health and Wellbeing. Addiction. 2010; 105(12): 2085-94. doi: 10.1111/j.1360-0443.2010.03096.x.

[14] Oakley Browne MA, Wells JE, Scott KM, McGee MA. Lifetime prevalence and projected lifetime risk of DSM-IV disorders in Te Rau Hinengaro: the New Zealand Mental Health Survey. Aust NZ J Psychiatry. 2006; 40(10): 865-74. doi: 10.1080/j.1440-1614.2006.019.

[15] Bromet EJ, Gluzman SF, Paniotto VI, Webb CP, Tintle NL, Zakhozha V, et al. Epidemiology of psychiatric and alcohol disorders in Ukraine: findings from the Ukraine World Mental Health survey. Soc Psychiatry Psychiatr Epidemiol. 2005; 40(9): 681-90. doi: 10.1007/

s00127-005-0927-9.

[16] Vicente B, Kohn R, Rioseco P, Saldivia S, Levav I, Torres S. Lifetime and 12-month prevalence of DSM-III-R disorders in the Chile psychiatric prevalence study. Am J Psychiatry. 2006 Aug;163(8):1362-70. doi: 10.1176/ajp.2006.163.8.1362.

[17] National Statistical Office. The smoking and drinking behaviour survey 2011. Bangkok: National Statistical Office, Ministry of Information and Communicaiton; 2015.

[18] Center for Alcohol Studies. Alcohol consumption and its impact in Thailand 2013. Nonthaburi: Center for Alcohol Studies; 2013.

[19] Witvorapong N, Ratisukpimol W, Tharnpanich N, Lephilibert N, Sianphanit C. A decade of Center for Alcohol Studies: knowledge for controlling alcohol consumption. Witvorapong N, editor. Bangkok: Rong Pim Duen Tula; 2014.

[20] Gulliver SB, Kamholz BW, Helstrom AW. Smoking cessation and alcohol abstinence: What do the data tell us? Alcohol Res Health. 2006; 29: 208-12. PMID: 17373411, PMCID: PMC6527036.

[21] Jansson L. Association between alcohol consumption and dental health. J Clin Periodontol. 2008; 35(5): 379-84. doi: 10.1111/j.1600-051X.2008.01210.x.

[22] Silapakit P, Somjai C, Thammawong P, Chutha W, Wongwan P, Puangmalai K. Thai National Mental Health Survey 2013: prevalence and factors related to alcohol use disorders in the North Chiangmai: Suanprung Psychiatric Hospital; 2016.

[23] Rooban T, Vidya K, Joshua E, Rao A, Ranganathan S, Rao UK, et al. Tooth decay in alcohol and tobacco abusers. JOMFP. 2011; 15(1): 14-21. doi: 10.4103/0973-029X.80032.

Journal of Associated Medical Sciences 2025; 58(3): 237-247

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 performance evaluation of the Dirui MUS3600 and FUS3000Plus automated urine analyzers utilized in the AMS CMU EQA unit's urinalysis proficiency testing program

Tanyarat Jomgeow*, Jirapat Narkdee, Nattawinan Piewlueng

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

ARTICLE INFO

Article history:
Received 6 February 2025
Accepted as revised 3 July 2025
Available online 9 July 2025

Keywords:
Urinalysis analyzer, automated urinalysis, Dirui model, MUS3600, FUS3000Plus, method validation

ABSTRACT

Background: Urinalysis is one of the essential laboratory tests for health checks, which requires highly skilled personnel. In medical laboratories, automated urine analyzers are crucial for reducing workloads compared to human methods. However, quality control is necessary to ensure their accuracy and precision. The Dirui MUS3600 and FUS3000Plus automated urinalysis analyzer models are also interested in taking part in the proficiency testing program, and the AMS CMU EQA unit has developed urine control materials for the Thai medical laboratory's urinalysis EQA program.

Objectives: The goal of this study was to evaluate the performance of the Dirui 3600 and FUS3000Plus urine analyzer models for usage in the AMS CMU EQA unit laboratory using method validation items.

Materials and methods: Total of 724 urine samples were collected, and three in-house urine control levels were prepared. Imprecision, accuracy, comparison, and diagnostic performance tests were determined by using MUS3600 and FUS3000Plus compared with the microscopic examination as the standard method.

Results: Both automated models provided excellent results of %CV of within day and between day running shows less than 10%. The agreement between automation and standard methods in physical, chemical, and sediment evaluation was 70-100%. The correlation coefficient for the RBC and WBC parameters compared with the manual microscope method ranged from r=0.88 to 0.93. Linearity results of both models show r=0.99 for RBCs and WBCs, respectively. Results of carry-over also show good reliability results below 0.005%. Moreover, the results of sensitivity and specificity in important urinary diseases indicate sediment for example RBC and WBC show more than 83% of overall results.

Conclusion: The Dirui MUS3600 and FUS3000Plus models of urinalysis analyzers yielded good results, indicating that they are suitable for use in the future for the collection and preparation of specimens for QC material used in the proficiency program that should follow the standard procedure for EQA participants. In addition, both models can produce good urinalysis results in clinical laboratories.

Introduction

Urinalysis is one of the most requested laboratory tests used for diagnosing kidney diseases and other disorders. Recently, automated urinalysis analyzers were used instead of manual methods for better performance and shorter turnaround times, especially in large-scale labs.[1,2] However, a good quality control strategy is necessary for the process in every test from automation. Even though more clinical laboratories in Thailand are

* **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: tanyarat.jomgeow@cmu.ac.th
doi: 10.12982/JAMS.2025.093
E-ISSN: 2539-6056

using urine analyzers, the automated results are not entirely trustworthy. Numerous investigations reveal that all the automated physical, chemical, and sediment evaluation results contain mistakes, so skilled personnel must carefully verify and validate the results of all process especially confirmed sediment evaluation under manual microscope.[3,4,5]

For these reasons, in 2020, we launched the urinalysis proficiency testing program under the AMS CMU EQA unit, Faculty of Associated Medical Sciences, Chiang Mai University. We provide 2 rounds a year with artificial urine to process all standard physical, chemical, and microscopic examination procedures. We currently have over 500 participants in Thailand for routine urinalysis proficiency testing. Our program's urine sediment quality control materials are made from patient urine samples that were discarded, and the sediments were treated using our proprietary formula fixative solution. Since more than 1,200 samples were collected for sediment control material preparation in each round of proficiency testing, the evaluation of the specimens by the manual method created an excessive workload for staff. Additionally, certain participants in the proficiency program employ an automated urine analyzer. Because the manual approach alone does not meet automation performance criteria for all participants, our lab needs to automate all procedures. Since the huge sediment collection process is essential in our lab, we are interested in the imaging technology of the automated urinalysis analyzer, which could deliver accurate and clear screening images of sediment for the collection process.

There are a variety of automated urinalysis analyzers used in clinical laboratories. Almost all chemical examinations use the principle of the photoelectric colorimetric method, whereas sediment evaluation uses imaging technology or the flow cytometry principle.[6,7] Several studies reported the advantages of imaging technology in the screening of sediment evaluation.[8,9,10] Cho EJ *et al*. revealed that a combination using automated urinalysis systems based on flow cytometry or digital imaging techniques could efficiently replace manual microscopic examination.[11] This research team also reported on the comparison of 5 different principles of automated urinalysis analyzers that could reduce the manual process when utilized in an appropriate way.[12] The comparative study of Dirui FUS200 and Urised 3 with manual microscopic examination reported by Yalcinkaya E. *et al*., in 2019 shows strong correlations existed between FUS200 and manual microscopy and are helpful for the diagnosis of pathological urine samples.[13] Moreover, according to a study by Benovska M. *et al*. in 2018, the FUS-2000 hybrid analyzer's microscopic component performs well analytically and closely matches light microscopy of urine sediment.[14]

Based on the earlier articles, we are considering the new Dirui instrument versions. The first model is the FUS3000Plus, a hybrid model with all-in-one modules, and the other one is the MUS3600 model, which integrates two modules of an automated urine analyzer. The instruments contain similar principles in the part of the microscopic examination with FUS2000, which had been shown good analytical performance of the analyzer and highly agreement with light microscopy of urine sediment.[14] Both models are based on dry chemistry: photoelectric colorimetry and flow-type micro-imaging technology covering 15 chemistry parameters and urine sediment image recognition, respectively.[15] To evaluate the performance of the Dirui MUS3600 and FUS3000Plus automated urine analyzers for future use in quality control sample processing in proficiency programs, we present the method validation items here, including imprecision, accuracy, comparison, and diagnostic performance tests.

Materials and methods

Sample collection and process

A total of 724 urine samples were collected from the Laboratory of Clinical Microscopy and the Laboratory of Kidney Disease Unit, Maharaj Nakorn Chiang Mai Hospital, from June 2023 to March 2024. The ethical committee of the Faculty of Associated Medical Sciences, Chiang Mai University, has permitted ethical proof documents (AMSEC-67EM-012). Collected urine samples were separated into two groups. The first group was used for the method evaluation test items, whereas the second group was used to prepare in-house sediment control material for sediment evaluation (3 levels: low, medium, and high level). The results of physical examination of control materials were identified with 3 medical technologist experts while the chemical examination confirmed by chemistry analyzer (VITROS®XT 7600 Integrated system) and cell count were confirming by hematologic analyzer (Sysmex XT1000i). After determining the urinalysis were done by using MUS3600 and FUS3000Plus, urine sediment was collected, fixed and stored at 2-8 °C. The sediment was evaluated under a microscope by using wet preparation of 22x22 mm. cover glass (number/HPF or number/LPF) and the improved Neubauer hemacytometer (cells/µL) by 3 expert medical laboratory staffs. Specimens that were stored for more than 2 hrs at room temperature or specimens with turbid containing numerous amorphous, debris, squamous epitheliums or microorganisms that obscured vision to detect sediment type clearly under the microscope were excluded.

Imprecision test

Method validation was done based on ICH guideline Q2 (R2): validation of analytical procedure, 2023.[16] Imprecision tests of physical, chemical, and sediment evaluation of Dirui MUS3600 and FUS3000Plus were done using in-house control materials (Table 1). The reproducibility of each control level was determined by analyzing 20 times within-day running and 20 times between-days running. Imprecision tests of physical and chemical examination results have been calculated by the agreement of results using % concordance of qualitative or semi-quantitative results. The imprecision of WBCs and RBCs evaluation was analyzed by mean, SD, and %CV by using sediment that confirmed concentration by counting

with an improved Neubauer hemacytometer (cells/µL) before determining by automation. The acceptance of imprecision in physical and chemical examination is more than 80% concordance, while in RBCs and WBCs parameters accepted by %CV are less than 10.

Table 1. Characteristics of in-house three-level urine control materials.

Urinalysis parameters		L1: Low level	L2: Medium level	L3: High level
Physical examination	Color	Yellow	Amber/brown/green	Orange/red
	Appearance	Clear	Slightly cloudy	Cloudy/heavy cloudy
	pH	5.0-6.5	6.6-7.5	7.6-8.5
	Specific gravity	1.001-1.010	1.011-1.020	1.021-1.030
Chemical examination	Protein	Negative (0 mg/dL)	0.1-0.3 gm/L	1.0-3.0 gm/L
	Glucose	Negative (0 mg/dL)	100-200 mg/dL	500-1,000 mg/dL
	Blood	Negative	5-10 cells/HPF	100-200 cells/HPF
	Nitrite	Negative	Positive	NA
	Leukocyte esterase	Negative	50-100 cells/HPF	20-30 cells/HPF
	Ketone	Negative	Positive	NA
	Urobilinogen	Normal	Abnormal	NA
	Bilirubin	Negative	Positive	NA
Microscopic examination	WBCs	Lot 1: 0-1 cells/HPF	Lot 1: 5-10 cells/HPF	Lot 1: 20-30 cells/HPF
		Lot 2: 0-1 cells/HPF	Lot 2: 30-50 cells/HPF	Lot 2: >100 cells/HPF
	RBCs	Lot 1: 0-1 cells/HPF	Lot 1: 5-10 cells/HPF	Lot 1: 20-30 cells/HPF
		Lot 2: 0-1 cells/HPF	Lot 2: >100 cells/HPF	Lot 2: >100 cells/HPF
	Epithelial cells	0-1 cell/HPF	5-10 cells/HPF	NA
	Yeast/budding yeast	0-1 cell/HPF	5-10 cells/HPF	NA
	Cast	Negative	0-1/LPF	NA
	Crystals	Negative	0-1/LPF	NA

Note: NA: not available

Accuracy test

Recovery spike tests were done by using non-spiked samples at the concentration of RBCs (350, 700, 1,500, 2,500, 5,000, 10,000 cells/µL) and WBCs (100, 200, 400, 800, 1,000 cells/µL), which were prepared by counting with an improved Neubauer hemacytometer (cells/µL). Then spiked samples were created by adding equal volumes of non-spiked samples, making 2 folds of each concentration (100% added concentration) as shown in Table 2. Then 5 spiked samples were evaluated by Dirui MUS3600 and FUS3000Plus before calculating % recovery. The formula used to calculate the %recovery is

% recovery = (concentration of spiked sample - concentration of non-spiked sample) x 100 concentration of added sample

The acceptance of % recovery is between 90-110.

Table 2. Accuracy assessment of the Dirui MUS3600 and FUS3000Plus analyzers based on the percent recovery (% recovery) of spiked urine samples. Acceptable recovery is defined within the range of 90% to 110%, according to standard validation criteria (N=20).

Parameters	Dirui MUS3600				Dirui FUS3000Plus			
	Non-spiked concentration	Spiked concentration	Added concentration	%Recovery	Non-spiked concentration	Spiked concentration	Added concentration	%Recovery
RBCs	5,000	9,783	5,000	95.66	5,000	10,150	5,000	103
	2,500	4,793	2,500	91.72	2,500	4,815	2,500	92.60
	1,250	2,461	1,250	96.88	1,250	2,572	1,250	105.76
	700	1,456	700	108	700	1,368	700	95.43
	350	725	350	107.14	350	696	350	98.86
WBCs	1,000	2,044	1,000	104.40	1,000	1,960	1,000	96
	800	1,540	800	92.50	800	1,615	800	101.88
	400	791	400	97.75	400	818	400	104.50
	200	417	200	108.50	200	411	200	105.50
	100	203	100	103	100	195	100	95

Comparison test

Urine samples (N=40) were determined twice, and physical, chemical, and sediment examination results were compared between two Dirui automation models, the Siemens strip reader (Clinitek Advantus® urine chemistry analyzer) as a current instrument and the manual microscopic examination as the standard method. The percentage of concordance and the correlation coefficient between methods were calculated. The acceptance of % concordance is the agreement of the result between 2 methods within ±1 different grade.

Linearity test

High concentrations of WBCs and RBCs were diluted into 5 dilutions (0-10,000 cells/cu.mm for RBCs and 0-2,500 cells/cu.mm. for WBCs), and each dilution was determined by duplicate sediment analysis by both automation models. Then the linear regression of detected results and expected concentration was analyzed.

Carryover test

Triplicate high control levels (H1, H2, H3) and triplicate low control levels (L1, L2, L3) were determined by both Dirui automation models. Quantitative parameters, RBCs and WBCs, were calculated %carryover. The percentage of carryover was calculated by using the formula:

$$\% \text{ Carryover} = (L1-L3)/(H3-L3) \times 100.$$

The acceptance of % carryover is not over than 0.01%.

Diagnostic performance test

Urine sediments (N=270) were collected and identified. The identified sediment type was separated into 14 categories containing normal and pathognomonic sediment types: RBC, WBC, non-squamous epithelial cells, bacteria, yeast cells or budding yeast or pseudo-hyphae, infectious crystal (struvite), RBC cast, dysmorphic RBC, calcium oxalate crystal, uric acid crystal, fatty cast, hyaline cast, granular cast, and waxy cast**.** Then each group was evaluated by 3 expert laboratory staffs under a microscope and compared with results detected by MUS3600 and FUS3000Plus. Then, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated.

Statistical analysis

Microsoft Excel 2016 was used to analyze calculations of the mean, SD, %CV, %agreement, %carryover, linear regression, correlation coefficient, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

Results

The within-day and between-day running results for the imprecision test of MUS3600 and FUS3000Plus were done using the three levels of in-house control materials shown in Table 3. For quantitative analysis available parameters, the %coefficient of variation of RBCs or WBCs evaluation shows from 1.97 to 4.71. in lot 1 (low to high level) and from 3.79 to 9.23 in lot 2 (low to extremely high level), while the qualitative and semi-quantitative parameters as physical and chemical examination results show 85-100% concordance in overall parameters (data not shown). The accuracy test results of MUS3600 and FUS3000Plus in RBCS and WBCs detection are shown in Table 2. The percentage of recovery tests in spiked samples displayed an acceptable range from 91.72 to 108.50%.

Table 3. Imprecision analysis of urinary sediment evaluation by the Dirui MUS3600 and FUS3000Plus analyzers.

QC materials			%CV of within-day running		%CV of between-day running	
			MUS3600 (N=20)	FUS3000Plus (N=20)	MUS3600 (N=20)	FUS3000Plus (N=20)
Lot 1	Level 1	RBCs	0	0	0	0
		WBCs	0	0	0	0
	Level 2	RBCs	2.38	1.93	3.49	2.97
		WBCs	1.97	2.07	3.64	3.75
	Level 3	RBCs	2.77	2.35	4.28	4.71
		WBCs	2.85	2.23	3.91	4.22
Lot 2	Level 1	RBCs	0	0	0	0
		WBCs	0	0	0	0
	Level 2	RBCs	3.99	3.79	4.06	5.48
		WBCs	4.73	6.01	7.66	9.23
	Level 3	RBCs	4.05	3.53	7.52	8.96
		WBCs	5.21	8.18	8.43	8.71

Note: The coefficient of variation (%CV) was calculated for both within-day (N=20) and between-day (N=20) runs to assess reproducibility.

The comparison test revealed that a chemical analysis of 40 urine samples revealed 92.5-100% agreement between the Siemen strip reader and the two Dirui automation models. The percentage concordance within one grade difference ranges from 81.4 to 92.4%, while the correlation coefficient for the RBC and WBC parameters ranges from r=0.88 to 0.93. Other sediment parameters, on the other hand, exhibit correlation coefficients from r=0.65 to 0.83 and percentage concordance within one grade difference ranging from 65.6 to 86.6% (Tables 4, 5, and 6).

With five dilutions of RBC and WBC samples, MUS3600 and FUS3000Plus were used to assess the concentration detectable performance of tests on the various specified concentrations by automation. MUS3600 and FUS3000Plus offer linearity concentration at r=0.99 for RBCs and WBCs, respectively, according to the data displayed in Figure 1A-1D. Carry-over testing was used to demonstrate that a high concentration of the prior sample could not contaminate the subsequent sample; the overall percentage carry-over for both models is 0.001-0.003% (data not shown).

Table 4. The Comparison test in % concordance of chemical examination and sediment evaluation from Dirui MUS3600 and FUS3000Plus compared with the previously used equipment, Siemen strip reader (duplicate experiments).

Chemical parameters	**%Concordance of detection when compared automation results with the Siemen strip reader (N=40)**	
	Dirui MUS3600	Dirui FUS3000Plus
pH	95	97.5
Specific gravity	97.5	100
Protein	97.5	100
Glucose	100	100
Blood	92.5	95
Leukocyte esterase	95	97.5
Nitrite	100	100
Ketone	97.5	95
Bilirubin	100	100
Urobilinogen	100	100

Table 5. Percent concordance of urine sediment evaluation results between automated analyzers (Dirui MUS3600 and FUS3000Plus) and manual microscopic examination (N=619). Comparison of sediment parameters evaluated by automated systems versus manual microscopy.

Sediment parameters	**%Concordance of sediment evaluation when compared automation results with microscopic examination(N=619)**	
	Dirui MUS3600	Dirui FUS3000Plus
RBCs	92.4	89.5
WBCs	82.7	81.4
Squamous epithelial cells	84.7	86.6
Non-squamous epithelial cells	77.1	73.2
Hyaline casts	72.7	69.6
Granular casts	65.9	67.0
Crystals	72.9	71.7
Bacteria	78.0	79.6
Yeast cells/budding yeast	65.6	66.4
Sperm	76.7	80.1

Note: Percent concordance refers to the agreement between results obtained from the automated analyzers and those from the reference microscopic method.

Table 6. Correlation coefficients (r) of urinary sediment evaluation results obtained from the Dirui MUS3600 and FUS3000Plus analyzers compared with manual microscopic examination performed by expert medical technologists.

Sediment parameters	Correlation coefficient (r) when compared automation results with microscopic examination (N=619), 95% Confidence intervals	
	Dirui MUS3600	Dirui FUS3000Plus
RBCs	0.93	0.91
WBCs	0.88	0.87
Squamous epithelial cells	0.72	0.75
Non-squamous epithelial cells	0.82	0.78
Hyaline casts	0.67	0.73
Granular casts	0.71	0.65
Crystals	0.69	0.71
Bacteria	0.75	0.83
Yeast cells/budding yeast	0.75	0.74
Sperm	0.77	0.72

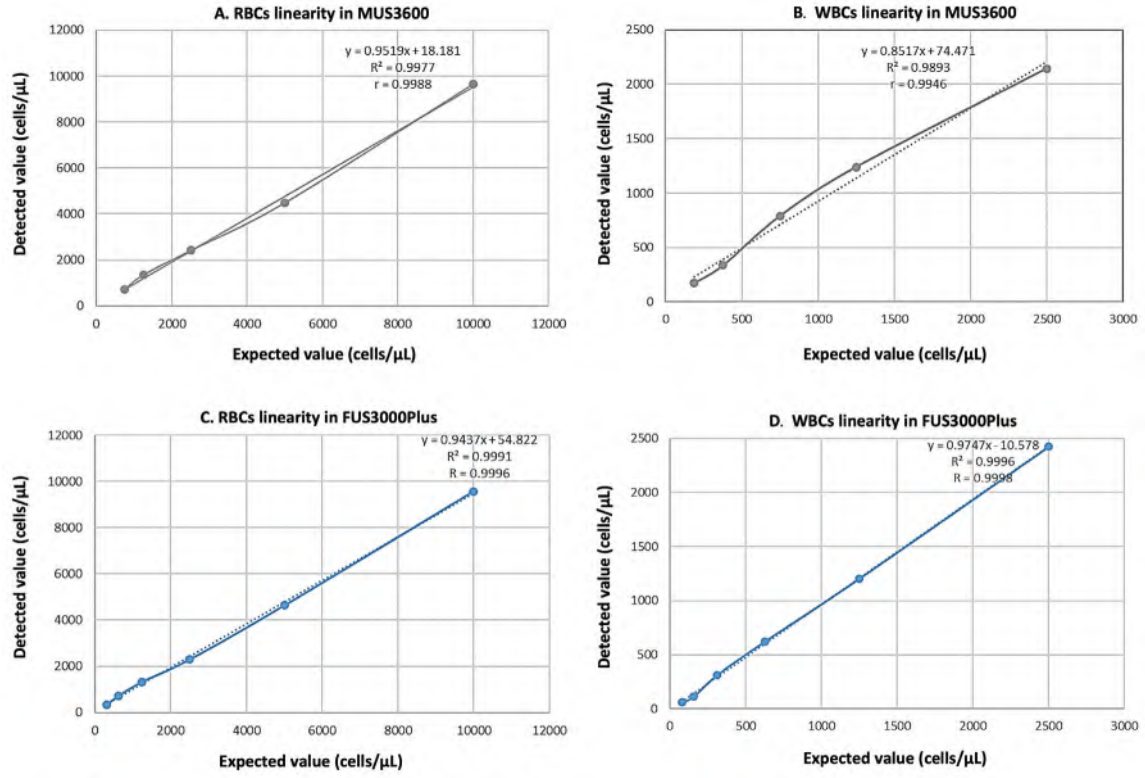


Figure 1. Linearity analysis of red blood cells (RBCs) and white blood cells (WBCs) using the Dirui MUS3600 and FUS3000Plus automated urine analyzers. A and B: linear correlation for RBCs and WBCs, respectively, on the MUS3600, C and D: corresponding linearity for the FUS3000Plus

Prior to testing the diagnosis performance of sediment evaluation with both automation models, 3 expert MTs classified urine sediments, and Table 7 displayed the percentages of sensitivity, specificity, PPV, and NPV. Overall sediment group sensitivity and specificity ranged from 50.0 to 100%. WBCs, RBCs, and non-squamous epithelial cells show over 80% of cellular sediments. Good specificity and sensitivity, above 80% and 70% respectively, were demonstrated in the microorganism group. The results indicate that the pathological cast has 80-100% specificity and about 60% sensitivity, whereas the hyaline cast has good sensitivity and specificity. Calcium oxalate crystals exhibit high sensitivity and specificity, whereas infectious crystals (struvite) and uric acid crystals only exhibit 60% sensitivity and over 95% specificity.

Table 7. Diagnostic performance evaluation of the Dirui MUS3600 and FUS3000Plus analyzers in the detection of urinary sediment, compared with manual microscopic examination as the reference standard.

Sediment parameters	**MUS3600**				**FUS3000Plus**				**N=270**
	%sensitivity	%specificity	PPV	NPV	%sensitivity	%specificity	PPV	NPV	
RBC	92.22	91.94	94.32	89.06	92.31	95.08	96.55	89.23	152
WBC	96.43	83.33	98.54	66.67	96.38	85.71	98.52	70.59	
Non-squamous epithelial cells	80.30	81.33	79.10	82.43	79.10	82.35	77.94	83.33	
Bacteria	72.73	97.40	96.55	78.13	73.75	97.22	96.72	76.92	
Yeast cells or budding yeast or pseudo-hyphae	71.43	87.38	72.92	86.54	71.74	86.79	70.21	87.62	
Infectious crystal (struvite)	64.00	98.43	88.89	93.28	62.50	97.66	83.33	93.28	
RBC cast	66.67	100.00	100.00	95.83	66.67	100.00	100.00	95.83	26
Dysmorphic RBC	41.18	88.89	87.50	44.44	58.82	88.89	90.91	53.33	
Calcium oxalate crystal	95.45	84.62	91.30	91.67	95.00	80.00	86.36	92.31	35
Uric acid crystal	64.29	95.24	90.00	80.00	64.29	90.48	81.82	79.17	
Fatty cast	50.00	100.00	100.00	80.00	50.00	100.00	100.00	80.00	12
Hyaline cast	81.48	83.33	88.00	75.00	80.00	75.00	80.00	75.00	45
Granular cast	66.67	84.62	75.00	78.57	66.67	80.77	70.59	77.78	
Waxy cast	57.14	97.37	80.00	92.50	60.00	97.37	75.00	94.87	

Discussion

The analytical performance of automated urine analyzers has been documented in some investigations using method evaluation items. The manufacturer's QC materials are utilized as control materials in the general method evaluation. Certain elements of the QC material, such as the type of sediment, the quantity of chemicals, or the concentration of sediments, cannot provide enough information. In this study, we create several levels of internal quality control materials to be used in the automated urinalysis analyzer technique assessment. These QC materials are available in various appearances for physical, chemical, and sediment concentrations from low to very high levels. They are appropriate for all aspects of method evaluation, particularly imprecision, accuracy, linearity, and carry-over testing.

To verify the stability and homogeneity of the control material in the future, it is first essential to prove the imprecision test. Since the findings of the physical and chemical tests are qualitative and semi-quantitative, the agreement percentage should be applied to these criteria. Upon physical examination, both models' within-day and between-day runs exhibited 100% concordance in terms of color and appearance found in normal and also abnormal urine (data not shown). Although there is no general research on these two characteristics, our lab thinks they are important for participant proficiency testing and also influence clinical changes in patients, such as hematuria or pyuria, which are manifestations of glomerular diseases and show changes in the color and appearance of urine. Yang WS's study investigation confirmed our findings.[17] The author concluded that urine RBC counting using the UF-1000i or Cobas 6500 urine analyzers underestimates the severity of hematuria in glomerular diseases because dysmorphic RBCs are prone to hemolysis and/or are not sufficiently recognized.

The result of specific gravity values exhibits 85–100% concordance while the pH imprecision test likewise indicates 100% concordance. Except for the blood and leukocyte esterase parameters, which are 95% and 90% in agreement, all parameters for the chemical analysis exhibit 100% concordance. We found that several variables are impacted, including sediment debris during the final week of QC material storage, which can lead to inaccurate data. Even when determining the exceptionally high concentrations of RBC and WBC displayed in Table 3, the sediment evaluation of both Dirui models demonstrates good precision at less than 10%. These findings are consistent with other reports in several Dirui automation models, such as the 2019 reports by Kucukgergin C. and staff and Yalcinkaya E. *et al*.[18,19]

We have performed a recovery test using the spiked samples to confirm the test's accuracy. There was a good percentage of recovery shown. Additionally, for inter-laboratory comparison, we introduced our own QC materials for physical, chemical, and sediment examination. The accuracy test results for Dirui MUS3600 and FUS3000Plus in physical and chemical evaluation revealed 90-100% concordance with the target value from professional MT and chemistry processes. Except for the percentage of agreement in the sediment determination of QC material No. 2 from the first round, which was only 66.70% of precise grading, the results show more than 80% of all parameters in both models. Examining this error, we found that MUS3600 models evaluated sediment higher than FUS3000Plus models for a single grade, although the clinical feature difference was insignificant. Since our laboratory's MUS3600 is the only instrument in Thailand that is compared among FUS3000Plus, the mode data for the 15-19 sites in the inter-laboratory comparison group are somewhat higher in MUS3600 but perfectly in agreement with FUS3000Plus. Since this procedure is not typically used for method validation, these results have not been included in this article. But for us, the statistics validated the test's correctness.

The comparison test was analyzed by % concordance, and the correlation coefficient revealed excellent results compared with the strip reader, and that was compared with microscopic examination. Although some parameters, such as cast, crystal, and microorganism, show lower % concordance than cellular components, both automation reports these sediments in unclassified groups to allow users to correct the results. These comparison results suggested that MUS3600 and FUS3000Plus perform as well as the method used in our laboratory.

The linearity test shows reliable results of RBCs and WBCs concentration between expected results and detected results within the wide range of concentration from 0-10,000 cells/µL and 0-2,500 cells/µL for RBCs and WBCs, respectively. The carry-over test of both parameters shows 0.001-0.003%, indicating no interference from previous high-concentration specimens to the next sample. We also determined the carry-over test six months later, and the results also show a carry-over of less than 0.005% (data not shown).

Additionally, we identified the Dirui MUS3600 and FUS3000Plus diagnostic performance tests. In most sediment parameters, the specificity in both models is greater than 80%. On the other hand, the sensitivity is reduced in cast and crystal characteristics and exhibits great results in detecting cellular components. The results for the detection of fatty casts and dysmorphic RBCs, which have only 41.18% and 50% specificity, respectively, may be due to the sediment types of high light reflection and irregular shape. However, there are some reports regarding the performance of the Dirui FUS-200 analyzer. For instance, Yuksel *et al*. found that the FUS-100's sensitivity for RBCs and WBCs is 73% and 68%, respectively, and Kocer D. *et al*. found that the analytical sensitivity for bacterium recognition and quantification was insufficient.[19,20] Moreover, the study reported by Bartosova K et al. also suggested that a manual microscope is still required for sediment confirmation in a variety of automation models.[21] It was also found that the diagnostic performance improved upon the Dirui MUS3600 and FUS3000Plus's imaging software update version.

There were the several limitations of this study: first, the automation was unable to recognize certain large-sized

sediment kinds or when it is clumping, such as uric acid crystal, triple phosphate crystals, and epithelial cells clumping but stilled good to identified WBC clumping. Second, the automation was unable to run a urine sample that was less than three milliliters in volume, but it still lost some volume and produced no results. This issue blocked us from having adequate volume to operate and caused us to switch to the manual technique.

For some sediment types that were varied and complicated in shape and dimension, for example, pathological casts, crystals, and dysmorphic RBCs, we found that the automation also reported them in the unclassified category or the other of each type (other crystal or other cast). Therefore, users must check these categories carefully. However, from our results, we suggested that the microscopic examination is still necessary in case of confirmation of pathognomonic sediments.

Conclusion

Our results indicated that the MUS36000 and FUS3000Plus urine automated analyzers are sufficient, accurate, precise, and dependable enough to screen urine sediment in the clinical laboratory rather than using manual microscopy in the AMS CMU EQA Center Unit. However, for clinical diagnosis purposes, certain pathologically specific urine sediments should be confirmed by manual microscopy when used in the hospital laboratory.

Ethical approval

The ethical committee of the Faculty of Associated Medical Sciences, Chiang Mai University, has permitted ethical proof documents (AMSEC-67EM-012).

Funding

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

Conflict of interest

The authors declare no conflict of interest.

CReDIT authorship contribution

Tanyarat Jomgeow: conceived the study, designed the experiments, and performed the experiment, data analysis and manuscript preparation; **Jirapat Narkdee** and **Nattawinan Piewlueng**: sample collection, perform automated and manual urinalysis data interpretation. All authors approved the final version.

Acknowledgements

The authors thank AMS CMU EQA unit staffs for their valuable helping on distribution of quality control material to hospital laboratories.

References

[1] Oyaert M, Delanghe JR. Semiquantitative, fully automated urine test strip analysis. J Clin Lab Anal. 2019; 33(5): e22870. doi: 10.1002/jcla.22870.

[2] Tantisaranon P, Dumkengkhachornwong K, Aiadsakun P, Hnoonual A. A comparison of automated urine analyzers cobas 6500, UN 3000-111b and iRICELL 3000 with manual microscopic urinalysis. Pract Lab Med. 2021; 24: e00203. doi: 10.1016/j.plabm.2021.e00203.

[3] Hannemann-Pohl K. Automation of urine sediment examination: a comparison of the Sysmex UF-100 automated flow cytometer with routine manual diagnosis (microscopy, test strips, and bacterial culture) Clin Chem Lab Med. 1999; 37: 753-64. doi: 10.1515/CCLM.1999.116.

[4] Sterry-Blunt RE, Randall KS, Doughton MJ, Aliyu SH, Enoch DA. Screening urine samples for the absence of urinary tract infection using the sediMAX automated microscopy analyser. J Med Microbiol. 2015; 64(6): 605-9. doi: 10.1099/jmm.0.000064.

[5] Shayanfar N, Tobler U, von Eckardstein A, Bestmann L. Automated urinalysis: first experiences and a comparison between the Iris iQ200 urine microscopy system, the Sysmex UF-100 flow cytometer and manual microscopic particle counting. Clin Chem Lab Med. 2007; 45(9): 1251-6. doi: 10.1515/CCLM.2007.503.

[6] Oyaert M, Delanghe J. Progress in automated urinalysis. Ann Lab Med. 2019; 39 (1): 15-22. Available from: https://doi: 10.3343/alm.2019.39.1.15.

[7] Larkey NE, Obiorah IE. Advances and progress in automated urine analyzers. Clin Lab Med. 2024; 44(3): 409-21. doi: 10.1016/j.cll.2024.04.003.

[8] Lamchiagdhase P, Preechaborisutkul K, Lomsomboon P, Srisuchart P, Tantiniti P, Khan-u-Ra N, Preechaborisutkul B. Urine sediment examination: a comparison between the manual method and the iQ200 automated urine microscopy analyzer. Clin Chim Acta. 2005; 358(1-2): 167-74. doi: 10.1016/j.cccn.2005.02.021.

[9] Anderlini R, Manieri G, Lucchi C, Raisi O, Soliera AR, Torricelli F, Varani M, Trenti T. Automated urinalysis with expert review for incidental identification of atypical urothelial cells: An anticipated bladder carcinoma diagnosis. Clin Chim Acta. 2015; 451:252-6. doi: 10.1016/j.cca.2015.10.005.

[10] Zaman Z, Fogazzi GB, Garigali G, Croci MD, Bayer G, Kránicz T. Urine sediment analysis: Analytical and diagnostic performance of sediMAX - a new automated microscopy image-based urine sediment analyser. Clin Chim Acta. 2010; 411(3-4): 147-54. doi: 10.1016/j.cca.2009.10.018.

[11] Cho EJ, Ko DH, Lee W, Chun S, Lee HK, Min WK. The efficient workflow to decrease the manual microscopic examination of urine sediment using on-screen review of images. Clin Biochem. 2018; 56: 70-4. doi: 10.1016/j.clinbiochem.2018.04.008.

[12] Cho J, Oh KJ, Jeon BC, Lee SG, Kim JH. Comparison of five automated urine sediment analyzers with manual microscopy for accurate identification of urine sediment. Clin Chem Lab Med. 2019; 57(11): 1744-53. doi: 10.1515/cclm-2019-0211.

[13] Yalcinkaya E, Erman H, Kirac E, Serifoglu A, Aksoy A, Isman FK. Comparative performance analysis of Urised 3 and DIRUI FUS-200 automated urine sediment analyzers and manual microscopic method.

Medeni Med J. 2019;34(3):244-51. doi: 10.5222/ MMJ.2019.23169.

[14] Benovska M, Wiewiorka O, Pinkavova J. Evaluation of FUS-2000 urine analyzer: analytical properties and particle recognition. Scand J Clin Lab Invest. 2018; 78(1-2): 143-8. doi: 10.1080/ 00365513. 2017. 1423108.

[15] Dirui MUS3600 product brochure. [cited 2024 Dec 5]. Available from: http://en.dirui.com.cn/list-20-1.html.

[16] ICH Q2(R2) Guideline on validation of analytical procedures [cited 2024 Dec 20]. Available from: https://www.ema.europa.eu/en/documents/ scientific-guideline/ich-q2r2-guideline-validation-analytical-procedures-step-5-revision-1_en.pdf.

[17] Yang WS. Automated urine sediment analyzers underestimate the severity of hematuria in glomerular diseases. Sci Rep. 2021; 11(1): 20981. doi: 10.1038/ s41598-021-00457-6.

[18] Kucukgergin C, Ademoglu E, Omer B, Genc S. Performance of automated urine analyzers using flow cytometric and digital image-based technology in routine urinalysis. Scand J Clin Lab Invest. 2019; 79(7): 468-74. doi: 10.1080/00365513.2019.1658894.

[19] Yuksel H, Kiliç E, Ekinci A, Evliyaoglu O. Comparison of fully automated urine sediment analyzers H800-FUS100 and LabUMat-UriSed with manual microscopy. J Clin Lab Anal. 2013; 27(4): 312-6. doi: 10.1002/ jcla.21604.

[20] Kocer D, Sarıguzel FM, Karakukcu C. Cutoff values for bacteria and leukocytes for urine sediment analyzer FUS200 in culture-positive urinary-tract infections. Scand J Clin Lab Invest. 2014; 74(5): 414-7. doi: 10.3109/00365513.2014.900189.

[21] Bartosova K, Kubicek Z, Franekova J, Louzensky G, Lavrikova P, Jabor A. Analysis of Four Automated Urinalysis Systems Compared to Reference Methods. Clin Lab. 2016; 62(11): 2115-23. doi: 10.7754/Clin. Lab.2016.160316.

Journal of Associated Medical Sciences 2025; 58(3): 248-253

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

Deep neural network-based prediction of RNA aptamers targeting E6 protein of high-risk human papilloma virus

Bundit Promraksa[1]* Yingpinyapat Kittirat[1], Dujdao Boonyod[1], Chawisa Phetumpai[1], Malinee Thanee[2], Anchalee Techasen[3]

[1]Regional Medical Sciences Center 2 Phitsanulok, Department of Medical Sciences, Ministry of Public Health, Phitsanulok Province, Thailand.
[2]Department of Pathology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand.
[3]Centre for Research and Development of Medical Diagnostic Laboratories (CMDL), Faculty of Associated Medical Science, Khon Kaen University, Khon Kaen Province, Thailand.

ARTICLE INFO

Article history:
Received 30 March 2025
Accepted as revised 13 July 2025
Available online 15 July 2025

Keywords:
RNA aptamer, human papillomavirus, deep neural network, cervical cancer, E6 protein

ABSTRACT

Background: The Human papilloma virus is the primary cause of cervical cancer. The virus integrates with the human genome to produce the E6 oncoprotein. Therefore, the E6 oncoprotein is a crucial molecular target for cancer progression or treatment. The development of aptamers is beneficial for interacting with the target protein and serves as a new strategy for detection or delivery systems.

Objectives: We aim to explore the candidate aptamer sequence against E6 oncoprotein using a computational-based method.

Materials and methods: This study designed the candidate aptamer against the target protein based on computational approaches using the AptaTrans pipeline. After obtaining the candidate aptamer sequences, the minimum free energy was calculated using the RNAfold web server. The tertiary structure was then generated using RNAComposer. Next, the molecular docking score was acquired from the GRAMM web server.

Results: The aptamer sequences with the best stability, as indicated by minimum free energy (MFE), are Sq3_16E6, Sq3_Actn, and Sq3_18E6, respectively. The aptamer sequences of Sq3_16E6 and Sq2_18E6 showed potential interactions with 8GCR and 6SJV, respectively.

Conclusion: Sq3_16E6 and Sq2_18E6 are appropriate for the development of the detection of the E6 protein in cervical swabs. Further investigation should be performed.

Introduction

High risk human papilloma virus (HPV) is non enveloped DNA virus, can cause the abnormalities of cervical epithelial cells, which primarily cause of cervical cancer. The individuals with risky behaviors, such as frequently changing sexual partners, engaging in sexual activity at a young age, and direct contact with lesions of an infected person, are at significant risk of developing cervical cancer [1]. Currently, HPV DNA testing is promoted with regular screening of cervical cancer in Thailand. This policy is recommended in response to these second most common cancers so that the detection of HPV DNA collected by self-cervical swabs have been promoted to prevent the development of these cancers in early stages[2]. Furthermore, the vaccination program in age under 12-13 years old significantly reduced caused of cervical cancer[3]. Regional Medical Sciences Center 2 Phitsanulok offers cervical cancer screening services using the HPV DNA test, which detects 14 high-risk

* **Corresponding contributor.**
Author's Address: *Regional Medical Sciences Center 2 Phitsanulok, Department of Medical Sciences, Ministry of Public Health, Phitsanulok Province, Thailand.*
E-mail address: bundit.p@dmsc.mail.go.th
doi: 10.12982/JAMS.2025.094
E-ISSN: 2539-6056

strains in its responsible area within Health Region 2. If patients test positive for HPV types 16 and 18, confirmation with colposcopy is recommended. For other strains, likely types 52 and 58, patients are taken liquid-based cytology to diagnose the progression of cervical cancer[4]. Notably,HPV DNA encodes several oncoproteins, including the oncogenic E6 and E7 proteins, which promote cervical carcinogenesis and are overexpressed during cervical transformation[5]. Therefore, an oncoprotein-based diagnostic test holds promise for being especially specific in detecting precancerous lesions that have progressed to a high-grade CIN stage or to cancer. The E6 protein binds to a short LxxLL consensus sequence within the cellular ubiquitin ligase E6-associated protein (E6AP), subsequently leading to degrad pro-apoptotic tumour suppressor p53. The meta-analysis suggested that the detection of E6 oncoprotein may be useful for triaging HPV-positive women by predicting the risk of developing cervical pre-cancer and cancer[6]. In consistency with the study by Ferrera et al. (2019), the detection of E6 oncoprotein is highly sensitive and serves as a specific marker for HPV16/18-related High-grade Squamous Intraepithelial Lesion (HSIL) lesions. Moreover, the development of lateral flow assays for E6 oncoprotein is useful for direct triage to treatment in resource-limited settings[7]. In summary, detecting HPV DNA, along with E6 and E7 proteins, is beneficial for identifying high-risk HPV infections. This approach is particularly important for assessing cervical cancer progression before the patient undergoes colposcopy by an obstetrician. Moreover, E6 is also the molecular target for treatment of cervical cancer[8]. Aptamers are the short nucleotides of DNA and RNA molecules which can bind to specific targets such as ions, small molecules or specific proteins. Aptamers are generally developed by the conventional method known as systematic evolution of ligands by exponential enrichment (SELEX)[9]. The limitation of SELEX is labor- and time- consuming. Moreover, this method rarely yields the number of effective candidate aptamers for further performance evaluation and validation. The use of a computational based method is alternative precision to predict the candidate aptamer sequence at the monomer level. AptaTrans, a deep neural network (DNN) model, was developed to utilize Monte-Carlo tree search (Apta-MCTS) for the exploration of the recommending RNA aptamer candidates[10,11]. This pipeline pretrain structural representation pretrained encoders to generates the deep neural network model. Then, the binding capabilities of aptamer and protein targets can investigate by the molecular docking tools. The predicted 3D model results explore different poses of the aptamer-protein interaction and identify the complexes with the lowest binding energies.

In this study, we aim to investigate the candidate aptamer sequence using a computational-based method. Following the AptaTrans pipeline, we predict candidate RNA aptamers that interact with the protein. We are initially exploring these RNA aptamers using computational methods for HPV detection (Figure 1).

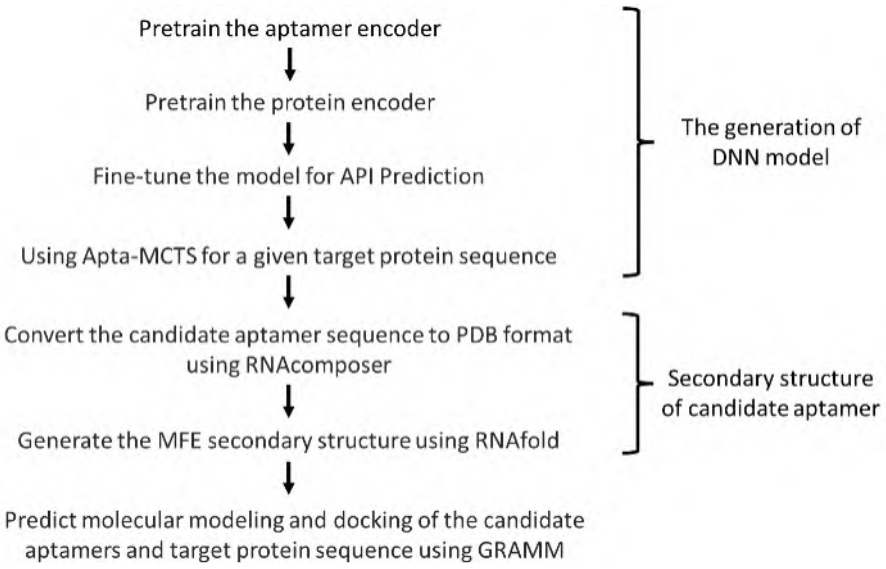


Figure 1. The workflow of this study to generate the candidate aptamer against E6 oncoprotein

Materials and methods

Generating a deep neural network model

The deep neural network model was executed followed by AptaTrans pipeline https://github.com/PNUMLB/AptaTrans) in the Windows subsystem for Linux environment. The core of the model is based on transformer-based encoders, which effectively capture the complex interactions between RNA aptamers and their target proteins. To generate neural network model, we used a pre-trained encoders of aptamers[12] and proteins[13]. A batch size of 16 was used, and the model was trained for 20 epochs. The neural network model was then fine-tuned for to further enhance performance.

Predicting candidate RNA aptamer using computational methods in deep neural networks

To obtain the RNA aptamer at the monomer level, the candidate aptamer was encoded by specifying the targeting

protein sequence and the depth of the Monte-Carlo search tree. The E6 oncoprotein sequence of HPV type 16 and 18 was retrieved from Uniprot (UniProt IDs: P06463 for HPV16 and P03255 for HPV18). 8GCR and 6SJV response to the HPV16 E6-E6AP-p53 complex and E6AP-LXXLL motifs of HPV type 16 and 18, respectively. Whereas, 3BYH represent actin. The number of iterations for the MCTS was set to 50, ensuring a thorough search for potential aptamer candidates.

Table 1. Candidate aptamer sequence

Candidate aptamer			MFE (-kcal/mol)
HPV16 (8GCR)	Sq1_16E6	GUUUAGCGAAUGCCCUUCAGUCUCUAACAAGAUGA	-1.90
	Sq2_16E6	CGCGGACCCCCUACAUUCCGCCGGGAUAUUACUAAAAGCGCUUUAAUCGU	-7.10
	Sq3_16E6	GAGGCGCAAGGCCGAACUGUAGAUUUAUAGGGGUGAACCAAGGACAUGC CGCGCGACUCC	-14.70
HPV18 (6SJV)	Sq1_18E6	GGUGAGCGAGCCCAUAGGUGGCUUACAGAGUUUUG	-10.70
	Sq2_18E6	GCGAAUUCGCCUUGGCACGAGAUCCGUAGGCAGGAGACGAAUUCGCGAUU	-19.10
	Sq3_18E6	CGCCGCGAGGUUAGCCUUACCGCACCCCCCGUCGACGAGAGGACGGGGCGG UCUCACGAA	-23.40
Actin (3BYH)	Sq1_Actn	AGAACACAUCCUUUAGUGCCAGUCCGACUUCUCGUUAGGUUUGACUGGGU	-14.50
	Sq2_Actn	AACCGCGCGGCCAUGUAUGCACGGAGUGUAGCCCUACUGUAGUCAAACUGA AACCGCCGC	-11.70
	Sq3_Actn	UAGACGCAGUUCAUACGAGCAGUUCGAUGUAUACUGGAACCCAAGUGAGU GGUGCAGAUUGUACAGCAAAUGGCUCGCCGCCUCCACGUCCGAAUCAAAG	-20.90

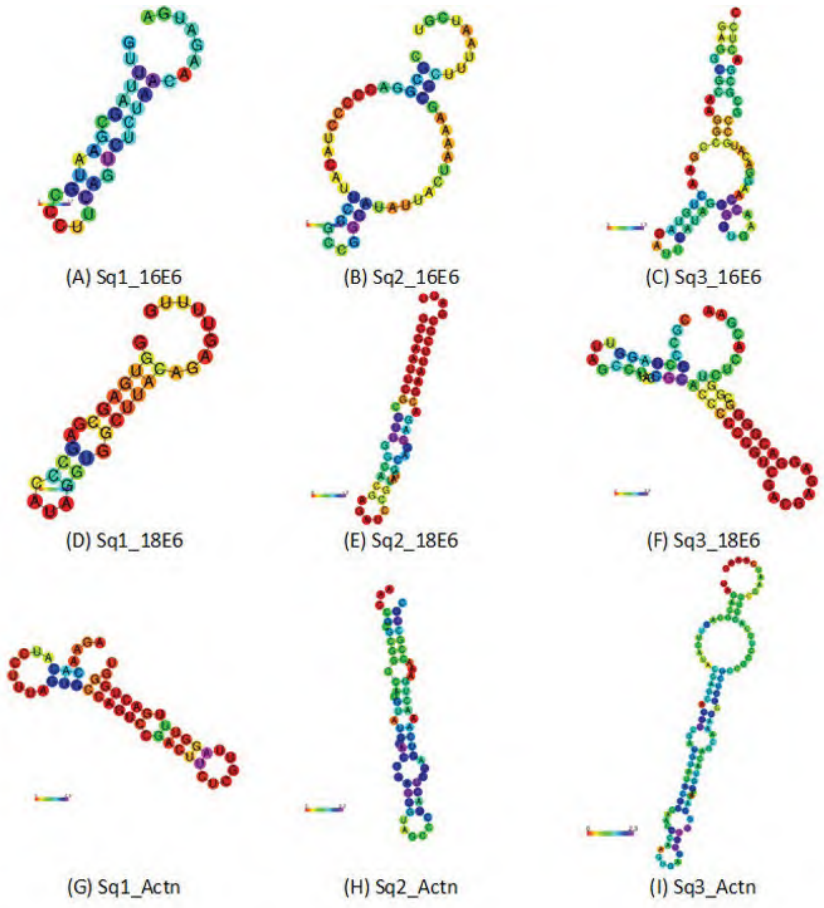


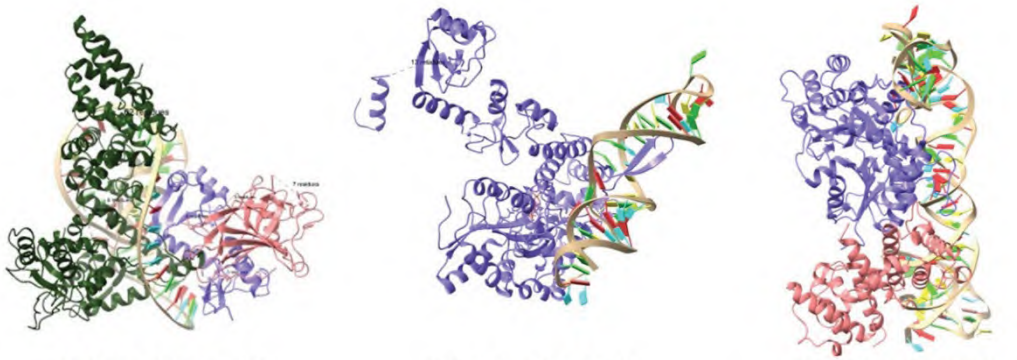
Figure 2. MFE secondary structure of the candidate aptamer sequence by RNAfold. Nucleotides are colored according to their positional entropy, as shown on the horizontal bar. Red colors indicate lower entropy, while blue colors indicate higher entropy

Molecular modeling and docking of the candidate aptamers

RNA aptamer sequences were converted to PDB format using RNAComposer web servers, to be used as ligands in subsequent analyses. The minimum free energy (MFE) secondary structure was obtained from RNAfold web server. The PDB format of target proteins, including 8GCR, 6SJV, and 3BYH, was downloaded from the Protein Data Bank (PDB). To predict the molecular interaction poses, RNA aptamer-protein dockings were performed using the GRAMM web server, which did not specify a binding site. The clustering threshold was set to 5 Å. Then, the 3D model structures were generated using the ChimeraX program (UCSF, USA).

Table 2. GRAMM's docking scores of candidate aptamers to target protein

HPV types	Candidate Aptamer sequences	Docking scores		
		8GCR	6SJV	3BYH
HPV16	Sq1_16E6	-630	-523	-634
	Sq2_16E6	-686	-654	-649
	Sq3_16E6	-802	-719	-712
HPV18	Sq1_18E6	-558	-613	-573
	Sq2_18E6	-661	-682	-659
	Sq3_18E6	-670	-670	-635
Actin	Sq1_Actn	-616	-640	-726
	Sq2_Actn	-670	-802	-818
	Sq3_Actn	-709	-664	-780



(A) Sq3_16E6-8GCR (B) Sq2_18E6-6SJV (C) Sq3_Actn-3BYH

Figure 3. The representative structure of the interaction illustrates the binding interfaces between the aptamer and target protein

Results

Candidate aptamers generation from DNN model

In this study, we performed deep neural networks to predict the sequences of aptamers targeting the proteins 8GCR, 6SJV, and 3BYH, which represent the E6 protein of HPV types 16 and 18, and actin, respectively. The candidate aptamer sequences are shown in Table 1. Based on the MFE results, the folding stability of each aptamer depends on sequence length. Additionally, the use of actin has been iterated as an internal control in the study. For this purpose, the RNA aptamer sequence most specific for this protein is Sq3_Actn, as it shows the highest binding affinity, indicated by the lowest docking scores and demonstrates lack of cross-reactivity with other proteins studied. The secondary structure of aptamer was shown in Figure 2.

The molecular docking of candidate aptamer to protein

The binding affinities of the candidate aptamers across all three proteins were evaluated using GRAMM web server (Table 2). Between candidate RNA aptamer against 8GCR protein, Sq3_16E6 exhibited the highest binding affinity based on its docking score. The docking scores of Sq2_18E6 and Sq1_18E6 against 6SJV protein were -682 and -613, respectively. Sq3_Actn showed a specific interaction with 3BYH protein, and demonstrated lower cross reactivity with other studied proteins. The interaction between representative candidate aptamer and the target protein has binding interfaces as shown Figure 3.

Discussion

HPV infection is a major risk factor for cervical cancer. An early region of HPV E6 and E7 genes integrate into the human genome, encode the oncoproteins playing a significant role during cervical cancer progression. The detection of the E6 oncoprotein serves as a specific marker for HPV-induced cancer progression. However, the development of detection methods may be problematic due to issues with the storage of cervical swab samples and the freeze-thaw cycles. Aptamers not only improve detection methods but also play a crucial role in the

development of next-generation drugs[14,15]. The process of obtaining specific recognition sequences, such as antibodies or high-throughput SELEX (HT-SELEX) aptamers, is still time consuming. Therefore, computational approaches are an alternative choice for improving the successful identification of aptamer sequence.

We obtained the aptamer sequence after inputting the target protein sequence into a DNN model, which predicted the candidate aptamer sequence based on Monte-Carlo Tree Search (MCTS)[11]. The general size of aptamer is between 30-100 nucleotides, so the length of candidate aptamer set up in these intervals. The selection and production of aptamers have been mentioned in which be challenges. Muhammad et al., 2022 have been generate the novel aptamer against NT-3 growth factor receptor based on computational approach with a relatively stable structure[16]. Moreover, the use of RNA aptamer, which was designed in silico, also demonstrated the ability to adhere on the surface of MCF-7 and MDA-MB-231 cells, which have been beneficial for cancer cell imaging[17]. In our study, Sq3_16E6 and Sq3_Actn expressed the low MFE, which relate to their stable positional entropy in the secondary structure. The tertiary structure was obtained from RNAcomposer webserver. This is the first study to apply the DNN predictive abilities using Aptatrans for aptamer sequence generation on cervical oncoprotein.

The molecular docking has been used as bioinformatics tool to predict the RNA aptamer-protein interactions[18]. Recently, the development of apta-sensors performed the molecular docking to predict the interaction between calcium/calmodulin-dependent serine protein kinase (CASK) protein, which propose to be breast cancer screening methods, against candidate aptamers[19]. Therefore, the molecular docking is the powerful tools to identify how the aptamer interacts with the target before further verification with other molecular techniques. In this study, we use GRAMM web server for protein docking to predict docking poses based on the Fourier transformation of the possible interaction sites of macromolecules and proteins[20]. The binding interfaces of the Sq3_16E6 groove can interact with the 8GRC protein, which represents a high docking score compared to other studied proteins. This interaction of Aptamers against proteins involves hydrogen bonding, electrostatic interactions, hydrophobic interactions and van der Waals forces[21]. However, the low binding affinity of RNA aptamers for HPV18E6 might result from poor shape complementarity, which is due to rigid and geometric-based scoring. RNA aptamers are highly flexible, and this limitation can cause inaccurate or low docking scores. Our result noticed that candidate aptamers, which expressed low MFE, had high potential interact with E6 oncoprotein.

Limitation

This study relies on computational approaches using the existing capabilities of the DNN model, which may not always accurately predict the complex structures and behaviors of aptamers. Next, the experimental validation is necessary to confirm these predictions.

Conclusion

The computational approach is a powerful tool for generating candidate aptamer. The molecular docking revealed the potential on the candidate aptamer and protein interactions. This study suggests that Sq3_16E6 and Sq2_18E6 are appropriate for further evaluation of their ability on E6 HPV type 16 and 18 oncoproteins, respectively.

Ethical approval

Not applicable

Funding

This work received no specific grant from any funding agency.

Conflict of interest

The authors declare no conflict of interest.

CReDIT authorship contribution statement

Bundit Promraksa: conceptualization, methodology, visualization, investigation, data curation, validation, writing original draft preparation; **Yingpinyapat Kittirat:** conceptualization, data curation; **Dujdao Boonyod:** conceptualization, data curation, writing-reviewing and editing; **Chawisa Phetumpai:** investigation, data curation; **Malinee Thanee:** data curation, writing-reviewing and editing; **Anchalee Techasen:** data curation, writing-reviewing and editing.

Acknowledgements

I would also like to thank Regional Medical Sciences Center 2 Phitsanulok for providing the resources and facilities needed to conduct my research. I would like to express my deepest gratitude to Ms. Amornrat Tassanakit, Director of Regional Medical Sciences Center 2 Phitsanulok for support and encouragement.

References

[1] Woodman CBJ, Collins SI, Young LS. The natural history of cervical HPV infection: unresolved issues. Nat Rev Cancer. 2007; 7(1): 11-22. doi: 10.1038/nrc2050.

[2] Ploysawang P, Rojanamatin J, Prapakorn S, Jamsri P, Pangmuang P, Seeda K, et al. National cervical cancer screening in Thailand. Asian Pac J Cancer Prev. 2021; 22(1): 25-30. doi: 10.31557/APJCP.2021.22.1.25.

[3] Falcaro M, Castañon A, Ndlela B, Checchi M, Soldan K, Lopez-Bernal J, et al. The effects of the national HPV vaccination programme in England, UK, on cervical cancer and grade 3 cervical intraepithelial neoplasia incidence: a register-based observational study. The Lancet. 2021; 398(10316): 2084-92. doi: 10.1016/S0140-6736(21)02178-4.

[4] Huh WK, Ault KA, Chelmow D, Davey DD, Goulart RA, Garcia FAR, et al. Use of primary high-risk human papillomavirus testing for cervical cancer screening: Interim clinical guidance. Obstet Gynecol. 2015; 125(2): 330-7. doi: 10.1097/AOG.0000000000000669.

[5] Derbie A, Mekonnen D, Woldeamanuel Y, Van Ostade X, Abebe T. HPV E6/E7 mRNA test for the detection of

high grade cervical intraepithelial neoplasia (CIN2+): a systematic review. Infect Agent Cancer. 2020; 15(1): 9. doi: 10.1186/s13027-020-0278-x.

[6] Downham L, Jaafar I, Rol ML, Nyawira Nyaga V, Valls J, Baena A, et al. Accuracy of HPV E6/E7 oncoprotein tests to detect high-grade cervical lesions: a systematic literature review and meta-analysis. Br J Cancer. 2024; 130(4): 517-25. doi: 10.1038/s41416-023-02490-w.

[7] Schweizer J, Lu PS, Mahoney CW, Berard-Bergery M, Ho M, Ramasamy V, et al. Feasibility Study of a Human Papillomavirus E6 Oncoprotein Test for Diagnosis of Cervical Precancer and Cancer. J Clin Microbiol. 2010; 48(12): 4646-8. doi: 10.1128/JCM.01315-10.

[8] Tan S, G.E. De Vries E, G.J. Van Der Zee A, De Jong S. Anticancer drugs aimed at E6 and E7 activity in HPV-positive cervical cancer. Curr Cancer Drug Targets. 2012; 12(2): 170-84. doi: 10.2174/156800912799095135.

[9] Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. Nature. 1990; 346(6287): 818-22. doi: 10.1038/346818a0.

[10] Shin I, Kang K, Kim J, Sel S, Choi J, Lee JW, et al. AptaTrans: a deep neural network for predicting aptamer-protein interaction using pretrained encoders. BMC Bioinformatics. 2023; 24(1): 447. doi: 10.1186/s12859-023-05577-6.

[11] Lee G, Jang GH, Kang HY, Song G. Predicting aptamer sequences that interact with target proteins using an aptamer-protein interaction classifier and a Monte Carlo tree search approach. Nebel JC, editor. PLOS ONE. 2021; 16(6): e0253760. doi: 10.1371/journal.pone.0253760.

[12] Danaee P, Rouches M, Wiley M, Deng D, Huang L, Hendrix D. bpRNA: large-scale automated annotation and analysis of RNA secondary structure. Nucleic Acids Res. 2018; 46(11): 5381-94. doi: 10.1093/nar/gky285.

[13] Berman HM. The protein data bank. Nucleic Acids Res. 2000; 28(1): 235-42. doi: 10.1093/nar/28.1.235.

[14] Khabbazian M, Jabbari H. AI-powered aptamer generation. Nat Comput Sci. 2022; 2(6): 356-7. doi: 10.1038/s43588-022-00253-w.

[15] Keefe AD, Pai S, Ellington A. Aptamers as therapeutics. Nat Rev Drug Discov. 2010; 9(7): 537-50. doi: 10.1038/nrd3141.

[16] Muhammad AM, Zari A, Alsubhi NH, Al-Zahrani MH, Alghamdi RA, Labib MM. Novel design of RNA aptamers as cancer inhibitors and diagnosis targeting the tyrosine kinase domain of the NT-3 growth factor receptor using a computational sequence-based approach. Molecules. 2022; 27(14): 4518. doi: 10.3390/molecules27144518.

[17] Bavi R, Liu Z, Han Z, Zhang H, Gu Y. In silico designed RNA aptamer against epithelial cell adhesion molecule for cancer cell imaging. Biochem Biophys Res Commun. 2019; 509(4): 937-42. doi: 10.1016/j.bbrc.2019.01.028.

[18] Navien TN, Thevendran R, Hamdani HY, Tang TH, Citartan M. In silico molecular docking in DNA aptamer development. Biochimie. 2021; 180:5 4-67. doi: 10.1016/j.biochi.2020.10.005.

[19] De Araújo NS, Moreira ADS, Abreu RDS, Junior VV, Antunes D, Mendonça JB, et al. Aptamer-based recognition of breast tumor cells: A new rra for breast cancer diagnosis. Int J Mol Sci. 2024; 25(2): 840. doi: 10.3390/ijms25020840.

[20] Singh A, Copeland MM, Kundrotas PJ, Vakser IA. GRAMM web server for protein docking. In: Gore M, Jagtap UB, Editors. Computational rug Discovery and Design (Methods in Molecular Biology; Vol. 2714). New York, NY: Springer US, pp 101-12. doi: 10.1007/978-1-0716-3441-7_5.

[21] Ji C, Wei J, Zhang L, Hou X, Tan J, Yuan Q, et al. Aptamer–protein interactions: From regulation to biomolecular detection. Chem Rev. 2023; 123(22): 12471-506. doi: 10.1021/acs.chemrev.3c00377.

Journal of Associated Medical Sciences 2025; 58(3): 254-262

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 of line probe assay and phenotypic drug susceptibility testing in detecting drug-resistant tuberculosis

Phu Thien Truong[1], Tran Ngoc Minh Le[2*], Van Thi Hue Tran[2,3,4], Tung Thanh Phan[1]

[1]*Microbiology Department, Cho Ray Hospital, Ho Chi Minh City, Vietnam.*
[2]*Medical Laboratory Technology Department, Faculty of Nursing and Medical Technology, University of Medicine and Pharmacy, Ho Chi Minh City, Vietnam.*
[3]*Microbiology and Parasitology Department, Faculty of Medicine, University of Medicine and Pharmacy, Ho Chi Minh City, Vietnam.*
[4]*Quality Control Center for Medical Laboratory, University of Medicine and Pharmacy, Ho Chi Minh City, Vietnam.*

ARTICLE INFO

Article history:
Received 1 May 2025
Accepted as revised 18 July 2025
Available online 22 July 2025

Keywords:
Drug-resistant tuberculosis, drug susceptibility testing, line probe assay, multidrug-resistant tuberculosis.

ABSTRACT

Background: Tuberculosis (TB) still threatens human beings when drug-resistant tuberculosis (DR-TB), such as rifampicin-resistant TB, isoniazid-resistant TB, multidrug-resistant TB (MDR-TB), pre-extensively drug-resistant TB, and extensively drug-resistant TB increases continuously. The drug susceptibility testing (DST) is important to detect DR-TB for TB treatment.

Objectives: The study aimed to assess first-line line probe assay (FL-LPA) performance of screening MDR-TB and detecting DR-TB on phenotypic drug susceptibility testing.

Materials and methods: A laboratory-based study was performed at Cho Ray Hospital from August 2023 to August 2024. The sputum samples of presumptive TB were inoculated in *Mycobacterium* growth indicator tube (MGIT). Positive inoculum was examined in acid-fast bacilli (AFB) by Ziehl-Neelsen microscope. Cord-forming AFB were yielded to FL-LPA to identify *Mycobacterium tuberculosis* complex (MTBC); detect rifampicin-resistant TB, isoniazid-resistant TB, and MDR-TB. The identified MTBC was subjected to FL phenotypic DST (performed by BACTEC MGIT 960) with SIRE kit, considering gold standard to assess FL-LPA performance. The detected multidrug and/or rifampicin-resistant TB (MDR/RR-TB) were subjected to the second-line MGIT DST including ethionamide, amikacin, levofloxacin, and linezolid to screen pre-extensively drug-resistant TB and extensively drug-resistant TB.

Results: Among 1853 samples inoculated, 621 positive MGIT tubes seen cord-forming AFB on Ziehl-Neelsen smear were performed to FL-LPA. Out of 621 LPA tests, 304 MTBC (61 isoniazid-resistant TB, 20 rifampicin-resistant TB, and 243 susceptible TB) were detected and compared to FL phenotypic DST. The excellent agreements between FL-LPA and FL phenotypic DST for detecting rifampicin-resistant TB, isoniazid-resistant TB, and MDR-TB were greater than 98%; kappa at 0.89 and above (p<0.001); with sensitivity values at 88.9% and above; specificity values at greater than 99%. For FL-MGIT DST, 101 (33.2%) were drug-resistant to at least one anti-TB agent, 81 (26.6%) to streptomycin, 60 (19.7%) to isoniazid, 20 (6.6%) to rifampicin. Among 20 MDR/RR-TB (2 rifampicin mono-resistant-TB and 18 MDR-TB) performed second line phenotypic DST, 25% resistance to ethionamide, and 100% susceptibility to amikacin, levofloxacin, and linezolid.

Conclusion: The performance of FL-LPA to detect rifampicin-resistant TB, isoniazid-resistant TB, and MDR-TB agreed perfectly with phenotypic DST. The reaffirmed critical concentration of isoniazid, rifampicin and levofloxacin would be used to screen DR-TB on population.

* **Corresponding contributor.**
Author's Address: *Medical Laboratory Technology Department, Faculty of Nursing and Medical Technology, University of Medicine and Pharmacy, Ho Chi Minh City, Vietnam.*
E-mail address: lnmtran.ncs21@ump.edu.vn
doi: 10.12982/JAMS.2025.095
E-ISSN: 2539-6056

Introduction

TB incidence increased to the highest cases of 7.5 million since 1995.[1] Furthermore, DR-TB such as rifampicin-resistant TB (RR-TB), isoniazid-resistant TB (HR-TB), MDR-TB (TB resistance to both RIF and INH), pre-extensively drug-resistant TB (MDR/RR-TB resistance to any fluoroquinolone),[2] and extensively drug-resistant TB (MDR/RR-TB resistance to any fluoroquinolone and at least one additional Group A drug such as bedaquiline or linezolid) still causes severe obstacles to TB treatment. MDR/RR-TB have reported 410,000 new cases and 160,000 deaths in 2022.[1] On the End TB strategy, presumptive TB were recommended to receive the rapid tests to detect TB and DR-TB from DST playing an important role for chemotherapy, treatment response and surveillance of emerge drug resistance.[3] FL-LPA is also recommended as an initial rapid test to detect resistance to RIF and INH within 48 hours.[4] LPA of Genoscholar NTM + MDRTB Detection Kit (NIPRO Corporation, Osaka, Japan) had sensitivity, specificity at 96.5, 97.5 for detecting RR-TB, and 94.9, 97.6 for HR-TB while Genotype MTBDR*plus* was reported the higher sensitivity, specificity at 98.2, 97.8 for detecting RR-TB and 95.4, 98.8 for HR-TB.[5] Genotype MTBDR*plus* identifies the most significant mutations of the *rpoB* gene (coding for β-sub-unit of the ribonucleic acid polymerase); the *katG* gene (coding for the catalase-peroxidase) and promoter region of the *inhA* gene (coding for nicotinamide adenine dinucleotide enoyl-acyl carrier protein reductase) to detect RR-TB, high and low level HR-TB, respectively.[6,7] The second-line (SL) LPA test has not been recommended for DST due to their detecting injectable anti-TB drugs which are not including the shorter oral regimen for treating MDR/RR-TB recently.[8,9] Whereas phenotypic culture-based DST with turnaround time about 2 weeks as the gold standard which is available for new and repurposed Group A drugs to treat MDR/RR-TB and detect pre-extensively drug-resistant TB (pre-XDR-TB), and 3.8% as extensively drug-resistant TB (XDR-TB).[10]

Vietnam was ranked in two of three WHO global lists of high-burden countries for TB and MDR/RR-TB with 9200 MDR/RR-TB incident cases in 2020.[1] Wrohan *et al*. reported 88% MDR-TB, 8.2% as pre-extensively drug-resistant TB, and 3.8% as XDR-TB on high-risk populations in Ha Noi and Thanh Hoa, Northern, Vietnam that was performed by Xpert MTB/RIF, FL-LPA, and DST in 2022.[11] So, an effort to improve testing and diagnosis of DR-TB is the prior challenge worldwide and in Vietnam particularly. We analyzed data between August 2023 and August 2024 for a retrospective laboratory-based study to detect drug-resistant TB at the Microbiology department in Cho Ray Hospital. The FL-LPA evaluated the detection performance of rifampicin, isoniazid, and multi-drug resistance of TB based on the gold standard of phenotypic conventional culture-based DST that screened drug resistance, multi-drug resistance, pre-extensive drug resistance, and extensive drug resistance of TB, as well. The data of our study reported the practice of screening DR-TB according to the updated definition of drug resistance.[12]

Materials and methods

A laboratory-based study was performed at Cho Ray Hospital, from August 2023 to August 2024. Positive MGIT cultures of sputum were performed by Ziehl-Neelsen staining. Smear of inoculum was covered with hot 0.3% Carbol Fuchsin-Phenol in 10 minutes, de-colored with 3% acid alcohol for 3 minutes, and counterstained with 0.3% Methylene Blue for 1 minute.[13] Rinsing slightly was performed after each steps before microscopic examination for cord formation of AFB (Figure 1).[14,15]

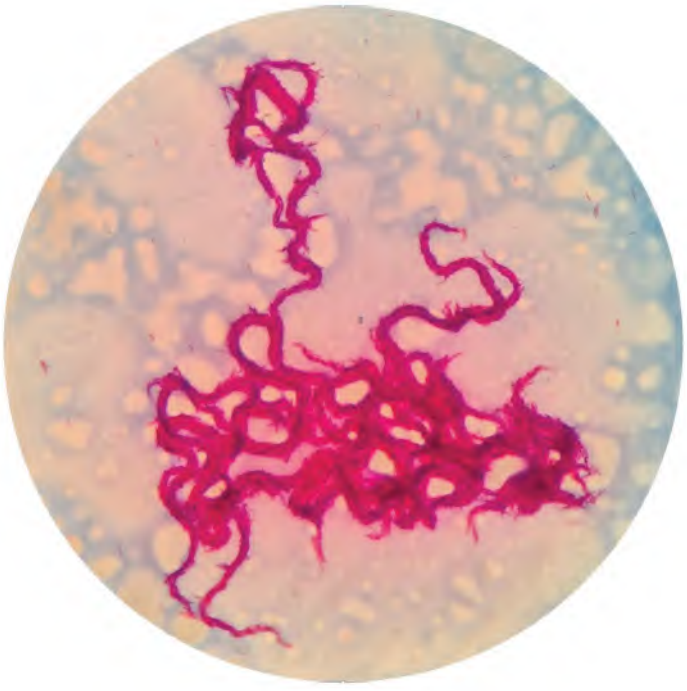


Figure 1. *Cord-forming acid-fast bacilli in liquid media on Ziehl-Neelsen microscopy.*

The inoculum with AFB cord formation was subjected to FL-LPA of GenoType MTBDR*plus* version 2.0 (Hain Life science, Nehren, Germany) to identify MTBC and detect RR-TB, HR-TB, and MDR-TB according to the procedure of manufacture.[6] Genolyse DNA extraction prepared with 1 mL of inoculum from positive MGIT was centrifuged for sediment and suspended with 100 µL of Lysis Buffer at 95 °C for 5 minutes, centrifuged at 13,000 rpm with 100 µL Neutralization Buffer for 100 µL of supernatant. A mixture including 10 µL of Amplification Mix A, 35 µL of Amplification Mix B and 5 µL of extracted DNA was amplified on CFX96 Real-Time system (Bio-rad, USA) in 15 min at 95 °C for denaturation; 10 cycles of 30 seconds at 95 °C and 2 minutes at 65 °C; 20 cycles including 30 seconds at 90 °C, 40 seconds at 50 °C, 40 seconds at 70 °C, and 8 minutes at 70 °C. Reversed hybridization was performed on membrane strip with 20 µL of PCR products with 20 µL of DEN solution, 1.0 mL of HYB solution, 1.0 mL of STR solution, 1.0 mL of RIN solution, 1.0 mL of conjugate and 1.0 mL of subtract solution for binding to probes targeting the most commonly occurring mutations (MUT), and wild-type (WT) probes.

Results were interpreted based on presence of bands for loci of *rpoB*, *katG* and *inhA* genes; presence of all WT probe bands for a sensitive classification of gene; absence of one or more WT bands indicating the strain has resistance to a specific drug and the absence of a WT band accompanying by the presence of a MUT probe band; presence of bands at both WT and MUT probe sites indicating either a heterogeneous test strain with partial resistance; or a mixed culture where at least one of the strains harbors a mutation. Deviation from the WT banding pattern for *rpoB, katG* and *inhA* probes indicated rifampicin resistance, high- and low-level Isoniazid resistance, respectively.

Detected MTBC by FL-LPA was yielded to DST on the MGIT BACTEC 960 against FL anti-TB agents of SIRE kit (Becton Dickinson, USA) considered as gold standard to evaluate FL-LPA performance. For preparing Mycobacterial suspension from MGIT positive, a tube observed positivity of day one to two was inverted and stood within ten minutes for sediment. For a positive observation of day three to five, tube was diluted five-fold.[16] Each MGIT tube for DST was added 0.8 mL of SIRE Supplement, 0.5 mL of inoculum suspension and 100 µL of reconstituted drug solution for the critical concentration of Streptomycin (STR) 1.0 µg/mL, INH 0.1 µg/mL, RIF 0.5 µg/mL, Ethambutol (EMB) 5 µg/mL.[3,17,2] The growth-control tube was added 800 µL of SIRE Supplement and 500 µL of inoculum suspension diluted 1:100 ratio (100 µL of inoculum : 9.9 mL BACTEC Diluting Fluid) without anti-TB drugs. After inoculation, the tube was incubated at 36±1°C in BACTEC MGIT 960 where fluorescence is detected automatically based on the growth of the bacteria in the presence of the drug about 4 to 13 days for both line DST.[16] The result was compared to a growth control (400 growth unit) when growth unit value in the drug-containing tube was less than 100 as susceptible (S) and greater than or equal 100 as resistant (R).

Detected MDR/RR-TB were subjected to the second line (SL) anti-TB agents (Sigma-Aldrich, Germany) including the critical concentration of amikacin (AMK) 1.0 µg/mL, ethionamide (EDT) 5.0 µg/mL, linezolid (LZD) 1.0 µg/mL, levofloxacin (LVX) 1.0 µg/mL (a later generation fluoroquinolone) and one GC tube.[2] The rates of drug resistance to anti-TB agents were evaluated and interpreted RR-TB, MDR-TB, pre-XDR-TB (MDR/RR-TB combined with LVX resistance), XDR-TB (MDR/RR-TB combined with both resistance to LVX and LZD).[2]

Quality assurance

The assays in this study including Genotype MTBDR*plus* version 2, first- and second-line MGIT DST testing were performed the internal control with negative control for each sample and positive control of MTB H37Ra for each batch;[18] accredited external quality assurance performance by Integrated Quality Laboratory Services, France; and certified method verification by the laboratory center of AIDS Clinical Trials Group and International Maternal Pediatric Adolescent AIDS Clinical Trials.

Statistical analysis

STATA 17.0 (StataCorp, College Station, Texas, USA) was used in this study. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of LPA were evaluated performance assay. The respective 95% confidence interval (CI) was computed using the Clopper–Pearson method. The agreement of resistant detection of two assays compared with kappa; p-value < 0.05 was considered statistically significant.

Results

Among 1853 tubes alarmed culture positives from MGIT 960 system, 621 were seen with cord formation on Ziehl-Neelsen AFB smears and performed to FL-LPA. Out of 621 LPA tests, 304 MTBC including 61 isoniazid-resistant-TB, 20 rifampicin-resistant TB, and 243 susceptible TB were detected and compared to FL-MGIT DST which was subjected to SIRE kit. Among 304 FL-MGIT DST, 203 were susceptible whereas 60, 20, and 18 were resistant to INH, RIF, and MDR-TB, respectively. A total of 18 MDR-TB and 2 rifampicin mono-resistant-TB were performed further in the SC-MGIT DST. There were only 5 ETD-resistant TB whereas AMK-resistant TB, LVX-resistant TB, LZD-resistant TB, pre-XDR-TB, and XDR-TB were not found (Figure 2).

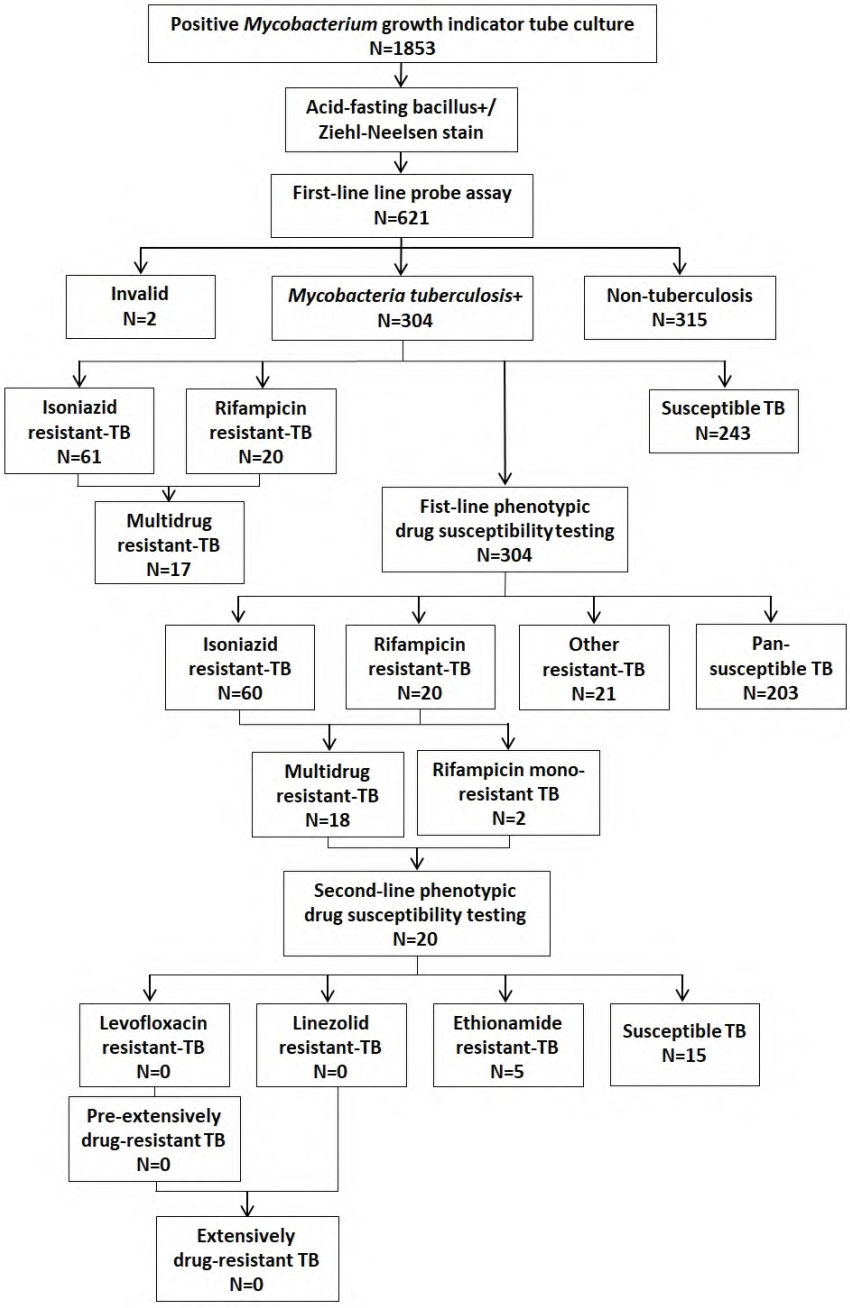


***Figure 2**. Flowchart of screening drug-resistant tuberculosis by line probe assay and phenotypic drug susceptibility testing.*

LPA was compared to the gold standard of FL-MGIT DST for the performance of detecting RR-TB, HR-TB, and MDR-TB. All high values of sensitivity at 88.9% and above; with specificity values being greater than 99%. HR-TB shows the highest values at 98.3% (95%CI, 91.1-99.9) for sensitivity, 99.2 (95%CI, 97.1-99.9) for specificity, 96.7% (95%CI, 88.7-99.6) for PPV, and 99.6 (95%CI, 99.7-99.9) for NPV. While RR-TB showed 90.0% (95%CI, 70.0-97.2) for both sensitivity and PPV; and 99.3 (95%CI, 97.5-99.8) for both specificity and NPV. MDR-TB was found at 88.9% (95%CI, 65.3-98.6), 99.6 (95%CI, 99.7-99.9), 94.1% (95%CI, 71.3-99.9, and 99.2 (95%CI, 97.1-99.9) for sensitivity, specificity, PPV, and NPV; respectively. There were excellent agreements at approximately 99% between LPA and MGIT DST with kappa values at 0.89, 0.97, and 0.91 ($p<0.001$) for RR-TB, HR-TB, and MDR-TB, respectively (Table 1).

Table 1. Performance of LPA compared with phenotypic DST for detecting TB resistance to Rifampicin, Isoniazid, and multidrug in TB.

	Anti-TB drugs	Susceptibility	First line MGIT DST		Sensitivity (95% CI)	Specificity (95% CI)	PPV% (95% CI)	NPV% (95% CI)	Agreement %	Kappa (*p* value)
			Resistant	Susceptible						
LPA	RIF	Resistant	18	2	90.0	99.3	90.0	99.3	98.7	0.89
		Susceptible	2	282	(70.0-97.2)	(97.5-99.8)	(70.0-97.2)	(97.5-99.8)		(<0.001)
	INH	Resistant	59	2	98.3	99.2	96.7	99.6	99.0	0.97
		Susceptible	1	242	(91.1-99.9)	(97.1-99.9)	(88.7-99.6)	(99.7-99.9)		(<0.001)
	RIF and INH (MDR-TB)	Resistant	16	1	88.9	99.6	94.1	99.2	98.8	0.91
		Susceptible	2	240	(65.3-98.6)	(97.7-99.9)	(71.3-99.9)	(97.1-99.9)		(<0.001)

***Note**: CI: confidence interval, DST: drug susceptibility testing, INH: isoniazid, LPA: line probe assay, MDR-TB: multidrug-resistant tuberculosis, MGIT: mycobacterium growth indicator tube, NPV: negative predictive value, PPV: positive predictive value, RIF: rifampicin.*

The mono-resistant and multidrug resistant strains were shown the banding pattern by LPA in Table 2. The frequency of inferred mutation was *rpoB* WT3 (2), *rpoB* WT4 (2), *rpoB* WT7 (4), *rpoB* WT8 (7) for RR-TB and *katG* WT (2), *inhA* WT1 (2) for HR-TB. For mono HR-TB, the mutation S315T1 associated with *katG* MUT1 of 39 for high level HR-TB was higher than mutation C15T associated with *inhA* MUT1 of 19 for low level HR-TB. While the frequency of *rpoB* mutation was D516V (5), H526Y (2), H526D (2), S531 (1) for RR-TB.

Table 2. Gene mutation pattern detected in drug resistant *Mycobacterium tuberculosis* strains by first-line line probe assay of Genotype MTBDR*plus* version 2.

Gene	Gene regions or associated mutations	Band MUT (mutation) WT (wild Type)	Rifampicin monoresistance	Isoniazid monoresistance	Multidrug resistance
rpoB					
	506-509	*rpoB* WT1			
	510-513	*rpoB* WT2			
	513-517	*rpoB* WT3			2
	516-519	*rpoB* WT4			2
	518-522	*rpoB* WT5			
	521-525	*rpoB* WT6			
	526-529	*rpoB* WT7			4
	530-533	*rpoB* WT8	2		5
	D516V	*rpoB* MUT1	1		4
	H526Y	*rpoB* MUT2A			2
	H526D	r*poB* MUT2B			2
	S531L	*rpoB* MUT3			1
katG					
	315	*katG* WT		2	
	S315T1	*katG* MUT1		28	11
	S315T2	*katG* MUT2		2	1
inhA					
	0.9375	*inhA* WT1		2	
	-8-	*inhA* WT2			
	C15T	*inhA* MUT1		13	6
	A16G	*inhA* MUT2			
	T8C	*inhA* MUT3A			
	T8A	*inhA* MUT3B			

Among 304 MTBC subjected to MGIT DST against the first-line drugs including STR, INH, RIF, and EMB; the overall drug resistance to at least one anti-TB agent was 101 (33.2%). The highest resistance proportion of STR mono-resistant-TB was 81 (26.6%) whereas 60 (19.7%) to INH, 20 (6.6%) to RIF and 1 (0.3%) to EMB. INH mono-resistant-TB was higher than RIF mono-resistant-TB at 16 (5.2%) and 2 (0.7%). There were 18 (5.9%) MDR-TB including 1 (0.3%) pan-resistance. Of 20 MDR/RR-TB yielded to SL-DST, 75% susceptible to panel of AMK, ETD, LVX and LZD were found. Only 25% ETD-resistant TB were detected from 5 MDR-TB with resistance to STR. MDR-TB was detected resistance to neither LVX nor LZD. Among 18 MDR-TB, there was neither pre-XDR TB nor XDR-TB found. (Table 3).

Table 3. Drug resistant tuberculosis detected by the first-and second-line phenotypic drug susceptibility testing.

Resistance classification	Number of detections N (%)	First-line MGIT DST (N=304)				Second-line MGIT DST (N=20)				Number of detections N (%)
		STR	INH	RIF	EMB	AMK	ETD	LVX	LZD	
Drug-resistance	101 (33.2)	81 (26.6)	60 (19.7)	20 (6.6)	1 (0.3)	0 (0)	5 (25)	0 (0)	0 (0)	
4 pan-susceptible	203 (66.8)	S	S	S	S					
Rifampicin mono-resistance	2 (0.7)	S	S	R	S	S	S	S	S	2 (10)
Isoniazid mono-resistance	16 (5.2)	S	R	S	S					
Mono-resistance, not INH/RIF	38 (12.5)	R	S	S	S					
Poly-resistance with INH resistance	27 (8.9)	R	R	S	S					
Multidrug resistance	15 (4.9)	R	R	R	S	S	S	S	S	10 (50)
		R	R	R	S	S	R	S	S	5 (25)
	2 (0.7)	S	R	R	S	S	S	S	S	2 (10)
	1 (0.3)	R	R	R	R	S	S	S	S	1 (5)
Pre-extensively resistance		S/R	R	R	S/R	S	S	R	S	0
Extensively resistance		S/R	R	R	S/R	S	S	R	R	0

***Note**: MGIT: mycobacterium growth indicator tube, DST: drug susceptibility testing, STR: streptomycin, INH: isoniazid, RIF: rifampicin, EMB: ethambutol, AMK: amikacin, ETD: ethionamide, LVX: levofloxacin, LZD: linezolid, R: resistant, S: susceptible.*

Discussion

The importance of DST for treatment was raised globally,[10] and reaffirmed the critical concentrations of isoniazid (INH) at 0.1 µg/mL, rifampicin (RIF) at 0.5 µg/mL and Levofloxacin (a fluoroquinolone) at 1.0 µg/mL for MGIT DST to reduce the rate of risk of being misclassified susceptibility of anti-TB agents .[3,16] Nevertheless, capacity of DST for MTB requires sophisticated laboratory infrastructure, proficient staff, good practice of quality assurance.

The study described the high LPA sensitivity of 90% and specificity of 99.3% for detecting RR-TB which are comparable to the previous report of Shah et al. in 2009 in Vietnam and a study in Uganda.[19] However, our findings were lower than sensitivity but higher than specificity in other studies on samples with smear positives in Uganda,[20] Peru,[21] India;[22] and from South Africa where LJ was used on the MDR-TB population;[23] Namibia [24] due to differences in their high-risk population or previously treated. This study showed higher sensitivity and specificity than studies reported by Mohamed et al. in 2020, [25] Hussain *et al*. in 2024,[26] in Ethiopia on popular including positive smear and the proportion LJ used as a gold standard.[22] Yadav et al. reported sensitivity and specificity were greater than 97% on smear-positive in 2013,[22] and compared with sequencing testing in 2021.[27] The current study described the LPA with very high sensitivity and specificity of detecting HR-TB of 98.3% and 99.2% which were higher than previous studies in South Africa,[25] India,[28] where LJ was considered the gold standard. The LPA sensitivity and specificity for MDR-TB of 88.9% and 94.1%, respectively seen similarly with studies in Peru.[21] However, Meaza et al. had a perfect sensitivity and specificity of 100% and some previous studies also showed higher values in Ethiopia,[22, 29] in India on presumptive MDR-TB size,[22] Uganda on smear positives.[20] MDR-TB (5.9%) detection was lower in Northwest Ethiopia.[29] Our study showed excellent agreements between FL-LPA and FL-MGIT DST at 98.7% (k=0.89), 99% (k=0.97), and 98.8% (k=0.91) for detecting RR-TB, HR-TB, and MDR-TB; respectively which were higher than a studies Wondale *et al*. for detecting RR-TB (k=0.49) and HR-TB (k=0.66), and lower for MDR-TB (k=1).[30]

The mono-resistant and multidrug resistant strains were shown the banding pattern by LPA. The Inferred mutation was 19 due to absence of wild type probes. The high mutations associated with *katG* MUT1 and *inhA* MUT1 were 39 and 19, respectively. For mono HR-TB, the mutation S315T1 associated with *katG* MUT1 was higher than mutation C15T associated with *inhA* MUT1. While the frequency of *inhA* mutation was D516V (5), H526Y (2), H526D (2), S531 (1). This also seen in the previous studies in high TB burden countries where DR strains transmitted continuously.[19,31]

Among 304 MTBC were subjected to MGIT DST against the first-line drugs including STR, INH, RIF, and EMB; The overall drug resistance to at least one anti-TB agent was 101 (33.2%), the highest value of resistance to

STR at 81 (26.6%) whereas 60 (19.7%) to INH, 20 (6.6%) to RIF, at least 1 (0.3%) to EMB and 5.9% for MDR-TB. INH mono-resistant-TB was higher than RIF mono-resistant-TB at 16 (5.2%) and 2 (0.7%). Our study can be comparable with previous studies in South Vietnam which found 26.3%, 19.4%, 16.6%, 2%, 1.1%, and 1.8% for resistance to at least one drug, STM, INH, RIF, EBM, and MDR in 2006;[32] and 19.8%, 3.4%, 2.5% and 3% for resistance to INH, RIF, EBM, and MDR in 2022 in Ca Mau province, Vietnam.[33] The slightly lower prevalence than their study may have come from 20 years ago when MDR-TB incident cases had not increased at the time of this study and Cho Ray located at a more crowded than Ca Mau may have detected a higher prevalence. In Ethiopia, RIF mono-resistant-TB, pan-susceptible was approximately 0.9% and 66.0% in Northwest Ethiopia, 2021.[29] Tutik *et al*. reported high proportions of DR-TB at 64.4%, 78%, and 14% for INH, RIF, and EMB while STR is the same at 13%.[34] This difference came from their study on patients diagnosed with DR-TB at entry. Their studies yielded SL-LPA, and MGIT to injectable anti-TB agents however WHO recommended drug groups for the treatment of RR/MDR-TB and all-oral regimens based on the susceptible TB and their benefit and harm.[35]

Proportions of DR-TB in a study in Ethiopia were higher than this study at 66%, 16%, and 17.9% for pan-susceptible, mono-resistance, and poly-resistance; respectively for newly diagnosed due to high incidence in this country. Nguyen et al. reported that central Vietnam (29.2%) was resistant to both antibiotics for phenotypically INH-resistant isolates at 46.3% had the Ser315Thr mutation. There were 8 different *rpo*B mutations in 22 (68.8%) of the RIF-resistant isolates with resazurin microtiter assay and polymerase chain reaction TaqMan.[17]

In this study, there were 18 (5.9%) MDR-TB including 1 (0.3%) pan-resistance. Of 20 MDR/RR-TB subjected to SC MGIT DST, only one quarter detected EDT-resistant TB while 100% were sensitive to AMK, LVX, and LZD. Neither pre-XDR-TB nor XDR-DR was detected due to LVX/LZD resistant-TB not being detected. In 2022, Wrohan *et al*. reported 88% MDR-TB, 8.2% as pre-XDR-TB, and 3.8% as XDR-TB in Ha Noi and Thanh Hoa, Northern, Vietnam performed by Xpert MTB/RIF, LPA, and DST.[11] The difference in MDR-TB detected came from the MDR-TB treated population and the old definition of pre-XDR-TB, XDR-TB used in the 2014 to 2016 period of study while our study conducted on presumptive TB and resistance classification was in line with WHO updated definitions from 2021. Minsk and Copenhagen had higher rates at 26.7% and 10%, 16.7% and 30.0%, 16.7% and 13.3% for MDR-TB, pre-XDR-TB, and XDR-TB detected among Belarusian HIV-positive patients.[36] However, the updated definition of pre-XDR-TB, and XDR-TB requires more studies of drugs for the TB-treating group A and B.[18]

Limitation

In this study, indirect LPA was only performed from inoculum culture of sputum, direct LPA test had not been studied from sputum sediment. Moreover, some DST discordant samples between LPA and MGIT should be detected by target next-generation sequencing. This study has not screened on previously treated TB patients yet.

Conclusion

Based on the gold standard of updated phenotypic DST, FL-LPA performance was compared in detecting RR-TB, HR-TB, and MDR-TB perfectly. Pre-XDR-TB and XDR-TB had not been found. We would keep practicing phenotypic culture-based DST with a diversity of recommended anti-TB agents and reaffirmed critical concentrations of INH, RIF and LVX of INH, RIF in the first line and LFV in the second line in screen DR-TB.

Ethical approval

This study was approved by the Ethics Committee in Biomedical Research of Cho Ray Hospital, Vietnam (No. 34-25/CN-HDDD).

Funding

None

Conflict of interest

None

CRedit authorship contribution statement

Phu Thien Truong: conceptualization, editing, validation; **Tran Ngoc Minh Le**: methodology, writing-original draft preparation; **Van Thi Hue Tran**: visualization, investigation, supervision; **Tung Thanh Phan**: data curation and software.

Acknowledgements

We appreciate the support of Microbiology Department in Cho Ray for this study.

References

[1] Stephenson J. WHO Report: Years of Progress in Global Tuberculosis Upset by COVID-19 Pandemic. JAMA Health Forum. 2022; 3(11): e224994. doi: 10.1001/jamahealthforum.2022.4994.

[2] World Health Organization. WHO operational handbook on tuberculosis. Module 3: Diagnosis - Rapid diagnostics for tuberculosis detection, third edition. Geneva: World Health Organization. 2024. Licence: CC BY-NC-SA 3.0 IGO. Available from: https://iris.who.int/bitstream/handle/10665/376155/9789240089501-eng.pdf?sequence=1.

[3] World Health Organization. Technical Report on critical concentrations for drug susceptibility testing of isoniazid and the rifamycins (rifampicin, rifabutin and rifapentine). Geneva: World Health Organization. 2021. Licence: CC BY-NC-SA 3.0 IGO. Available from: https://www.who.int/publications/i/item/9789240017283.

[4] World Health Organization. WHO consolidated guideline on tuberculosis. Module 3: Diagnosis - rapid diagnosis for tuberculosis detection, 2021 update. Geneva: World Health Organization. 2021. Licence: CC BY-NC-SA 3.0 IGO. Available from: https://www.

who.int/publishcations/i/item/9889240029415.

[5] Naidoo K, Perumal R, Ngema SL, Shunmugam L, Somboro AM. Rapid Diagnosis of Drug-Resistant Tuberculosis-Opportunities and Challenges. Pathogens. 2023; 13(1). doi: 10.3390/pathogens13010027.

[6] Hain Lifescience. GenoType MTBDR*plus* VER 2 - Instruction for use [Internet]. 2015 [cited 2025 April 29]. Available from: https://www.hain-lifescience.de/include_datei/kundenmodule/packungsbeilage/download.php?id=936.

[7] Traore AN, Rikhotso MC, Mphaphuli MA, Patel SM, Mahamud HA, Kachienga LO, et al. Isoniazid and Rifampicin Resistance-Conferring Mutations in *Mycobacterium tuberculosis* Isolates from South Africa. Pathogens. 2023; 12(8). doi: 10.3390/pathogens 12081015.

[8] World Health Organization. WHO operational handbook on tuberculosis. Module 4: treatment - drug-resistant tuberculosis treatment, 2022 update. Geneva: World Health Organization. 2022. Available from: https://iris.who.int/bitstream/handle/10665/365333/9789240065116-eng.pdf?sequence=1.

[9] Sinha P, Jacobson KR, Horsburgh CR, Jr., Acuna-Villaorduna C. At Long Last: Short, All-Oral Regimens for Multidrug-Resistant Tuberculosis in the United States. Open Forum Infect Dis. 2023; 10(4): ofad177. doi: 10.1093/ofid/ofad177.

[10] World Health Organization. Manual for selection of molecular WHO-recommended rapid diagnostic tests for detection of tuberculosis and drug-resistant tuberculosis. Geneva: World Health Organization. 2022. Licence: CC BY-NC-SA 3.0 IGO. Available from: https://www.who.int/publications/i/item/9789240042575.

[11] Wrohan I, Nguyen TA, Nguyen VN, Nguyen BH, Hoang TTT, Nguyen PC, et al. Predictors of treatment outcomes among patients with multidrug-resistant tuberculosis in Vietnam: a retrospective cohort study. BMC Infect Dis. 2022; 22(1): 68. doi: 10.1186/s12879-021-06992-x.

[12] World Health Organization. WHO announces updated definitions of extensively drug-resistant tuberculosis. Geneva: World Health Organization. 2021 [cited 2025 Feb 13]. Available from: https://www.who.int/news/item/27-01-2021-who-announces-updated-definitions-of-extensively-drug-resistant-tuberculosis.

[13] World Health Organization. Laboratory Diagnosis of Tuberculosis by Sputum Microscopy- The Handbook. SA Pathology. 2013. Available from: https://www.stoptb.org/sites/default/files/imported/document/TB_MICROSCOPY_HANDBOOK_FINAL.pdf.

[14] Mannan A, Iram S, Ahmad A, Hussain S, Ahmad BM. Identification of TBc: Using MTP 64 protein and cord formation. J Pak Med Assoc. 2017; 67(10): 1600-3. Available from: https://www.academia.edu/92703961/Identification_of_TBc_Using_MTP_64_protein_and_cord_formation.

[15] Pinhata JMW, Felippe IM, Gallo JF, Chimara E, Ferrazoli L, de Oliveira RS. Growth characteristics of liquid cultures increase the reliability of presumptive identification of *Mycobacterium tuberculosis* complex. J Med Microbiol. 2018; 67(6): 828-33. doi: 10.1099/jmm.0.000734.

[16] World Health Organization. Technical manual for drug susceptibility testing of medicines used in the treatment of tuberculosis. Geneva: World Health Organization. 2018. Licence: CC BY-NC-SA 3.0 IGO. Available from: https://www.who.int/publications/i/item/9789241514842.

[17] Nguyen H, Nguyen H, Ha D, Huong D, Trung V, Ngoc K, et al. Rifampicin resistant *Mycobacterium tuberculosis* in Vietnam, 2020-2022. J Clin Tuberc Other Mycobact Dis. 2024; 35: 100431. doi: 10.1016/j.jctube.2024.100431.

[18] European Centre for Disease Prevention and Control. Handbook on tuberculosis laboratory diagnostic methods in the European Union - updated 2022. Stockholm, STHML: European Centre for Disease Prevention and Control. 2023. doi: 10.2900/433652.

[19] Albert H, Bwanga F, Mukkada S, Nyesiga B, Ademun JP, Lukyamuzi G, et al. Rapid screening of MDR-TB using molecular Line Probe Assay is feasible in Uganda. BMC Infect Dis. 2010; 10: 41. doi: 10.1186/1471-2334-10-41.

[20] Pisal NS, Shah NC, Gandhi NN, Rao AS, Dedania MS, Pisal NS. A 1-year comparative evaluation of clinical performance of conventional direct composite restoration technique with a novel “custom shield” technique in class I compound lesions - A randomized clinical study. J Conserv Dent. 2022; 25(2): 135-9. doi: 10.4103/jcd.jcd_309_21.

[21] Puyen ZM, Acosta J, Obregon G, Pacheco E, Ramirez H, Mendoza A, et al. Use and evaluation of a line probe assay in patients with tuberculosis in Peru: 2011-2013. Rev Panam Salud Publica [serial online]. 2016; 39(1): 19-25. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27754534

[22] Meaza A, Kebede A, Yaregal Z, Dagne Z, Moga S, Yenew B, et al. Evaluation of genotype MTBDR*plus* VER 2.0 line probe assay for the detection of MDR-TB in smear positive and negative sputum samples. BMC Infect Dis. 2017; 17(1): 280. doi: 10.1186/s12879-017-2389-6.

[23] Maningi NE, Malinga LA, Antiabong JF, Lekalakala RM, Mbelle NM. Comparison of line probe assay to BACTEC MGIT 960 system for susceptibility testing of first and second-line anti-tuberculosis drugs in a referral laboratory in South Africa. BMC Infect Dis. 2017; 17(1): 795. doi: 10.1186/s12879-017-2898-3.

[24] Gunther G, Saathoff E, Rachow A, Ekandjo H, Diergaardt A, Marais N, et al. Clinical Evaluation of a Line-Probe Assay for Tuberculosis Detection and Drug-Resistance Prediction in Namibia. Microbiol Spectr. 2022; 10(3): e0025922. doi: 10.1128/spectrum.00259-22.

[25] Mahomed S, Mlisana K, Cele L, Naidoo K. Discordant line probe genotypic testing vs culture-based drug

susceptibility phenotypic testing in TB endemic KwaZulu-Natal: Impact on bedside clinical decision making. J Clin Tuberc Other Mycobact Dis. 2020; 20: 100176. doi: 10.1016/j.jctube.2020.100176.

[26] Hussain S, Sultan S, Riaz S, Hussain H, Javed H, Mazhar R. Synergy of Xpert (MTB/RIF) and Line probe assay for detection of rifampicin resistant strains of Mycobacterium tuberculosis. J Infect Dev Ctries. 2024; 18(8): 1241-8. doi: 10.3855/jidc.18945.

[27] Yadav RN, Kumar Singh B, Sharma R, Chaubey J, Sinha S, Jorwal P. Comparative Performance of Line Probe Assay (Version 2) and Xpert MTB/RIF Assay for Early Diagnosis of Rifampicin-Resistant Pulmonary Tuberculosis. Tuberc Respir Dis (Seoul). 2021; 84(3): 237-44. doi: 10.4046/trd.2020.0171.

[28] Raizada N, Sachdeva KS, Chauhan DS, Malhotra B, Reddy K, Dave PV, et al. A multi-site validation in India of the line probe assay for the rapid diagnosis of multi-drug resistant tuberculosis directly from sputum specimens. PLoS One. 2014; 9(2): e88626. doi: 10.1371/journal.pone.0088626.

[29] Yigzaw WB, Torrelles JB, Wang SH, Tessema B. Magnitude of Phenotypic and MTBDR*plus* Line Probe Assay First-Line Anti-Tuberculosis Drug Resistance Among Tuberculosis Patients; Northwest Ethiopia. Infect Drug Resist. 2021; 14: 497-505. doi: 10.2147/IDR.S292058.

[30] Wondale B, Medhin G, Abebe G, Tolosa S, Mohammed T, Teklu T, et al. Phenotypic and genotypic drug sensitivity of Mycobacterium tuberculosis complex isolated from South Omo Zone, Southern Ethiopia. Infect Drug Resist. 2018; 11: 1581-9. doi: 10.2147/IDR.S165088.

[31] Madukaji L, Okohu I, Usman S, Oyedum U, Enagi A, Usman A, et al. Early detection of Pre-XDR TB with line probe assay in a high TB burden country. Afr Health Sci. 2021; 21(3): 968-74. doi: 10.4314/ahs.v21i3.2.

[32] Huyen MN, Tiemersma EW, Lan NT, Cobelens FG, Dung NH, Sy DN, et al. Validation of the GenoType MTBDR*plus* assay for diagnosis of multidrug resistant tuberculosis in South Vietnam. BMC Infect Dis. 2010; 10: 149. doi: 10.1186/1471-2334-10-149.

[33] Callum J, Nguyen PTB, Martinez E, Nguyen VT, Garden F, Nguyen NV, et al. Prevalence and genetic basis of first-line drug resistance of Mycobacterium tuberculosis in Ca Mau, Vietnam. ERJ Open Res. 2022; 8(4). doi: 10.1183/23120541.00122-2022.

[34] Tutik K, Kristin PD, Yulia DS, Soedarsono. Discordance between Genexpert, line probe assay and drug susceptibility testing in assessing Drug-resistant tuberculosis. Indian J Forensic Medicine & Toxicology. 2021; 15(1): 1613-20. doi: https://doi.org/10.37506/ijfmt.v15i1.13642.

[35] World Health Organization Global Tuberculosis Programme. Rapid communication: key changes to the treatment of drug-resistant tuberculosis. Geneva, GE: World Health Organization. 2022 [cited 2025 Feb 13]. Available from: https://www.who.int/publications/i/item/WHO-UCN-TB-2022-2.

[36] Podlekareva DN, Folkvardsen DB, Skrahina A, Vassilenko A, Skrahin A, Hurevich H, et al. Tuberculosis Drug Susceptibility, Treatment, and Outcomes for Belarusian HIV-Positive Patients with Tuberculosis: Results from a National and International Laboratory. Tuberc Res Treat. 2021; 2021: 6646239. doi: 10.1155/2021/6646239.

Journal of Associated Medical Sciences 2025; 58(3): 263-268

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

Proficiency testing results of CD4+ T cell determination in Thailand

Saowanit Chairatanapiwong[1*], Panida Kulawong[1], Puwadon Lawapakul[1], Panida Pongpunyayuen[1], Phennapha Klangsinsirikul[2]

[1]AMS CMU EQA Unit, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Division of Transfusion Science, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 10 February 2025
Accepted as revised 18 July 2025
Available online 23 July 2025

Keywords:
Proficiency testing, CD4, CD4+ cells, Thailand.

ABSTRACT

Background: CD4+ T cell number is an important marker for assessing immune status, differentiating the severity of infections, and monitoring treatment in patients infected with human immunodeficiency virus (HIV). The determination of the percentage and absolute count of CD4+ T cells plays an essential role in treatment decisions and in ongoing patient care. Consequently, this analysis must be highly effective, with proper quality control processes. Participation in proficiency testing is a key component of the quality system.

Objectives: This study aimed to develop a proficiency testing program for CD4+ T cell analysis in Thailand.

Materials and methods: Two proficiency testing materials were prepared and assigned the numbers CD4-01, which was low level, and CD4-02, which was normal level. PT materials were tested for the four following parameters: %CD3+ cells, CD3+ cell count, %CD4+ T cells, and CD4+ T cell count in 39 participating laboratories. Performance assessment of the laboratories was shown as a z score according to the ISO13528:2022 standard.

Results: Proficiency testing materials met the acceptable criteria for homogeneity and stability according to ISO13528:2022. As due to uncertainty of assigned values are smaller than $0.3\sigma_{pt}$, the performance of participating laboratories was assessed using a z score. More than 89 percent of participated laboratories demonstrated acceptable performance across all four parameters, while less than six percent were classified as within the warning signal range, and less than eight percent were deemed unacceptable. The causes of unacceptable performance included inappropriate instrument settings, inadequate compensation, and issues with gating techniques.

Conclusion: Proficiency testing materials for CD4+ T cell determination were satisfactorily prepared and met the criteria for homogeneity and stability according to ISO13528:2022. The results of the pilot study conducted in 39 laboratories showed that more than 87 percent of the results were within acceptable range.

Introduction

CD4+ T cell enumeration is important for assessing immune status, differentiating the severity of infections, and monitoring treatment in patients infected with human immunodeficiency virus (HIV). This test plays an essential role in treatment decisions and ongoing patient care.[1] Thus, this analysis must be highly effective, with proper quality control processes. Moreover, it is essential to conduct proficiency testing to compare the results of CD4+ T cell enumeration across different instruments to ensure consistency. Participation in proficiency testing is a key component of the quality system according to the international

∗ **Corresponding contributor.**
Author's Address: *AMS CMU EQA Unit, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: saowanit.ch@cmu.ac.th
doi: 10.12982/JAMS.2025.096
E-ISSN: 2539-6056

standard for medical laboratory quality management, ISO 15189:2022, clause 7.3.7.3, which states that laboratories should engage in external quality assessment programs that comply with the ISO/IEC 17043 standard.[2]

CD4+ T cells are reported as both percentage and absolute count. The percentage of CD4+ T cells is analyzed using flow cytometry. The absolute count of CD4+ T cells can be estimated using either the dual-platform or single-platform technique. In the dual-platform technique, the absolute count is derived by multiplying the lymphocyte count (obtained from a blood cell counter) with the percentage of CD4+ T cells which is determined by flow cytometry. In contrast, the single-platform technique uses only a flow cytometer to calculate the absolute CD4+ T cell count. This method involves adding a known concentration of polystyrene beads to the stained blood sample for analysis.

The Associated Medical Sciences, Chiang Mai University External Quality Assessment (AMS CMU EQA) unit is dedicated to developing proficiency testing (PT) materials for laboratory tests that meet user needs, enabling participating laboratories to receive evaluation results that reflect the accuracy and precision of their laboratory test results. This ensures that results are reported effectively and uniformly, serving as an indicator for quality development and the ongoing maintenance of laboratory standards.

This study aimed to develop PT materials for CD4+ T cell enumeration to facilitate proficiency testing services and subsequently apply for ISO/IEC 17043 accreditation. The processes included the development of quality control materials, testing appropriate storage conditions, assessing the homogeneity and stability of the PT materials, and conducting practical trials in at least 30 medical laboratories across Thailand. This research was expected to lead to the development of PT materials for CD4+ T cell enumeration, enabling proficiency testing services in the future and providing knowledge for developing PT materials for other tests.

Materials and methods

Proficiency testing material preparation

Proficiency testing materials were prepared from rejected donor blood bags that did not meet volume standards from the Blood Bank Section, Maharaj Nakorn Chiang Mai Hospital. The study was reviewed by the ethics committee, Faculty of Associated Medical Sciences, Chiang Mai University (AMSEC-67EM-005). Fixed and preserved whole blood samples were prepared as two proficiency testing materials, which were then stored in a refrigerator.[3-8] The proficiency testing materials were assigned the numbers CD4-01, for low level, and CD4-02, for normal level of CD4+ T cells. In this proficiency testing program, PT materials were tested for the four following parameters: %CD3+ cells, CD3+ cell count, %CD4+ T cells, and CD4+ T cell count.

Assessment for homogeneity and stability of PT materials

For each lot of proficiency testing materials, ten tubes were randomly chosen and tested for %CD3+ cells, CD3+ cell count, %CD4+ T cells, and CD4+ T cell count in duplicate to assess the between-sample standard deviation (Ss).[9,10] The acceptance criteria for homogeneity, as outlined in ISO 13528:2022, involve comparing Ss to the σ_{pt} values used for laboratory assessment. The proficiency testing material was considered adequately homogeneous when $Ss \leq 0.3\sigma_{pt}$. Additionally, for each lot of proficiency testing materials, two tubes were randomly selected and tested in duplicate to evaluate stability during transport, and another two tubes were tested for stability at the end of the program. To evaluate stability during transportation, two tubes of each PT material were sent to the participating laboratory in Yala province, which was requested to return the materials to the PT provider. For stability at the end of the program, two tubes of each PT material were stored in a refrigerator until reporting deadline. These stability results (y_2) were compared to those of the homogeneity test (y_1). The σ_{pt} values used for laboratory assessment served as the acceptance criteria for the stability test. Proficiency materials were considered adequately stable when $|\bar{y}_1-\bar{y}_2| \leq 0.3\sigma_{pt}$, and sufficiently stable when $|\bar{y}_1-\bar{y}_2| \leq 0.3\sigma_{pt}+2\sqrt{u^2_{(y_1)}+u^2_{(y_2)}}$

Operation of the pilot study

The proficiency testing materials were packed in a box with ice packs and sent to 43 medical laboratories in Thailand that participated voluntarily. The laboratories were required to return their results to the provider within 15 days. The %CD3+ cells and %CD4+ T cells parameters were designated to be reported with one decimal place, while the CD3+ cell count and CD4+ T cell count were required to be reported as absolute numbers. The provider then assessed laboratory performance using z scores, which compared the analytical results to the consensus values from all participating laboratories. Assigned values (X_{pt}) and standard deviations (σ_{pt}) from the consensus values of results submitted by the laboratories were calculated using Algorithm A.[10] Subsequently, the uncertainties of the assigned values ($u_{(x_{pt})}$) were proved to be less than $0.3\sigma_{pt}$ thus validating the use of z scores. The z score was calculated using the equation: (participant value - X_{pt}) / σ_{pt} according to ISO 13528:2022 standard. For interpretation, the absolute value of the z score was considered as follows: $|z\ score| \leq 2.0$ indicated that the results were within acceptable limits, $2.0 < |z\ score| < 3.0$ indicated that the results were in a warning signal range, and $|z\ score| \geq 3.0$ indicated that the results were unacceptable.

Results

Assessment of homogeneity and stability of PT materials

The homogeneity test results of both PT materials in all parameters met the acceptance criteria for adequately homogeneous according to ISO 13528:2022, confirming their suitability for proficiency testing. The results of the homogeneity test for PT materials number CD4-01 and CD4-02 are shown in Table 1 and 2, respectively.

The stability test results indicated that PT materials number CD4-01 and CD4-02 remained stable throughout the transportation and at the end of the program. The results of the stability test for the PT materials were shown in Table 3-4.

Table 1. Results of the homogeneity test for PT material number CD4-01.

Test	Homogeneity means	Between-samples deviation (S_s)	$0.3\sigma_{pt}$	$S_s \leq 0.3\sigma_{pt}$	Result
%CD3+ cell	64.28%	0.29	0.93	Pass	Adequately homogeneous
%CD4+ T cell	1.48%	0.10	0.13	Pass	Adequately homogeneous
CD3+ cell (cells/ µL)	382.4	0.0	19.5	Pass	Adequately homogeneous
CD4+ T cell (cells/ µL)	8.8	0.6	0.9	Pass	Adequately homogeneous

Table 2. Results of the homogeneity test for PT material number CD4-02.

Test	Homogeneity means	Between-samples deviation (S_s)	$0.3\sigma_{pt}$	$S_s \leq 0.3\sigma_{pt}$	Result
%CD3+ cell	70.31%	0.47	0.68	Pass	Adequately homogeneous
%CD4+ T cell	40.03%	0.42	0.45	Pass	Adequately homogeneous
CD3+ cell (cells/ µL)	1,116.8	14.2	39.2	Pass	Adequately homogeneous
CD4+ T cell (cells/ µL)	635.7	0.0	22.4	Pass	Adequately homogeneous

Table 3. Results of the Stability Test for PT material number CD4-01.

Test	y_1	y_2	$\lvert \bar{y}_1 - \bar{y}_2 \rvert$	$0.3\sigma_{pt}$	Criteria*	Expanded criteria+	Result
%CD3+ cell	64.28%	64.53%^{T}	0.25	0.93	Pass	Not done	Adequately stable
		65.95%^{E}	1.68		Not pass	Pass	Sufficiently stable
%CD4+ T cell	1.48%	1.55%^{T}	0.07	0.13	Pass	Not done	Adequately stable
		1.30%^{E}	0.18		Not pass	Pass	Sufficiently stable
CD3+ cell (cells/ µL)	382.4	380.5^T	1.9	19.5	Pass	Not done	Adequately stable
		403.3^E	20.9		Not pass	Pass	Sufficiently stable
CD4+ T cell (cells/ µL)	8.8	9.0^T	0.3	0.9	Pass	Not done	Adequately stable
		7.8^E	1.0		Not pass	Pass	Sufficiently stable

***Note**: y_1: mean of homogeneity test, y_2: mean of stability test, ^{T}stability of transportation, ^{E}stability at the end of round, $^*\lvert \bar{y}_1 - \bar{y}_2 \rvert \leq 0.3\sigma_{pt}$, $^+\lvert \bar{y}_1 - \bar{y}_2 \rvert \leq 0.3\sigma_{pt} + 2\sqrt{u^2_{(y_1)} + u^2_{(y_2)}}$*

Table 4. Results of the Stability Test for PT material number CD4-02.

Test	y_1	y_2	$\lvert \bar{y}_1 - \bar{y}_2 \rvert$	$0.3\sigma_{pt}$	Criteria*	Expanded criteria+	Result
%CD3+ cell	70.31%	69.05 %^{T}	1.25	0.68	Not pass	Pass	Sufficiently stable
		69.63 %^{E}	0.68		Pass	Not done	Adequately stable
%CD4+ T cell	40.03%	38.50 %^{T}	1.53	0.45	Not pass	Pass	Sufficiently stable
		38.28 %^{E}	1.76		Not pass	Pass	Sufficiently stable
CD3+ cell (cells/ µL)	1,116.8	1,103.5^T	13.3	39.2	Pass	Not done	Adequately stable
		1,108.3^E	8.5		Pass	Not done	Adequately stable
CD4+ T cell (cells/ µL)	635.7	615.8^T	20.0	22.4	Pass	Not done	Adequately stable
		609.5^E	26.2		Not pass	Pass	Sufficiently stable

***Note**: y_1: mean of homogeneity test, y_2: mean of stability test, ^{T}stability of transportation, ^{E}stability at the end of round, $^*\lvert \bar{y}_1 - \bar{y}_2 \rvert \leq 0.3\sigma_{pt}$, $^+\lvert \bar{y}_1 - \bar{y}_2 \rvert \leq 0.3\sigma_{pt} + 2\sqrt{u^2_{(y_1)} + u^2_{(y_2)}}$*

Participating laboratories

Proficiency testing materials were sent to 43 medical laboratories, and results from 39 of those laboratories were returned to the proficiency testing provider. The techniques and analytical instruments used in the 39 laboratories are shown in Table 5.

Statistical data analysis of the participating laboratories

The data reported by a total of 39 laboratories included 39 data points for %CD3+ cells and %CD4+ T cells, but only 35 and 36 data points for CD3+ cell count and CD4+ T cell count for PT materials number CD4-01 and CD4-02 respectively. Four laboratories using the dual-platform technique did not report CD3+ cell count and CD4+ T cell count. Additionally, one laboratory reported CD3+ cell count and CD4+ T cell count only for PT material number CD4-01 because the differential WBC count for PT material number CD4-02 could not be analyzed by blood cell counter.

The distribution of all data was first analyzed using a kernel density plot as recommended by ISO13528:2022 and all plots showed only one high peak that revealed no large difference was enough to warrant a separate group of method assessment.[9,10] Subsequently, the assessment of laboratory performance was presented using z scores, which compared the analytical results with the consensus values from all laboratories. The number of laboratories in each performance score range for the pilot study is shown in Table 6.

Table 5. The techniques and analytical instruments used in participating laboratories.

No	Instrument	Number of laboratories (%)		
		Dual-platform technique	Single-platform technique	Total
1	Beckman Coulter FC500	10	9	19 (48.7)
2	Beckman Coulter DxFLEX	1	7	8 (20.5)
3	BD FACSCalibur	2	4	6 (15.4)
4	Beckman Coulter Navios	1	2	3 (7.7)
5	BD FACSLyric	2	0	2 (5.1)
6	BD FACSVia	0	1	1 (2.6)
	Total	**16 (41)**	**23 (59)**	**39 (100)**

Table 6. Number of participating laboratories in each performance level.

Test	PT material number	N	Laboratory performance Number of laboratories (%)		
			Acceptable	Warning signal	Unacceptable
%CD3+ cell	CD4-01	39	37 (94.9)	0 (0.0)	2 (5.1)
	CD4-02	39	35 (89.7)	1 (2.6)	3 (7.7)
%CD4+ T cell	CD4-01	39	35 (89.8)	2 (5.1)	2 (5.1)
	CD4-02	39	34 (87.2)	2 (5.1)	3 (7.7)
CD3+ cell count	CD4-01	35	32 (91.4)	2 (5.7)	1 (2.9)
	CD4-02	36	34 (94.4)	2 (5.6)	0 (0.0)
CD4+ T cell count	CD4-01	35	32 (91.4)	1 (2.9)	2 (5.7)
	CD4-02	36	34 (94.4)	2 (5.6)	0 (0.0)

Discussion

As stated in ISO 15189:2022, section 7.3.7.3, to monitor the performance of the examination methods, laboratories should participate in External Quality Assessment (EQA) programs that are appropriate to the examinations performed. This section also emphasizes the importance of selecting EQA programs that are both suitable and competent. Where possible, EQA programs that meet the requirements of ISO/IEC 17043 should be selected. In alignment with this requirement, this study was developed in compliance with ISO/IEC 17043:2023. Furthermore, the statistical analysis of this study was conducted in accordance with ISO 13528:2022, as recommended by ISO/IEC 17043:2023 for ensuring the validity and reliability of performance evaluation.

Both PT materials were adequately homogeneous ($Ss \leq 0.3\sigma_{pt}$) in all four parameters. The difference in the average values of %CD3+ cells and %CD4+ T cells between the homogeneity mean (when prepared) compared to the stability mean of transportation and stability mean at the end of the program did not exceed 2%, and the maximum difference for absolute CD3+ cells and CD4+ T cells count was only 26.2 cells/µL. Some parameters were not classified as adequately stable (as shown in Table 3 and Table 4). However, when the uncertainty of measurement was considered to apply expanded criteria,

the results were still considered acceptable and classified as "sufficiently stable" for the assessment. These data shown that both PT materials for all parameters were suitable for this proficiency program.

Although around 5-10 percent of the participating laboratories assessed z score evaluation results outside the acceptable range (absolute z score ≤2.0).

Potential causes of analytical errors that laboratories with z score evaluation results outside the acceptable range should investigate were as follows.[11-14]

First, the voltage settings of the detector and compensation settings were inappropriate; this might result in the lymphocyte populations not being clearly distinguished. Additionally, the placement of the quadrant boundaries may not be suitable.

Secondly, the gating of lymphocyte populations was inadequate. When analyzing CD3 and CD4 in lymphocyte populations, the largest group consisted of T cells (CD3+) at approximately 61-85%. These cells had a higher CD45 expression as compared to B and NK cells (CD3-, CD4-), with NK cells exhibiting higher SSC properties than T cells. If the gating was not comprehensive, it might lead to inaccuracies in analyzing %CD3+ cells and %CD4+ T cells. Additionally, care should be taken to ensure that the proportion of monocytes (CD3-, CD4+) does not exceed 5%.

In some cases, the number of lymphocytes analyzed in the sub-population was less than 2,500 cells. If using an auto-stop setting, it should be set at more than 2,500 cells (typically set to 3,000 lymphocytes). It was found that some laboratories using the single-platform technique which set the auto stop at 3,000 for the bead population. However, in blood samples with a low cell count, the bead population may reach 3,000 events before the lymphocyte count reaches 2,500 cells.

Moreover, the selection of beads for absolute cell count calculation was inappropriate. Beads located at the top may represent clusters of more than one bead, while those at the bottom may be broken beads or non-bead particles.

The last cause was the PT materials may not be suitable for analyzing differential cell counts using certain blood cell analyzers, as there may be discrepancies in results across different analytical principles. Additionally, some analyzers may not be capable of performing differential cell counts. When analyzing the absolute count data from all laboratories using a Kernel density plot, no distinct peak separation was observed, leading to the analysis of z scores from the overall data. This may result in some laboratories using the Dual-platform technique having z scores for Absolute CD3+/CD4+ T cell counts that fall outside the acceptable range. However, if a laboratory's z score for %CD3+/%CD4+ T cell counts is within the acceptable range, it indicates that the laboratory's performance in the analysis is considered acceptable.

Conclusion

PT materials for CD4+ T cell determination were satisfactorily been prepared and passed the criteria for homogeneity and stability according to ISO13528:2022. The results of the pilot study in 39 laboratories were within acceptable range more than 87 percent. This study may offer valuable insights for evaluating the performance of CD4+ T cell determination in clinical laboratories and could serve as a framework for developing future proficiency testing programs according to ISO17043.

Ethical approval

The study was reviewed by the ethics committee, Faculty of Associated Medical Sciences, Chiang Mai University (AMSEC-67EM-005).

Funding

This study was granted by the Faculty of Associated Medical Sciences, Chiang Mai University.

Conflict of interest

The authors declare no conflicts of interest regarding this manuscript.

CRedit authorship contribution statement

Saowanit Chairatanapiwong: conceptualization, methodology, formal analysis, writing - original draft; **Panida Kulawong**: investigation; **Puwadon Lawapakul**: investigation; **Panida Pongpunyayuen**: resources; **Phennapha Klangsinsirikul**: supervision, writing - review and editing.

References

[1] Centers for Disease Control and Prevention [Internet]. Recommendations for HIV prevention with adults and adolescents with HIV in the United States. [updated 2014 Dec 11; cited 2024 Oct 2]. Available from: https://stacks.cdc.gov/view/cdc/44064.

[2] ISO 15189 (2022) Medical laboratories - Requirements for quality and competence. ISO, Geneva; 2022

[3] Arnold JE, Matsche MA, Rosemary K. Preserving whole blood in formalin extends the specimen stability period for manual cell counts for fish. Vet Clin Pathol. 2014; 43(4): 613-20. doi: 10.1111/vcp.12214.

[4] Chow S, Hedley D, Grom P, Magari R, Jacobberger JW, Shankey TV. Whole blood fixation and permeabilization protocol with red blood cell lysis for flow cytometry of intracellular phosphorylated epitopes in leukocyte subpopulations. Cytometry A. 2005; 67(1): 4-17. doi: 10.1002/cyto.a.20167.

[5] Ding T, Bergeron M, Seely P, Yang X, Diallo TO, et al. Compatibility of stabilized whole blood products with CD4 technologies and their suitability for quality assessment programs. PLOS ONE 2014; 9(8): e103391. doi: 10.1371/journal.pone.0103391

[6] Noulsri E, Lerdwana S, Pattanapanyasat K. Low-CD4+ T-lymphocyte blood samples for external quality assessment of CD4 testing in resource-poor settings. Accred Qual Assur. 2014; 19: 423-31. doi: 10.1007/s00769-014-1086-z.

[7] Pobkeeree V, Lerdwana S, Siangphoe U, et al. External Quality Assessment Program on CD4+ T-Lymphocyte

Counts for Persons with HIV/AIDS in Thailand: History and Accomplishments. APJAI 2009; 27: 225-32. PMID: 20232577.

[8] Yibalih NK, Wolday D, Kinde S, Weldearegay GM. External quality assessment on CD4+ T-cell count using in-house proficiency testing panels for CD4 count laboratories in Addis Ababa, Ethiopia Ethiop J Health Sci. 2019; 29(3): 309-20. doi: 10.4314/ejhs.v29i3.3.

[9] ISO/IEC 17043 (2023) Conformity assessment-general requirements for proficiency testing. ISO, Geneva; 2023.

[10] ISO 13528 (2022) Statistical methods for use in proficiency testing by interlaboratory comparison. ISO, Geneva; 2022

[11] Illingworth A, et.al. [Internet]. Assay Development and Validation of T, B, and NK Lymphocyte Subset Enumeration. [updated 2024 Apr 15; cited 2024 Oct 2]. Available from: https://www.cytometry.org/web/modules/Module_27.pdf.

[12] Centers for Disease Control. Guidelines for the performance of CD4+ T-cell determinations in persons with human immunodeficiency virus infection. MMWR Recomm Rep. 1992; 41(RR-8): 1-17. doi: 10.1037/e545992006-001

[13] Centers for Disease Control. Guidelines for performing single-platform absolute CD4+ T-cell determinations with CD45 gating for persons infected with human immunodeficiency virus. MMWR Recomm Rep. 2003; 52(RR-2): 1-13. PMID: 12583540.

[14] Centers for Disease Control. [Internet]. Revised guidelines for performing CD4+ T-cell determinations in persons infected with human immunodeficiency virus (HIV). [updated 2001 Feb 5; cited 2024 Oct 2]. Available from: https://www.cdc.gov/mmwr/preview/mmwrhtml/00045580.htm.

Journal of Associated Medical Sciences 2025; 58(3): 269-277

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

Implications of bilirubin testing and the prevalence of abnormal liver function tests among rural adults in Mae Ka Subdistrict, Chiang Mai, Thailand

Fahsai Kantawong[1*], Khanittha Punturee[1], Suwatsin Kittikunnathum[1], Pharisa Nanthawong[1], Kanya Preechasuth[2], Chayada Sitthidet Tharinjaroen[2], Witida Laopajeon[3], Rujirek Chaiwongsa[4], Nutjeera Intasai[5], Tanyarat Jomgeow[5]

[1]Division of Clinical Chemistry, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Division of Clinical Microbiology, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[3]Division of Clinical Immunology, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[4]Division of Transfusion Sciences, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[5]Division of Clinical Microscopy, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 16 February 2025
Accepted as revised 18 July 2025
Available online 27 July 2025

Keywords:
Bilirubin, liver function test, ALBI score, medication.

ABSTRACT

Background: Liver function tests (LFTs) measure the levels of specific enzymes and proteins in the blood that can change when the liver is damaged. Bilirubin, a byproduct of heme degradation, primarily from red blood cell breakdown (approximately 80%), is commonly included in these assessments. Although bilirubin and liver enzymes are well-established markers of liver function, they do not always indicate the presence of liver lesions. Therefore, for accurate diagnosis, changes in liver enzyme and bilirubin concentrations should be interpreted in the context of baseline characteristics such as pesticide exposure, alcohol consumption, and medication use.

Objectives: To investigate the prevalence of abnormal LFTs and alterations in bilirubin concentrations among adults in Mae Ka Subdistrict, Chiang Mai, Thailand.

Materials and methods: A descriptive cross-sectional study was conducted at the Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, from January to March 2024. Adults from Mae Ka Subdistrict participated in the study. A total of 102 subjects, aged 23 to 75 years (mean age 56.88±11.26 years), were included. Liver function tests, including ALT, AST, ALP, ALB, total bilirubin, and direct bilirubin, were performed using an automated analyzer (BA 400). The ALBI score and Pearson correlation were used to evaluate the relationships between liver enzymes, bilirubin levels, and baseline characteristics. A $p \leq 0.05$ was considered statistically significant. Data analysis was performed using SPSS version 17 and Microsoft Excel.

Results: A strong correlation was observed between the use of medications for diabetes mellitus, hypertension, or hyperlipidemia and abnormal LFTs. Medication use was associated with elevations in total bilirubin (TB), ALBI score, ALT, and ALP. However, not all individuals with elevated bilirubin levels had abnormal liver enzymes, and not all individuals with abnormal liver enzymes had elevated bilirubin levels. The ALBI score identified two individuals with normal liver enzyme and albumin (ALB) levels who may require closer monitoring and further investigation for potential liver disease.

Conclusion: Medication use was associated with abnormal LFTs, potentially serving as an early warning sign. The underlying non-communicable diseases (NCDs) that necessitate medication use may be the primary contributors. Early diagnosis and treatment of liver disease, along with prevention of NCDs and metabolic syndrome, may improve health outcomes in this community.

∗ **Corresponding contributor.**
Author's Address: *Division of Clinical Chemistry, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: fahsai.k@cmu.ac.th
doi: 10.12982/JAMS.2025.097
E-ISSN: 2539-6056

Introduction

Liver disease remains one of the leading causes of death in Thailand. In rural areas, the most common causes of liver abnormalities include alcohol consumption, metabolic syndrome related to overweight and obesity, and chronic viral hepatitis B and C.[1] Despite medical advances, significant challenges remain in the diagnosis and management of liver disease. Early and regular monitoring through liver function tests (LFTs) is essential for preventing or reducing the risk of disease progression.[2] Recent technological developments, such as the application of artificial intelligence (AI), have enhanced the diagnostic capabilities for liver diseases.[3] However, in rural and low- to middle-income settings, the adoption of advanced healthcare technologies remains limited, primarily due to a shortage of trained healthcare professionals.[4] As a result, routine clinical chemistry tests, including LFTs, continue to be a practical and accessible tool for liver disease screening in these communities.[2]

In routine health check-ups, bilirubin testing is often not performed unless clinical symptoms such as jaundice, dark urine, pale stools, or abdominal pain are present. Some studies suggest that elevated total bilirubin could serve as a reflex indicator, reducing the need for direct (conjugated) bilirubin testing. Specifically, conjugated bilirubin measurement may not be necessary in apparently healthy males with total bilirubin levels below 21.0 µmol/L (<1.23 mg/dL) and females with levels below 17.0 µmol/L (1.0 mg/dL).[5] Bilirubin testing also has wider diagnostic and prognostic applications. For instance, Yamashita et al. introduced the albumin-bilirubin (ALBI) score as a simple method to assess primary biliary cholangitis in a large Japanese cohort. [6] In this context, bilirubin levels less than or equal to 0.6 times the upper limit of normal, or normal alkaline phosphatase (ALP), were associated with the lowest risk of liver transplantation or death.[7] Beyond its role in liver function testing, the albumin-to-bilirubin (ALB/BIL) ratio has been explored as a predictor of disease severity in patients with glaucoma.[8] Although an isolated increase in conjugated bilirubin with normal total bilirubin may not always be clinically significant, it can occasionally indicate underlying liver disease. Such isolated conjugated hyperbilirubinemia is often linked to adverse drug reactions.[9]

Although abnormalities in LFTs can be non-specific and occur in a variety of clinical conditions, they may represent the earliest indication of potentially serious disease in otherwise asymptomatic individuals. It is essential that such findings be interpreted within the context of the patient's overall clinical presentation and supported by other laboratory investigations. When LFT abnormalities are detected, further evaluation may be warranted to determine the underlying cause. This may include repeat liver function testing, imaging studies, and, in some cases, liver biopsy for definitive diagnosis.

Monitoring liver function abnormalities at the community level is especially important in populations with known risk factors such as pesticide exposure in agricultural settings or prevalent alcohol consumption. Early identification of individuals at risk for advanced fibrosis or cirrhosis could enable timely lifestyle interventions and treatment, thereby potentially halting disease progression and reducing both morbidity and mortality.[10]

This study aims to investigate the prevalence of abnormal LFTs among adults in Mae Ka Subdistrict, Chiang Mai, Thailand. The findings may inform public health strategies and educational initiatives aimed at reducing exposure to liver disease risk factors and promoting liver health in the community.

Materials and methods

Study population

A cross-sectional study was conducted with 102 participants residing in 14 villages within the Mae Ka Subdistrict, Chiang Mai Province, Thailand. Mae Ka Subdistrict is geographically located at approximately 19°19'58"N latitude and 98°57'00"E longitude. Inclusion criteria included individuals aged 18 years or older who were native residents of the selected villages. Individuals with a prior diagnosis of liver disease were excluded. Sample collection was conducted on Sunday afternoons in February 2024, and participant interviews were carried out by internship Medical Technology students. The purpose of the study was explained to all participants in the local language, and written informed consent was obtained in accordance with the ethical principles outlined in the Declaration of Helsinki (1975). The study protocol was approved by the Clinical Research Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University.

Data collection

A structured questionnaire was administered to collect information on participants' demographic and health-related characteristics. Data collected included age, sex, weight, height, occupation, alcohol consumption, exposure to pesticides, family history of disease, known medical conditions, and details of current medication use. The questionnaire was administered through face-to-face interviews conducted by internship Medical Technology students.

Blood chemistry analysis

A total of 5 mL of non-fasting venous blood was collected from each participant. Serum was separated by centrifugation at 2,500 rpm for 10 minutes on the same afternoon as the sample collection. All liver function tests were performed on non-fasting serum samples using an automated chemistry analyzer (BA 400). Six biochemical parameters related to liver function were assessed in this study: albumin (ALB), albumin-bilirubin (ALBI) score, total bilirubin (TB), direct bilirubin (DB), alanine aminotransferase (ALT), aspartate aminotransferase (AST), and alkaline phosphatase (ALP). The analytical principles of these assays were summarized in Table 1.

Table 1. Assay principles.

Test	**Principle** (BioSystems)
ALP (alkaline phosphatase)	Hydrolysis of 4-nitrophenylphosphate to 4-nitrophenol and 2-amino-2-methyl-1-propanol; absorbance of 4-nitrophenol measured at 405 nm.
ALT (alanine aminotransferase)	Transfer of the amino group from alanine to 2-oxoglutarate to form pyruvate and glutamate; pyruvate is reduced to lactate by lactate dehydrogenase with concomitant oxidation of NADH to NAD^+; rate of NADH depletion measured at 340 nm.
AST (aspartate aminotransferase)	Transfer of the amino group from aspartate to 2-oxoglutarate to form oxaloacetate and glutamate; oxaloacetate is reduced to malate by malate dehydrogenase with concomitant oxidation of NADH to NAD^+; rate of NADH depletion measured at 340 nm.
ALB (albumin)	Formation of a colored complex between albumin and bromocresol green; absorbance measured at 640 nm.
DB (direct bilirubin)	Reaction of conjugated bilirubin with 3,5-dichlorophenyl diazonium salt to form a diazo complex; absorbance measured at 535 nm.
TB (total bilirubin)	Coupling of both unconjugated and conjugated bilirubin with diazonium reagent in the presence of cetrimide; absorbance measured at 535 nm.

ALBI calculations

The ALBI score was calculated using a previously established formula: $(\log_{10}[\text{bilirubin}] \times 17.1 \times 0.66) + (\text{albumin} \times 10 \times -0.085)$, where bilirubin was measured in mg/dL and albumin in g/dL. ALBI grade cut-off points were defined as follows: a score ≤-2.60 was classified as grade 1; a score >-2.60 to ≤-1.39 was classified as grade 2; and a score >-1.39 was classified as grade 3. [6, 11]

Statistical analysis

Liver function test values were presented as mean±SD, and baseline characteristics were summarized as percentages. The Kolmogorov–Smirnov test was used to assess the normality of continuous data. Pearson correlation analysis was employed to evaluate associations between LFT parameters and questionnaire-derived variables. Group comparisons of LFT parameters by baseline characteristics were conducted using the independent Student's t-test and the Mann–Whitney U test. All statistical analyses were performed using SPSS version 17 (IBM Corp., Armonk, NY, USA) and Microsoft Excel (Microsoft Corp., Redmond, WA, USA), with p values ≤ 0.05 considered statistically significant.

Results

Clinical characteristics of participants

The clinical and baseline characteristics of the 102 participants were summarized in Table 2. The mean age at enrollment was 57.6 years for men and 55.8 years for women, with men comprising 61.8% of the cohort. All participants were asymptomatic for liver disease at the time of sampling. Sixty-three percent reported regular alcohol consumption, 47% reported occupational or environmental pesticide exposure, and 39% were receiving medications for diabetes mellitus, hypertension, or hyperlipidemia.

Table 2. Demographic data (N=102).

Characteristic	**N** (%)	**BMI** (mean±SD)
Gender		
Male	63 (61.8)	23.7 (±2.8)
Female	39 (38.2)	24.2 (±4.4)
Male age		
<50	16 (15.69)	24.6 (±2.8)
51-60	17 (16.67)	24.4 (±3.2)
>60	30 (29.41)	22.7 (±2.2)
Female age		
≤50	12 (11.76)	24.9 (±6.2)
51-60	12 (11.76)	24.1 (±3.4)
>60	15 (14.71)	23.8 (±3.4)
Behavior characteristics and underlying disease		
Agriculture exposed to pesticide	46 (45.1)	-
Alcohol consumption	63 (61.8)	-
Medication (diabetes mellitus/ hypertension/ hyperlipidemia)	39 (38.2)	-

Frequency distribution of self-reported alcohol consumption among study participants (N=63) was shown in Table 3. Most drinkers reported consuming alcohol one to three times per week (N=20; 31.7%), followed by occasional consumption (N=18; 28.6%). Less frequent patterns included three to five times per week (N=10; 15.9%), daily intake (N=9; 14.3%), and almost daily intake (N=6; 9.5%).

Table 3. Alcohol consumption frequency.

Alcohol consumption frequency	Count (N)
1-3 times per week	20
Occasionally	18
3-5 times per week	10
Daily	9
Almost daily	6
Total	63

The biochemical results for the cohort were presented in Table 4. Mean±SD values for each parameter were as follows: ALT 29±16 U/L, AST 26±12 U/L, ALP 94 ± 29 U/L, total bilirubin 0.50±0.27 mg/dL, direct bilirubin 0.22±0.14 mg/dL, and albumin 4.60±0.26 g/dL.

Table 4. Biochemical data of the subjects (N=102).

Parameters	Mean (±SD)		
	Males	Females	Total
ALT (U/L)	34 (±18)	21 (±8)	29 (±16)
AST (U/L)	28 (±13)	21 (±6)	26 (±12)
ALP (U/L)	97 (±32)	89 (±24)	94 (±29)
TB (mg/dL)	0.5 (±0.28)	0.4 (0.21)	0.5 (±0.27)
DB (mg/dL)	0.24 (±0.16)	0.17 (± 0.09)	0.22 (±0.14)
ALB (mg/dL)	4.6 (±0.27)	4.6 (±0.24)	4.6 (±0.26)

ALBI calculations

ALBI scores were calculated as described previously, with a mean score of -8.0535 (Table 5). Most participants (98.0%, N=98) were classified as ALBI grade 1 (score ≤-2.60), whereas two subjects (2.0%) fell into ALBI grade 2 (score >-2.60 to ≤-1.39). Both grade 2 subjects were male and had normal AST, ALT, ALP, and albumin levels but elevated total and direct bilirubin.

Table 5. ALBI score of the patients (N=102).

Parameters	Mean (±SD)
ALBI score	-8.0535 (±2.4498)
Male ALBI score	-7.4288 (±2.2582)
Female ALBI score	-9.0625 (±2.4369)
Minimum	-12.8328
Maximum	-2.1397

The subject with the highest ALBI score (-2.1397) was overweight (body mass index [BMI] 26.5 kg/m²), reported hypertension and hyperlipidemia but was not on related medications, and had histories of alcohol consumption and pesticide exposure. The second subject (ALBI -2.5887; BMI 22.5 kg/m²) was a 36-year-old male with diabetes mellitus who was receiving treatment and exhibited no liver enzyme elevations. ALBI grade cut-off values were defined as follows: ≤-2.60 (grade 1), >-2.60 to ≤-1.39 (grade 2), and >-1.39 (grade 3).[6]

Correlation between liver function tests

TB exhibited a strong positive correlation with DB, with a Pearson correlation coefficient of r=0.937 (p<0.001; Figure 1A). Given that the ALBI score was derived in part from TB, a similarly strong correlation was observed between ALBI score and DB (r=0.935; p<0.001; Figure 1B).

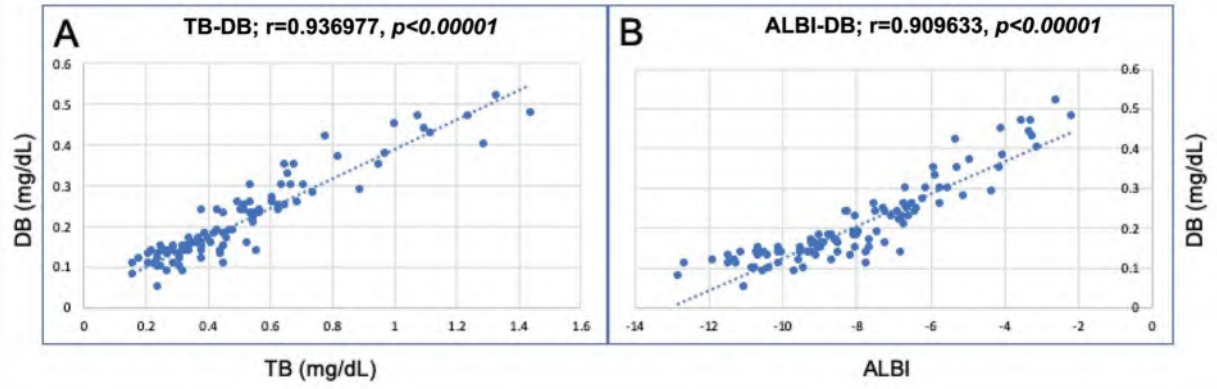


Figure 1. *Scatter plots with fitted regression lines illustrating the relationships among total bilirubin (TB), direct bilirubin (DB), and albumin-bilirubin (ALBI) score in 102 participants. A: TB versus DB (r=0.937, p<0.0001), B: ALBI score versus DB (r=0.910, p<0.0001).*

The relationship between serum AST and ALT levels in 102 adult participants from Mae Ka Subdistrict, Chiang Mai, Thailand was shown in Figure 2. Each point represents an individual's paired AST and ALT values, with the dashed line indicating the linear regression fit (r=0.783, *p*<0.0001). The moderate positive correlation suggests that elevations in AST tend to coincide with elevations in ALT, although several outliers were present at higher values of AST (up to approx. 190 U/L) and ALT (up to ~145 U/L). This association reflects concurrent hepatocellular enzyme release in response to liver injury or stress, reinforcing the utility of assessing both enzymes together in clinical screening.

In contrast, weak positive associations were found (Figure 3) between DB and AST (r=0.215; *p*=0.028), DB and ALT (r=0.198; *p*=0.041), ALT and ALP (r=0.184; *p*=0.053), and AST and ALP (r=0.162; *p*=0.075;). ALB levels were inversely related to ALP (r=-0.238; *p*=0.015) and showed a negligible negative association with AST (r=-0.031; *p*=0.742).

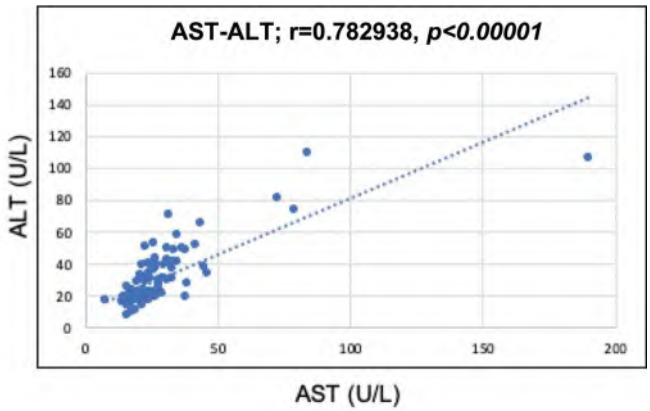


Figure 2. *Scatter plot showing the moderate positive correlation between serum AST and ALT levels in 102 participants (r=0.783, p<0.0001).*

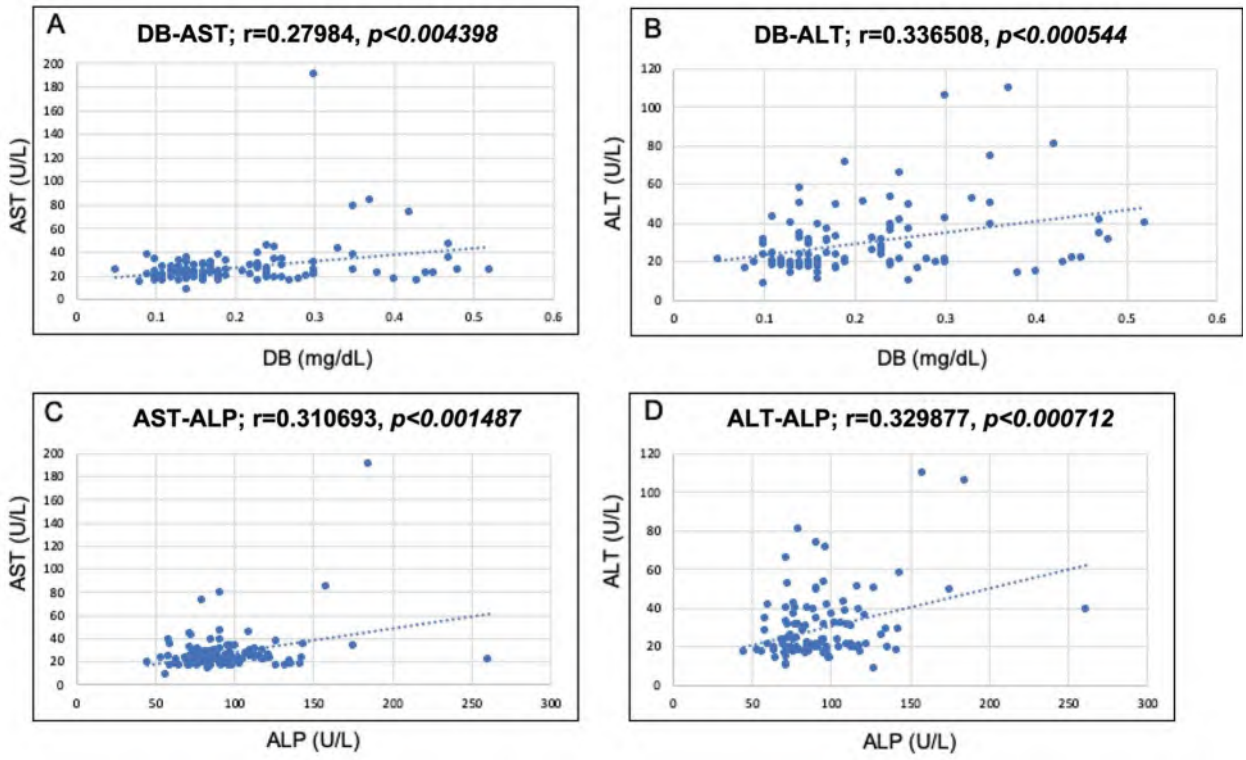


Figure 3. *Scatter plots with fitted regression lines illustrating weak positive associations among liver function parameters in 102 participants. Panel A: DB versus AST (r=0.280, p=0.004), Panel B: DB versus ALT (r=0.337, p=0.001), Panel C: AST versus ALP (r=0.311, p=0.001), Panel D: ALT versus ALP (r=0.330, p=0.001).*

Association between baseline characteristics and abnormal biochemical data

The frequency and distribution of abnormal biochemical parameters among the 102 participants were summarized in Table 6. Elevated ALT (>41 U/L) was observed in 14.7% of participants (N=15), while 6.9% (N=7) had AST levels above the reference threshold (>40 U/L). Elevated ALP was found in 11.8% of male participants (cut-off >115 U/L) and 7.8% of female participants (cut-off >105 U/L). Additionally, 6.9% (N=7) showed albumin levels exceeding 5.0 g/dL.

For bilirubin parameters, 5.9% (N=6) had elevated total bilirubin (>1.0 mg/dL), and 41.2% (N=42) had elevated direct bilirubin (>0.2 mg/dL). Notably, 14.7% (N=15) of participants showed elevated direct bilirubin while their total bilirubin remained within the normal range, suggesting a pattern of isolated conjugated hyperbilirubinemia. A total of six participants had both total and direct bilirubin elevated.

Table 6. Frequency of abnormal biochemical parameters among 102 participants. Cut-off values were based on standard laboratory reference ranges.

Biochemical (cut-off value)	**N (%)**
High ALT (>41 U/L)	15 (14.71)
High AST (>40 U/L)	7 (6.86)
High ALP male (>115U/L)	12 (11.76)
High ALP female (>105 U/L)	8 (7.84)
High ALB (>5.0 mg/dL)	7 (6.86)
High TB (>1.0 mg/dL)	6 (5.88)
High DB (>0.2 mg/dL)	42 (41.18)
High TB & high DB	6 (5.88)
High DB & other parameters & normal TB	15 (14.71)

***Note**: ALP values were sex specific. "High DB & other parameters & normal TB" refers to participants with elevated direct bilirubin and at least one other abnormal parameter, while total bilirubin remained within the normal range.*

Association between baseline characteristics and abnormal biochemical data were analyzed. Participants were stratified by alcohol consumption, pesticide exposure, and use of medications for noncommunicable diseases (NCDs: diabetes mellitus, hypertension, or hyperlipidemia). No significant differences in any liver function parameters were observed between alcohol drinkers and non-alcohol drinkers, nor between participants with and without pesticide exposure. In contrast, participants receiving NCD medications (N=40) exhibited significantly higher total bilirubin (TB), ALBI score, ALT, and ALP levels compared with those not on such medications (N=62) (Figure 4). Albumin and AST were normally distributed, so both the independent Student's t-test (for normally distributed variables) and the Mann-Whitney U test (for non–normally distributed variables) were applied to assess differences between groups. All tests with $p \leq 0.05$ considered statistically significant.

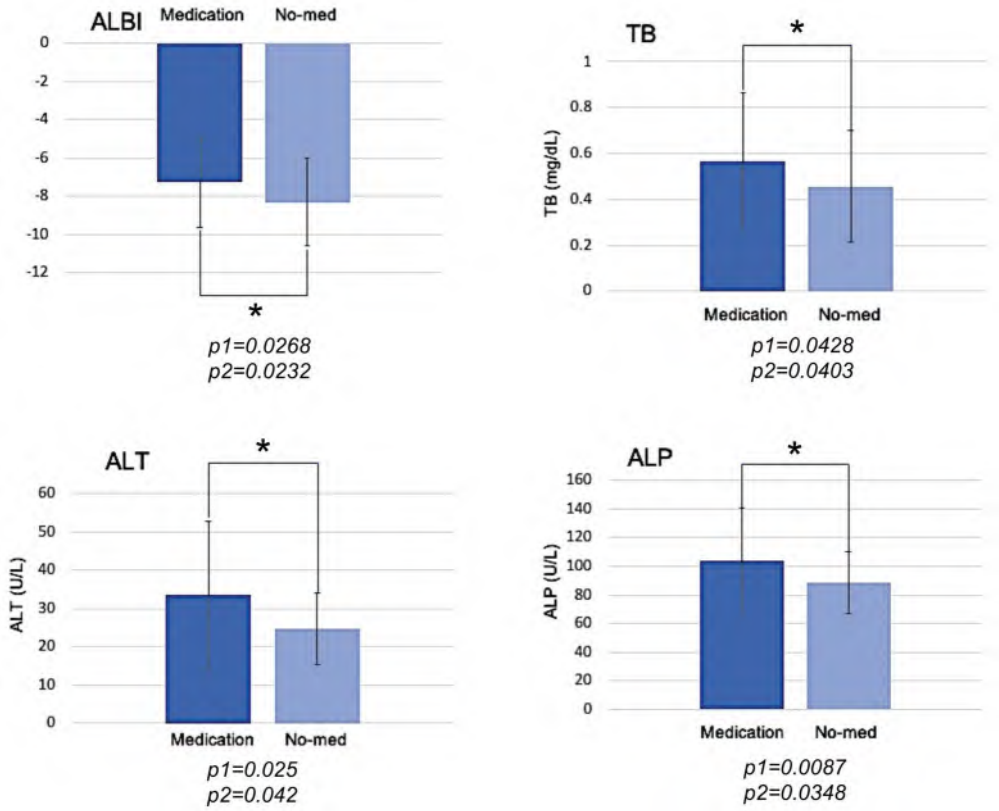


Figure 4*. Comparison of liver function parameters between participants receiving medications for diabetes mellitus, hypertension, or hyperlipidemia (Medication) and those not receiving such medications (No-med).*

Note*: Bars represent mean±SD for ALBI score (top left), total bilirubin (TB; top right), alanine aminotransferase (ALT; bottom left), and alkaline phosphatase (ALP; bottom right).* p_1 *indicates the independent Student's t-test and* p_2 *the Mann-Whitney U test.*

Discussion

The prevalence of asymptomatic abnormal LFTs varies by population and the specific assays employed.[12] In the general population, up to 8% may exhibit elevated LFTs, with rates increasing with age and in the presence of risk factors such as alcohol abuse, obesity, and chronic viral hepatitis.[2,13] Persistent LFT abnormalities are associated with an elevated risk of serious liver complications, including cirrhosis and hepatocellular carcinoma,[14] as well as extrahepatic sequelae such as fluid retention and cognitive impairment.[15,16] Identification of abnormal LFTs is therefore critical to prompt investigation of underlying causes and to guide interventions ranging from lifestyle modification to pharmacotherapy.[17] Notably, several commonly prescribed medications including antihypertensives (e.g., amlodipine, losartan),[18-20] lipid-lowering agents, and antihyperglycemic drugs,[21-23] have been implicated in drug-induced liver enzyme elevations. In addition, hypertension itself may contribute to increased levels of ALT and ALP.[24]

The liver plays a crucial role in processing and eliminating bilirubin from the body. Elevated bilirubin levels, known as hyperbilirubinemia, can indicate underlying hepatic conditions such as hepatitis, inflammation, cirrhosis, biliary obstruction, or hemolytic anemia. Occasionally, a paradoxical pattern is observed in which conjugated (direct) bilirubin is elevated while total bilirubin remains within the normal range. Typically, an increase in total bilirubin reflects elevations in both unconjugated and conjugated fractions. However, in some instances, an isolated rise in direct bilirubin may occur despite a normal total bilirubin level.

This paradoxical result can be difficult to interpret and often requires further investigation. While the exact mechanisms are not fully understood, possible contributing factors include mild cholestasis, early hepatocellular dysfunction, and analytical variability. A previous report described cases in which total bilirubin appeared normal while direct bilirubin levels were higher than total bilirubin, attributed to assay interference by paraproteins in patients with multiple myeloma.[25] It is important to note that conjugated bilirubin is produced within hepatocytes. Previous studies have suggested that direct bilirubin may be more useful than total bilirubin in predicting clinical outcomes in patients with liver cirrhosis.[26] In the absence of analytical interference such as paraproteins, elevated direct bilirubin may serve as an early indicator of liver dysfunction, even when other liver enzymes such as ALT, AST, ALP, and albumin remain within the reference range.

Based on our findings, it is recommended that both total and direct bilirubin be included in routine LFT screening, particularly in populations with known risk factors such as noncommunicable diseases, pesticide exposure, or alcohol consumption. In this study, 14.71% of participants showed elevated direct bilirubin levels despite having total bilirubin within the normal range, reinforcing the clinical value of assessing both parameters.

In this study, the ALBI score was calculated based on a previously established model originally developed for grading the risk of hepatocellular carcinoma.[27] It has also been validated as a prognostic tool in patients with primary biliary cholangitis.[6,28] At first glance, the ALBI score appeared to have limited applicability in this cohort, as all participants were asymptomatic for liver disease. However, two notable cases were identified with ALBI scores within grade 2.

The participant with the highest ALBI score was a 60-year-old male with a score of -2.1397 and a body mass index (BMI) of 26.5 kg/m². He showed no elevation in liver enzymes, yet his questionnaire responses indicated a history of hypertension and hyperlipidemia, though he was not receiving any medications. The second case was a 36-year-old male with an ALBI score of -2.5887 and a BMI of 22.5 kg/m². He also had normal liver enzyme levels but reported having diabetes and was on antidiabetic medication. Both individuals demonstrated elevated total and direct bilirubin levels. These findings highlight the potential utility of bilirubin-related markers, such as the ALBI score, in detecting sub-clinical liver dysfunction, particularly in individuals with relevant baseline characteristics, even when traditional liver enzyme levels are normal.

Furthermore, this study found no significant association between pesticide exposure and abnormal LFTs. This may reflect effective training and preventive practices promoted by Agricultural Extension Officers in the region. Similarly, alcohol consumption was not significantly associated with abnormal liver function parameters. This could be due to variability in alcohol type, quantity, and frequency of intake (Table 3), which may have attenuated any observable biochemical effects in this cohort.

Limitations

This study had several key limitations: its cross-sectional design offered only a one-time snapshot and could not establish cause and effect; enrolling just 102 people from one rural subdistrict meant the results had limited statistical power and generalizability; reliance on a single, non-fasting blood draw meant enzyme and bilirubin levels could vary based on recent meals; self-reported data on alcohol use, pesticide exposure, and medication intake were vulnerable to recall bias; and the absence of screening for hepatitis, fatty liver, or other liver diseases meant undiagnosed conditions could have confounded the results. Future studies should employ larger, more diverse cohorts, fasting samples, comprehensive health screenings, and multivariable analysis.

Conclusion

The use of medications for the treatment of diabetes mellitus, hypertension, or hyperlipidemia was the primary factor associated with abnormal liver function tests (LFTs) in adults residing in Mae Ka Subdistrict, Chiang Mai, Thailand. The ALBI score successfully identified two asymptomatic cases with elevated bilirubin levels despite normal liver enzyme values. Both individuals had underlying metabolic conditions, specifically hyperlipidemia in one case and diabetes in the other. These findings highlight the potential utility of the ALBI score and bilirubin testing in detecting subclinical liver dysfunction. A comprehensive clinical evaluation and appropriate follow-up investigations remain essential to determine the clinical significance of such abnormalities in individual patients.

Ethical approval

This study is part of the Medical Technology internship for students enrolled in the Community Medical Technology course (510408) and forms part of the MOU between the Faculty of Medical Technology and the Mae Ka community, for which ethical approval has already been obtained.

Funding

This study was supported by the Challenge Fund 2023, Faculty of Associated Medical Sciences, Chiang Mai University.

Conflict of interest

The authors declare no conflicting interests.

CRedit authorship contribution statement

Fahsai Kantawong: conceptualization, methodology, investigation (field work), formal analysis, writing – original draft, writing – review and editing, manuscript submission; **Khanittha Punturee:** investigation (field work), project administration, volunteer recruitment, coordination; **Suwatsin Kittikunnathum:** investigation (sample analysis); **Pharisa Nanthawong:** investigation (sample analysis), **Kanya Preechasuth:** manuscript editing, investigation (field work); **Chayada Sitthidet Tharinjaroen, Witida Laopajeon, Rujirek Chaiwongsa, Nutjeera Intasai, and Tanyarat Jomgeow**: investigation (field work).

Acknowledgements

We wish to express our sincere gratitude to the Faculty of Associated Medical Sciences for funding through the Challenge Fund 2023, and to the Medical Technology internship students enrolled in the Community Medical Technology course (510408) in 2023 for their invaluable assistance.

References

[1] Poovorawan K, Treeprasertsuk S, Thepsuthammarat K, Wilairatana P, Kitsahawong B, Phaosawasdi K. The burden of cirrhosis and impact of universal coverage public health care system in Thailand: Nationwide study. Ann Hepatol. 2015; 14(6): 862-8. doi: 10.5604/16652681.1171773.

[2] Ramakrishnan A, Velmurugan G, Somasundaram A, Mohanraj S, Vasudevan D, Vijayaragavan P, et al. Prevalence of abnormal liver tests and liver fibrosis among rural adults in low and middle-income country: A cross-sectional study. EClinical Medicine. 2022; 51: 101553. doi: 10.1016/j.eclinm.2022.101553.

[3] Lu F, Meng Y, Song X, Li X, Liu Z, Gu C, et al. Artificial intelligence in liver diseases: Recent Advances. Adv Ther. 2024; 41(3): 967-90. doi: 10.1007/s12325-024-

02781-5.

[4] Ngoc Dinh M, Nygate J, Hoang Minh Tu V, Thwaites CL, Group GGCEV. New technologies to improve healthcare in low- and middle-income countries: Global Grand Challenges satellite event, Oxford University Clinical Research Unit, Ho Chi Minh City, 17th-18th September 2019. Wellcome Open Res. 2020; 5: 142. doi: 10.12688/wellcomeopenres.16008.2.

[5] Zhang GM, Hu ZD. Conjugated bilirubin as a reflex test for increased total bilirubin in apparently healthy population. J Clin Lab Anal. 2018; 32(2): e22233. doi: 10.1002/jcla.22233.

[6] Yamashita Y, Umemura T, Kimura T, Joshita S, Hirohara J, Nakano T, et al. Prognostic utility of albumin-bilirubin grade in Japanese patients with primary biliary cholangitis. JHEP Rep. 2023; 5(4): 100662. doi: 10.1016/j.jhepr.2022.100662.

[7] Murillo Perez CF, Harms MH, Lindor KD, van Buuren HR, Hirschfield GM, Corpechot C, et al. Goals of treatment for improved survival in primary biliary cholangitis: Treatment target should be bilirubin within the normal range and normalization of alkaline phosphatase. Am J Gastroenterol. 2020; 115(7): 1066-74. doi: 10.14309/ajg.0000000000000557.

[8] He C, Zhang G, Fu J, Zhang R, Li A, Liu D, et al. Clinical significance of albumin- and bilirubin-based biomarkers in glaucoma: A retrospective case-control study. Oxid Med Cell Longev. 2022; 2022: 8063651. doi: 10.1155/2022/8063651.

[9] Cohen SD, Devant C, Delaval L, Charlier C, Canouï E, Chouchana L. Isolated conjugated hyperbilirubinemia with rifampicin and cross-reaction with rifabutin: A drug-endogenous substance interaction case report. Therapie. 2024; 79(4): 479-82. doi: 10.1016/j.therap.2023.09.006.

[10] Macpherson I, Abeysekera KWM, Harris R, Mansour D, McPherson S, Rowe I, et al. Identification of liver disease: why and how. Frontline Gastroenterol. 2022; 13(5): 367-73. doi: 10.1136/flgastro-2021-101833.

[11] Johnson PJ, Berhane S, Kagebayashi C, Satomura S, Teng M, Reeves HL, et al. Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach-the ALBI grade. J Clin Oncol. 2015; 33(6): 550-8. doi: 10.1200/JCO.2014.57.9151.

[12] Radcke S, Dillon JF, Murray AL. A systematic review of the prevalence of mildly abnormal liver function tests and associated health outcomes. Eur J Gastroenterol Hepatol. 2015; 27(1): 1-7. doi: 10.1097/MEG.0000000000000233.

[13] Lala V, Zubair M, Minter DA. Liver Function Tests. StatPearls Publishing, Tampa.2022.

[14] Pinter M, Trauner M, Peck-Radosavljevic M, Sieghart W. Cancer and liver cirrhosis: implications on prognosis and management. ESMO Open. 2016;1(2) :e000042. doi: 10.1136/esmoopen-2016-000042.

[15] Sun T, Du H, Li Z, Xiong J, Liu Y, Li Y, et al. Decoding the contributions of gut microbiota and cerebral metabolism in acute liver injury mice with and without cognitive dysfunction. CNS Neurosci Ther. 2023; 29(Suppl 1): 31-42. doi: 10.1111/cns.14069.

[16] Sun T, Feng M, Manyande A, Xiang H, Xiong J, He Z. Regulation of mild cognitive impairment associated with liver disease by humoral factors derived from the gastrointestinal tract and MRI research progress: a literature review. Front Neurosci. 2023; 17: 120 6417. doi: 10.3389/fnins.2023.1206417.

[17] Kalas MA, Chavez L, Leon M, Taweesedt PT, Surani S. Abnormal liver enzymes: A review for clinicians. World J Hepatol. 2021; 13(11): 1688-98. doi: 10.4254/wjh.v13.i11.1688.

[18] Varghese G, Madi L, Ghannam M, Saad R. A possible increase in liver enzymes due to amlodipine: A case report. SAGE Open Med Case Rep. 2020; 8: 2050313X20917822. doi: 10.1177/2050313X20917822.

[19] Diogo J, Monteiro R, Coelho C, Ghiletchi A, Leão R, Loureiro C. Drug-induced liver injury due to Losartan. Eur J Case Rep Intern Med. 2021; 8(11): 002856. doi: 10.12890/2021_002856.

[20] Choi WJ, Kim GA, Park J, Jang S, Jung WJ, Shim JJ, et al. Incidence and pattern of aminotransferase elevation during anti-hypertensive therapy with angiotensin-II receptor blockers. J Korean Med Sci. 2022; 37(33): e255. doi: 10.3346/jkms.2022.37.e255.

[21] Bhardwaj SS, Chalasani N. Lipid-lowering agents that cause drug-induced hepatotoxicity. Clin Liver Dis. 2007;11(3):597-613, vii. doi: 10.1016/j.cld.2007.06.010.

[22] Ashraf J, Ali Khan M, Minhaj S, Khatti S, Aarij KM, Shehzad M, et al. Statins and abnormal liver function tests: Is there a correlation? Cureus. 2020; 12(8): e10145. doi: 10.7759/cureus.10145.

[23] Siddiqui MT, Amin H, Garg R, Chadalavada P, Al-Yaman W, Lopez R, et al. Medications in type-2 diabetics and their association with liver fibrosis. World J Gastroenterol. 2020; 26(23): 3249-59. doi: 10.3748/wjg.v26.i23.3249.

[24] Baeradeh N, Seif M, Rezaianzadeh A, Hosseini SV. Investigating the relationship between liver enzymes and incidence of hypertension: A population-based cohort study in Kharameh, a city in the South of Iran. Health Sci Rep. 2023; 6(10): e1601. doi: 10.1002/hsr2.1601.

[25] Yilmaz NS, Sen B, Gulbahar O. Contribution of the laboratory to a diagnosis process by sequential reflective testing: Paraprotein interference on a direct bilirubin assay. Biochem Med (Zagreb). 2021; 31(2): 020801. doi: 10.11613/BM.2021.020801.

[26] Lee HA, Jung JY, Lee YS, Jung YK, Kim JH, An H, et al. Direct bilirubin is more valuable than total bilirubin for predicting prognosis in patients with liver cirrhosis. Gut Liver. 2021; 15(4): 599-605. doi: 10.5009/gnl20171.

[27] Toyoda H, Johnson PJ. The ALBI score: From liver function in patients with HCC to a general measure of liver function. JHEP Rep. 2022; 4(10): 100557. doi: 10.1016/j.jhepr.2022.100557.

[28] Feng J, Xu JM, Fu HY, Xie N, Bao WM, Tang YM. Prognostic scores in primary biliary cholangitis patients with advanced disease. World J Gastrointest Surg. 2023; 15(8): 1774-83. doi: 10.4240/wjgs.v15.i8.1774.

Journal of Associated Medical Sciences 2025; 58(3): 278-287

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

Enhanced detection of *Trypanosoma evansi* in Cattle: Superior performance of LAMP compared to PCR and CATT/*T. evansi* test

Wuttichote Jansaento[1], Mallika Osiriphan[2], Nattharinee Kongta[3], Wittawat Modethed[4], Thanakorn Rompo[4], Suwit Duangmano[2,5*]

[1]*Faculty of Medical Technology, Nation University, Lampang Province, Thailand.*
[2]*Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
[3]*Department of Parasitology, Faculty of medicine, Chiang Mai University, Chiang Mai Province, Thailand.*
[4]*The Fifth Regional Livestock Office, Chiang Mai Province, Thailand.*
[5]*Hematology and Health Technology Research Center, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*

ARTICLE INFO

Article history:
Received 17 June 2025
Accepted as revised 24 July 2025
Available online 29 July 2025

Keywords:
Trypanosomiasis, Trypanosoma evansi, loop mediated isothermal amplification, RoTat1.2

ABSTRACT

Background: *Trypanosoma evansi*, the causative agent of surra, poses a major veterinary concern in tropical regions, particularly affecting cattle and buffalo. The disease leads to reproductive failures, including abortion, and significant economic losses. Early and accurate diagnosis is crucial for effective control, especially in endemic, resource-limited areas.

Objectives: This study aimed to develop a loop-mediated isothermal amplification (LAMP) assay for rapid detection of *T. evansi* and evaluate its diagnostic performance in comparison with conventional polymerase chain reaction (PCR) and the CATT/*T. evansi* card agglutination test.

Materials and methods: Four LAMP primers were designed to target the RoTat 1.2 variant surface glycoprotein (VSG) gene of *T. evansi*. Optimal reaction parameters, including temperature and incubation time, were established. The LAMP assay, conventional PCR, and the CATT/*T. evansi* card agglutination test were performed on 79 blood samples collected from cattle with suspected trypanosomiasis in northern Thailand (Lamphun and Chiang Mai). Diagnostic sensitivity, specificity, and agreement between tests were statistically analyzed.

Results: The LAMP assay detected *T. evansi* in 32 (40.5%, 95% CI: 29.8-51.9%) samples, slightly outperforming PCR, which detected 30 (37.9%, 95% CI: 27.6-49.0%). However, this difference was not statistically significant (McNemar's test, p=0.724). The CATT/*T. evansi* test yielded 45 (56.9%) positives but lacked the ability to differentiate active infection from prior exposure.

Conclusion: The LAMP assay demonstrated high sensitivity, specificity, and rapid detection capabilities under simplified conditions, making it highly suitable for field applications. When paired with colorimetric or lateral flow readouts, LAMP offers a promising point-of-care diagnostic tool for improving trypanosomiasis control in endemic regions.

Introduction

Trypanosoma evansi, the causative agent of trypanosomiasis (commonly known as surra), represents a significant threat to mammalian livestock health and is classified as a notifiable pathogen under the Animal Epidemic Act B.E. 2499 in Thailand. This extracellular hemoprotozoan parasite proliferates within the circulatory and lymphatic systems of vertebrate hosts, frequently resulting in reproductive complications including abortion and breeding failure. The economic impact of *T. evansi*

∗ **Corresponding contributor.**
Author's Address: *Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: suwit.du@cmu.ac.th
doi: 10.12982/JAMS.2025.098
E-ISSN: 2539-6056

infection on the veterinary industry is substantial, primarily due to its rapid transmission dynamics within cattle populations and the subsequent livestock mortality and productivity losses.

The epidemiological significance of *T. evansi* has expanded beyond traditional endemic regions, with documented emergence in previously unaffected geographical areas and infection of novel host species.[1] Recent comprehensive reviews have highlighted the continued global spread of surra, with new cases reported across tropical and subtropical regions, emphasizing the ongoing threat to livestock industries worldwide.[2] While some infected cattle exhibit clinical manifestations, others remain asymptomatic carriers, serving as reservoirs for transmission to diverse animal species including equines, elephants, large felids, canines, and swine.[3,4] Although human infection with animal trypanosomes is typically prevented by trypanolytic factors present in human serum, certain *T. evansi* isolates have demonstrated resistance to human plasma under specific conditions, raising concerns about potential zoonotic transmission.[5] Molecular characterization has identified two principal serotypes of *T. evansi* based on the presence or absence of the *RoTat 1.2 VSG* gene: serotype A, which expresses this gene, and serotype B, which lacks this genetic element. Notably, the majority of *T. evansi* isolates characterized in Thailand possess the *RoTat 1.2 VSG* gene, making this target an important diagnostic marker for serotype A identification in the region.[6] The specificity and prevalence of this genetic marker provide an excellent foundation for the development of molecular diagnostic assays tailored to the Thai epidemiological context.

The diagnosis of *T. evansi* infection encompasses various methodological approaches, including direct microscopic examination, serological testing, animal inoculation, and DNA-based detection techniques.[7–9] In Thailand, the standard diagnostic protocol relies primarily on microscopic examination as the first-line detection method. The hematocrit centrifuge technique (HCT), commonly referred to as the Woo test, represents a simple and cost-effective diagnostic approach for animal trypanosomiasis; however, this method is constrained by inherently low sensitivity due to its dependence on the direct visualization of motile parasites in EDTA-anticoagulated blood specimens. Multiple factors significantly influence the diagnostic performance and reliability of microscopic detection methods. Low parasitemia levels, particularly during early infection stages or chronic disease phases, frequently result in false-negative outcomes due to insufficient parasite density for microscopic detection. Pre-analytical variables also critically impact diagnostic accuracy: elevated storage and transportation temperatures adversely affect parasite viability by accelerating glucose consumption, leading to energy depletion and subsequent loss of motility—a key criterion for microscopic identification. Transportation delays exceeding 24 hours further compromise parasite viability and motility, thereby reducing detection sensitivity.[10] Additionally, examination of Giemsa-stained blood films using conventional microscopy frequently fails to detect patent infections and chronic manifestations of the disease, where parasitemia levels may be below the threshold of microscopic detection.[11] Serological diagnostic approaches, particularly the card agglutination test for *T. evansi* (CATT/*T. evansi*® kit), represent the current reference standard for antibody-based detection as recommended by the World Organisation for Animal Health (OIE). This assay utilizes a freeze-dried suspension of purified, fixed, and stained bloodstream-form trypanosomes expressing the variable surface glycoprotein RoTat 1.2 as the detection antigen. Despite its recognized diagnostic value and international endorsement, the widespread implementation of CATT/*T. evansi* faces significant practical limitations, including restricted importation and commercialization due to varying national regulatory frameworks, usage constraints, and prohibitive costs that limit accessibility in resource-constrained settings.

DNA-based detection methods, particularly polymerase chain reaction (PCR) techniques, demonstrate superior sensitivity and specificity for diagnosing active *T. evansi* infections. The stability of parasitic DNA persists for 24–48 hours post-mortem, allowing reliable detection even after parasite death, with conventional PCR turnaround times of 3–4 hours. However, PCR methodology remains constrained by its requirement for expensive thermal cycling equipment, limiting its implementation in budget-restricted laboratories and remote geographical areas, particularly in developing countries where trypanosomiasis is endemic.[12] Recent advances in LAMP technology have demonstrated successful point-of-care applications in veterinary medicine, including rapid antimicrobial susceptibility testing for bacterial pathogens in dogs[13] and multiplex detection systems for swine pathogens using microfluidic platforms.[14] Contemporary developments in LAMP methodology have focused on improving field deployability through smartphone-based detection systems and colorimetric readouts.[15] These innovations have shown particular promise for parasitic disease detection, with recent studies demonstrating successful point-of-care applications for various parasitic organisms.[16]

Loop-mediated isothermal amplification (LAMP) offers a promising alternative that reduces dependence on sophisticated instrumentation by utilizing a single amplification temperature of 60–65°C throughout the reaction. This technique achieves remarkable amplification efficiency, generating 10^9–10^{10} fold increases in target DNA within one hour without requiring expensive automation or thermal cycling equipment.[17,18] LAMP has demonstrated versatility across veterinary diagnostic applications, including pathogen detection in aquaculture species such as shrimp[19] and livestock including swine.[20] Furthermore, LAMP methodology has proven successful for diagnosing both human and animal trypanosomiasis in previous studies.[21,22] The implementation of LAMP-based diagnostics for *T. evansi* detection could significantly mitigate economic losses in the veterinary sector through early pathogen identification, enabling rapid veterinary

intervention before widespread epidemic development. However, comprehensive comparative studies evaluating LAMP performance against established PCR and ELISA methodologies for *T. evansi* diagnosis remain limited in the literature. Recent advances in loop-mediated isothermal amplification (LAMP) technology have included smartphone-based fluorescence detection, lyophilized reagent formulations for field deployment, and multiplexed assays for simultaneous pathogen detection.[23,24] However, limited comparative data exists for LAMP versus conventional methods specifically for *T. evansi* detection in Southeast Asian cattle populations, representing a significant knowledge gap for regional disease control strategies.

Therefore, the objective of this study was to compare the diagnostic efficiencies of conventional PCR and CATT/*T. evansi* test with LAMP technique to determine the most suitable method for rapid, field-deployable detection of *Trypanosoma evansi* based on amplification of the *RoTat 1.2 VSG* gene target.

Materials and methods

Sample collection and preparation

Archived whole blood specimens stored at -20 °C from dairy cattle were utilized for this comparative diagnostic study. All blood samples represented residual specimens from routine annual health surveillance programs conducted by veterinarians from The Fifth Regional Livestock Office, Chiang Mai, Thailand, during 2020. Due to budget limitations, 79 cattle from Lamphun and Chiang Mai provinces were selected using convenience sampling based on clinical signs suggestive of trypanosomiasis and herd owners' approval for participation. A confirmed positive control sample from cattle with microscopically confirmed *T. evansi* infection was included in the study design to validate assay performance. For detecting a 15% difference in sensitivity between methods (80% vs 65%) with 80% power and α=0.05, approximately 74 samples per group were required. Our sample of 79 meets this minimum requirement for this preliminary evaluation.

Extraction of genomic DNA and PCR amplification using TR3/4 primers

DNA was extracted from blood samples by using Nucleospin blood extraction kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany). Briefly, 200 µL EDTA blood was mixed with 25 µL of proteinase K and 200 µL buffer. Then, the mixture was vortexed vigorously for cell lysis. After that, the mixture was incubated at 70°C for 10-15 minutes. A total of 210 µL absolute ethanol was added to the mixture and vortexed again. Then, the mixture was transferred into a Nucleospin Blood Column and centrifuged at 12,000 rpm for 2 minutes. After that, the silica membrane was washed two times. The dried silica membrane was centrifuged in the column at 12,000 rpm for 2 minutes. Then, 100 µL elution buffer was added, incubated for 1 minute and centrifuged in the column at 12,000 rpm for 2 minutes to elute pure DNA. For PCR amplification, TR3/4 primers were derived from a trypanosome-specific repetitive nucleotide sequence fragment that amplified 257 bp amplicon, Forward primer (TR3: 5'-GCGCGGATTCTTTGCAGACGA-3'), Reverse primer (TR4: 5'-TGCAGACACTGGAATGT-3'). Conventional PCR was performed by using 1x Quick Taq HS Dye mix reagent (Toyobo Life Science, Tokyo, Japan). The PCR conditions were as follows: 30 cycles of denaturation at 95 °C for 30 seconds, annealing at 55 °C for 30 second and extension at 72 °C for 30 seconds. The PCR products were separated by electrophoresis through 2% agarose gels containing Redsafe nucleic staining dye, and the PCR products were visualized under UV illumination.

LAMP assay

LAMP primers recognizing six sections of the *T. evansi RoTat 1.2 VSG* gene (Accession no. AF317914) were designed using Primer Explorer version 3 software (http:// primerexplorer.jp/lampv3e/index.html), F3 primer (5'-AATTCTGCCCGCAGTTGC-3'), B3 primer (5'-CCCTCTAG GTAGCTGTCTCC-3'), FIP primer GGCTTTGCCCACACAAAA CTTGCAGGGGCGATTTCATCG-3'), BIP primer GCACAAATGC CGACGGTAACGCCGAGGTGTCATAGTTGG-3').

LAMP was carried out in a total 25 µL reaction mixture containing 1x isothermal amplification buffer, 0.8 µM each FIP and BIP, 0.1 µM each F3 and B3, 1.4 mM each dNTP, 8 mM $MgSO_4$, 320 U/ml Bst 2.0 WarmStart DNA polymerase (New England Biolabs, MA, USA) and 100 ng of DNA template. Thereafter, incubation at 65 °C for 30 minutes in Mini Heating Dry Bath incubator and heating at 80 °C for 5 minutes to terminate the reaction were performed. Aliquots of 2 µL of LAMP products were observed in a ladder-like pattern in 2% agarose gels containing Redsafe nucleic staining dye and the LAMP products were visualized under UV illumination. LAMP products were also detected with the naked eye by observing white turbidity in the reaction mixture.

Sensitivity of LAMP assay

The detection limit of the LAMP assay was determined using a 10-fold serial dilution of 100 ng of purified genomic DNA from infected cattle and using a 10-fold serial dilution of lysed parasite by double-distilled water to determine the sensitivity of the analytical assay. The sensitivity of the LAMP assay was compared with that of conventional PCR by using TR3/4 primers.

Analysis of cattle samples

A total of 79 genomic DNA from cattle with suspected trypanomiasis were used. The DNA was prepared and stored at -20°C before use. A total of 100 ng of the DNA template was added to compose 25 µL of LAMP reaction mixture, and amplification was performed as in section 2.4.

Serological analysis

Serodiagnosis of animal trypanosomosis due to *T. evansi* was performed by using CATT/*T. evansi* test kit (Free University of Brussels and the Institute of Tropical Medicine, Belgium). The assay was performed following the manufacturer's instruction. Briefly, 25 µL of plasma

or serum were mixed with approximately 45 µL of homogenized CATT antigen (variable surface antigen VSA common to all *T. evansi* stocks) in the supplied plastic card. A stirring rod was used to mixed and spread out the reaction mixture to about 1 mm form the edge of test area. The plastic card was rotated on a flatbed orbital rotator for 5 minutes at 70 rpm. After rotation the agglutination results were compared with the positive control.

Statistical analysis

McNemar's test was used to compare paired proportions between diagnostic methods. Cohen's kappa coefficient was calculated to assess agreement between methods. Statistical significance was set at $p<0.05$. Sensitivity, specificity, and 95% confidence intervals were calculated using PCR as the reference standard. Analyses were performed using SPSS version 20 (IBM Corp.).

Results

Detecting T. evansi by conventional PCR and LAMP assay

Conventional PCR amplification yielded the expected 257 bp amplicon from whole-blood genomic DNA, confirming the presence of *T. evansi* infection as shown in Figure 1. Among the 79 blood specimens analyzed, 30 samples (37.9% , 95% CI: 27.6-49.0%) tested positive for *T. evansi* DNA using conventional PCR methodology.

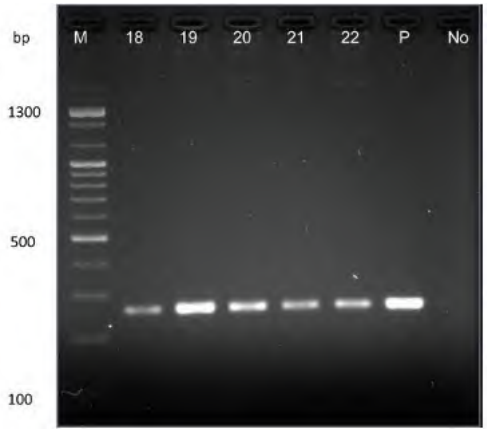


Figure 1. *Detection of T. evansi by conventional PCR using TR3/4 primers. Agarose gel electrophoresis showing 257 bp amplicons. Lanes 18-22: cattle DNA samples with varying band intensities indicating different parasite loads, Lane P: positive control, Lane M: 100 bp DNA marker, Lane No: negative control.*

Optimal LAMP reaction conditions were established using 0.1 µM of each outer primer (F3 and B3), 0.8 µM of each inner primer (FIP1 and BIP1), and 8 mM $MgSO_4$, combined with 1× isothermal buffer, 1.4 mM dNTPs, 320 U/ml Bst 2.0 WarmStart DNA polymerase, and 100 ng of template DNA. The reaction was maintained at 65°C for 30 minutes, with positive and negative reactions clearly distinguishable by visual inspection (Figure 2). Agarose gel electrophoresis analysis revealed distinct patterns characteristic of LAMP amplification products. Positive samples exhibited a DNA ladder-like pattern with multiple bands and smear formations, while negative samples showed no such ladder-like banding pattern. These amplification characteristics were consistent with the positive control (PC), whereas the negative control (NC) and non-infected blood samples displayed single band patterns (Figure 3). Using the optimized LAMP assay, 32 blood samples (40.5% , 95% CI: 29.8-51.9%) tested positive for *T. evansi* DNA, while 47 samples (59.5%) were negative (Table 1). Although LAMP detected 2 additional positive samples compared to PCR (40.5% vs 37.9%), this difference was not statistically significant ($p=0.724$). The clinical relevance of this minimal difference requires evaluation in larger studies.

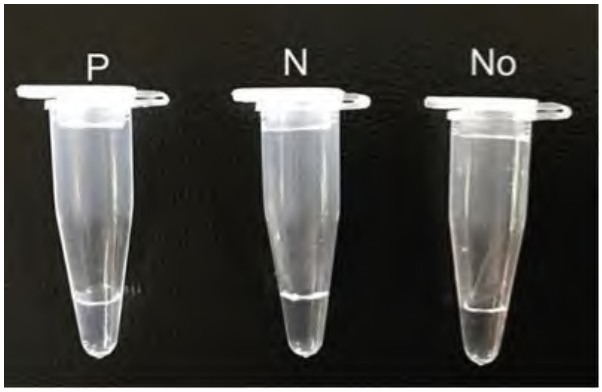


Figure 2. *Detection of T. evansi by LAMP assay. LAMP products detected by naked eye. (P) Positive LAMP products; (N) Negative LAMP products; (No) No template control.*

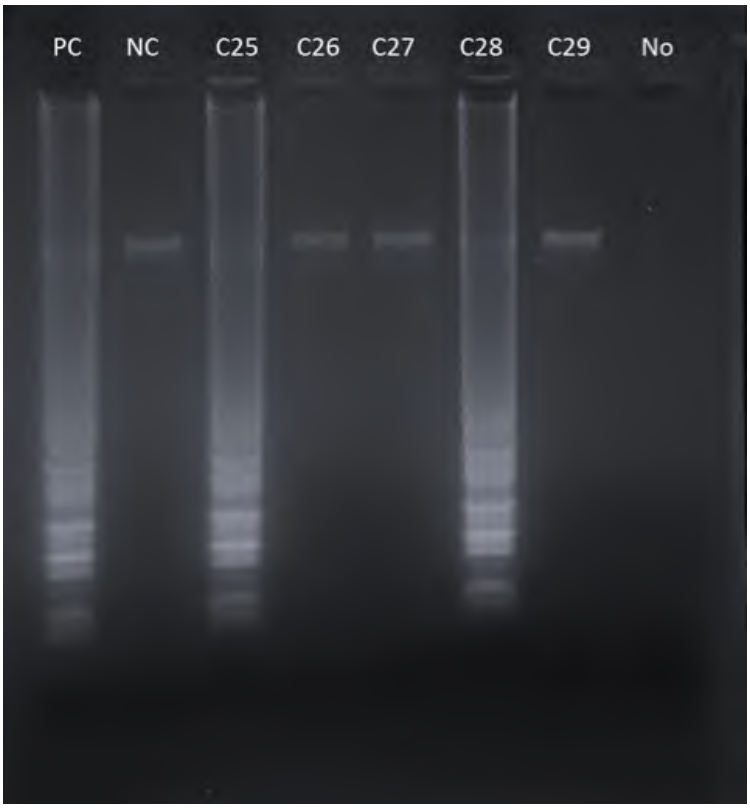


Figure 3. *LAMP assay results showing characteristic ladder-like pattern in positive samples. Multiple bands and smearing indicate successful isothermal amplification. Lane PC: positive control, Lane NC: negative control, Lanes C25-C29: cattle samples.*

Table 1. Detection of *Trypanosoma evansi* infections in cattle using PCR, LAMP assay, and CATT/*T. evansi* kit.

Experiment	Positive N (%)	Negative N (%)	Statistical comparison
PCR	30 (37.9%)	49 (62.1%)	Reference method
LAMP	32 (40.5%)	47 (59.5%)	vs PCR: p=0.724†
CATT/*T. evansi*	45 (56.9%)	34 (43.1%)	vs molecular methods: p<0.001†

Note*: †McNemar's test for paired comparisons agreement between LAMP and PCR, κ=0.89 (95% CI: 0.78–0.95).*

Statistical comparison between diagnostic methods revealed excellent agreement between LAMP and PCR. McNemar's test revealed no statistically significant difference between LAMP and PCR detection rates (χ^2=0.125, p=0.724). Agreement analysis showed excellent concordance between methods (κ = 0.89, 95% CI: 0.78-0.95). Using PCR as reference standard, LAMP demonstrated sensitivity of 96.7% (95% CI: 82.8-99.9%) and specificity of 93.9% (95% CI: 83.1-98.7%).

Analytical sensitivity determination

To evaluate the analytical performance of the LAMP assay, the detection limits of both conventional PCR and LAMP methodologies were determined using 10-fold serial dilutions of purified genomic DNA extracted from *T. evansi*-infected cattle. Comparative analysis by 2% agarose gel electrophoresis demonstrated equivalent analytical sensitivity between conventional PCR (Figure 4A) and the LAMP assay (Figure 4B), indicating comparable detection thresholds for both molecular diagnostic approaches. Both methods showed similar detection limits down to the 10^{-3} dilution (approximately 0.1 ng/µL genomic DNA). However, precise limit of detection in terms of parasite equivalents per µL cannot be determined without standardized parasite quantification. Visual assessment suggests comparable analytical sensitivity, but quantitative confirmation using qPCR standards is recommended for definitive LOD determination.

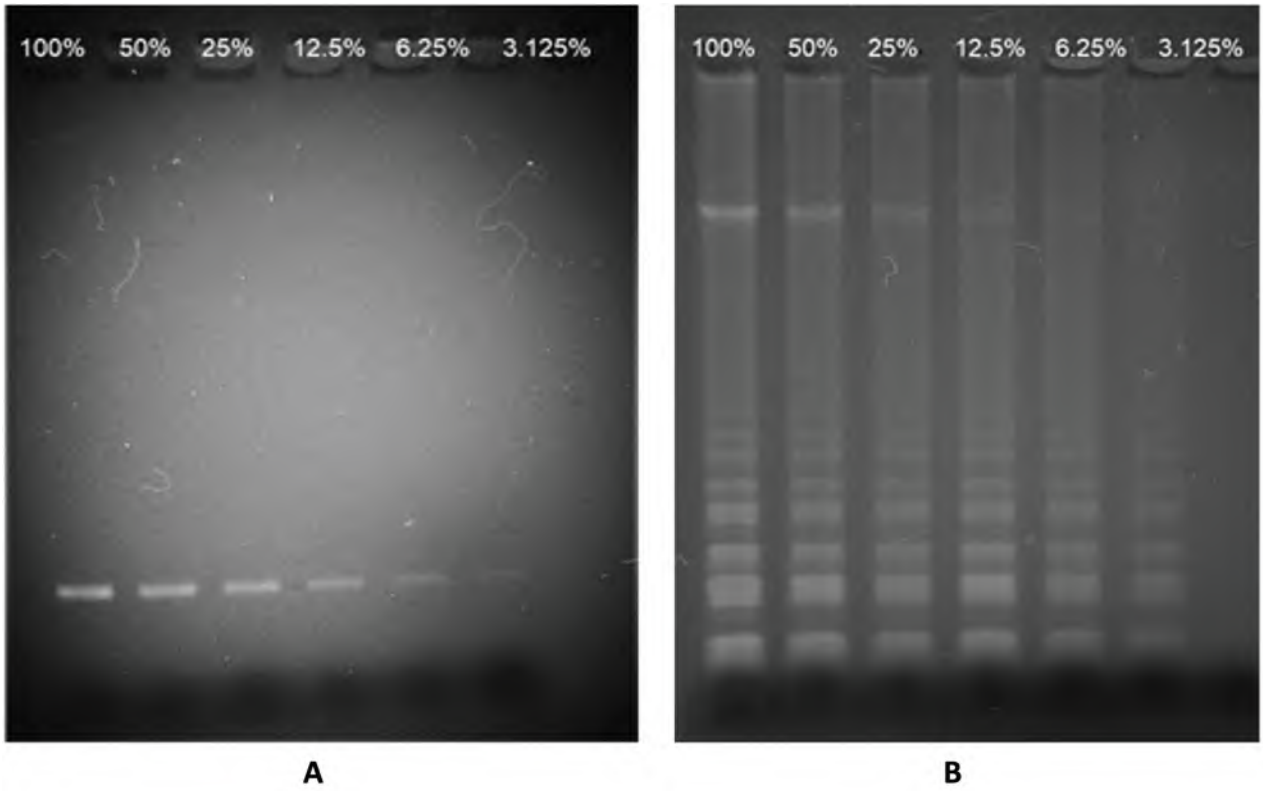


Figure 4. *Comparative sensitivity analysis of PCR and LAMP assays for Trypanosoma evansi detection using serial dilutions of T. evansi DNA template from cattle. A: conventional PCR products, B: LAMP amplification products.*

Serological analysis

Serological testing was performed using the CATT/*T. evansi* kit, which employs a direct agglutination principle for detecting anti-*T. evansi* antibodies in bovine blood specimens. The assay involves mixing plasma or serum samples with variable surface glycoprotein (VSG) RoTat 1.2 antigen reagent, with antigen-antibody complexes observed as aggregates by naked eye examination within 5 minutes. Positive results are characterized by dark-colored clumping (agglutination) around the edges of sample drops, as compared to the positive control (Figure 5). Among the 79 specimens tested, 45 blood samples (57.0%) were positive by CATT/*T. evansi* serological testing. Notably, seven samples that tested positive by CATT showed negative results in both PCR and LAMP molecular assays. This discordance may be attributed to several factors: (1) previous *T. evansi* infection with subsequent parasite clearance, resulting in persistent antibodies but undetectable parasitic DNA; (2) parasitemia levels below the analytical detection limits of the molecular assays employed; or (3) sampling limitations, as the 200 µL sample volume utilized in this study may have contributed to false-negative results in specimens with low parasite density.

Figure 5. *Serological analysis using CATT/T. evansi kit showing agglutination reactions (characterized by dark-colored clumping around the edges of sample drops). Pos: Positive control, Neg: Negative control, B8-B15: cattle plasma samples.*

Discussion

Trypanosoma evansi represents the most geographically widespread pathogenic trypanosome species globally, necessitating accurate diagnostic capabilities and effective therapeutic interventions for successful disease control, particularly in regions characterized by low-intensity infections. As the prevalence of trypanosomiasis declines through control measures, enhanced diagnostic sensitivity and specificity become increasingly critical to prevent disease re-emergence from undetected outbreaks within animal reservoir populations. Addressing the epidemiological significance of *T. evansi* in Thailand, we developed a rapid and simplified diagnostic assay capable of detecting *T. evansi* genomic DNA using loop-mediated isothermal amplification (LAMP) methodology. Comparative analysis of 79 bovine blood specimens revealed that 32 samples (40.5% , 95% CI: 29.8-51.9%) tested positive by LAMP, compared to 30 samples (37.9%, 95% CI: 27.6-9.0%) positive by conventional PCR, demonstrating superior sensitivity of the LAMP assay over traditional PCR methodology. These findings align with previous studies reporting enhanced LAMP sensitivity for trypanosome detection.[21,25,26] However, the relatively small sample size (n=79) may limit the statistical power for definitive comparisons between diagnostic methods. Future studies with larger sample sizes would strengthen the statistical validity of sensitivity comparisons.

Recent field applications of LAMP technology have demonstrated its practical utility in veterinary diagnostics. Pirolo *et al.* validated a colorimetric LAMP test for bacterial pathogen identification in dogs, achieving near 100% sensitivity and specificity under laboratory conditions, though field performance showed some reduction in accuracy.[13] Ji *et al.* developed a microfluidic-LAMP chip capable of simultaneously detecting multiple swine pathogens with 100% specificity and excellent stability, demonstrating the potential for multiplexed detection systems.[14] These developments highlight both the promise and practical challenges of implementing LAMP technology in real-world veterinary settings. The superiority of LAMP over conventional PCR was further evidenced by more distinct visualization of amplification products on 2% agarose gel electrophoresis, facilitating clearer interpretation of results. Our study demonstrated a detection limit of 3.125% for the LAMP assay, although higher sensitivities have been reported in other investigations.[25,26,27] The inherent advantages of LAMP methodology—including rapid amplification kinetics, single-temperature incubation requirements, and equipment simplicity—position it as a superior alternative to conventional PCR and other amplification strategies such as nested PCR and real-time quantitative PCR for field-based pathogen detection without sophisticated instrumentation requirements.[28]

The practical advantages of LAMP extend beyond sensitivity improvements, with results obtainable within 30 minutes using standard dry bath incubation. To address limitations associated with UV-based result interpretation,

recently developed colorimetric-fluorometric indicators (CFI) enable real-time fluorescence monitoring with visible color changes observable by the naked eye. The progressive consumption of magnesium ions during LAMP amplification causes hydroxynaphthol blue (HNB) within the CFI to undergo a characteristic color transition from purple to sky blue, eliminating the need for gel electrophoresis.[29]

The versatility of LAMP methodology has been demonstrated across diverse parasitological applications in veterinary diagnostics. Kumagai *et al.* demonstrated LAMP's capability to detect schistosomal DNA from primary infections as early as one day post-infection in snails across epidemic areas, establishing it as a robust screening method.[30] Similarly, Plutzer and Karanis successfully applied LAMP for Giardia duodenalis detection in fecal specimens.[31] Kong *et al.* reported excellent specificity of LAMP for Toxoplasma gondii detection in murine blood samples, with no cross-reactivity observed against other parasites and positive detection within one day of experimental infection. [32] Park *et al.* demonstrated that reverse transcription-LAMP (RT-LAMP) provided superior performance compared to conventional LAMP for detecting European and North American porcine reproductive and respiratory syndrome viruses, with reduced turnaround times.[20]

These studies collectively demonstrate the expanding application of LAMP methodology for infectious disease detection in veterinary medicine over the past decade. The enhanced specificity achieved through dual primer pairs binding to six distinct regions on target DNA, combined with single-temperature amplification using strand-displacing Bst DNA polymerase, contributes to the method's robustness, high sensitivity, and specificity while maintaining rapid turnaround times. Furthermore, the reduced requirement for expensive reagents and sophisticated instrumentation significantly decreases the cost of *T. evansi* detection, making it particularly suitable for resource-limited settings.

Examination of 47 cattle revealed that all animals tested negative by LAMP assay. However, serological analysis detected antibodies against *T. evansi* in 13 animals (27.7%), indicating previous exposure to the parasite. Serological testing was performed using the CATT/*T. evansi* kit, which employs the direct agglutination method for detection of anti-*T. evansi* antibodies in bovine blood samples. Test results are interpreted visually by direct observation of agglutination patterns on reaction cards, with accuracy dependent upon the expertise of trained technicians.

The CATT/*T. evansi* kit presents inherent limitations that must be considered in diagnostic applications. Notably, this serological method cannot differentiate between active and resolved infections, rendering it unsuitable as a test-of-cure. This limitation stems from the persistence of anti-Trypanosoma spp. antibodies, which typically appear 4-8 days post-infection and may remain detectable for extended periods following parasite clearance or therapeutic intervention.[33,34]

In contrast, loop-mediated isothermal amplification (LAMP) represents a sensitive and specific molecular diagnostic tool particularly suited for early detection of *T. evansi* infection during the acute phase, when antibody titers may be low or absent.[35] The LAMP assay offers several technical advantages over conventional PCR and ELISA methodologies, including isothermal amplification at a constant temperature of 65 °C, eliminating the requirement for expensive thermocycling equipment. Furthermore, LAMP demonstrates detection limits equivalent to or superior to conventional PCR while providing reduced processing time.[36] While LAMP offers several advantages including rapid results and isothermal amplification, certain limitations must be acknowledged. The high amplification efficiency of LAMP can increase the risk of cross-contamination if strict laboratory protocols are not followed. Additionally, LAMP may be prone to nonspecific amplification if primers are not carefully optimized, potentially leading to false-positive results. Future validation studies should include rigorous specificity testing and standardized protocols to minimize these risks. Despite these considerations, LAMP remains a promising tool for field-based diagnosis when appropriate quality control measures are implemented.

Limitation

Based on agarose gel visualization, both methods showed similar detection limits in this preliminary comparison. However, without quantitative measurement of amplification products or standardized parasite concentrations, we cannot definitively conclude equivalent analytical sensitivity. Additionally, all blood samples were archived specimens stored at -20°C, and prolonged storage may have resulted in DNA degradation, potentially affecting the sensitivity of molecular detection methods and leading to underestimation of true diagnostic performance. Future studies should employ quantitative methods (qPCR, fluorescence measurement) for precise sensitivity comparison.

Conclusion

Based on our findings, the LAMP assay represents a viable alternative to conventional PCR-based diagnostic methods for *T. evansi* detection. The technique's inherent advantages—including rapid processing, high sensitivity, and isothermal amplification requirements—make it particularly suitable for field applications. These characteristics position LAMP as a promising point-of-care testing (POCT) platform for *T. evansi* diagnosis in clinical settings and livestock screening programs, supporting broader trypanosomiasis control and elimination initiatives. The complementary use of both molecular (LAMP) and serological (CATT) methods provides comprehensive diagnostic coverage, with LAMP offering superior performance for acute infection detection and CATT providing epidemiological insights into exposure history within populations.

Ethical approval

Ethical approval was not required as this study utilized archived samples from routine veterinary health surveillance conducted by the Fifth Regional Livestock Office. All samples were anonymized, and no additional animal procedures were performed specifically for this research. The use of archived samples for diagnostic method evaluation falls under institutional guidelines for secondary analysis of surveillance data.

Funding

This research was supported by Agricultural Research Development Agency (ARDA), grant number CRP6105021850. The authors would like to thank the Faculty of Associated Medical Sciences for the research facility support.

Conflicts of Interest

The authors declare no conflict of interest.

CRedit authorship contribution statement

Wuttichote Jansaento: conceptualization, methodology, formal analysis; **Mallika Osiriphan**: formal analysis, data curation, writing – original draft preparation; **Natthariinee kongta**: formal analysis, data curation, writing – original draft preparation; **Wittawat Modethed**: resources, writing – review & editing**; Thanakorn Rompo:** resources, writing – review and editing; **Suwit Duangmano**: conceptualization, methodology, writing – review & editing, project administration, funding acquisition.

Data availability

The corresponding author may provide the data supporting the results of this study upon reasonable request.

Acknowledgements

The authors would like to thank the Faculty of Associated Medical Sciences for the research facility support.

References

[1] Lun ZR, Fang Y, Wang CJ, Brun R. Trypanosomiasis of domestic animals in China. Parasitol Today. 1993; 9(2): 41-5.

[2] Aregawi WG, Agga GE, Abdi RD, Büscher P. Systematic review and meta-analysis on the global distribution, host range, and prevalence of *Trypanosoma evansi*. Parasit Vectors. 2019; 12(1): 67.

[3] Davison HC, Thrusfield MV, Husein A, Muharsini S, Partoutomo S, Rae P, et al. The occurrence of *Trypanosoma evansi* in buffaloes in Indonesia, estimated using various diagnostic tests. Epidemiol Infect. 2000; 124(1): 163-72.

[4] Desquesnes M, Holzmuller P, Lai DH, Dargantes A, Lun ZR, Jittaplapong S. *Trypanosoma evansi* and surra: a review and perspectives on origin, history, distribution, taxonomy, morphology, hosts, and pathogenic effects. Biomed Res Int. 2013; 2013: 194176.

[5] Joshi PP, Shegokar VR, Powar RM, Herder S, Katti R, Salkar HR, et al. Human trypanosomiasis caused by *Trypanosoma evansi* in India: the first case report. Am J Trop Med Hyg. 2005; 73(3): 491-5.

[6] Ngaira JM, Olembo NK, Njagi EN, Ngeranwa JJ. The detection of non-RoTat 1.2 *Trypanosoma evansi*. Exp Parasitol. 2005; 110(1): 30-8.

[7] Monzón CM, Mancebo OA, Roux JP. Comparison between six parasitological methods for diagnosis of *Trypanosoma evansi* in the subtropical area of Argentina. Vet Parasitol. 1990; 36(1-2): 141-6.

[8] Omanwar S, Rao JR, Basagoudanavar SH, Singh RK, Butchaiah G. Direct and sensitive detection of *Trypanosoma evansi* by polymerase chain reaction. Acta Vet Hung. 1999; 47(3): 351-9.

[9] Verloo D, Magnus E, Büscher P. General expression of RoTat 1.2 variable antigen type in *Trypanosoma evansi* isolates from different origin. Vet Parasitol. 2001; 97(3): 183-9.

[10] Reid SA, Husein A, Copeman DB. Evaluation and improvement of parasitological tests for *Trypanosoma evansi* infection. Vet Parasitol. 2001; 102(4): 291-7.

[11] Zayed AA, Habeeb SM, Allam NAT, Ashry HM, Mohamed AH, Ashour AA, et al. A critical comparative study of parasitological and serological differential diagnostic methods of *Trypanosoma evansi* infections in some farm animals in Egypt. Am-Eurasian J Agric Environ Sci. 2010; 8(6): 633-42.

[12] Holland WG, My LN, Dung TV, Thanh NG, Tam PT, Vercruysse J, et al. The influence of ambient temperature on the performance of the polymerase chain reaction for the diagnosis of *Trypanosoma evansi* infections in experimentally infected water buffalo. Vet Parasitol. 2001; 102(4): 331-40.

[13] Pirolo M, Menezes M, Poulsen M, Søndergaard V, Damborg P, Poirier AC, et al. A LAMP point-of-care test to guide antimicrobial choice for treatment of Staphylococcus pseudintermedius pyoderma in dogs. Vet J. 2024; 304: 106105.

[14] Ji C, Zhou L, Chen Y, Fang X, Liu Y, Du M, et al. Microfluidic-LAMP chip for the point-of-care detection of gene-deleted and wild-type African swine fever viruses and other four swine pathogens. Front Vet Sci. 2023; 10: 1116352.

[15] Silva-Pereira TC, Martín-Hernández R, Higes M, Camacho-Nuez M, Brandariz S, Graça A, et al. SMART-LAMP: a smartphone-operated handheld device for real-time colorimetric point-of-care diagnosis of infectious diseases via loop-mediated isothermal amplification. Biosens Bioelectron. 2022; 214: 114051.

[16] Hartle-Mougiou K, Gubili C, Xanthopoulou P, Kasapidis P, Valiadi M, Gizeli E. Development of a quantitative colorimetric LAMP assay for fast and targeted molecular detection of the invasive lionfish Pterois miles from environmental DNA. Front Mar Sci. 2024; 11: 1358793.

[17] Notomi T, Okayama H, Masubuchi H, Yonekawa

T, Watanabe K, Amino N, et al. Loop-mediated isothermal amplification of DNA. Nucleic Acids Res. 2000; 28(12): E63.

[18] Mori Y, Nagamine K, Tomita N, Notomi T. Detection of loop-mediated isothermal amplification reaction by turbidity derived from magnesium pyrophosphate formation. Biochem Biophys Res Commun. 2001; 289(1): 150-4.

[19] Kiatpathomchai W, Jaroenram W, Arunrut N, Jitrapakdee S, Flegel TW. Shrimp Taura syndrome virus detection by reverse transcription loop-mediated isothermal amplification combined with a lateral flow dipstick. J Virol Methods. 2008; 153(2): 214-7.

[20] Park JY, Park S, Park YR, Kang DY, Kim EM, Jeon HS, et al. Reverse-transcription loop-mediated isothermal amplification (RT-LAMP) assay for the visual detection of European and North American porcine reproductive and respiratory syndrome viruses. J Virol Methods. 2016; 237: 10-13.

[21] Kumar B, Maharana BR, Brahmbhatt NN, Thakre BJ, Parmar VL. Development of a loop-mediated isothermal amplification assay based on RoTat1.2 gene for detection of *Trypanosoma evansi* in domesticated animals. Parasitol Res. 2021; 120(5): 1873-82.

[22] Njiru ZK, Mikosza AS, Matovu E, Enyaru JC, Ouma JO, Kibona SN, et al. African trypanosomiasis: sensitive and rapid detection of the sub-genus Trypanozoon by loop-mediated isothermal amplification (LAMP) of parasite DNA. Int J Parasitol. 2008; 38(5): 589-99.

[23] Kim J, Álvarez-Rodríguez A, Li Z, Radwanska M, Magez S. Recent progress in the detection of surra, a neglected disease caused by *Trypanosoma evansi* with a one health impact in large parts of the tropic and sub-tropic world. Microorganisms. 2023; 12(1): 44.

[24] Matthew MA, Christie J, Yang N, Yao C. A loop-mediated isothermal amplification (LAMP) assay specific to *Trichomonas tenax* is suitable for use at point-of-care. Microorganisms. 2022; 10(3): 594.

[25] Salim B, Bakheit MA, Kamau J, Nakamura I, Sugimoto C. Molecular epidemiology of camel trypanosomiasis based on ITS1 rDNA and RoTat 1.2 VSG gene in the Sudan. Parasit Vectors. 2011; 4: 31.

[26] Tong Q, Chen R, Kong Q, Goossens J, Radwanska M, Lou D, et al. DNA detection of *Trypanosoma evansi*: diagnostic validity of a new assay based on loop-mediated isothermal amplification (LAMP). Vet Parasitol. 2018; 250: 1-6.

[27] Ordóñez N, Salacinas M, Mendes O, Seidl MF, Meijer HJG, Schoen CD, et al. A loop-mediated isothermal amplification (LAMP) assay based on unique markers derived from genotyping by sequencing data for rapid in planta diagnosis of Panama disease caused by Tropical Race 4 in banana. Plant Pathol. 2019; 68(9): 1682-93.

[28] Kuboki N, Inoue N, Sakurai T, Di Cello F, Grab DJ, Suzuki H, et al. Loop-mediated isothermal amplification for detection of African trypanosomes. J Clin Microbiol. 2003; 41(12): 5517-24.

[29] Hayashida K, Kajino K, Hachaambwa L, Namangala B, Sugimoto C. Direct blood dry LAMP: a rapid, stable, and easy diagnostic tool for Human African Trypanosomiasis. PLoS Negl Trop Dis. 2015; 9(3): e0003578.

[30] Kumagai T, Furushima-Shimogawara R, Ohmae H, Wang TP, Lu S, Chen R, et al. Detection of early and single infections of *Schistosoma japonicum* in the intermediate host snail, Oncomelania hupensis, by PCR and loop-mediated isothermal amplification (LAMP) assay. Am J Trop Med Hyg. 2010; 83(3): 542-8.

[31] Plutzer J, Karanis P. Rapid identification of Giardia duodenalis by loop-mediated isothermal amplification (LAMP) from faecal and environmental samples and comparative findings by PCR and real-time PCR methods. Parasitol Res. 2009; 104(6): 1527-33.

[32] Kong QM, Lu SH, Tong QB, Lou D, Chen R, Zheng B, et al. Loop-mediated isothermal amplification (LAMP): early detection of *Toxoplasma gondii* infection in mice. Parasit Vectors. 2012; 5: 2.

[33] Hilali M, Abdel-Gawad A, Nassar A, Abdel-Wahab A, Magnus E, Büscher P. Evaluation of the card agglutination test (CATT/T. evansi) for detection of *Trypanosoma evansi* infection in water buffaloes (Bubalus bubalis) in Egypt. Vet Parasitol. 2004; 121(1-2): 45-51.

[34] Aquino LP, Machado RZ, Lemos KR, Marques LC, Garcia MV, Borges GP. Antigenic characterization of *Trypanosoma evansi* using sera from experimentally and naturally infected bovines, equines, dogs, and coatis. Rev Bras Parasitol Vet. 2010; 19(2): 112-8.

[35] Verma R, Das G, Singh AP, Kumar S, Nath S, Sengupta PP, et al. Molecular and genetic diversity in isolates of *Trypanosoma evansi* from naturally infected horse and dogs by using RoTat 1.2 VSG gene in Madhya Pradesh, India. Mol Biol Rep. 2023; 50(9): 7347-56.

[36] Thekisoe OM, Inoue N, Kuboki N, Tuntasuvan D, Bunnoy W, Borisutsuwan S, et al. Evaluation of loop-mediated isothermal amplification (LAMP), PCR and parasitological tests for detection of *Trypanosoma evansi* in experimentally infected pigs. Vet Parasitol. 2005; 130(3-4): 327-30.

Journal of Associated Medical Sciences 2025; 58(3): 288-297

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Icaritin Delivered by Hyaluronic Acid-modified Liposome Enhanced Apoptosis and Anti-Metastasis of Huh7 Liver Cancer Cells

Xiaoduan Sun[1,2,3], Xin Long[4], Ruilin Lu[5], Sawitree Chiampanichayakul[1,6,7], Songyot Anuchapreeda[1,6,7], Siriporn Okonogi[7], Yan Lin[3], Zhirong Zhong[3*], Singkome Tima[1,6,7*]

[1]*Hematology and Health Technology Research Center, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Province Thailand.*
[2]*Department of Pharmacy, The Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China*
[3]*Key Laboratory of Medical Electrophysiology, Ministry of Education, School of Pharmacy, Southwest Medical University, Luzhou, Sichuan, China.*
[4]*School of Clinical Medical Sciences, Southwest Medical University, Luzhou, Sichuan, China.*
[5]*Suining First People's Hospital, Suining, Sichuan, China.*
[6]*Cancer Research Unit of Associated Medical Sciences (AMS-CRU), Chiang Mai University, Chiang Mai Province, Thailand.*
[7]*Center of Excellence in Pharmaceutical Nanotechnology, Faculty of Pharmacy, Chiang Mai University, Chiang Mai Province, Thailand.*

ARTICLE INFO

Article history:
Received 1 June 2025
Accepted as revised 3 August 2025
Available online 7 August 2025

Keywords:
Hepatocellular carcinoma, drug delivery systems, hyaluronic acid, icaritin, reactive oxygen species, metastasis, apoptosis; ARPC1B.

ABSTRACT

Objectives: The advancement of tumor-targeted drug delivery systems, particularly those utilizing hyaluronic acid (HA)-modified nanocarriers, presents a promising strategy for enhancing cancer therapy by specifically targeting CD44-overexpressing tumor cells. In this study, HA-functionalized liposomes (HA-Lip) were employed to deliver the anti-cancer drug icaritin (ICT) to liver cancer cells to overcome its poor water solubility and low bioactivity.

Materials and methods: The efficacy of HA-Lip-ICT was evaluated through a combination of *in vitro* assays, including reactive oxygen species (ROS) production measurement, cell migration and invasion assays using Huh7 liver cancer cells. RNA sequencing and Western blot analysis were used to elucidate the molecular mechanisms underlying ICT's anti-cancer effects.

Results: HA-Lip-ICT significantly increased the production of reactive oxygen species (ROS) induced by icaritin and effectively inhibited the migration and invasion of Huh7 liver cancer cells *in vitro* through HA/CD44 interactions. RNA sequencing and Western blot analysis revealed that icaritin inhibited liver cancer progression by promoting apoptosis, inducing cell cycle arrest, and reducing metastasis. Notably, HA-Lip-ICT demonstrated a tumor migration inhibition rate of 55.6% and down-regulated the metastasis closely related protein ARPC1B. Moreover, it significantly upregulated proteins associated with apoptosis and cell cycle regulation, including caspase 3a, p53, and p21.

Conclusion: The research highlights that HA-Lip-ICT presents considerable promise as a targeted drug delivery system for tumors, particularly in enhancing liver cancer treatment by reducing tumor spread and facilitating cell death.

Introduction

Hepatocellular carcinoma (HCC) is the predominant form of primary liver cancer, representing around 80% to 90% of all cases. It is associated with a poor prognosis, having significantly low survival rates.[1] The main treatments are surgery and liver transplantation, but many HCC cases advance to inoperable stages, especially in regions like China, where HCC rates remain at high levels.[1, 2] Even after surgical resection, the long-term outlook remains poor due to the persistent risk of cancer recurrence or metastasis.[3-5] Conventional chemotherapies like cisplatin

∗ **Corresponding contributor.**
Author's Address: *Hematology and Health Technology Research Center, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Province Thailand.*
E-mail address: singkome.tima@cmu.ac.th and zhongzhirong@126.com
doi: 10.12982/JAMS.2025.099
E-ISSN: 2539-6056

show limited efficacy in advanced HCC, and their long-term use is hampered by acquired resistance and side effects.[6] Therefore, it is urgent to explore natural plant-derived compounds as potential anti-HCC agents to develop more effective cancer therapies.[7, 8]

Icaritin (ICT), a prenylated flavonoid derived from Epimedium plants, particularly the Berberidaceae family, has shown significant potential in cancer treatment. Extracts from *Herba Epimedii*, traditionally utilized in Chinese medicine for their tonic and aphrodisiac effects, serve as the source of icaritin.[9] Icaritin is currently undergoing evaluation in phase III clinical trials for advanced hepatocellular carcinoma, bolstered by significant preclinical and early-stage clinical data showcasing its therapeutic efficacy.[10] Icaritin demonstrates potent anti-tumor efficacy against HCC by modulating multiple signaling pathways and cellular processes, including inhibition of CXCR4, NF-κB and STAT3, regulation of various microRNAs, induction of ROS production, modulation of sphingosine kinase-1, and interaction with the RAGE-HMGB1 pathway.[10] Icaritin also plays a crucial role in modulating the interplay between apoptosis and autophagy, leading to significant alterations in the tumor microenvironment that promote an immune response.[10] These multifaceted actions make icaritin a promising therapeutic option for HCC.[10] However, icaritin faces challenges in clinical use due to its poor water solubility and lower bioavailability. Therefore, developing new strategies to increase its aqueous solubility and enhance its bioavailability is essential.[11,12]

Nano drug delivery systems have improved the efficacy of drugs in the treatment of hepatocellular carcinoma. Nanoparticles as drug delivery systems are a popular strategy that can effectively deliver poorly soluble drugs to target tissues and enhance drug dissolution.[13] These systems allow for lower drug dosages, reducing systemic toxicity risk, extending drug release, and improving targeted delivery to cancer cells.[14] Liposomes are one of the most promising nanoparticles for clinical applications, enhancing drug dissolution by encapsulating hydrophobic drugs within their hydrophobic bilayer and exhibiting favorable biodegradability and biocompatibility.[15,16]

Functionalized liposomes with target ligands offer additional abilities. Hyaluronic acid (HA) is a natural material with strong affinity for CD44 receptors, which are overexpressed in cancer cells, allowing targeted delivery through HA-CD44 interactions.[17] Hyaluronic acid has gained considerable attention in the formulation of liposomes, primarily due to its unique properties of hydrophilicity and biocompatibility.[18-20] Therefore, applying an HA layer to liposomes provides stability and enhances active targeting ability. This modification facilitates its uptake into cancer cells via the HA/CD44 ligand-receptor interaction.

Studies have shown that CD44 upregulation is closely related with poor prognosis of HCC and a shortened survival time, highlighting CD44's role in maintaining cancer stem cells and regulating reactive oxygen species (ROS) in human HCC cells like Huh7, which are aggressive and invasive liver cancer cells.[21-23] Leveraging the interaction between CD44 and hyaluronic acid presents a promising strategy for designing targeted drug delivery systems specifically for hepatocellular carcinoma treatment.

In our previous study, a cholesterol-tri(ethyleneglycol)-hyaluronic acid conjugate (HA-Chol) was synthesized and used for preparation of HA-functionalized liposomes (HA-Lip) to enable the precise delivery of icaritin.[24] Cholesterol was chosen for conjugation with HA to maintain liposome fluidity, utilizing its free secondary hydroxyl group for chemical modifications.[25] It was found that icaritin-loaded HA-coated liposomes (HA-Lip-ICT) significantly enhanced icaritin uptake by Huh7 cells.[24] The HA-modified liposomes exhibit several advantageous characteristics that enhance their potential for effective drug delivery, particularly for encapsulating icaritin.

In vitro, HA-Lip-ICT significantly enhanced intracellular uptake, increased cytotoxicity, and arrested cell cycle compared with Lip-ICT. In addition, HA-Lip-ICT increased the apoptotic rate three-fold in Huh7 cells compared with Lip-ICT.[24] According to our findings, we hypothesized that HA-Lip-ICT could address icaritin's solubility and bioavailability challenges, induce cellular apoptosis, and enhance its *in vitro* tumor migration inhibition capability on liver cancer cells via CD44-mediated endocytosis. Huh7 cells, which are more aggressive and invasive than HepG2 cells, were selected for this study because they express higher levels of CD44, a key target for hyaluronic acid-based drug delivery, as confirmed by previous studies and our prior research.[23, 24, 26] Thus, we plan to explore the potential mechanisms of its anti-metastasis and apoptosis induction. We investigated ROS generation by HA-Lip-ICT and assessed its metastasis inhibition efficacy on HCC cell models overexpressing CD44, evaluating wound healing, migration and invasion assays. To address the molecular mechanism of icaritin in HA-Lip-ICT, we explored intrinsic molecular mechanisms through RNA sequencing and performed bioinformatic analysis to understand how HA-Lip-ICT induced cell cycle arrest, apoptosis and metastasis inhibition, confirmed by Western blotting for key protein involvement.

Materials and methods

Materials

Icaritin was purchased from Plant Origin Biological Co. (Nanjing, China; 98% purity). Fetal bovine serum (FBS) and trypsin were sourced from Gibco-BRL (New York, USA). TRIzol® reagent was obtained from Invitrogen (Carlsbad, CA, USA). Antibodies against ARPC1B, p21, caspase-3a, GAPDH and p53 were procured from ABclonal Technology (Wuhan, China).

Cell culture

The human liver cancer cell line, Huh7, was obtained from the Chinese Academy of Sciences' Cell Bank (Shanghai, China) and cultured in DMEM medium supplemented with 10% FBS, 100 IU/mL penicillin and 100 µg/mL streptomycin at 37 °C in a 5% CO_2 atmosphere.

ROS generation measurement

ROS induced by icaritin has been proven to promote Huh7 cell apoptosis and cell cycle arrest.[27] The ROS production induced by icaritin and liposomal formulations were measured using the fluorescent dye hydroethidine 2',7'-dichlorodihydrofluorescein diacetate (DCFH-DA, Molecular Probes, Beyotime). After treatment with various icaritin formulations for 24 hrs, cells were incubated with 10 µM DCFH-DA at 37°C for another 30 mins. Subsequently, the cells were washed with PBS buffer and analyzed for ROS levels using flow cytometry (Becton Dickinson, NJ, USA).

HA-Lip-ICT inhibits cell migration and invasion in vitro

For the wound healing assay, Huh7 cells were seeded in 6-well plates and scratched using a pipette tip after 24 hrs of incubation. Icaritin (10 µM) formulations were then added to the wells. The wound area was then imaged under an inverted microscope (Olympus, Tokyo, Japan) at 0 and 12 hrs post-treatment. For migration assay, Huh7 cells were treated with icaritin (10 µM) formulations for 24 hrs. Then, 6×10^4 pretreated cells in 100 µL of serum-free DMEM were placed into the upper chamber of a transwell (pore size: 8 µm; Beyotime, China). Simultaneously, 600 µL DMEM with FBS was added to the lower chamber. After incubation for 24 hrs, cells were washed with PBS and the remaining cells on the upper chamber were removed. The migrated cells on the lower chamber were fixed with paraformaldehyde and stained for 30 mins using crystal violet. The stained transwell membranes were observed via microscope. Subsequently, the crystal violet stained cells were dissolved in 33% acetic acid, and the optical density was measured at 570 nm. For cell invasion, 6×10^4 pretreated cells were placed into the upper chamber of the transwell coated with Matrigel (Beyotime, China) and incubated for 24 hrs. The subsequent steps were identical to those described for the migration assay.

RNA sequencing and bioinformatic analysis

RNA-seq of control and icaritin-treated (10 µM, 24 hrs) Huh7 cells was performed using the BGISEQ-500 sequencing system (Beijing Genomics Institute, Shenzhen, Guangdong, China). Gene expression was quantified and normalized using the RSEM tool. Differentially expressed genes (DEGs) between control and icaritin-treated Huh7 were screened by NOISeq method. Analysis of DEGs was conducted using Gene Ontology (GO) function classification annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment.[28]

Western blotting

After treatment with free icaritin, Lip-ICT and HA-Lip-ICT at IC_{50} or IC_{20} concentrations of free icaritin, RIPA buffer was used to prepare whole cell lysates. Protein quantification was conducted using the BCA protein assay. Samples containing equal protein amounts (30 µg) from the lysates were subjected to separation via SDS-PAGE and then transferred onto nitrocellulose membranes. Following a block with 5% non-fat milk, the blots were incubated overnight at 4 °C with primary antibodies against glyceraldehyde 3-phosphate dehydrogenase (GAPDH), p21, caspase 3a, p53 and ARPC1B, and then with peroxidase-conjugated secondary antibodies. The membranes were stripped and the immunoblot bands were visualized using enhanced chemiluminescence (Millipore). GAPDH served as a control for protein loading.

Statistical analysis

All results were obtained from three independent experiments, and the data are presented as mean±SD. Student's t-test or one-way analysis of variance (ANOVA) was used to test the significance of the experiment. Statistical differences were considered significant when $p<0.05$ and very significant when $p<0.01$.

Results

HA-Lip-ICT enhanced ROS generation in Huh7 cells

It has been observed that ROS promoted Huh7 cell apoptosis and inhibited migration ability.[29, 30] ROS production was detected after various icaritin formulation treatments by using the redox-sensitive fluoroprobe, DCFH-DA. Flow cytometry was employed for measurement the fluorescent signal. As shown in Figure 1, after treatment with icaritin formulations for 24 hrs, the ROS levels in the Huh7 cells were significantly increased compared with control group. However, the HA-Lip-ICT group has a significantly higher ROS level than Lip-ICT; this may suggest that HA-Lip-ICT enhanced the uptake of icaritin by targeting CD44 receptors which has been confirmed by our previous study.[24] The ROS level in the HA-Lip-ICT group was lower than that of the free icaritin group, providing an important insight into the mechanism, whereby free icaritin had a rapid cellular uptake via passive diffusion.[24, 31, 32]

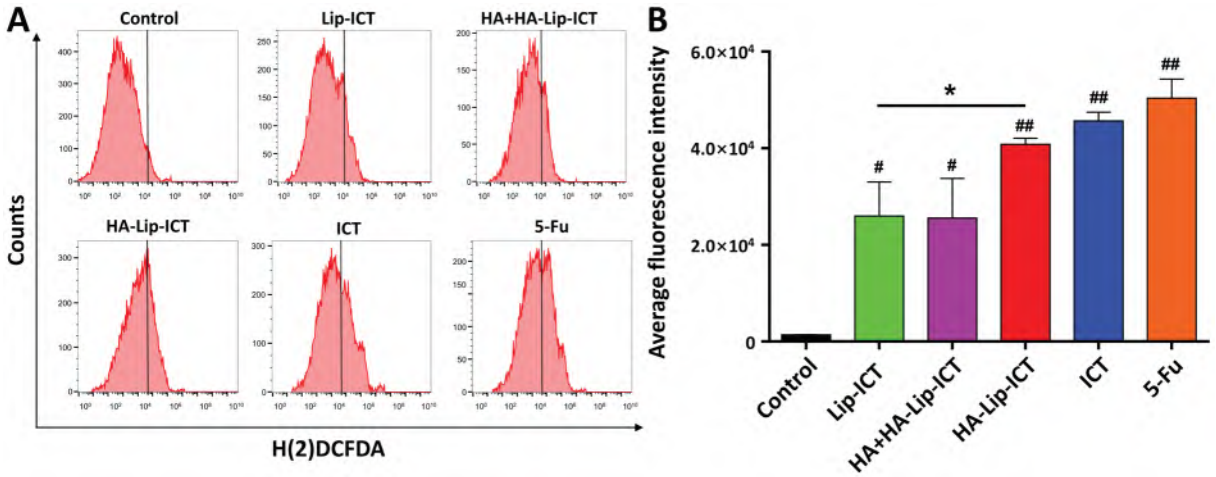


Figure 1. *ROS generation levels by various icaritin formulations and their quantification. Data are shown as mean±SD, N=3. *p<0.05 between HA-Lip-ICT and Lip-ICT; #p<0.05, ##p<0.01 vs. control group.*

In vitro cell migration and invasion inhibition of HA-Lip-ICT

The cell migration was first assessed by wound healing assay (Figure 2). In the control group, the wound healing was around 50% after 12 hrs, suggesting that Huh7 cells had a strong migration ability. The free icaritin group and 5-Fu as positive control group exhibited the best inhibition ability on cell migration. The wound healing ratio for Lip-ICT and HA-Lip-ICT were 44.6% and 35.5%, respectively, implying the higher inhibition of cell migration for HA-Lip-ICT (Figure 2A and B). These results may be due to the HA coated on liposomes, which led to increased cellular uptake of HA-Lip-ICT. As for migration assay, the same pattern of HA-Lip-ICT inhibition effect was observed. The migrated cells in the HA-Lip-ICT group were significantly lower than in the Lip-ICT group and the migration ratio was 45.5% for HA-Lip-ICT, whereas it reached 74.6% for Lip-ICT (Figure 2C). As shown in Figure 2D, the invasion of Huh7 cells through the Matrigel-coated membrane was notably inhibited by icaritin formulations compared with control group. Moreover, the cell invasion rates for Lip-ICT, HA-Lip-ICT and free icaritin were 67.3%, 30.1% and 8.5%, respectively. The results from the cell migration and invasion assays demonstrated that HA-modified liposomes loaded with icaritin significantly inhibited the migratory and invasive capabilities of Huh7 liver cancer cells compared with normal liposomes.

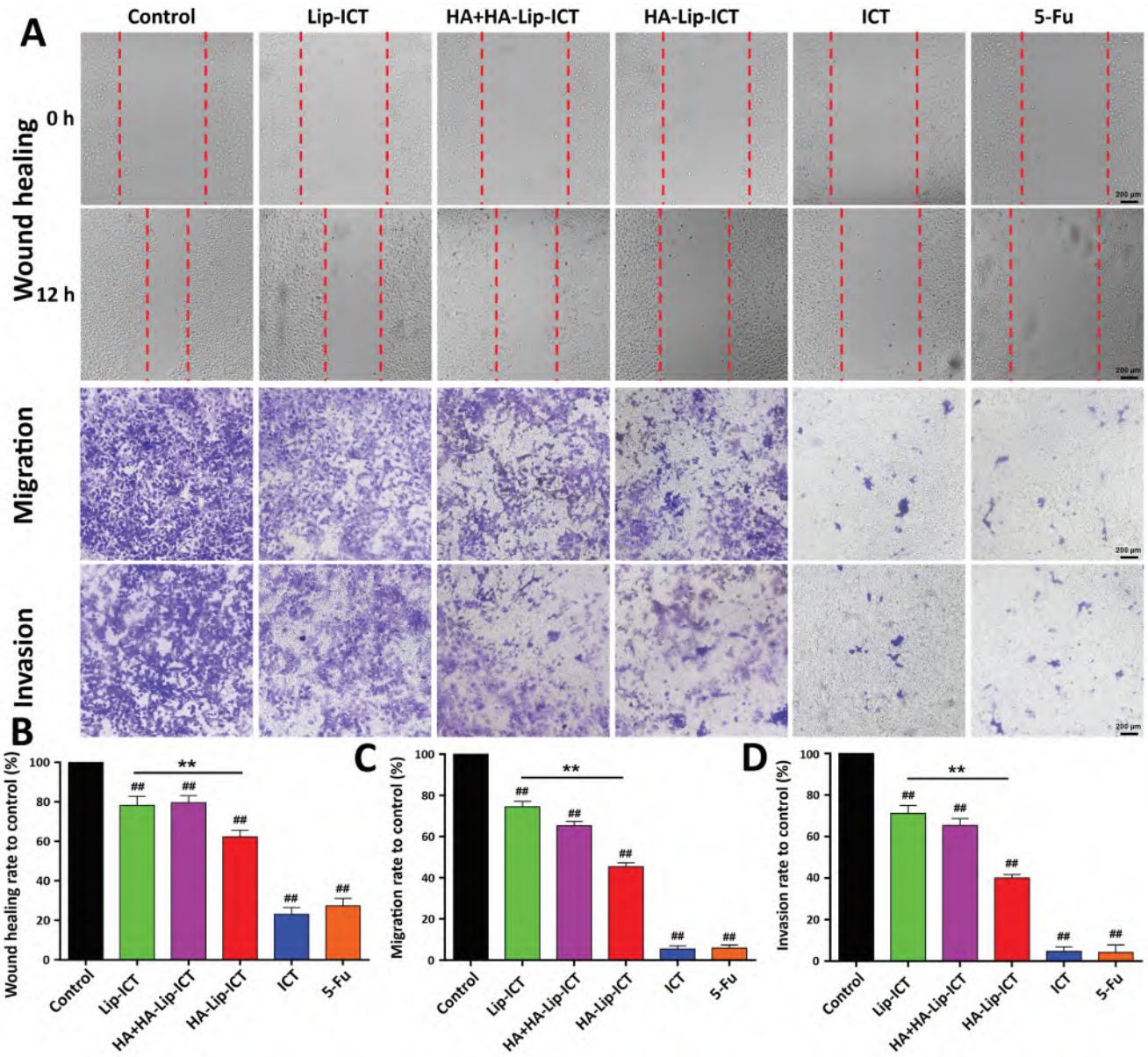


Figure 2. *In vitro inhibitory effects of HA-Lip-ICT on cell migration and invasion abilities of Huh7 cells. A: typical images of wound healing, migration and invasion assessments of various icaritin formulations (10 μM) on Huh7 cells, B: quantified wound healing inhibitory rate of various icaritin formulations in comparison with negative control (**p<0.01 between HA-Lip-ICT and Lip-ICT), C: quantified migration, D: invasion inhibitory rates of various icaritin formulations in comparison with negative control (**p<0.01 between HA-Lip-ICT and Lip-ICT). ##p<0.01 vs control group.*

Regulation of apoptosis, cell cycle and metastasis in Huh 7 cells by HA-Lip-ICT

In the present study, 1,358 kinds of up-regulated genes and 1,092 kinds of down-regulated genes were identified based on quantitative genomics, and the data were presented as volcano plots (Figure 3A) and histograms (Figure 3B). KEGG pathway analysis revealed differentially expressed genes. More specifically, the top 10 genes were enriched in multiple biological processes, such as cell cycle, apoptosis, p53 signaling pathway, cellular senescence, necroptosis, etc. (Figure 3C). Meanwhile, GO enrichment analysis indicated that differential genes were involved in the cellular response to the regulation of cell cycle, apoptotic process, cellular response to DNA damage stimulus, regulation of cyclin-dependent proteins, DNA replication, etc. (Figure 3D).

To elucidate the underlying anti-tumor mechanisms of HA-Lip-ICT, the expression levels of key markers associated with cell cycle arrest (p21), apoptosis activator (caspase 3) and the tumor suppressor (p53), were examined in Huh7 cells treated with different icaritin formulations using Western blot analysis. The icaritin formulations notably increased the expression of the cyclin-dependent kinase inhibitor p21 in Huh7 cells (Figure 3E), as observed in the Western blotting. Figure 3F shows that all icaritin formulations elevated the expression of activated caspase 3. Moreover, Huh7 cells treated with HA-Lip-ICT exhibited higher levels of p53 expression compared with those treated with Lip-ICT (Figure 3G).

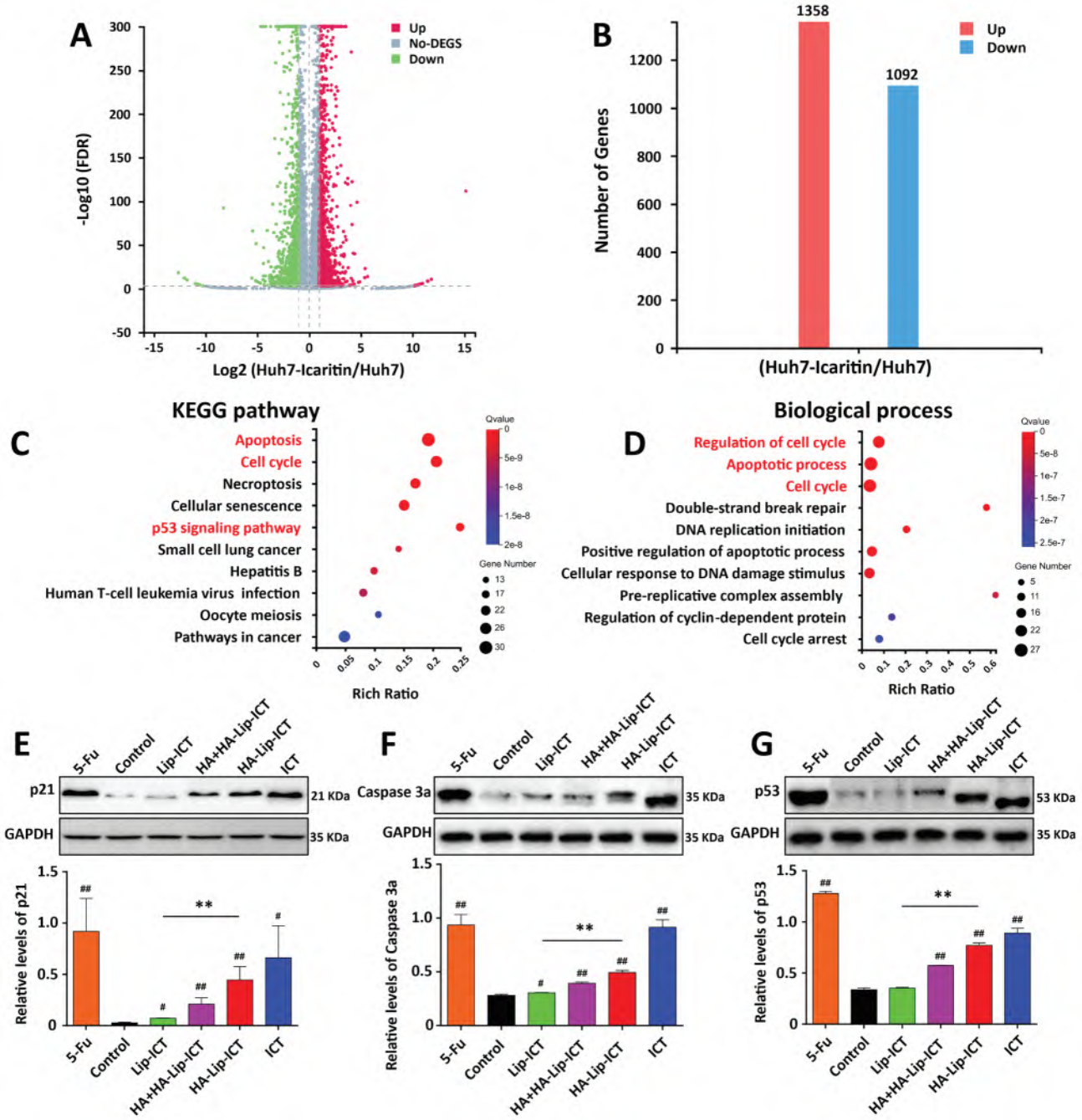


Figure 3. *Regulatory mechanisms of HA-Lip-ICT on the apoptosis and cell cycle process of Huh7 cells. A: volcano map of differentially expressed genes, B: histogram displaying the number of differentially expressed genes, C: KEGG pathway analysis of differentially expressed genes, D: GO analysis of differentially expressed genes. Western blot analysis of p21 (E), active caspase 3 (F), and p53 protein levels (G) in Huh7 cells after treatment for 48 hrs with various Lip-ICT formulations. GAPDH was the loading control. Data are shown as mean±SD, N=3. **p<0.01 between HA-Lip-ICT and Lip-ICT; #p<0.05, ##p<0.01 vs. control group.*

KEGG pathway analysis revealed differentially expressed genes in the Huh7 migration process. More specifically, the top 10 genes were enriched in multiple biological processes, with the top gene regulating actin cytoskeleton (Figure 4A). Meanwhile, GO enrichment analysis indicated that differential genes were involved in the cellular response to the regulation of actin filament polymerization, actin cytoskeleton organization, positive regulation of cell migration, etc. (Figure 4B). In addition, a heatmap of differentially expressed genes between icaritin and Huh7 cell migration was constructed (Figure 4C). To investigate the potential anti-metastasis mechanisms of HA-Lip-ICT, the expression levels of key markers associated with migration of ARPC1B were examined in Huh7 cells treated with different icaritin formulations using Western blot analysis. The icaritin formulations notably decreased the expression of ARPC1B in Huh7 cells, as observed in the Western blotting (Figure 4D). While there was no decreased level of ARPC1B in the 5-Fu group, it is possible 5-Fu inhibited Huh7 cell migration without involving this protein.[33, 34] In addition, a protein-protein interaction (PPI) network of differentially expressed genes connected with ARPC1B was constructed using the STRING database (Figure 4E).

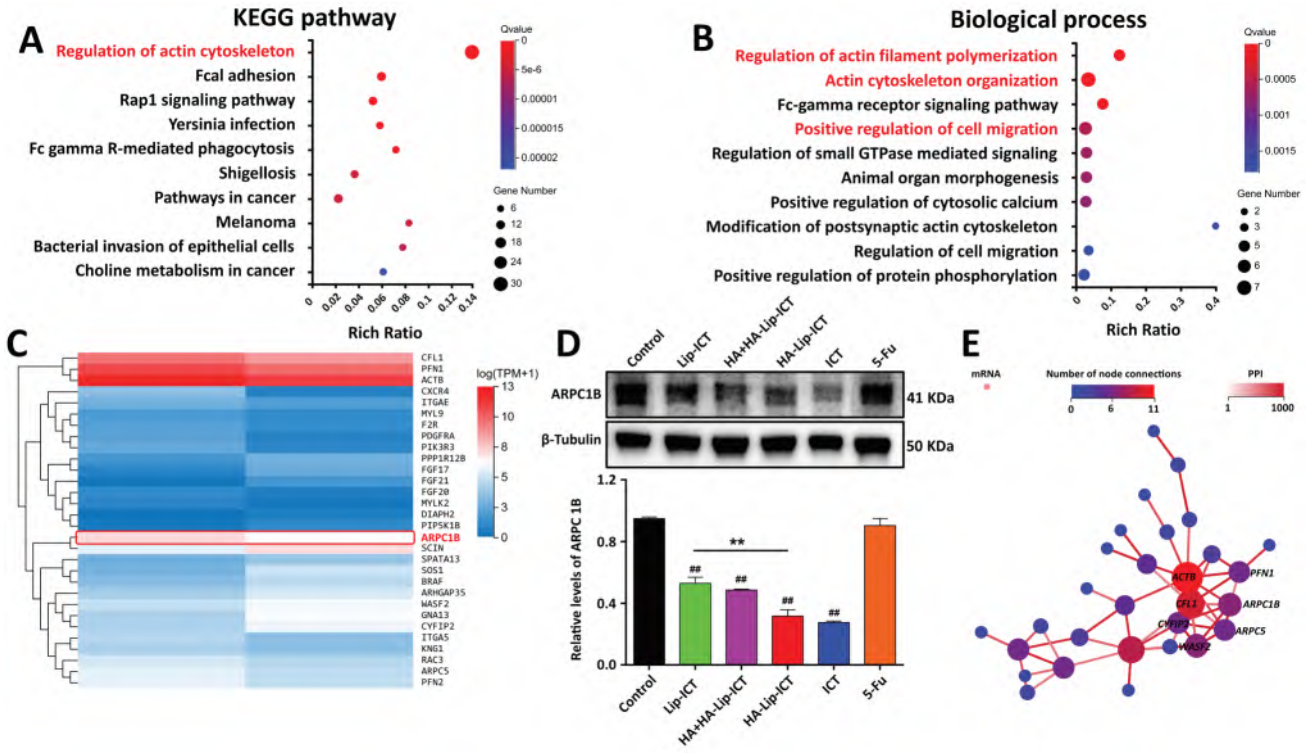


Figure 4. *Regulatory mechanisms of HA-Lip-ICT on the migration process of Huh7 cells. A: KEGG pathway analysis of differentially expressed genes, B: GO analysis of differentially expressed genes, C: heat map of differentially expressed genes, D: expression level of ARPC1B and its quantification, E: PPI network of differentially expressed genes connected with ARPC1B. Data are shown as mean±SD, N=3. **p<0.01 between HA-Lip-ICT and Lip-ICT; ##p<0.01 vs. control group.*

Discussion

The higher uptake level of HA-Lip-ICT in Huh7 cells can indeed be attributed to the elevated expression of CD44 in those cells. Our previous study demonstrated that HA-CD44 binding between liposomes and Huh7 cells promoted increased cellular uptake of HA-Lip-ICT.[24] Previous studies have demonstrated that the presence of competing hyaluronic acid in medium can significantly reduce the cellular uptake of HA-modified nanoparticles. Therefore, a receptor competition group (HA+HA-Lip-ICT) was designed.[35] In our study, the pre-treatment condition with free HA (HA+HA-Lip-ICT) was performed along with HA-Lip-ICT treatment. The increase in ROS generation induced by HA-Lip-ICT was significantly reduced in the presence of free HA, suggesting a competitive interaction between free HA and HA-Lip-ICT over binding to the CD44 receptor. This result suggested that HA-CD44 interaction plays a key role in promoting enhanced cellular uptake.

According to our previous findings on apoptosis induction, HA-modified liposomes loaded with icaritin significantly enhanced the apoptosis-inducing capacity of icaritin, particularly by targeting cells that overexpressed the CD44 receptor.[24] In addition, the apoptosis rate of cells in the HA-Lip-ICT group was lower than the free icaritin group. This may be due to the fast uptake of free icaritin by cells through passive diffusion.[31, 32] Our previously reported cell cycle distribution indicated that HA-modified liposomes loaded with icaritin induced a significant increase in the percentage of cells in the G0/G1 phase compared with normal liposomes, which is believed to be a result of the enhanced internalization of icaritin facilitated by the HA-CD44 interaction.[24] The anti-migration results indicated

that HA-Lip-ICT reinforced the metastasis-inhibiting ability of icaritin, especially by targeting CD44.

Icaritin has been demonstrated to enhance oxidative stress and activate the p53-p21 pathway, which in turn inhibits cyclin-dependent kinases (CDKs), causing cell cycle arrest and apoptotic induction in tumor cells.[36] The protein p21 is a pivotal regulator of G1/S phase transition and these alterations in p21 expression likely favored cell cycle arrest at the G0/G1 phase, corroborating the findings from flow cytometry analysis.[27] The protein results suggested that icaritin possibly induced caspase-dependent apoptosis and inhibited hepatocarcinogenesis through the IL-6/Jak2/Stat3 signaling pathway,[37] and the well-established tumor suppressor, p53, played pivotal roles in both cell cycle arrest and apoptosis regulation.[38] On the other hand, icaritin and its nanoformulations exhibited an anti-metastatic effect by inhibiting the migration and invasion capacity of various cancer cell types via signaling pathways, including Akt/mTOR, JAK/STAT3 pathways.[39-41]

It has been reported that the overexpression of ARPC1B promoted the migration, invasion and epithelial-mesenchymal transition in a variety of cancer cells.[42-44] The *ARPC1B* gene, along with its closely related gene *ARPC1A*, operates in a regulatory loop to control tumor metastasis, with the ARPC1A/B axis being a potential key regulator of this process.[45] High expression of ARPC1B correlates with immune infiltration and poor outcomes in glioblastoma.[46] In this study, both free ICT and HA-Lip-ICT decreased ARPC1B expression levels leading to the attenuation of HCC migration and invasion. Based on our knowledge, this might be the first publication that reported the anti-metastatic activity of ICT by inhibiting ARPC1B expression in HCC. Taken together, these observations strongly suggest that HA-Lip-ICT exerts its anti-tumor activity by promoting apoptosis, inducing cell cycle arrest, and inhibiting tumor metastasis. Additionally, future *in vivo* migration inhibition experiments are necessary to further validate these observations and before clinical application, systemic stability of the liposomal formulation and the risk of immunogenicity or unexpected immune responses *in vivo* should be addressed.

Limitations

The study primarily lacks *in vivo* data to verify the systemic stability, pharmacokinetics, long-term toxicity, immunogenicity, and anti-metastatic efficacy of the HA-Lip-ICT formulation. Potential immune responses from the liposomal carrier or HA modification remain unassessed, so comprehensive *in vivo* validation and safety evaluation are needed to confirm its clinical potential.

Conclusion

A novel approach using hyaluronic acid (HA)-modified liposomal nanoplatforms to boost the absorption of an antitumor drug for hepatocellular carcinoma was described. *In vitro*, HA-Lip-ICT significantly improved ROS generation, inducing apoptosis and causing cell cycle arrest, and inhibited migration compared with Lip-ICT. This enhancement could be attributed to hyaluronic acid promoting CD44 receptor-mediated endocytosis. Liposomes functionalized with HA significantly enhanced the ability of icaritin to inhibit Huh7 cell migration and invasion. In conclusion, HA-Lip-ICT has promising potential as an effective targeted drug delivery strategy to enhance the therapeutic effect of liver cancer treatment. However, further *in vivo* studies need to be evaluated like systemic stability, potential long-term toxicity, and immunogenicity of HA-Lip-ICT before clinical application.

Ethical approval

No ethical approval was required for this study as it did not involve human or animal subjects.

Funding

This work was supported by grants from the Science and Technology Strategic Cooperation Project of Suining First People's Hospital and Southwest Medical University (2023SNXNYD05) and the Sichuan Science and Technology Program (2024YFFK0380).

Conflict of interest

The authors declare that they hold no competing interests.

CRedit authorship contribution statement

Xiaoduan Sun, Zhirong Zhong and **Singkome Tima:** conceptualized and designed the study; **Xin Long** and **Ruilin Lu:** conducted the formal analysis and drafted the manuscript; **Sawitree Chiampanichayakul, Songyot Anuchapreeda, Siriporn Okonogi** and **Yan Lin:** provided supervision and critical revisions. All authors approved the final version of the manuscript.

Acknowledgements

All authors are thankful for getting help and support from the Key Laboratory of Medical Electrophysiology, Ministry of Education, School of Pharmacy, Southwest Medical University, Luzhou, Sichuan, China.

References

[1] Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, et al. Global cancer statistics, 2012. CA Cancer J Clin. 2015; 65: 87-108. doi: 10.3322/caac.21262.

[2] Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA Cancer J Clin. 2016; 66: 115-32. doi: 10.3322/caac.21338.

[3] El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology. 2007; 132: 2557-76. doi: 10.1053/j.gastro.2007.04.061.

[4] Jing P, Luo H, Tan J, Liao C, Zhang S. Natural polyphenol-loaded cross-linked lipoic acid vesicles treat triple-negative breast cancer by cancer cell killing and metastasis inhibition. Mater Des. 2023; 236: 112461. doi: 10.1016/j.matdes.2023.112461.

[5] Tan J, Jing P, Xiao X, Liao Y, Liao C, Zhang S. Cross-linked lipoic acid nanocapsules serve as H2O2 amplifier to strengthen the H2O2-sensitive prodrug activation.

Sci China Chem. 2023; 66: 2654-63. doi: 10.1007/s11426-022-1647-2.
[6] Singh S, Singh PP, Roberts LR, Sanchez W. Chemo-preventive strategies in hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol. 2014; 11: 45-54. doi: 10.1038/nrgastro.2013.143.
[7] Cragg GM, Newman DJ. Plants as a source of anti-cancer agents. J Ethnopharmacol. 2005; 100: 72-9. doi: 10.1016/j.jep.2005.05.011.
[8] Newman DJ. Natural products as leads to potential drugs: an old process or the new hope for drug discovery? J Med Chem. 2008; 51: 2589-99. doi: 10.1021/jm0704090.
[9] Sze SC, Tong Y, Ng TB, Cheng CL, Cheung HP. Herba Epimedii: anti-oxidative properties and its medical implications. Molecules. 2010; 15: 7861-70. doi: 10.3390/molecules15117861.
[10] Bailly C. Molecular and cellular basis of the anticancer activity of the prenylated flavonoid icaritin in hepatocellular carcinoma. Chem Biol Interact. 2020; 325: 109124. doi: 10.1016/j.cbi.2020.109124.
[11] Li Y, Sun S, Chang Q, Zhang L, Wang G, Chen W, et al. A strategy for the improvement of the bioavailability and antiosteoporosis activity of BCS IV flavonoid glycosides through the formulation of their lipophilic aglycone into nanocrystals. Mol Pharm. 2013; 10: 2534-42. doi: 10.1021/mp300688t.
[12] Chang Q, Wang GN, Li Y, Zhang L, You C, Zheng Y. Oral absorption and excretion of icaritin, an aglycone and also active metabolite of prenylflavonoids from the Chinese medicine Herba Epimedii in rats. Phytomedicine. 2012; 19: 1024-8. doi: 10.1016/j.phymed.2012.05.017.
[13] Sabit H, Abdel-Hakeem M, Shoala T, Abdel-Ghany S, Abdel-Latif MM, Almulhim J, et al. Nanocarriers: A reliable tool for the delivery of anticancer drugs. Pharmaceutics. 2022; 14(8): 1566. doi: 10.3390/pharmaceutics14081566.
[14] Mitchell MJ, Billingsley MM, Haley RM, Wechsler ME, Peppas NA, Langer R. Engineering precision nanoparticles for drug delivery. Nat Rev Drug Discov. 2021; 20: 101-24. doi: 10.1038/s41573-020-0090-8.
[15] Lee MK. Liposomes for enhanced bioavailability of water-insoluble drugs: *In Vivo* evidence and recent approaches. Pharmaceutics. 2020; 12(3): 264. doi: 10.3390/pharmaceutics12030264.
[16] Cheng R, Liu L, Xiang Y, Lu Y, Deng L, Zhang H, et al. Advanced liposome-loaded scaffolds for therapeutic and tissue engineering applications. Biomaterials. 2020; 232: 119706. doi: 10.1016/j.biomaterials.2019.119706.
[17] Kesharwani P, Chadar R, Sheikh A, Rizg WY, Safhi AY. CD44-targeted nanocarrier for cancer therapy. Front Pharmacol. 2021; 12: 800481. doi: 10.3389/fphar.2021.800481.
[18] Su S, Kang P. Recent advances in nanocarrier-assisted therapeutics delivery systems. Pharmaceutics. 2020; 12(9): 837. doi: 10.3390/pharmaceutics12090837.
[19] Majumder J, Taratula O, Minko T. Nanocarrier-based systems for targeted and site specific therapeutic delivery. Adv Drug Deliv Rev. 2019; 144: 57-77. doi: 10.1016/j.addr.2019.07.010.
[20] Patra JK, Das G, Fraceto LF, Campos EVR, Rodriguez-Torres MDP, Acosta-Torres LS, et al. Nano based drug delivery systems: recent developments and future prospects. J Nanobiotechnology. 2018; 16: 71. doi: 10.1186/s12951-018-0392-8.
[21] Endo K, Terada T. Protein expression of CD44 (standard and variant isoforms) in hepatocellular carcinoma: relationships with tumor grade, clinicopathologic parameters, p53 expression, and patient survival. J Hepatol. 2000; 32: 78-84. doi: 10.1016/s0168-8278(00)80192-0.
[22] Asai R, Tsuchiya H, Amisaki M, Makimoto K, Takenaga A, Sakabe T, et al. CD44 standard isoform is involved in maintenance of cancer stem cells of a hepatocellular carcinoma cell line. Cancer Med. 2019; 8: 773-82. doi: 10.1002/cam4.1968.
[23] Cannito S, Bincoletto V, Turato C, Pontisso P, Scupoli MT, Ailuno G, et al. Hyaluronated and PEGylated liposomes as a potential drug-delivery strategy to specifically target liver cancer and inflammatory cells. Molecules. 2022; 27(3): 1062. doi: 10.3390/molecules27031062.
[24] Sun X, He Z, Lu R, Liu Z, Chiampanichayakul S, Anuchapreeda S, et al. Hyaluronic acid-modified liposomes Potentiated in-vivo anti-hepatocellular carcinoma of icaritin. Front Pharmacol. 2024; 15: 1437515. doi: 10.3389/fphar.2024.1437515.
[25] Ohvo-Rekila H, Ramstedt B, Leppimaki P, Slotte JP. Cholesterol interactions with phospholipids in membranes. Prog Lipid Res. 2002; 41: 66-97. doi: 10.1016/s0163-7827(01)00020-0.
[26] Song M, Liang Y, Li K, Zhang J, Zhang N, Tian B, et al. Hyaluronic acid modified liposomes for targeted delivery of doxorubicin and paclitaxel to CD44 over-expressing tumor cells with improved dual-drugs synergistic effect. J Drug Deliv Sci Technol. 2019; 53: 101179. doi: 10.1016/j.jddst.2019.101179.
[27] Wang S, Wang Q, Wang H, Qin C, Cui X, Li L, et al. Induction of ROS and DNA damage-dependent senescence by icaritin contributes to its antitumor activity in hepatocellular carcinoma cells. Pharm Biol. 2019; 57: 424-31. doi: 10.1080/13880209.2019.1628073.
[28] Ren F, Li J, Yuan X, Wang Y, Wu K, Kang L, et al. Dandelion polysaccharides exert anticancer effect on Hepatocellular carcinoma by inhibiting PI3K/AKT/mTOR pathway and enhancing immune response. J Funct Foods. 2019; 55: 263-74. doi: 10.1016/j.jff.2019.02.034.
[29] Munakarmi S, Chand L, Shin HB, Hussein UK, Yun BS, Park HR, et al. Anticancer effects of Poncirus fructus on hepatocellular carcinoma through regulation of apoptosis, migration, and invasion. Oncol Rep. 2020; 44: 2537-46. doi: 10.3892/or.2020.7790.
[30] Siqueira EDS, Concato VM, Tomiotto-Pellissier F, Silva TF, Bortoleti B, Goncalves MD, et al. Trans-

chalcone induces death by autophagy mediated by p53 up-regulation and beta-catenin down-regulation on human hepatocellular carcinoma HuH7.5 cell line. Phytomedicine. 2021; 80: 153373. doi: 10.1016/j.phymed.2020.153373.

[31] Liu D, Zhang Q, Wang J, Guan S, Cai D, Liu J. Inhibition of growth and metastasis of breast cancer by targeted delivery of 17-hydroxy-jolkinolide B via hyaluronic acid-coated liposomes. Carbohydr Polym. 2021; 257: 117572. doi: 10.1016/j.carbpol.2020.117572.

[32] Wang J, Liu D, Guan S, Zhu W, Fan L, Zhang Q, et al. Hyaluronic acid-modified liposomal honokiol nanocarrier: Enhance anti-metastasis and antitumor efficacy against breast cancer. Carbohydr Polym. 2020; 235: 115981. doi: 10.1016/j.carbpol.2020.115981.

[33] Hu X, Ding J, Wang G, Zhang X. The combination of ulinastatin and 5-fluorouracil synergistically inhibits hepatocellular carcinoma growth. J Int Med Res. 2020; 48(3): 0300060520909776. doi: 10.1177/0300060520909776.

[34] Ghafouri-Fard S, Abak A, Tondro Anamag F, Shoorei H, Fattahi F, Javadinia SA, et al. 5-Fluorouracil: A narrative review on the role of regulatory mechanisms in driving resistance to this chemotherapeutic agent. Front Oncol. 2021; 11: 658636. doi: 10.3389/fonc.2021.658636.

[35] Tran TH, Choi JY, Ramasamy T, Truong DH, Nguyen CN, Choi HG, et al. Hyaluronic acid-coated solid lipid nanoparticles for targeted delivery of vorinostat to CD44 overexpressing cancer cells. Carbohydr Polym. 2014; 114: 407-15. doi: 10.1016/j.carbpol.2014.08.026.

[36] Jeong SH. Inhibitory effect of phytol on cellular senescence. Biomed Dermatol. 2018; 2: 1-9. doi: 10.1186/s41702-018-0025-8.

[37] Yu Z, Guo J, Hu M, Gao Y, Huang L. Icaritin exacerbates mitophagy and synergizes with doxorubicin to induce immunogenic cell death in hepatocellular carcinoma. ACS Nano. 2020; 14: 4816-28. doi: 10.1021/acsnano.0c00708.

[38] Li H, Liu Y, Jiang W, Xue J, Cheng Y, Wang J, et al. Icaritin promotes apoptosis and inhibits proliferation by down-regulating AFP gene expression in hepatocellular carcinoma. BMC Cancer. 2021; 21: 318. doi: 10.1186/s12885-021-08043-9.

[39] Gao L, Ouyang Y, Li R, Zhang X, Gao X, Lin S, et al. Icaritin inhibits migration and invasion of human ovarian cancer cells via the Akt/mTOR signaling pathway. Front Oncol. 2022; 12: 843489. doi: 10.3389/fonc.2022.843489.

[40] Li X, Li C, Zhou P, Chen S. Inhibitory effect of icaritin on proliferation, migration, and invasion of human nasopharyngeal carcinoma cell CNE2 by regulating STAT3 activation. Pharmazie. 2019; 74: 685-7. doi: 10.1691/ph.2019.9632.

[41] Xiao Y, Yao W, Lin M, Huang W, Li B, Peng B, et al. Icaritin-loaded PLGA nanoparticles activate immunogenic cell death and facilitate tumor recruitment in mice with gastric cancer. Drug Deliv. 2022; 29: 1712-25. doi: 10.1080/10717544.2022.2079769.

[42] Huang J, Zhou H, Tan C, Mo S, Liu T, Kuang Y. The overexpression of actin related protein 2/3 complex subunit 1B(ARPC1B) promotes the ovarian cancer progression via activation of the Wnt/beta-catenin signaling pathway. Front Immunol. 2023; 14: 1182677. doi: 10.3389/fimmu.2023.1182677.

[43] Gamallat Y, Zaaluk H, Kish EK, Abdelsalam R, Liosis K, Ghosh S, et al. ARPC1B is associated with lethal prostate cancer and its inhibition decreases cell invasion and migration *In vitro*. Int J Mol Sci. 2022; 23: 1476. doi: 10.3390/ijms23031476.

[44] Liu T, Zhu C, Chen X, Wu J, Guan G, Zou C, et al. Dual role of ARPC1B in regulating the network between tumor-associated macrophages and tumor cells in glioblastoma. Oncoimmunology. 2022; 11: 2031499. doi: 10.1080/2162402X.2022.2031499.

[45] Chang D, Du H, Chen X, Bian X, Tian W, Shen J, et al. A controlled random gene perturbation method identifies ARPC1B gene as a key regulator of cancer metastasis. Genes Dis. 2023; 10: 687-9. doi: 10.1016/j.gendis.2022.06.006.

[46] Liao C, Chen W, Xu G, Wang J, Dong W. High expression of ARPC1B correlates with immune infiltration and poor outcomes in glioblastoma. Biochem Biophys Rep. 2024; 37: 101619. doi: 10.1016/j.bbrep.2023.101619.

Journal of Associated Medical Sciences 2025; 58(3): 298-306

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Prevalence and associated factors of pre-frailty and frailty among Thai community-dwelling older adults aged 65 years and above

Sukwida Manorangsan[1,2], Pawan Chaiparinya[1], Sirima Mongkolsomlit[3], Sujitra Boonyong[1*]

[1]Human Movement Performance Enhancement Research Unit, Department of Physical Therapy, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand.
[2]Department of Physical Therapy, Faculty of Allied Health Sciences, Thammasat University, Pathum Thani Province, Thailand.
[3]Faculty of Public Health, Thammasat University, Pathum Thani Province, Thailand.

ARTICLE INFO

Article history:
Received 10 March 2025
Accepted as revised 30 July 2025
Available online 14 August 2025

Keywords:
Pre-frailty, physical frailty, Thai older adults, community-dwelling.

ABSTRACT

Background: As Thailand's aging population grows rapidly, frailty status may develop and have serious implications. Early detection or prevention can be encouraged, especially prior to the emergence of advanced frailty and functional deterioration. Pre-frailty and frailty should be identified and investigated.

Objectives: To explore the prevalence of frailty statuses and their associations with sociodemographic characteristics among Thai older adults.

Materials and methods: This study was a cross-sectional design. Two hundred sixty-eight Thai community-dwelling older adults aged ≥65 years were recruited. Their physical frailty status was determined using the Fried frailty phenotypic criteria, which comprise weakness, slowness, low physical activity, weight loss, and exhaustion. Multinomial (Polytomous) logistic regression was used to detect associations between frailty status and sociodemographics.

Results: Among participants, the prevalences of pre-frailty, frailty, and robustness were 38.4%, 42.5%, and 19.0%, respectively. In the multivariate-adjusted model, pre-frailty was significantly associated with age (relative risk ratio(RRR)=1.10; 95% confidence interval (CI)=1.01-1.18, p=0.020), occupation as governor or owner (RRR=0.37; 95%CI=0.15-0.92, p=0.032), and housewife (RRR=0.29; 95%CI=0.09-0.93, p=0.037). Likewise, frailty was associated with age (RRR=1.28; 95%CI=1.18-1.39, p<0.001), gender (RRR=0.34; 95%CI=0.14-0.85, p=0.021), BMI (RRR=0.88; 95%CI=0.79-0.98, p=0.019), and education (RRR=0.23l; 95%CI=0.09-0.58, p=0.002). Additionally, comorbidity with ≥3 diseases was significantly associated with frailty (RRR=16.34; 95%CI=1.64-162.46, p=0.017).

Conclusion: Our findings reveal comparable prevalence rates of pre-frailty and frailty among Thai older adults. While age and occupation were associated with pre-frailty, a broader range of factors, including age, gender, BMI, comorbidities, and education, showed associations with frailty. These results underscore the importance of considering multiple factors collectively when screening or assessing an elderly individual's risk of developing pre-frailty or frailty.

Introduction

Pre-frailty describes a transitional health state beyond normal aging, characterized as a complex, multifactorial condition linked to gradual physiological decline.[1] Though often subclinical, it may present as fatigue or weakness and includes distinct bio-psychosocial domains. Pre-frailty increases the risk of frailty and adverse health outcomes.[1,2] Previous studies have identified a high prevalence of pre-frailty in community-dwelling older adults, highlighting its role as a key transitional stage and target for early intervention.[3-8]

∗ **Corresponding contributor.**
Author's Address: *Human Movement Performance Enhancement Research Unit, Department of Physical Therapy, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand.*
E-mail address: sujitra.b@chula.ac.th
doi: 10.12982/JAMS.2025.100
E-ISSN: 2539-6056

Although pre-frailty has received less research attention, its prevalence remains consistently higher than that of frailty, ranging from 43.1% to 49.5% globally[3-6] and from 50.9% to 57.4% in Thailand,[7,8] whereas frailty affects only 8% to 22% globally[3-6] and 13.9% to 15.9% in Thailand.[7,8] The higher prevalence of pre-frailty underscores the risk of progression without timely intervention. As a reversible stage between robustness and frailty, pre-frailty offers a crucial opportunity for preventive action.[1,9-11]

Identifying factors linked to pre-frailty enables early detection and prevention. As a reversible stage, pre-frailty can be addressed before advanced decline.[1] However, research on its prevalence and associated factors in Thai community-dwelling older adults remains limited. Age, education, marital status, and comorbidities are key contributors to pre-frailty.[4,10,11] A systematic review of studies conducted in Europe and China identified several risk factors associated with both frailty and pre-frailty, including advanced age, female sex, rural residence, low educational attainment, poor marital status, living alone, financial hardship, and limited access to medical reimbursement.[4,10]

Understanding the prevalence of pre-frailty and frailty is essential for guiding health policy and planning effective responses to the challenges posed by an aging population. Accurate prevalence data can assist policymakers in evaluating the capacity of health and social service systems to manage the growing burden of geriatric pre-frailty, particularly in the context of Thailand's rapidly aging society. Furthermore, identifying factors associated with these conditions is critical to gaining deeper insight into the underlying causes of geriatric vulnerability within the Thai population. This knowledge is pivotal for developing timely, targeted interventions aimed at preventing or delaying the progression to frailty and reducing its adverse health outcomes. To support these objectives, it is necessary to update existing prevalence estimates and reassess contributing sociodemographic and health factors. The prevalence of frailty, as classified by the Fried physical frailty criteria, was explored in elderly individuals aged 65 years and older who were living in a community.[9] Therefore, the purpose of this study was to investigate the prevalence of pre-frailty and frailty among community-dwelling individuals aged 65 years and older. It also aimed to examine the associations between frailty status and sociodemographic characteristics in this population.

Materials and methods

Study design

A cross-sectional study was conducted among older Thai adults living in communities in Pathum Thani Province, Thailand, from November 2021 to August 2022. The recruitment location was chosen based on potential participants' willingness to engage in the study and the cooperation of the health-promoting hospital, which provided space for the assessment station. The sample size was determined using population proportion estimates from the previous study.[7] A sampling error fixed at five percentage points, and a 95% confidence level (α=0.05) were considered. Due to the potential for losses or refusals, the sample size was increased by 10%, resulting in a sampling plan of 268 older individuals. This study was approved by the Research Ethics Committee of Chulalongkorn University, Thailand (COA No. 089/2020) and conducted according to the Declaration of Helsinki. All participants provided written informed consent before starting data collection.

Participants

This study enrolled 268 participants using the convenience sampling method. Inclusion criteria were age 65 years or older, residence in the home for at least 1 year, ability to read and communicate in Thai, and no uncorrected hearing or visual impairments. Exclusion criteria included a medical history suggestive of Parkinson's or Alzheimer's disease, identified through a standard screening questionnaire administered by the main researcher.

Procedures

For each eligible participant, personal sociodemographic data were recorded, including age, sex, weight, height, body mass index (BMI), educational level, occupation, marital status, residential status, and health status, such as underlying diseases and other medical conditions. Participants were then evaluated for physical performance and frailty phenotypes based on the 2001 Fried physical frailty criteria.[9] The assessment included five components; gait speed, muscle weakness, unintentional weight loss, low levels of physical activity (PA), and self-reported exhaustion; and was conducted by a physical therapist with 15 years of experience.

Gait speed was measured by timing participants as they walked at a comfortable speed along a 4.57-meter (15 feet) walkway.[9] One to two practice trials were conducted to help participants acclimate. The assessor issued the following instruction to participants: "Ready, set, go. When I say go, walk at your comfortable speed until I say stop." To eliminate the effect of acceleration and deceleration, walking speed was timed in the middle of the walkway. Timing commenced when the participant's foot strike passed the first meter mark and stopped when the toe-off passed the last 1-meter mark. Data were recorded for three trials and then averaged. If the average walking speed was in the lowest 20% (by sex and height), a score of one point was recorded.[9]

Muscle weakness was measured by the grip strength (Takei Grip Strength Dynamometer, model T.K.K. 5401 GRIP-D, Japan) of the dominant hand in a standing position. Participants were allowed to practice once before data collection. The assessor instructed participants to "squeeze with as much force as you can." After two recorded trials, the data were averaged. If the average hand grip strength fell within the lowest 20% (by sex and BMI), a score of one point was assigned.[9]

Weight loss was identified by asking participants whether they had unintentionally lost more than 4.5 kg

during the past year. They were asked, "How much did you weigh last year, and how much do you weigh now?" The comparison between their weight on the day of data collection and their weight from the previous year determined the score. A weight loss exceeding 4.5 kg resulted in a score of one point.[9]

The Thai version of the International Physical Activity Questionnaire-Short Form (IPAQ-SF) was used to evaluate PA level, as it demonstrated strong test–retest reliability (intraclass correlation coefficient, ICC=0.69).[12] The IPAQ-SF consists of questions about the amount of time spent engaging in PA during the preceding 7 days. The assessor administered the questionnaire by reading each item aloud to each participant individually, thereby ensuring uniform comprehension and minimizing potential misinterpretation. Participants were then prompted to provide their responses. If the total kcal per week fell in the lowest 20% (<383 kcal/week for males or <270 kcal/week for females), a score of one point was given.[9]

Exhaustion was assessed by participant self-report using structured questions administered by the assessor. Participants were asked whether they had experienced extreme physical fatigue or exhaustion that prevented them from participating in activities. Specifically, they were asked, "Do you feel so exhausted or extremely physically fatigued that you could not perform something afterward?" and "Do you feel that all activities are done with difficulty?" Responses were rated based on the frequency of these feelings during the past week using a 4-point Likert scale: 0 = rarely or none of the time (<1 day), 1 = some or a little of the time (1-2 days), 2 = a moderate amount of the time (3-4 days), and 3 = most of the time (>5 days).[13] Participants who responded with a score of 2 or 3 were classified as meeting the frailty criteria.[9]

Statistical analysis

All statistical analyses were conducted using STATA version 17. For continuous variables such as age, BMI, walking velocity, grip strength (kg), total PA, and frailty score, the mean and standard deviation (SD) were presented. Categorical variables, including sex, marital status, residential status, educational level, occupation before retirement, comorbidities, PA level, and physical frailty characteristics (slowness, weakness, low PA, weight loss, and exhaustion), were reported as percentages.

Based on data normality, one-way analysis of variance (ANOVA) was used to compare continuous variables between the pre-frailty, frailty, and robust groups**.** Categorical variables were compared between groups using a chi-square or Fisher's exact test. Univariate analyses were used to describe associations between frailty status and sociodemographic characteristics (age, gender, BMI, marital status, residential status, educational level, occupation, and comorbidity). Factors with $p<0.25$ were retained and included as covariates in the multivariable analysis.[14] Prior to performing multivariable logistic regression, age and occupation before retirement (labor, governor, owner, and housewife), with labor as the reference category, were identified as key variables for the pre-frailty model. For the frailty model, relevant factors included age, gender, BMI, education, and comorbidity (none = no disease, 1-2 diseases, and ≥3 diseases), with "none" as the reference category. Because the proportional odds assumption for the ordinal logistic regression model did not hold, a multinomial logistic regression analysis was conducted instead. This method, suitable for modeling categorical outcomes with more than two levels (robust, pre-frailty, and frailty), with the robust group as the reference category, allowed for the estimation of relative risk ratios (RRRs) with 95% confidence intervals (CI) for each outcome. The significance level was set at p-value < 0.05.

Results

Prevalence of pre-frailty and frailty

The prevalence of each frailty status, along with participants' demographic characteristics, is shown in Table 1. Of the participating Thai community-dwelling older adults, 19.0% were in the robust group, 38.4% were in the pre-frailty group, and 42.5% were in the frailty group. There were more females than males in all groups. The mean frailty score in the frailty group (3.69±0.78) was significantly higher than pre-frailty (1.24±0.43), as shown in Table 2.

Table 1. Participant demographics and characteristics.

Variables	Robust	Pre-frailty	Frailty	*p* value for difference
Number of participants (N, %)	51 (19.0%)	103 (38.4%)	114 (42.5%)	
Gender (N, %)				
Women	33 (64.7%)	61 (59.2%)	83 (72.8%)	
Men	18 (35.3%)	42 (40.8%)	31 (27.2%)	
Age (years), mean±SD	69.9±4.1	72.4±5.5	79.0±7.3	<0.001*[a]
BMI (kg/m^2), mean±SD	25.5±3.8	24.5±3.8	23.9±3.9	0.040*[a]
Marital status (N, %)				0.707 [b]
Single	4 (7.8%)	6 (5.8%)	10 (8.8%)	
Married/widow/divorce	47 (92.2%)	97 (94.2%)	104 (91.2%)	
Living status (N, %)				0.211 [b]
Alone	2 (3.9%)	13 (12.6%)	14 (12.3%)	
With relatives	49 (96.1%)	90 (87.4%)	100 (87.7%)	
Education level (N, %)				<0.001*[b]
Elementary or no degree	30 (58.8%)	64 (62.1%)	94 (82.5%)	
Secondary or higher degree	21 (41.2%)	39 (37.9%)	20 (17.5%)	
Occupation before retirement (N, %)				0.009*[b]
Labor	10 (19.6%)	44 (42.7%)	26 (22.8%)	
Governor/owner	32 (62.7%)	47 (45.6%)	71 (62.3%)	
Housewife	9 (17.7%)	12 (11.7%)	17 (14.9%)	
Comorbidity diseases (N, %)				0.017*[b]
0 = None	6 (11.8%)	8 (7.8%)	1 (0.9%)	
1-2 diseases	28 (54.9%)	45 (43.7%)	56 (49.1%)	
≥3 diseases	17 (33.3%)	50 (48.5%)	57 (50.0%)	

Note: *BMI: body mass index SD: standard deviation, *significant at p≤0.05, [a]ANOVA, [b]Chi-square test.*

Table 2. Physical pre-frailty / frailty and percentage of positive item in each group.

Variables	Robust	Pre-frailty	Frailty	*p*-value for difference
Frailty score, mean±SD	0	1.24±0.43	3.69±0.78	<0.001*[a]
Min-max	0	1-2	3-5	
Walking velocity, mean±SD	1.07±0.14	1.02±0.21	0.66±0.27	<0.001*[a]
Slowness (N, %)	0	7 (6.8%)	78 (68.4%)	<0.001*[b]
Hand grip strength, mean±SD	27.5±7.5	23.6±6.8	15.8±5.4	<0.001*[a]
Weakness (N, %)	0	44 (42.7%)	104 (91.2%)	<0.001*[b]
Total Physical activity, mean±SD	2187.4±2181.0	904.4±984.5	336.5±592.7	<0.001*[a]
Low Physical activity (N, %)	0	46 (44.7%)	93 (81.6%)	<0.001*[b]
Weight loss (N, %)	0	15 (14.6%)	53 (46.5%)	<0.001*[b]
Exhaustion (N, %)	0	16 (15.5%)	93 (81.6%)	<0.001*[b]

Note: *SD: standard deviation, MET.min/week: Metabolic Equivalent of Task in minutes per week, *significant at p≤0.05; [a]ANOVA; [b]Chi-square test.*

Sociodemographic characteristics among frailty statuses

Table 1 presents participant characteristics by frailty status. Frail individuals were the oldest group (mean age: 79.0 years), had the lowest BMI (23.9 kg/m²), and showed the highest proportions of low education (82.5%) and multiple comorbidities (≥3 diseases: 50.0%). Pre-frail individuals showed intermediate characteristics, with moderate age (72.4 years), BMI (24.5 kg/m²), and comorbidity burden, reflecting a transitional state. Robust individuals were younger (69.9 years), had a higher BMI (25.5 kg/m²), better education, fewer comorbidities, and were more likely to have held professional occupations.

Physical frailty characteristics

Table 2 presents physical frailty characteristics and the percentage of participants with positive items for each frailty status. Frail individuals had the highest frailty scores (mean=3.69), significantly slower walking speed (0.66 m/s), the weakest grip strength (15.8 kg), and the lowest PA level (336.5 MET.min/week), with most reporting slowness (68.4%), weakness (91.2%), exhaustion (81.6%), and weight loss (46.5%) (p<0.001). Pre-frail individuals showed intermediate deficits across all measures, including moderate walking speed (1.02 m/s), grip strength (23.6 kg), and PA (904.4 MET.min/week), with 42.7% reporting weakness and 44.7% reporting low activity. In contrast, robust individuals had no frailty symptoms, demonstrating the highest physical performance and PA levels.

Factors associated with pre-frailty and frailty

Age and occupation prior to retirement were identified as significant predictors in the pre-frailty model, whereas age, gender, BMI, educational attainment, and comorbidity were relevant factors in the frailty model. Multinomial logistic regression analysis revealed that each additional year of age was significantly associated with an increased risk of pre-frailty, with a RRR of 1.10 (95%CI [1.01, 1.18], p=0.020), when compared to the robust reference group (Table 3). Furthermore, certain occupations—specifically governor, business owner, or housewife—were significantly associated with a reduced risk of pre-frailty, by approximately 63.0% and 71.0%, respectively, relative to the labor group (RRR=0.37; 95%CI [0.15, 0.92], p=0.032, and RRR=0.29; 95%CI [0.09, 0.93], p=0.037).

For the frailty outcome, multinomial logistic regression analysis demonstrated that each additional year of age was significantly associated with a 1.28-fold increase in the risk of frailty compared to the robust reference group (RRR=1.28; 95%CI [1.18, 1.39], p<0.001). Males exhibited a significantly lower risk of frailty, with a 66.0% reduction compared to females (RRR=0.34; 95%CI [0.14, 0.85], p=0.021). In terms of BMI, a higher BMI was associated with a significantly decreased risk of frailty, with an RRR of 0.88 (95%CI [0.79, 0.98], p=0.019), indicating a 12.0% risk reduction relative to those with lower BMI. Similarly, participants with at least secondary education had a 77.0% lower risk of frailty compared to those with lower educational attainment (RRR=0.23; 95%CI [0.09, 0.58], p=0.002). Finally, individuals with comorbidities involving three or more chronic conditions showed a markedly increased risk of frailty—approximately 16.34 times higher than those without any chronic conditions (RRR=16.34; 95%CI [1.64, 162.46], p=0.017), as presented in Table 3.

Table 3. Associated factors related with pre-frailty and frailty by multinomial logistic regression.

Outcome	Variable	RRR	SE	z	*p-value*	95% CI	
Robust		(base outcome)					
Pre-frailty	Age	1.10	0.04	2.33	0.020*	1.01	1.18
	Gender						
	Female	Ref					
	Male	0.81	0.32	-0.53	0.594	0.37	1.76
	Body mass index	0.92	0.04	-1.64	0.101	0.84	1.02
	Education						
	Elementary or no degree	Ref					
	Secondary or higher degree	0.96	0.38	-0.1	0.922	0.44	2.09
	Occupation						
	Labor	Ref					
	Governor/owner	0.37	1.17	-2.14	0.032*	0.15	0.92
	Housewife	0.29	1.72	-2.09	0.037*	0.09	0.93
	Comorbidity						
	None	Ref					
	1-2 diseases	1.55	0.98	0.7	0.487	0.47	5.14
	≥3 diseases	2.50	1.68	1.36	0.173	0.67	9.30
	Constant	0.05	1.09	-1.08	0.279	0.00	94.98
Frailty	Age	1.28	0.54	5.92	0.000*	1.18	1.39
	Gender						
	Female	Ref					
	Male	0.34	1.58	-2.32	0.021*	0.14	0.85
	Body mass index	0.88	0.05	-2.35	0.019*	0.79	0.98
	Education						
	Elementary or no degree	Ref					
	Secondary or higher degree	0.23	1.08	-3.1	0.002*	0.09	0.58
	Occupation						
	Labor	Ref					
	Governor/owner	1.42	0.76	0.65	0.514	0.50	4.05
	Housewife	0.75	1.50	-0.43	0.668	0.21	2.76
	Comorbidity						
	None	Ref					
	1-2 diseases	9.02	10.32	1.92	0.055	0.96	84.99
	≥3 diseases	16.34	19.15	2.38	0.017*	1.64	162.46
	Constant	0.00	0.00	-4.35	0.000	0.00	0.00

Note: *RRR: relative risk ratio, SE: standard errors, z: z-statistics, 95%CI: 95%confidence interval, *significant at p<0.05, Ref: reference.*

Discussion

This study examined pre-frailty and frailty prevalence and their links to sociodemographic factors among community-dwelling older adults aged 65 years and above. Prevalence rates were 38.4% for pre-frailty and 42.5% for frailty. Pre-frailty was associated with age and preretirement occupation, while frailty correlated with age, gender, BMI, comorbidity, and education.

The pre-frailty prevalence in this study (38.4%) was lower than in prior Thai studies (50.9%-57.4%)[7,8] and slightly below international findings (43.1%-49.5%).[3-6] In contrast, frailty prevalence (42.5%) was significantly higher than reported in both Thai[7,8] and global reports (6.3%-24.0%).[3-6] These discrepancies may partly stem from differences in participant age, as previous Thai studies included individuals aged ≥60 years[7,8] while the referenced review included individuals aged 50 and higher.[3] In contrast, this study included only adults aged ≥65 to align with the Fried frailty phenotype[9] and prior research.[3,6-8] This higher age threshold likely excluded younger pre-frail individuals under age 65, in whom pre-frailty is often first detected.[15] Given that frailty increases with age,[16] the higher minimum participant age in this study likely contributed to a greater proportion of older adults who had already developed frailty, leading to an increased overall prevalence.

This study found high rates of low PA (44.7%) and muscle weakness (42.7%) among pre-frail individuals. Frail participants exhibited nearly double these rates,

with 91.2% muscle weakness, 81.6% low PA, and 81.6% exhaustion. The current findings align with Shinohara et al., who identified muscle weakness as a precursor to frailty,[17] and Xue *et al*.'s framework, which suggested that behavioral changes emerge as physiological reserves diminish and the ability to meet environmental demands declines prior to the onset of frailty.[18] Physical inactivity accelerates the loss of type II muscle fibers, worsening weakness and musculoskeletal deterioration, which together drive frailty progression.[18] Low PA and muscle weakness were the most common components of pre-frailty, making them key indicators. Promoting regular PA and strengthening exercises is essential to maintain independence and may slow down the progression from pre-frailty to frailty, emphasizing the need for timely preventive interventions.

The current study found that 14.6% of pre-frail individuals experienced weight loss, reflecting age-related changes observed in previous research.[19] In addition, exhaustion was significantly less common in pre-frail individuals (15.5%) compared to frail individuals (81.6%). Age-related muscle loss and inactivity likely contribute to fatigue and slower gait speed, leading to reduced physical performance in everyday activities.[18] These effects likely reflect more comorbidities, lower PA, and greater fatigue in the frail group. Age-related sarcopenia further impairs mobility, fueling a cycle of physical inactivity and muscular deterioration.[9] The link between aging, muscle loss, and inactivity highlights the need for early detection and timely intervention to halt the progression from robustness to frailty.

In contrast, although slowness was more prevalent among the frailty group, it remained one of the less common frailty components. Previous studies have shown that slowness strongly predicts longevity[20] and future disability[5] in older adults. However, as a cross-sectional study, this research identified factors associated with frailty but was unable to predict longevity or disability. Table 2 shows that slowness was the least common characteristic in pre-frailty and the second least common in frailty. Previous studies have reported that a gait speed less than 0.8 meters per second in those ≥75 increases frailty risk.[21] Based on our findings, promoting PA and increasing walking speed should be prioritized for individuals with pre-frailty, as physical inactivity and slow gait accelerate the loss of fast-twitch muscle fibers and contribute to muscle weakness. Therefore, the primary focus for individuals with frailty should be active muscle contractions and strengthening exercises.

The multivariate-adjusted model showed age and preretirement occupation were significantly linked to pre-frailty. Consistent with prior studies, frailty increased with age.[18] Occupations like governor, business owner, or housewife had a 63% to 71% lower pre-frailty risk compared to labor-intensive jobs. This risk disparity may stem from the physical demands of manual labor, including heavy lifting, prolonged standing, and repetitive tasks, which increase musculoskeletal disorders and pre-frailty risk.[22] One study found those with musculoskeletal complaints in three to five bodily regions were 3.77 times more likely to be pre-frail or frail.[22] Laborers may also be more vulnerable due to limited access to preventive care and lower participation in health-promoting behaviors.[23] The cumulative physical demands of labor-intensive work may accelerate the decline of physiological reserves, reducing the body's ability to recover from illness or injury and contributing to the onset of pre-frailty. This evidence highlights that frail individuals face greater socioeconomic and occupational disadvantages. One study reported that frail individuals had higher rates of economic hardship, outdoor work, longer workweeks, and extended daily hours,[24] highlighting the need to consider occupational and socioeconomic factors when assessing frailty risk.

The multivariate-adjusted model identified comorbidity, age, BMI, gender, and education as key frailty factors. Individuals with ≥3 chronic conditions had a significantly higher frailty risk (RRR=16.34) than those without.[25] This aligns with evidence that age-related physiological decline and multiple comorbidities compound to accelerate frailty onset.[25] Furthermore, participants in this study with ≥3 comorbidities had a 16.34-fold higher frailty risk than those without, consistent with a higher comorbidity rate in the frailty group. Conditions like hypertension, diabetes, cardiovascular disease, and osteoarthritis impair mobility and increase fatigue, contributing to frailty symptoms. Multiple chronic conditions were linked to increased inactivity, exhaustion, and weakness, indicating that comorbidity reflects health burden and accelerates the progression from robustness to pre-frailty and frailty.[6,26]

Frailty was more prevalent among women, partly due to the hormonal changes that occur after menopause. These hormonal shifts contribute to muscle degeneration, sarcopenia, and age-related muscle loss, leading to reduced physical function.[27] Estrogen decline associated with menopause exacerbates sarcopenia, which, when combined with low activity, fatigue, and weight loss, accelerates the development of frailty.[28,29]

Educational level significantly affects frailty prevalence.[4,7,8,30] Lower education is linked to physical, social, and environmental disadvantages, reduced health literacy, and difficulty managing medical care.[31-33] It also correlates with socioeconomic challenges, such as low income, poor housing, limited access to nutrition, physical inactivity, and delayed treatment, all of which increase the risk of chronic disease.[26]

This study highlights the clinical importance of identifying key factors associated with pre-frailty, with an emphasis on age and preretirement occupation. While age is a well-known risk factor, the impact of occupational history represents a significant finding, emphasizing its role in screening and prevention measures. These insights support the integration of pre-frailty screening into geriatric care, particularly for community-dwelling and socioeconomically disadvantaged older adults. Assessing work history can guide targeted follow-up and educational interventions during the preretirement phase. The study also found that low PA was the most common frailty component among pre-frail participants, underscoring its

role as a key modifiable risk factor. Promoting regular PA is essential to preserve functional independence and reduce the risk of progression to frailty, emphasizing the need for early intervention. A key strength of this study is the use of the validated Fried frailty phenotype.[9]

Limitation

Several limitations should be acknowledged. First, the study's cross-sectional design limits causal inference, and the sample drawn from older adults in Pathum Thani may reduce generalizability. Future research involving a larger and more diverse cohort study is recommended to further clarify frailty determinants and guide targeted interventions.

Conclusion

Among community-dwelling adults aged ≥65, pre-frailty and frailty prevalence were 38.4% and 42.5%, respectively. Pre-frailty was linked to age and labor-intensive work, while frailty was associated with age, gender, low BMI, low education, and comorbidity. These findings highlight the need for early detection and targeted interventions to prevent frailty progression.

Ethical approval

This study was approved by the Research Ethics Committee of Chulalongkorn University, Thailand (COA No. 089/2020) and conducted according to the Declaration of Helsinki. All the participants provided written informed consent before starting data collection.

Funding

This study was supported by the 90th Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund; Grant numbers GCUGR1125632113D-no113).

Conflict of interest

The authors reveal no actual or potential conflicts of interest relevant to the content of this article.

CRedit authorship contribution statement

Sukwida Manorangsan: conceptualization, methodology, formal analysis, investigation, writing-original draft preparation, writing-review and editing, funding acquisition; **Pawan Chaiparinya**: writing-original draft preparation, writing-review and editing; **Sirima Mongkolsomlit**: formal analysis, writing-review and editing; **Sujitra Boonyong**: conceptualization, methodology, investigation, writing-original draft preparation, writing-review and editing, funding acquisition.

Acknowledgements

We would like to thank all Thai community older adults who participated in this study.

References

[1] Sezgin D, Liew A, O'Donovan MR, O'Caoimh R. Pre-frailty as a multi-dimensional construct: a systematic review of definitions in the scientific literature. Geriatr Nurs. 2020; 41(2): 139-46. doi: 10.1016/j.gerinurse.2019.08.004.

[2] Fernández-Garrido J, Ruiz-Ros V, Buigues C, Navarro-Martinez R, Cauli O. Clinical features of prefrail older individuals and emerging peripheral biomarkers: a systematic review. Arch Gerontol Geriatr. 2014; 59(1): 7-17. doi: 10.1016/j.archger.2014.02.008.

[3] O'Caoimh R, Sezgin D, O'Donovan MR, Molloy DW, Clegg A, Rockwood K, et al. Prevalence of frailty in 62 countries across the world: a systematic review and meta-analysis of population-level studies. Age Ageing. 2021; 50(1): 96-104. doi: 10.1093/ageing/afaa219.

[4] Zeng XZ, Meng LB, Li YY, Jia N, Shi J, Zhang C, et al. Prevalence and factors associated with frailty and pre-frailty in the older adults in China: a national cross-sectional study. Front Public Health. 2023; 11: 1110648. Published 2023 Jul 24. doi: 10.3389/fpubh.2023.1110648.

[5] Shimada H, Makizako H, Doi T, Tsutsumimoto K, Suzuki T. Incidence of disability in frail older persons with or without slow walking speed. J Am Med Dir Assoc. 2015; 16(8): 690-6. doi: 10.1016/j.jamda.2015.03.019.

[6] Rodríguez-Laso Á, García-García FJ, Rodríguez-Mañas L. Transitions between frailty states and its predictors in a cohort of community-dwelling Spaniards. J Am Med Dir Assoc. 2022; 23(3): 524.e1-e11. doi: 10.1016/j.jamda.2021.07.008.

[7] Wongtrakulruang P, Muangpaisan W, Panpradup B, Tawatwattananun A, Siribamrungwong M, Tomongkon S. The prevalence of cognitive frailty and pre-frailty among older people in Bangkok metropolitan area: a multicenter study of hospital-based outpatient clinics. J Frailty Sarcopenia Falls. 2020; 5(3): 62-71. doi: 10.22540/JFSF-05-062.

[8] Thinuan P, Siviroj P, Lerttrakarnnon P, Lorga T. Prevalence and potential predictors of frailty among community-dwelling older persons in Northern Thailand: a cross-sectional study. Int J Environ Res Public Health. 2020; 17(11): 4077. doi:10.3390/ijerph17114077.

[9] Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2001; 56(3): M146-56. doi: 10.1093/gerona/56.3.m146.

[10] O'Caoimh R, Galluzzo L, Rodríguez-Laso Á, Van der Heyden J, Ranhoff AH, Carcaillon-Bentata L, et al. Transitions and trajectories in frailty states over time: a systematic review of the European Joint Action ADVANTAGE. Ann Ist Super Sanita. 2018; 54(3): 246-52. doi: 10.4415/ANN_18_03_12.

[11] Sergi G, Veronese N, Fontana L, De Rui M, Bolzetta F, Zambon S, et al. Pre-frailty and risk of cardiovascular disease in elderly men and women: the Pro.V.A.

study. J Am Coll Cardiol. 2015; 65(10): 976-83. doi: 10.1016/j.jacc.2014.12.040.

[12] Rattanawiwatpong P, Khunphasee A, Pongurgsorn C, Intarakamhang P. Validity and reliability of the Thai version of short format international physical activity questionnaire (IPAQ). J Thai Rehabil. 2006; 16: 147-60.

[13] Orme JG, Reis J, Herz EJ. Factorial and discriminant validity of the center for epidemiological studies depression (CES-D) scale. J Clin Psychol. 1986; 42: 28-33. doi: 10.1002/1097-4679(198601)42:1<28::aid-jclp2270420104>3.0.co;2-t.

[14] Zhang Z. Model building strategy for logistic regression: purposeful selection. Ann Transl Med. 2016; 4(6): 111. doi: 10.21037/atm.2016.02.15.

[15] Gordon SJ, Baker N, Kidd M, Maeder A, Grimmer KA. Pre-frailty factors in community-dwelling 40-75 year olds: opportunities for successful ageing. BMC Geriatr. 2020; 20(1): 96. doi: 10.1186/s12877-020-1490-7.

[16] Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of frailty in community-dwelling older persons: a systematic review. J Am Geriatr Soc. 2012; 60(8): 1487-92. doi: 10.1111/j.1532-5415.2012.04054.x.

[17] Shinohara T, Saida K, Tanaka S, Murayama A, Higuchi D. Transition to frailty in older Japanese people during the coronavirus disease 2019 pandemic: a prospective cohort study. Arch Gerontol Geriatr. 2022; 98: 104562. doi: 10.1016/j.archger.2021.104562.

[18] Xue QL. The frailty syndrome: definition and natural history. Clin Geriatr Med. 2011; 27(1): 1-15. doi: 10.1016/j.cger.2010.08.009.

[19] Morley JE. Frailty and sarcopenia in elderly. Wien Klin Wochenschr. 2016; 128(Suppl 7): 439-45. doi: 10.1007/s00508-016-1087-5.

[20] Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, et al. Gait speed and survival in older adults. JAMA. 2011; 305(1): 50-8. doi: 10.1001/jama.2010.1923.

[21] Castell MV, Sánchez M, Julián R, Queipo R, Martín S, Otero Á. Frailty prevalence and slow walking speed in persons age 65 and older: implications for primary care. BMC Fam Pract. 2013; 14: 86. doi: 10.1186/1471-2296-14-86.

[22] Norheim KL, Bøggild H, Andersen JH, Omland Ø, Bønløkke JH, Madeleine P. Retirement status and frailty: a cross-sectional study of the phenotype of manual workers aged 50-70 years. Eur J Public Health. 2021; 31(1): 116-21. doi: 10.1093/eurpub/ckaa188.

[23] Yoshizawa Y, Tanaka T, Takahashi K, Fujisaki-Sueda-Sakai M, Son BK, Iijima K. Impact of health literacy on the progression of frailty after 4 years among community-dwelling older adults. Int J Environ Res Public Health. 2021; 19(1): 394. doi: 10.3390/ijerph19010394.

[24] Matsugaki R, Fujino Y, Zaitsu M, Saeki S, Matsuda S, Ogami A. Frailty is a risk factor for occupational falls among older workers: an internet-based prospective cohort study. J Occup Health. 2024; 66(1): uiae065. doi: 10.1093/joccuh/uiae065.

[25] Ferrucci L, Walston JD. Frailty. In: Halter JB, Ouslander JG, Studenski S, High KP, Asthana S, Supiano MA, et al, Editors. Hazzard's geriatric medicine and gerontology. 8th Ed. New York: McGraw-Hill; 2022; Chapter 42.

[26] Nejatinamini S, Campbell DJT, Godley J, Minaker LM, Sajobi TT, McCormack GR, et al. The contribution of modifiable risk factors to socioeconomic inequities in cardiovascular disease morbidity and mortality: a nationally representative population-based cohort study. Prev Med. 2023; 171: 107497. doi:10.1016/j.ypmed.2023.107497

[27] Khadilkar SS. Musculoskeletal disorders and menopause. J Obstet Gynaecol India. 2019; 69(2): 99-103. doi: 10.1007/s13224-019-01213-7.

[28] Greco EA, Pietschmann P, Migliaccio S. Osteoporosis and sarcopenia increase frailty syndrome in the elderly. Front Endocrinol (Lausanne). 2019; 10: 255. doi: 10.3389/fendo.2019.00255.

[29] Blumer J, Le B. Rethinking sarcopenia and frailty of the elderly. Post Reprod Health. 2025; 2053369 1251337173. doi: 10.1177/20533691251337173.

[30] Singhal S, Singh S, Dewangan GC, Dey S, Banerjee J, Lee J, et al. The prevalence of frailty and its relationship with sociodemographic factors, regional healthcare disparities, and healthcare utilization in the aging population across India. Aging Med (Milton). 2023; 6(3): 212-21. doi:10.1002/agm2.12263.

[31] Eyigor S, Kutsal YG, Duran E, Huner B, Paker N, Durmus B, et al. Frailty prevalence and related factors in the older adult-FrailTURK Project. Age (Dordr). 2015; 37(3): 9791. doi: 10.1007/s11357-015-9791-z.

[32] Lee YS, Chen YM, Chou YY, Lin CS, Chen JP, Kuo FH, et al. Prefrailty subtypes differentially predict 5-year mortality in the functionally independent geriatric population. Arch Gerontol Geriatr. 2023; 106: 104897. doi: 10.1016/j.archger.2022.104897.

[33] Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med. 2011; 155(2): 97-107. doi: 10.7326/0003-4819-155-2-201107190-00005.

Journal of Associated Medical Sciences 2025; 58(3): 307-312

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Frequencies of major blood group antigens and phenotypes among blood donors in Al-Jouf, Saudi Arabia.

Sanaa Hussein[1*], Maram Alruwili[1], Dina Alruwaili[1], Layali Alruwaili[1], Wasan Alblaihed[1], Joury Alsarrah[1], Abozer Elderdery[1], Motea Eskandar[2], Kamal Alruwaili[2], Hamoud Alruwaili[2], Fawaz Alenazy[1], Albadawi Talha[1], Fadia Alruwaili[2]

[1]Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Jouf University, Saudi Arabia.
[2]Central Blood Bank - Al-Jouf Regional Lab, Saudi Arabia.

ARTICLE INFO

Article history:
Received 29 April 2025
Accepted as revised 12 July 2025
Available online 20 August 2025

Keywords:
ABO, Rh, Kell, phenotypes, donors.

ABSTRACT

Background: Blood group antigens are a representation of some personal traits.

Objectives: To determine the frequency of ABO, Rhesus and Kell blood group phenotypes among blood donors in the Al-Jouf region, Saudi Arabia.

Materials and methods: 980 donors were included in retrospective study design. Data was collected from blood donor records from July 2022 to January 2023 at Al-Jouf Central Blood Bank. ABO, Rhesus phenotypes (D, C, E, e, c), and K antigen have been detected in the blood bank using the fully automated Bio-Rad IH-500.

Results: Regarding the ABO system, the O phenotype was the most common among the donors (approximately 45.7%), and followed by A (25.4%) and B (24.9%), while the AB group had the lowest number (4%). Analysis of Rhesus blood group phenotypes demonstrated that 89.3% of donors were Rh D positive and 10.7% were Rh D negative. Other Rhesus antigens were found in variable percentages; the e antigen was the most prevalent among donors (99.2%),followed by the c and C antigens (72.7% and 69.2%, respectively), and the E antigen was the least prevalent (22.8%). Only 0.3% of donors were positive for Cw. Accordingly, the DCce phenotype (24.9%) was the most common Rh phenotype among donors, followed by DCe and Dce (22.8% and 20.2%,respectively). CE and CEe phenotypes had the lowest prevalence among donors(0.1%). Out of 980 donors, 178 donor (18.2%) had the Kell antigen (K).

Conclusion: The most prevalent blood group among donors was O Rh D+ while the most prevalent Rhesus antigen was the e antigen. The most common Rhesus phenotype among donors was DCce. Reporting Rh phenotypes and Kell antigen (K) is extremely important from a clinical standpoint, especially during pregnancy and blood transfusions as it can improve maternal and fetal care and enhance safer blood transfusion.

Introduction

So far, 346 Red Blood Cell (RBC) antigens (Ags) have been found across 36 blood group systems. In the transfusion and transplant fields, the ABO blood group is still the most important, followed by Rhesus (Rh) and Kell.[1] According to reports, the frequency of D antigen in Saudi Arabia ranges from 84% to 91%.[2-5]

Approximately fifty distinct Rh group antigens have been found to date, with D, C, c, E, and e being the most clinically significant, creating IgG antibodies when transfusing incompatible blood or following abortion, and causing hemolytic disease of the newborn (HDN).[6,7] This happens in varied degrees of severity, with D Ag being the most immunogenic compared to antigens from Rhesus and other systems.[88] Similarly, Oladeinde and

* **Corresponding contributor.**
Author's Address: *Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Jouf University, Saudi Arabia.*
E-mail address: selfatih@ju.edu.sa
doi: 10.12982/JAMS.2025.101
E-ISSN: 2539-6056

colleagues found that D Ag accounts for half of all cases of maternal alloimmunization with higher immunogenicity,[9] followed by c, E, C, and e antigens, and that all of these antigens mediate hemolytic transfusion reactions (HTR), particularly delayed reactions in addition to HDN. [10, 11] RhCE and RhD encode eight distinct haplotypes of the Rh system, denoted as Dce, dce, DCe, dCe, DcE, dcE, DCE, and dCE, or in shorter nomenclature: R0, r, R1, r', R2, r'', Rz, and ry. [4,12]

The Kell blood group system is regarded clinically significant after the ABO and Rh systems in transfusion medicine due to its immunogenic features. This system comprises numerous antigens (Ags), with the most common being K and k (previously known as Kell and Cellano, respectively), Kp^a, and Kp^b, all of which have immunogenic features. In most ethnicities, the k antigen is more frequent than the K antigen; 98% of Blacks and 91% of Caucasians have the K-k+ phenotype.[10,11,13] The absence of any of these antigens may elicit an immunological response following blood transfusion or abortion, resulting in the development of alloantibodies that cause hemolytic transfusion responses (HTR) and hemolytic disease of the newborn (HDN).[14]

The first Kell system antibodies found were anti-K1 (Kelleher) and anti-K2 (Cellano).[15] The clinically most relevant antigen is K (KEL1). Anti-K antibodies are second to anti-D antibodies the most frequently encountered antibodies in severe HDFN. K0 (the null phenotype), which was identified alongside Kp^a, is one of the other antigens of the Kell system. While Kp^b (K4) has high frequency mutations, Kp^a (K3) was shown to have low frequency mutations.[16]

The prevalence of the k antigen in Saudi Arabia has been reported to be greater than 99%.[14,17,18] The heterozygous form (K, k) was found to be around 17% in the Riyadh region, which is significantly varied than its prevalence in other ethnic groups,[19] such as Indians (3.5%),[20] Africans (24%),[21] and Caucasians (9%).[22]

Most cases of abnormal hemoglobin (Hb), particularly sickle and thalassemic individuals, were detected that 12-18% with alloimmunization due to blood type incompatibility.[23-25] Compared to the RhD antigen, the K antigen is approximately 8-10 times less immunogenic.[26]

In general, the incidence of red cell alloimmunization due to blood group incompatibility ranges between 2% and 10%, but it can reach 60% in specific groups of multi-transfused individuals. Phenotyping of the ABO, Rhesus, and Kell systems can decrease or prevent probable subsequent hemolytic transfusion reactions.[27]

Accurate phenotyping can improve maternal and fetal care, patient outcomes and enhance safer blood transfusion practices. Identifying the most frequent blood phenotypes aids to understand genetic variation in the region and provide prediction of allele distribution within a community.

Our study aimed to determine the frequency of red blood cell antigens and phenotypes of ABO, Rh, and Kell systems, among donors in Sakaka city, Al-Jouf, Saudi Arabia.

Materials and methods

A retrospective analysis was undertaken on blood donors at the Al-Jouf Central Blood Bank in Saudi Arabia. A total of 980 records of donors who had matched the criteria of donation and had been accepted by the blood bank were included in the study excluded repeated donors. Concerning demographic and laboratory information of ABO, Rh, and Kell system, data from July 2022 to January 2023 was collected from the blood bank electronic database.

Fully automated system technique:Bio-Rad IH-500 analyzer has been used in the blood bank to detect blood group antigens. The ID-Card DiaClon ABO/D+ Reverse Grouping (Bio-Rad,Switzerland) was used to perform both forward and reverse ABO grouping as well as Rh D. The presence or absence of the C (RH2), Cw (RH8), c (RH4), E (RH3), e (RH5), and K (KEL1) antigens was examined using the ID-Card DiaClon Rh- Subgroups+ Cw+ K.

The findings of blood groups, including ABO, Rhesus phenotypes (D, C, E, e, c), and Kell antigen (K) were analyzed using the Statistical Package for Social Sciences version 23 to find descriptive statistics as frequencies for blood groups and means±SD for age and weight.

Results

Donor demographics are listed in Table 1. Near to half of the donors were O blood group while the lowest percentage of donors was found in the AB blood group Table 2.

As shown in Table 3, the D-antigen, which is represented by 875 donors (89.3%), was more common in donors of the O blood type and less common in donors of the A, B, and AB blood groups. According to Table 4, surveillance of the major Rhesus antigens revealed that antigen e was the most prevalent (99.2%), followed by D (89.3%), C (69.2%), c (72.7%), and E (22.8%). The Kell antigen (K) was present in 178 donors (18.2%), while it was not present in 81.8% of donors. The DCce was the most common Rh phenotype (24.9%) followed by DCe (22.8%) and Dce (20.2%) at shown in Table 5. Comparison of the Rhesus antigen with other regions in Saudi Arabia shows nearly similar distribution of Rhesus antigens in Saudi Arabia except for limited regions (Table 6).

Table 1. Characteristics of donors age and weight.

	N	Minimum	Maximum	Mean (±SD)
Age (years)	980	18.00	61.00	33.4 (9.3)
Weight (kg)	980	46.00	153.00	86.7(16.1)

Table 2. Distribution of ABO Blood group among donors (N=980).

ABO	Frequency (%)
A	249 (25.4)
B	244 (24.9)
AB	39 (4)
O	448 (45.7)
Total	980

Table 3. ABO blood group and D antigen coexistence among donors.

ABO Blood group	Rhesus D antigen		Total
	Negative	Positive	
A	29	220	249
B	24	220	244
AB	2	37	39
O	50	398	448
Total	105	875	980

Table 4. Expression of e, D, c, E, C, KELL-antigen expression in donors (N=980).

Antigen	Reaction	Frequency	%
e	neg	8	0.80
	pos	972	99.2
D	neg	105	10.7
	pos	875	89.3
c	neg	268	27.3
	pos	712	72.7
E	neg	757	77.2
	pos	223	22.8
C	neg	302	30.8
	pos	678	69.2
K	neg	802	81.8
	pos	178	18.2

Note: neg: abscence of Ag, pos: presence of Ag.

Table 5. The percentage of Rh phenotypes and possible genotypes**.**

Phenotype	Wiener	Fisher-Race	%
DCce	R1r	DCe/ce	24.9
DCe	*R 1R1, R1r-*	*DCe/DCe*	22.8
Dce	R0r	Dce/Dce	20.2
	R0R0	Dce/ce	
DCEce	Rzr	DCE/ce	11.5
ce	rr	ce/ce	4.9
DcEe	R2r	DcE/ce	4.6
Cce	r- r	Ce/ce	2.35
DCE	RzRz	DCE/DCE	2.35
Ce	r-r-	Ce/Ce	2.14
DcE	R2R2	DcE/DcE	1.7
CcEe	r-r--	Ce/cE	1.12

Table 5. The percentage of Rh phenotypes and possible genotypes. (Continued)

Phenotype	Wiener	Fisher-Race	%
DCEe	RzR1	*Rzr-, R1ry*	1.02
cEe	r--r	cE/ce	0.41
CcED	RzR2	*Rzr0, R2ry*	0.2
CE	rYrY	CE/CE	0.1
CEe	ryr-	Ce/CE	0.1

Table 6. The percentage of D, C, E, c, and e antigens in this study compared to other regions of Saudi Arabia and different ethanities.

Region/Area	Antigens (%)				
	C	E	c	e	D
Our study in Jouf	69.2	22.8	72.7	99.2	89.3
Jazan Province[7]	70.97	18.91	75.38	97.95	93.3
Women in Hail Province[36]	76.4	29.4	73.0	97.0	93.4
Eastern Region of Saudi Arabia[17]	59	21	86	97	80
Rhyadh[37]	62.3	23.5	74.3	95	84.8
Taif City[2]	62.3	31.3	81.7	95.8	87.8
In European[15]	68	29	80	98	85
Caucasian[16]	68	29	80	98	85
Blacks[21]	27	22	96	98	92
Asian[20]	93	39	47	96	99

Discussion

In terms of demographics, all members of our study population were men. Due to a number of social and cultural variables, men make up most blood donors. A comparable study conducted in central Saudi Arabia found that 82.98% of donors were men.[28]

According to our findings, the most prevalent blood group is O (45.7%), while the least common is AB (4.0%). This is consistent with a 2011 study conducted in Al-Jouf, which found that O was 39.8%, A was 27.5%, B was 25.5%, and AB was 7.3%.[29]

Additionally, our study concurred with other studies carried out in different Saudi Arabian districts.

O-RhD+ > A-RhD+ > B-RhD+ > O-RhD- > AB-RhD+ > A-RhD- > B-RhD- > AB-RhD- was the blood type frequency order in Asir.[30] The most prevalent blood group in Qassim was O+ (42%), which was followed by A+ (23.4%), B+ (20.9%), O- (5.45%), AB+ (3.4%), A- (2.8%), B- (2.1%), and AB- (0.5%).[7] The outcomes in the province of Jazan were quite similar in which A accounted for 29.44%, B for 10.44%, AB for 1.15 percent, and O for 58.97 percent.[31] Only 7.2% of ABO groups in southwestern Saudi Arabia were Rh-negative, with O accounting for 56.8%, A for 33.4%, B for 6%, and AB for 3.8%.[32] The frequency of the Rh antigens, D, C, E, c, and e is consistent with many studies in Saudi Arabia and contrasted to few ethnics worldwide (Table 6). The Rh haplotype DCe was most prevalent in Caucasians (42%), followed by Asians (70%), and Native Americans (44%). Most Black people (44%) have the Rh haplotype Dce.[33]

According to our research, DCce (24.3%) is the most common Rh phenotype. DCe was represented by about 22.8%, and Dce by 20.8%. CEe and CE phenotypes had the lowest occurrence among donors, with the latter being 0.102%. ce was the most common among Rhesus negative donors, accounting for 4.9% of all cases. Furthermore, these results are consistent with a study done in Hail found that, on average, DCcee was the most prevalent phenotype in the community (32.2%), followed by DCCee (26.6%), while Ccddee (0.8%) and CCDEe (0.2%) were the least prevalent. According to similar screening conducted at King Abdulaziz Medical City in Riyadh, DCcee was the most common Rh phenotype observed (28.7%), DCCee (24.5%), and dccee (13.7%) were the next most common Rh phenotypes, and other phenotypes were DccEe (11.5%), Dccee (8.0%), DCcEe (6.8%), DccEE (4.0%), DCCEE (0.5%) and DEE (0.5%).[35] Also, another study in Taif City found that DCce was the most prevalent phenotype (29.5%).[2] The incidence of Rh phenotypes was in line with a study in Riyadh that indicated the most prevalent phenotypes were R1r (CcDee) (31%) and R1R1 (CCDee) (22%).[35] Accordingly, it seems that the Rhesus phenotypes distribution is approximately similar in different regions of Saudi Arabia. Regarding the Cw phenotype, our study discovered that just 0.3% of donors were Cw positive, whereas 99.7% had a negative Cw phenotype.

In the current study, the frequentcy of KEL1 (K) antigen expression 178 from total donors 980 (18.2%) and this is consistent with other studies in Saudi Arabia as in Riyadh (14%),[36] Taif city (22.1%) [2] and 18.2% among

Saudi Donors at King Abdulaziz Medical City-Riyadh,[37] but this percentage decreased in other regions of the country such as the prevalence of K antigen is found to be (12.3%) among blood donors in Jeddah city[38] and (8%) in the eastern region of Saudi Arabia.[17] This modest variance between regions in Saudi Arabia could be attributed to ethnic diversity, consanguinity, or migratory patterns.

These findings provide prediction of allelic distribution in al-Jouf, Saudi Arabia.

Limitations

The study depends on records of the routine blood group testing so that phenotyping of Kell blood group was restricted to (K) antigen and other antigens not included in the study.

Conclusion

According to our findings, O-RhD+ blood types are the most common among donors in Al-Jouf, Saudi Arabia whereas AB blood types are the least common. The e antigen is the most common Rhesus antigen and the most common phenotype among donors was DCce.

Ethical approval

This study was approved by the Institutional Research Ethics Committee, Qurayyat Health Affairs in Aljouf. (Approval number/protocol number_2023-53.

Funding

This research received no external funding.

Conflicts of interest

The authors declare there are no conflicts of interest.

CRediT authorship contribution statement

Sanaa Hussein: research design and manuscript preparation; **Maram Alruwili, Dina Alruwaili, Layali Alruwaili, Wasan Alblaihed, Joury Alsarrah:** data collection, data analysis; **Motea Eskandar, Kamal Alruwaili, Hamoud Alruwaili, Fadia ALruwaili:** laboratory analysis, analysis coordinating and data reviewing; **Abozer Elderdery:** main role in editing the manuscript; **Fawaz Alenazy, Albadawi Talha:** manuscript writing and layout of results reviewing. Final version of the manuscript was reviewed and approved by All authors.

Acknowledgements

The authors would like to thank Jouf University and the colleagues at Al-Jouf Regional Lab and Central Blood Bank.

References

[1] Rahadi I, Choodoung M, Choodoung A. Studies on red blood cells (rbcs) and white blood cells (WBCS) detection by image processing. Newest Phys Sci Res Updates. 2021; 7(2): 50-8. doi.org/10.9734/bpi/nupsr/v7/8996D.

[2] Felimban R, Al-Ghamdi A, Elmissbah T, Almalki A, Alzoghaibi N, Hakami N, et al. ABO, Rh, and Kell blood group antigen frequencies in blood donors of Taif City, Saudi Arabia. Clin Lab. 2023; 69(7). doi: 10.7754/Clin.Lab.2023.230104.

[3] Qanash H, Alcantara JC, Alshammari AM, Aljuhani A, Binsaleh NK, Almashjary MN, et al. The frequencies of ABO and Rh phenotypes among male blood donors in northwestern Saudi Arabia. Clin Lab. 2022; 68(11). doi: 10.7754/Clin.Lab.2022.220123.

[4] Altayar MA, Jalal MM, Kabrah A, Qashqari FS, Jalal NA, Faidah H, et al. Prevalence and association of transfusion transmitted infections with ABO and Rh blood groups among blood donors in the western region of Saudi Arabia: A 7-year retrospective analysis. Medicina. 2022;58(7): 857. doi: 10.3390/medicina58070857.

[5] Hamali HA, Madkhali MM, Dobie G, Madkhali AM, Madkhali B, Hummadi Y, et al. Prevalence of Rh and K phenotypes among blood donors from different ethnicities in Samtah (Southwestern Region) Saudi Arabia. Int J Immunogenet. 2022; 49(3): 202-8. doi: 10.1111/iji.12577.

[6] Halawani AJ, Arjan AH. ABO, RH, and KEL1 antigens, phenotypes and haplotypes in Southwestern Saudi Arabia. Clin Lab. 2021; 67(2): 344-8. doi: 10.7754/Clin.Lab.2020.200633.

[7] Alabdulmonem W, Shariq A, Alqossayir F, AbaAlkhail FM, Al-Musallam AY, Alzaaqi FO, et al. Sero-prevalence ABO and Rh blood groups and their associated transfusion-transmissible infections among blood donors in the Central Region of Saudi Arabia. J Infect Public Health. 2020; 13(2): 299-305. doi: 10.1016/j.jiph.2019.12.004.

[8] Selleng K, Jenichen G, Denker K, Selleng S, Müllejans B, Greinacher A. Emergency transfusion of patients with unknown blood type with blood group O Rhesus D positive red blood cell concentrates: a prospective, single-centre, observational study. Lancet Haematol. 2017; 4(5): e218-24. doi: 10.1016/S2352-3026(17)30051-0.

[9] Oladeinde BH, Olaniyan MF, Muhibi MA, Uwaifo F, Richard O, Omabe NO, et al. Association between ABO and RH blood groups and hepatitis B virus infection among young Nigerian adults. J Prev Med Hyg. 2022; 63(1): E109. doi 10.15167/2421-4248/jpmh2022.63.1.1967.

[10] Harmening DM. Modern blood banking & transfusion practices. FA Davis; 2018.

[11] Dean L, Dean L. Blood groups and red cell antigens. Bethesda: NCBI; 2005.

[12] Avent ND, Reid ME. The Rh blood group system: a review. Blood, Blood. 2000; 95(2): 375-87. doi :10.1182/blood.V95.2.375.

[13] Ristovska E, Bojadjieva TM, Velkova E, Dimceva AH, Todorovski B, Tashkovska M, et al. Rare blood groups in ABO, Rh, Kell systems–Biological and clinical significance. Prilozi. 2022; 2022: 0021. doi: 10.2478/prilozi-2022-0021.

[14] Daniels G, Hadley A, Green CA. Causes of fetal anemia in hemolytic disease due to anti-K. Transfusion. 2003; 43(1): 115-6. doi: 10.1046/j.1537-2995.2003.00327.x.

[15] Daniels G. The molecular genetics of blood group

polymorphism. Hum Genet. 2009; 126(6): 729-42. doi: 10.1007/s00439-009-0738-2.
[16] Chown B, Lewis M, Kaita K. A new Kell blood-group phenotype. Nature. 1957; 180: 711. doi: 10.1038/180711a0.
[17] Owaidah AY, Naffaa NM, Alumran A, Alzahrani F. Phenotype frequencies of major blood group systems (Rh, Kell, Kidd, Duffy, MNS, P, Lewis, and Lutheran) among blood donors in the Eastern Region of Saudi Arabia. J Blood Med. 2020; 11: 59-65. doi 10.2147/JBM.S236834.
[18] Siransy Bogui L, Dembele B, Sekongo Y, Abisse S, Konaté S, Sombo M. Phenotypic profile of Rh and Kell blood group systems among blood donors in Cote d'Ivoire, West Africa. J Blood transfus. 2014; 2014(1): 309817. doi 10.1155/2014/309817.
[19] 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.
[20] Makroo R, Gupta R, Bhatia A, Rosamma NL. Rh phenotype, allele and haplotype frequencies among 51,857 blood donors in North India. Blood Transfus. 2014; 12(1): 36-9. doi: 10.2450/2013.0300-12.
[21] Adewoyin AS, Lee GM, Adeyemo TA, Awodu OA. Rh and Kell blood group antigen prevalence in a multi-ethnic cohort in Nigeria: implications for local transfusion service. Immunohematology. 2018; 34(2): 61-5. PMID: 29989421.
[22] Westhoff CM, Reid ME. the Kell, Duffy, and Kidd blood group systems. Immunohematology. 2004; 20(1): 37-49. PMID: 15373667.
[23] Pirenne F, Floch A, Diop S. Alloimmunisation against red blood cells in sickle cell disease: transfusion challenges in high-income and low-income countries. Lancet Haematol. 2023; 10(6): e468-76. doi: 10.1016/S2352-3026(23)00066-2.
[24] Kambale-Kombi P, Djang'eing'a RM, Alworong'a Opara JP, Minon JM, Sepulchre E, Bours V, Floch A, Pirenne F, Tshilumba CK, Batina-Agasa S. Red blood cell alloimmunisation in sickle cell disease patients in the Democratic Republic of the Congo. Transfus Med. 2023; 33(2): 137-46.. doi: 10.1111/tme.12939.
[25] Kuriri FA, Ahmed A, Alanazi F, Alhumud F, Ageeli Hakami M, et al. Red blood cell alloimmunization and autoimmunization in blood transfusion-dependent sickle cell disease and β-thalassemia patients in Al-Ahsa Region, Saudi Arabia. Anemia. 2023; 2023(1): 3239960. doi 10.1155/2023/3239960.
[26] 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. PMID: 12383334.
[27] Klein HG, Anstee DJ. Mollison's Blood Transfusion in Clinical Medicine. John Wiley & Sons; 2014.
[28] Alabdulmonem W, Shariq A, Alqossayir F, AbaAlkhail FM, Al-Musallam AY, Alzaaqi FO. et al. Sero-prevalence ABO and Rh blood groups and their associated transfusion-transmissible infections among blood donors in the Central Region of Saudi Arabia. J Infect Public Health. 2020; 13: 299-305. doi: 10.1016/j.jiph.2019.12.004.
[29] Eweidah MH, Rahiman S, Ali MH, Al-Shamary AM. Distribution of ABO and Rhesus (RHD) Blood Groups in Al-Jouf Province of the Saudi Arabia. Anthropol. 2011; 13(2): 99-102. doi: 10.1080/09720073.2011.11891182.
[30] Belali TM. Distribution of ABO and rhesus types in the Northern Asir Region in Saudi Arabia. J Blood Med. 2022: 13: 643-8. doi: 10.2147/JBM.S383151. eCollection 2022.
[31] Saboor M, Zehra A, Hamali HA, Halawani AJ, Mobarki AA, Madkhali AM, Abdullah S. Prevalence of A2 and A2B subgroups and anti-A1 antibody in blood donors in Jazan, Saudi Arabia. Int J Gen Med. 2020: 13: 787-90. doi: 10.2147/IJGM.S272698. eCollection 2020.
[32] Sarhan MA, Saleh KA, Bin-Dajem SM. Distribution of ABO blood groups and rhesus factor in Southwest Saudi Arabia. Saudi Med J. 2009; 30(1): 116-9. PMID: 19139784.
[33] Reid ME, Lomas-Francis C. Ii blood group collection, The Blood Group Antigen, Facts Book Series, 471-7.
[34] Alcantara JC, Alcantara TY, Al Shaghdali KF, Al Shammari FD, Al Shammari FG. Rh alleles and phenotypes among Saudi women in Hail Region, Saudi Arabia. Adv Life Sci. 2021; 8(2): 114-118.:13. doi doi.org/10.62940/als. v8i2.1001.
[35] Elsayid M, Al Qahtani FS, Al Qarni AM, Almajed F, Al Saqri F, Qureshi S. Determination of the frequency of the most immunogenic Rhesus antigens among Saudi donors in King Abdulaziz Medical City–Riyadh. J Nat Sci Biol Med. 2017; 8(1): 56. doi 10.4103/0976-9668.198361.
[36] Alalshaikh M, Almalki Y, Hasanato R, Almomen A, Alsughayir A, Alabdullateef A, et al. Frequency of Rh and K antigens in blood donors in Riyadh. Hematol Transfus Cell. 2022; 44(4): 555-9. doi: 10.1016/j.htct.2021.03.003.
[37] Elsayid m, alfaifi am, almutairi ak, almajed f, al saqri f, qureshi s. phenotypic profile of kell blood group system among saudi donors at king abdulaziz medical city-riyadh. J Med Sci Clin Res. 2017; 5(1): 15654-7. doi.org/10.18535/jmscr/v5i1.75.
[38] Felimban RI, Sumeda SM. Distribution of Kell antigens K, k, Kpa and Kpb among blood donors in Jeddah city of Western Saudi Arabia. Asian J Transfus Sci. 2021; 15(1): 75-81. doi: 10.4103/ajts.AJTS_109_19.

Journal of Associated Medical Sciences 2025; 58(3): 313-320

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 modifying respiratory muscle device: development and pilot testing

Raoyrin Chanavirut[1]*, Rattakarn Yensano[2], Ponlapat Yonglitthipagon[1], Saowanee Nakmareong[1], Lapasrada Chaipattaranun[1], Warisara Jaratpongtakul[1]

[1]School of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen Province, Thailand.
[2]Department of Physics, Faculty of Science, Khon Kaen University, Khon Kaen Province, Thailand.

ARTICLE INFO

Article history:
Received 5 May 2025
Accepted as revised 18 August 2025
Available online 22 August 2025

Keywords:
Respiratory device, incentive spirometry, maximum inspiratory pressure, physical therapy

ABSTRACT

Background: A respiratory muscle device was designed in various forms and was available at different price ranging. However, financially disadvantaged users may find it difficult to afford the device.

Objectives: The objective of the study was to develop a device utilizing readily available materials. Subsequently, a pilot study was conducted to evaluate maximal inspiratory pressure (MIP) among healthy adult subjects. Furthermore, feedback and suggestions regarding the device were collected using a questionnaire.

Materials and methods: The device's components were evaluated to ensure safety. Subsequently, laboratory tests were conducted to assess device performance, including measurements of windmill blade rotation and airflow, to determine and prescribe the exercise intensity. Healthy subjects aged 18 to 25 years were then recruited and randomly assigned into either the control group or the exercise group (N=16 per group). Subjects in the exercise group were instructed to use the device daily for four weeks, while those in the control group were asked to continue their usual daily activities. MIP was measured in all participants on a weekly basis throughout the study. Then, subjects in the exercise group were given a questionnaire to evaluate their opinions regarding the device.

Results: The result revealed that all components of the device complied with safety standards. There was no significant difference in MIP between the exercise group and the control group. However, by the third and fourth weeks of the training period, the exercise group exhibited a significant improvement in MIP compared to their values in the first week. Furthermore, user feedback indicated that the device's size, weight, and usability were very satisfactory.

Conclusion: The study concluded that the modified device did not have a significant effect on inspiratory muscle strength, as indicated by MIP, in healthy adult subjects. However, further research is warranted to more thoroughly evaluate the effectiveness of the device and in relevant clinical populations.

Introduction

A respiratory muscle device is frequently recommended for patients with respiratory muscle weakness due to various conditions; its aim is to improve respiratory muscle strength and enhance lung expansion.[1,2] The patient is supposed to inhale or exhale through the device, and its resistance can be adjusted with various methods.[2] For a respiratory muscle device to be classified as a medical device, several phases must be successfully completed, including prototyping, safety testing, and trials in humans. Regarding the latter, three types of trials must be conducted sequentially. The first one involves testing on a small group

* **Corresponding contributor.**
Author's Address: *School of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen Province, Thailand.*
E-mail address: raocha@kku.ac.th
doi: 10.12982/JAMS.2025.102
E-ISSN: 2539-6056

of people to assess safety and side effects. The second one entails a larger group of subjects to evaluate effectiveness and identify possible side effects. The third one involves testing on an even larger group of people to confirm effectiveness and safety, as well as compare the proposed device to existing treatments.[3]

In Thailand, respiratory muscle devices are sold by several companies at different prices. However, the Comptroller General's Department has not included them on the list of nondrug medical supplies.[4] As a result, patients who are advised to use these devices must purchase them themselves. This creates a problem of access for individuals with financial constraints. In addition, the availability of these products may also be limited to certain stores, such as those in urban areas. Therefore, patients who live in remote areas may face further barriers to accessing these devices. Based on these facts, the present study was conducted to develop a modified respiratory muscle device from readily available materials, and its performance was initially assessed in the laboratory. Then, a pilot study was conducted on a small group of healthy subjects to evaluate the device's effectiveness on respiratory muscle strengthening, which was indicated by maximum inspiratory pressure (MIP). The questionnaire also collected feedback regarding the device. The data thus obtained can help researchers enhance the device for optimal performance and inform future trials involving relevant clinical populations.

Materials and methods

The present study was divided into two stages, as illustrated in Figure 1. First, a modified device was developed. Then, the impact of the device on inspiratory muscle strength was evaluated. Furthermore, participants' feedback regarding the device was collected.

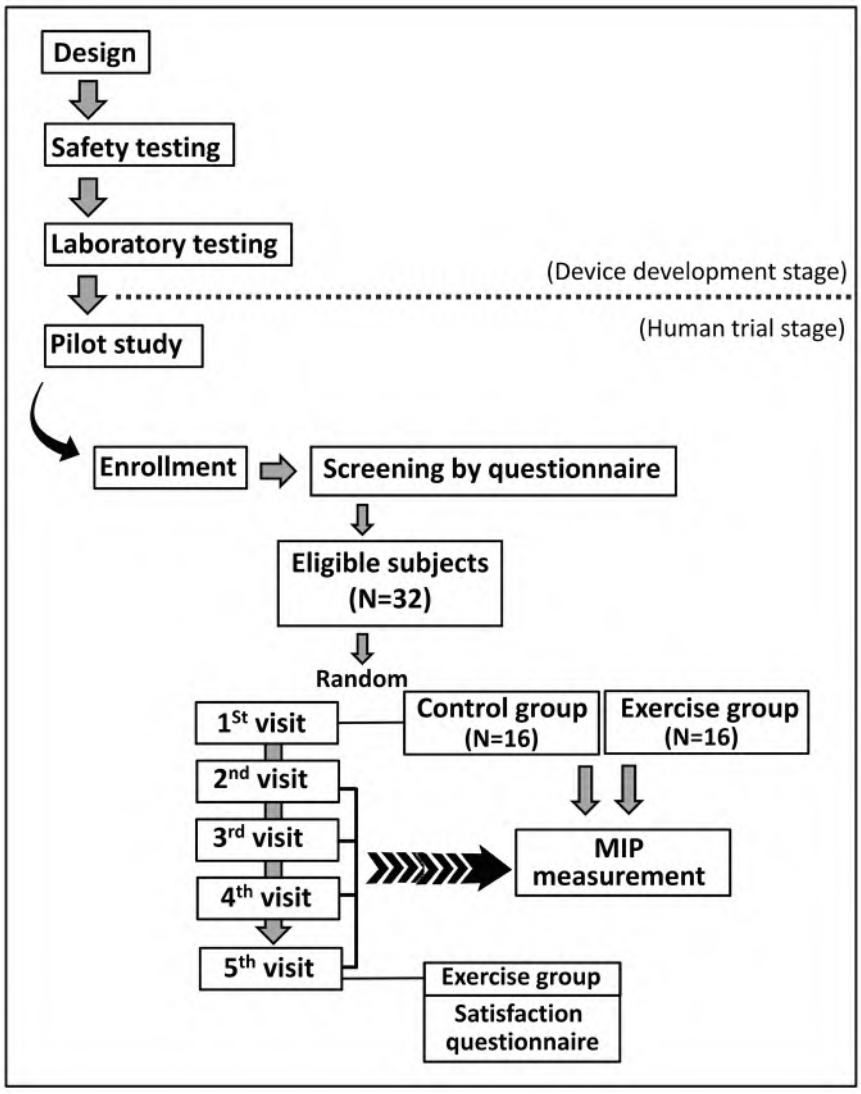


Figure 1. *Consort diagram of the study. MIP: maximum inpiratory pressure.*

Device development

The device was developed by modifying a 22-ounce plastic cup, a plastic cap, and a plastic windmill (Figure 2). Both the cup and the cap were made from food-grade polypropylene plastic, while the windmill was a plastic toy. All the components were submitted for safety-standard testing and assessment of plastic particle release, as detailed in Table 1. The test results demonstrated that the components complied with the relevant safety standards, which confirmed the suitability of the materials used in the device.

The windmill was installed inside the cup and fixed with double-sided adhesive tape (Figure 2A). Then, the cap was placed on the cup (Figure 2B). To let the inhaled air flow, three holes measuring seven millimeters in diameter were made in the bottom of the cup with a soldering iron (Figure 2C). The original diameter of the cap hole was 10 millimeters, which is larger than the typical silicone straw available on the market. Therefore, an eight-millimeter hole was made in the cap to allow the silicone straw, which served as the inhalation tube, to be properly fitted (Figure 2B).

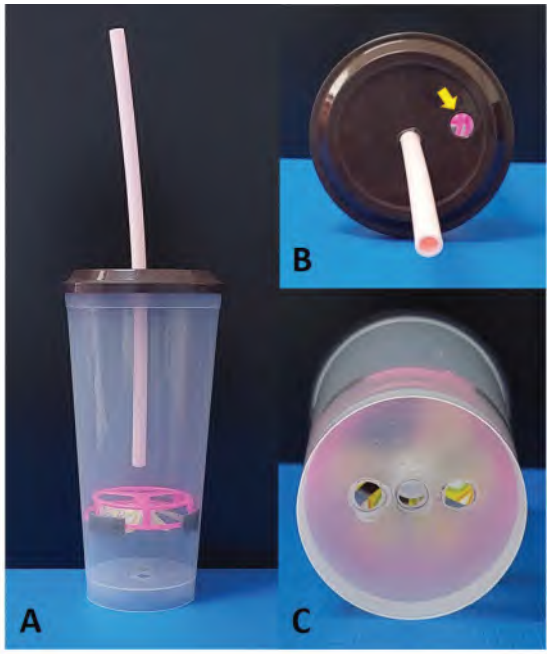


Figure 2. *Components of the modifying device; the ready-to-use device. A: cap with a silicone straw, B: three seven-millimeter diameter holes at the cup bottom, C: arrow indicated the original cap's hole.*

Table 1. Safety testing results of device components

Parameter (mg/kg)	Cup	Cap	Windmill
Antimony	NT	NT	54.0[B]
Arsenic	NT	NT	ND[B]
Barium	NT	NT	14[B]
Cadmium	ND[A]	ND[A]	ND[B]
Chromium	ND[A]	ND[A]	<1[B]
Copper	0.804[A]	2.079[A]	NT
Lead	ND[A]	ND[A]	ND[B]
Mercury	NT	NT	ND[B]
Selenium	NT	NT	ND[B]
Zinc	4.422[A]	2.324[A]	NT
Diameter sizes of plastic particle (μm)			
>10	ND[C]	ND[C]	ND[C]
>25	ND[C]	ND[C]	ND[C]

Note: Testing method; [A]: In-house method based on EPA 3052, [B]: IEC 62321-5:2013, [C]: Thai industrial standard 531-2558. NT: not tested, ND: not detected.

The next step involved evaluating the device characteristics. First, the device resistance was assessed using a phlegm suction machine (7E-C Yuwell). This assessment included the following steps: 1) marking one of the windmill blades with a magic pen and assembling the device, 2) connecting the silicone straw to the suction tube, 3) mounting a video camera at an angle that enabled the rotation of the marked windmill blade to be observed through the device's bottom holes, 4) opening or closing the bottom holes with transparent adhesive tape, 5) turning the regulator valve for maximum suction pressure, 6) switching on the phlegm suction machine and the video camera for a duration of five seconds, and 7) counting the number of spins of the windmill blades using the iMovie software. This procedure was meant to test the number of spins of the windmill blades under the same suction pressure but with varying conditions (open or closed) for the device's bottom holes. The number of windmill blade spins is shown in Table 2. Second, the Biopac Research MP35 Data Acquisition System was employed to evaluate the airflow measurement, and an offline analysis was conducted using the Biopac Student Lab software (Version 4.1). The results for the mean airflow are presented in Table 2.

Table 2. General characteristics of the modified device based on laboratory testing.

Condition	Number of spinning	Mean airflow (L/sec)	Level of difficulty
7 7 7	11	0.580	Hardest
7 7 0	18	0.619	Moderate
7 0 7	28	0.622	Easy
7 0 0	34	0.625	Easiest

Note: 7: bottom hole of the modified device is open, 0: the hole has been sealed with transparent adhesive tape.

Subjects

As the first human trial of the device, a pilot study was conducted with healthy subjects. The sample size was calculated based on data from the Nepomuceno study.[5] Then, 32 healthy individuals aged 18-25 years were recruited. Next, they were randomly divided into two groups: control (N=16) and exercise (N=16). Personal information was collected through a self-administered questionnaire. No medical or physical conditions affecting respiratory mechanics were present, such as scoliosis or kyphosis. The subjects' BMI values ranged from 18.5 to 22.9 kg/m². Individuals were excluded if they had a history of fainting while taking a deep breath or holding their breath, syncope of unknown origin, or a clavicle, rib, or sternum fracture within the previous six months.

MIP measurement

MIP was measured by a physical therapist. Ten samples at the same age range were included for intra-rater reliability measurement. Test-retest variability was determined by an intraclass correlation coefficient (ICC). The ICC value for MIP assessed by the therapist was 0.996, which indicates high reliability.[6] MIP was measured using the MicroRPM respiratory pressure meter in accordance with its guidelines.[7] The subjects were instructed to perform their maximum inspiratory efforts in a sitting position, with a one-minute break between efforts. Then, the maximum value of three efforts that varied by less than 10% was reported as the MIP value.

Program

The exercise group.

The data presented in Table 2 were used as the basis for classifying the levels of device difficulty. Four levels were determined, ranging from the easiest to the hardest, with the number of windmill blade spins serving as the primary criterion. The duration of four weeks for the training program was chosen to reflect these four levels. All participants in the exercise group began training with the easiest level. Then, each week, they were instructed to increase the difficulty level by one step, thus progressing through all four levels.

Transparent adhesive tape was used to close the bottom holes in accordance with the difficulty level. It was also used to close the original cap hole, as shown in Figure 2B. Then, the participants were instructed to use the device by taking a slow, deep breath and sustaining a maximal inspiratory effort for three seconds. They were instructed to use the device at home 10 times per set, 3 sets per round, with a 2-minute break between sets, and 2 rounds per day. The training duration was approximately 20 minutes per day.[5] The participants had to keep exercise diaries to record their exercise adherence. They were also monitored by phone twice a week. In addition, they were scheduled to return once a week to have their MIP measured. Then, after completing the training program, they were asked to rate their satisfaction with the device using a questionnaire. This questionnaire elicited responses regarding device structure, usability, cleaning, and storage, as well as overall satisfaction. The participants were asked to rate their satisfaction using a five-point Likert scale: 1 ("very dissatisfied"), 2 ("dissatisfied"), 3 ("neutral"), 4 ("satisfied"), or 5 ("very satisfied"). In the questionnaire, there was also space to provide additional suggestions or comments.

Control group

The control group was asked to carry out various daily routines as usual. They were also scheduled for MIP reassessment once a week throughout the study period.

Statistical analysis

Descriptive information (i.e., the general characteristics of the two groups and the opinions expressed in the questionnaires) is presented in the form of percentages. Data are presented in the mean±SD format. Statistical analysis was performed using Stata (version 18), with a significance level of $p<0.05$. Because the MIP data were not normally distributed, and since the focus was on estimating group-averaged effects of device-based training on the MIP outcome, rather than subject-specific effects, a generalized estimating equation model was used. In addition, the Bonferroni test was utilized to compare MIP at the second, third, and fourth weeks to MIP at the first week in each group.

Results

There were no significant differences between the general characteristics of the two groups, as shown in Table 3. The exercise group had a 100% exercise adherence rate, and no adverse events were reported. There was no significant difference in MIP between the two groups (Table 4). However, in the exercise group, the MIP values in the third and fourth weeks were significantly higher than in the first week (Table 4). Furthermore, when considering the overall effect of training with the device, the MIP value in the exercise group increased by 1.77 centimeters of water compared to the control group, in which this increase was not statistically significant.

The exercise group also provided their opinions. Almost all the respondents (94%) were very satisfied with the device's size and weight; however, regarding the device's material, strength, and configuration, the percentages of very satisfaction were only 69%, 56%, and 50%, respectively. All the participants were very satisfied with the device's usability due to its simplicity. However, the device was difficult to transport, as the following ratings were given concerning ease of transportation: 37.5% satisfied, 19% neutral, and 6% extremely dissatisfied. They were also very satisfied with the device's ease of cleaning (88%) and storage (81%). The rating for overall satisfaction was also considerable (81%). Three participants complained that the windmill blades spun unevenly.

Table 3. General characteristics of the subjects.

Characteristics	Control group (N=16)	Exercise group (N=16)
Male: Female (%)	4:12 (25%:75%)	6:10 (37%:63%)
Age (years)	20.8±1.00 (19-21)	20.6±1.15 (18-22)
Body weight (kg)	56.7±8.48 (47.0-68.9)	55.4±7.21 (47.2-69.3)
Height (cm)	165.5±8.95 (153.0-181.0)	165.1±7.87 (155.0-180.0)
BMI (kg/m^2)	20.6±1.28 (18.2-22.9)	20.2±1.18 (18.6-22.2)
Waist circumference (cm)	70.5±5.93 (63.4-85.0)	69.3±4.85 (60.5-77.0)
Hip circumference (cm)	89.7±7.91 (73.0-101.0)	90.2±6.38 (80.0-98.0)
Waist-to-hip ratio	0.79±0.07 (0.70-0.90)	0.77±0.08 (0.64-0.90)

Note: Data presented in mean±SD (min-max).

Table 4. The maximum inspiratory pressure measurement.

	Control group (N=16)	Exercise group (N=16)	Mean difference (95%CI)	*p* value
Overall			1.77 (-6.75, 10.30)	0.684[A]
At first week	102.31±19.51	113.06±19.62	1.77 (-9.09, 12.63)	>0.999[B]
At second week	101.13±24.21	118.63±23.92	8.52 (-2.34, 19.38)	0.200[B]
At third week	104.88±29.25	123.94±23.73[C]	10.08 (-0.78, 20.95)	0.082[B]
At fourth week	104.69±28.10	124.13±25.82[C]	10.46 (-0.40, 21.32)	0.065[B]

Note: Data are presented in centimeters of water as mean ± SD. [A]By generalized estimating equations model adjusted for baseline values measured at the first visit, [B]Comparison of MIP between the control and exercise groups at the first, second, third and fourth week was performed using the Bonferroni test, [C]Comparison of MIP at the second, third, and fourth weeks with the first week within each group was performed using the Bonferroni test, with significance set at p<0.05.

Discussion

The major result of the present study is the development of a modified respiratory muscle device (a type of medical device). All the components of the device were safe in accordance with the relevant standards. Furthermore, the findings of the pilot study conducted with healthy subjects show no significant differences in respiratory muscle strength, as indicated by MIP.

The modified device in this pilot study is classified as a non-in vitro diagnostic medical device according to the Thailand Medical Device Act.[8] Furthermore, it is comparable to the medical equipment standard number 93/42/EEC, class I medical devices.[9] As a result, all the components were tested to ensure compliance with safety standards before conducting the pilot study. The components were safe to use. Furthermore, the equipment to build the device costs approximately USD 1 per set. The device currently holds a petty patent (No. 23538) in Thailand.

In the pilot study, the effectiveness of the modified device on respiratory muscle strength was assessed using MIP. However, the MIP values in the exercise group did not increase significantly when compared to the control group. Several factors may account for this finding. First, the exercise protocol may have contributed to the lack of significant improvement. According to previous studies, the effective training period for improving respiratory muscle strength is typically 4-8 weeks.[10,11] Moreover, another study suggested that, in healthy individuals, the training intensity should exceed 50% of MIP, with a frequency of 1 to 2 sessions per day, 5 to 7 days per week, to achieve significant adaptation of muscle function within 3 weeks.[12] In the present pilot study, the lack of increase in the MIP values of the exercise group may be due to the fact that the modified device failed to provide sufficient load to elicit a training effect.[13] Second, in the pilot study, all the participants were required to start the exercise at the easiest level, without consideration of individual respiratory muscle strength. This methodological decision might account for the absence of significant changes in MIP. However, this approach allowed the participants to experience all four difficulty levels, which ensured the collection of detailed feedback on the device.

Previous studies conducted in young, healthy individuals have indicated that increases in diaphragm thickness are associated with improvements in MIP.[11, 14] They have suggested that these improvements are attributable not only to diaphragm hypertrophy but also to neural adaptations, including improved coordination of synergistic muscle contractions and an increased capacity to maximize the activation of individual muscles.[11,14] Similarly, in the present pilot study, the exercise group showed significant improvements in MIP at the third and fourth weeks of the training period compared to the first week. This finding may be attributed to enhanced coordination of synergistic muscle contractions through training, which contributed to the observed increases in MIP.

In a previous study, Das *et al.* reported an adaptive device consisting of a sterile surgical glove firmly tied with micropore medical tape to the broad end of a sterile injection syringe barrel, with the piston removed.[15] To exercise, patients were instructed to take as deep a breath as they could and exhale into the device via the free end of the syringe. They were encouraged to inhale enough air to inflate all the glove's fingers. The bulging of the gloves' fingers acted as an incentive. This type of device is simple to assemble and thus readily available. Its reuse was not specified by the authors. Because the device was built with medical-grade materials, which are typically designed for single use and disposal, the overall cost per unit is likely comparable to that of the modified device used in the present pilot study. Additionally, since the device in Das *et al*.'s study was used for expiratory breathing, its training effects on respiratory muscles may differ from those of the modified device in this pilot study, which was employed for inspiratory breathing.

The BreatheMAX device developed by Jones *et al*.[16,17] consists of a sealed water bottle and two tubes, a design that differs from that of other devices available in Thailand. The inlet tube sinks beneath the surface of the water, providing an incentive in the form of bursting bubbles and a bubbling sound. The BreatheMAX has both incentive spirometer (IS) and positive expiratory pressure (PEP) functions that adjust resistance based on the water level in the bottle. Its effectiveness has been studied to aid sputum clearance,[16,17] reduce resting blood pressure in isolated systolic hypertension patients,[18] and increase MIP in patients with COPD.[19] The training program to increase MIP lasted eight weeks at an intensity level of 30% of MIP in COPD patients. The authors found a significant increase (31%) in MIP.[19] This training duration is notably different from the one adopted in the present pilot study. The modified device examined here offers only four adjustable resistance levels; therefore, with a progressive increase in difficulty each week, it could be used for a training period of only four weeks. Furthermore, the BreatheMAX is manufactured using medical-grade plastic, which allows for repeated use. However, its retail price is around USD 23.5 per set, which may create an accessibility barrier for patients with limited financial resources.

Leelarungrayub *et al*. demonstrated a simple prototype device.[20] This was constructed using a PVC plastic water pipe and plastic caps. The main tube consisted of a 10-centimeter-long PVC water pipe with a diameter of 2 centimeters, and it featured a flat, 2-centimeter-wide mouthpiece at one end. Three plastic caps, each with a central hole of different diameters (2, 4, and 6 millimeters), were designed to be placed on the opposite end of the mouthpiece to provide variable resistance.[20] The training period with this simple prototype device lasted for three weeks, during which the aperture size of the device was progressively decreased. Participants began training with a 6-millimeter opening in the first week, progressed to a 4-millimeter opening in the second week, and completed the protocol with a 2-millimeter opening in the third week. Leelarungrayub *et al*. did not provide details regarding the type of plastic used or the cost per set. However, compared to the present pilot study, the study had a shorter training duration, corresponding to the number of plastic caps

designed by the researchers. Still, they found that their device was effective in improving MIP among healthy participants aged 20-23 years.[20]

The following aspect concerns the design modifications of the device used in the present study. The blades spin when air enters the cup through the holes in the bottom, which creates a feedback mechanism that encourages users to take deep breaths. This mechanism is not influenced by gravity, which is required in other respiratory devices, such as the BreatheMAX.[17] Hence, users can perform the exercise in any position with the modified device developed in this study. Furthermore, when comparing the modified device to the adaptive device created by Das *et al*.[15] and the simple prototype by Leelarungrayub *et al*.,[20] the spinning blades may serve to further motivate users to engage in deep breathing exercises or provide visual feedback on sustained MIP. Moreover, as the modified device was adapted from a windmill toy, its colorful blades offer an additional advantage, particularly for use with children. However, parameters such as lung volume and thoracic expansion were not assessed in this pilot study, so further research is needed to investigate these aspects.

The questionnaire results show that almost all the participants were very satisfied with the device's size and weight, as well as its usability. Moreover, the silicone straw can be reused by simply cleaning it with dishwasher detergent and thoroughly drying it. As a result, the participants were very satisfied with the ease of cleaning. However, some of the respondents commented on the uneven spinning of the blades. In response to this feedback, the researcher carefully inspected each windmill before assembling it into the device set. Incorporating suggestions from subjects and the data obtained from this study into future research will likely help confirm the device's efficacy and may also increase opportunities for its application in clinical contexts.

Limitation

This pilot study has several limitations. First, it was conducted exclusively in healthy subjects; therefore, the findings may not be generalizable to other populations. Future trials are needed in relevant clinical populations. Second, a standard protocol for resistance calibration was not available, as the device was modified and adapted from other equipment. Consequently, resistance was assessed using a modified approach, as detailed in the above section on device development. Third, since the cup, cap, and windmill were not made from medical-grade plastic, the ethics committee permitted their use for only one day. As a result, the subjects in the exercise group had to change the device daily, which increased the associated costs and may have caused variability in windmill blade rotation across different sessions. Future scholars should consider improving the device design and manufacturing it from medical-grade plastic to facilitate repeated use. In addition, the tool's validity and consistency of repeated performance for training will be confirmed by the technology acceptance model (TAM) of user experience (UX), and user interface (UI) for the standardized tool will be confirmed. Fourth, because the characteristics of the modified device limited the exercise intensity levels in this pilot study, the study duration could not be extended beyond four weeks. Thus, future study should consider determining exercise intensity based on each participant's MIP to ensure appropriate training loads.

Conclusion

The modified device developed in the present study did not have a significant effect on inspiratory muscle strength, as indicated by MIP, in healthy young subjects. However, further research is warranted to evaluate the effectiveness of the device more thoroughly.

Ethical approval

The study was approved by the Ethics Committee of the Khon Kaen University (HE654036), and all the participants gave written informed consent.

Funding

The study was supported by the Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen, Thailand. Grant number: 66030.1.8.1.3/608.

Conflicts of interest

The authors declare no conflicts of interest.

CRediT authorship contribution

Raoyrin Chanavirut: conceptualization, resources, writing-reviewing and editing; **Rattakarn Yensano**: laboratory testing; **Ponlapat Yonglitthipagon**: statistical analysis; **Saowanee Nakmareong**: methodology; **Lapasrada Chaipattaranun** and **Warisara Jaratpongtakul**: data collection. All authors have read and agreed to the final version of the manuscript.

Acknowledgements

The authors wish to express our gratitude to all volunteers who participated in this study. We also gratefully acknowledge the financial support from the Faculty of Associated Medical Sciences, Khon Kaen University.

References

[1] Kadu D. Incentive spirometer and respiratory muscle training devices: What do physiotherapists need to know? J Pulm Respir Med. 2023; 13: 620. doi: 10.37421/2161-105X.2023.13.620.

[2] Menezes KKP, do Nascimento LR, Avelino PR, Polese JC, Teixeira-Salmela LF. A review on respiratory muscle training devices. J Pulm Respir Med. 2018; 8: 451. doi: 10.4172/2161-105X.1000451.

[3] Bharti A. A review on phases of clinical trials and types of clinical trials. Der Pharmacia Lettre. 2021; 13: 59-64.

[4] Drug and medical supply medical center, Ministry of Public Health. Non-medical supplies. [serial online] 2019 [cited 2025 April 18]. Available from: https://dmsic.moph.go.th/index/detail/7951 (in Thai)

[5] Nepomuceno BRV Jr, Barreto MS, Almeida NC, Guerreiro CF, Xavier-Souza E, Neto MG. Safety and

efficacy of inspiratory muscle training for preventing adverse outcomes in patients at risk of prolonged hospitalization. Trials. 2017; 18: 626. doi: 10.1186/s13063-017-2372-y.

[6] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016; 15: 155-63. doi: 10.1016/j.jcm.2016.02.012.

[7] Laveneziana P, Albuquerque A, Aliverti A, Babb T, Barreiro E, Dres M, et al. ERS statement on respiratory muscle testing at rest and during exercise. Eur Respir J. 2019; 53: 1801214. doi: 10.1183/13993003.01214-2018.

[8] Food and drug administration. Re: Classes of medical devices or medical devices of which manufacturers or importers are required to declare specifications (No. 2), B.E. 2563 (2020) [serial online] 2020 [cited 2025 April 18]. Available from: https://shorturl.asia/cfNpX

[9] Food and drug administration. Re: Risk classification of medical devices, B.E. 2562 (2019) [serial online] 2020 [cited 2025 April 18]. Available from: https://en.fda.moph.go.th/cat2-health-products?ppp=10&page=1&kw=risk+classification&cate_2=

[10] Souza H, Rocha T, Pessoa M, Rattes C, Brandao D, Fregonezi G, et al. Effects of inspiratory muscle training in elderly women on respiratory muscle strength, diaphragm thickness and mobility. J Gerontol A Biol Sci Med Sci. 2014; 69: 1545-53. doi: 10.1093/gerona/glu182.

[11] Downey AE, Chenoweth LM, Townsend DK, Ranum JD, Ferguson CS, Harms CA. Effects of inspiratory muscle training on exercise responses in normoxia and hypoxia. Respir Physiol Neurobiol. 2007; 156: 137-46. doi: 10.1016/j.resp.2006.08.006.

[12] Romer LM, McConnell AK. Specificity and reversibility of inspiratory muscle training. Med Sci Sports Exerc. 2003; 35: 237-44. doi: 10.1249/01.MSS.0000048642.58419.1E.

[13] Sheel AW. Respiratory muscle training in healthy individuals: physiological rationale and implications for exercise performance. Sports Med. 2002; 32: 567-81. doi: 10.2165/00007256-200232090-00003.

[14] Enright SJ, Unnithan VB, Heward C, Withnall L, Davies DH. Effect of high-intensity inspiratory muscle training on lung volumes, diaphragm thickness, and exercise capacity in subjects who are healthy. Phys Ther. 2006; 86: 345-54. doi: 10.1093/ptj/86.3.345.

[15] Das AK, Sharma A, Sondhi R, Goyal S. A simple incentive spirometry device for morale-boosting for geriatric patients. Eur Geriatr Med. 2022; 13: 1495-6. doi: 10.1007/s41999-022-00683-9.

[16] Kluayhomthong S, Khrisanapant W, Chaisuksant S, Jones C. Effectiveness of a new breathing device "BreatheMAX®" to increase airway secretion clearance in patients with ventilatory dependence. J Med Tech Phy Ther. 2011; 23: 95-108. (in Thai)

[17] Jones CU, Kluayhomthong S, Chaisuksant S, Khrisanapant W. Breathing exercise using a new breathing device increases airway secretion clearance in mechanically ventilated patients. Heart Lung. 2013; 42: 177-82. doi: 10.1016/j.hrtlng.2012.12.009.

[18] Ubolsakka-Jones C, Tongdee P, Jones DA. The effects of slow loaded breathing training on exercise blood pressure in isolated systolic hypertension. Physiother Res Int. 2019; 24: e1785. doi: 10.1002/pri.1785.

[19] Wongsawat C, Ubolsakka-Jones C, Panwong W, Phimphasak C, Domthong P, Kiatboonsri S. Effectiveness of inspiratory muscle training with a new breathing pattern in elderly patients with chronic obstructive pulmonary disease. Arch AHS. 2020; 31: 467-80. doi: https://he01.tci-thaijo.org/index.php/ams/article/view/235114.

[20] Leelarungrayub J, Pinkaew D, Yankai A, Chautrakoon B, Kuntain R. Simple artificial training device for respiratory muscle strength and lung volumes in healthy young male and female subjects: A pilot study. J Bodyw Mov Ther. 2017; 21: 995-1002. doi: 10.1016/j.jbmt.2016.11.012.

Journal of Associated Medical Sciences 2025; 58(3): 321-328

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Efficacy of rehabilitation in improving social cognition and behavioral outcomes of pediatric traumatic brain injury: An integrative review

Bandita Gupta* **and Ishika Upadhyay**

Department of Occupational Therapy, Assistant Professor, NIMS University, Jaipur, Rajasthan- 303121, India.

ARTICLE INFO

Article history:
Received 18 February 2025
Accepted as revised 18 August 2025
Available online 22 August 2025

Keywords:
Adolescents with TBI, cognitive outcomes in pediatric TBI, pediatric rehabilitation in neurological cases, pediatric traumatic brain injury, psychological aspect of pediatric TBI.

ABSTRACT

Background: One of the main causes of long-term disability in kids and teenagers is pediatric traumatic brain injury (TBI), which frequently leads to deficiencies in behavioural and social cognition. Rehabilitation helps in enhancing these outcomes.

Objectives: The purpose of this integrative review is to assess the available data on how well rehabilitation works to improve behavioral and social cognition in children with traumatic brain injury.

Materials and methods: An extensive search of electronic databases and registries (PubMed, Scopus, Google Scholar, and Web of Science, among others) was conducted. Studies that looked at how rehabilitation therapies affected the social cognition and behavioural outcomes of pediatric TBI published between 2010 and 2024 were included.

Results: Eighteen studies met the inclusion criteria, which encompassed various rehabilitation interventions, including family-based interventions, social skills training, neurocognitive remediation, and cognitive-behavioural therapy (CBT). Overall, both behavioural outcomes, such as emotional regulation and adaptive behaviours, as well as social cognition, including theory of mind, emotional recognition, and social problem-solving abilities, were shown to improve with rehabilitation.

Conclusion: In children and adolescents with TBI, rehabilitation interventions—especially those centered on social skills training, cognitive-behavioural methods, and neurocognitive remediation—can greatly enhance social cognition and behavioural outcomes. The analysis underscores the necessity of customized, multidisciplinary rehabilitation programs and identifies research topics for the future, such as including neurobiological underpinnings and long-term monitoring to gauge how long therapy effects last.

Introduction

Children may be more susceptible to traumatic brain injury (TBI) because of the critical brain development that takes place between infancy and childhood. Early childhood brain injuries have the potential to impede or postpone core neurodevelopment.[1,2]

Each year, on average, 634,000 children in the US suffer from TBI. Children under the age of four and teenagers aged 15 and up have the highest rates of TBI-related ER visits.[3,4] About 40% of head injuries that need hospitalization for children under 4 years old may be caused by falls, which are a significant cause of TBI in young children (<4 years old).[3-6] Injuries caused by child maltreatment, such as shaken infant syndrome, may be the cause of 20-70% of hospitalizations for young children. Accidents involving motor vehicles are the most common cause of injuries among older children.[3]

* **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Assistant Professor, NIMS University, Jaipur, Rajasthan- 303121, India.*
E-mail address: banditagupta9@gmail.com
doi: 10.12982/JAMS.2025.103
E-ISSN: 2539-6056

Acquired brain injury (ABI) is often treated during the acute phase, which includes ER care, potential inpatient hospitalization, outpatient treatment, and ultimately reintegration into the home and school setting.[7] Children who have had brain injuries should be regularly watched for unresolved functional deficits since they typically have persistent, unmet functional demands across numerous domains. Unmet therapy demands, especially in the domains of cognition and communication, were linked to longer recovery times following an injury and a diagnosis of complicated mild TBI.[8,9]

Cognition seems to be significantly compromised in moderate-to-severe TBI approximately one month after the injury or soon after post traumatic amnesia (PTA) resolves.[10,11] Patients with slight brain injuries typically recover cognitively quickly, returning nearly to "usual level performance" in three months.[12,13] Effective cognitive rehabilitation strategies started after TBI have been shown in trials and articles to speed up restoration and reduce loss of function.[14]

Adolescents with traumatic brain injuries often exceed their learning capacity, therefore, cognitive understanding is crucial.[15,16] Although the impact and manner of the trauma determine the level of impairment, intellectual, emotional, and psychological impairments might not become apparent for years following the injury.[9,17]

Despite physical recovery after pediatric TBI, cognitive and behavioural challenges often go unnoticed by caregivers and rehabilitation teams.[7] These "invisible" impairments significantly impact long-term reintegration into daily life and society. Although existing literature highlights the importance of cognitive and behavioural interventions post-TBI, there remains a critical gap in standardized, effective rehabilitation guidelines tailored for pediatric populations.[18] This integrative review aims to gather information from different sources and provide a generalized guideline for effective rehabilitation and promote the well-being of children with TBI.

Objectives of the study

1. To understand how rehabilitation can improve emotional, social, and psychological outcomes in children with traumatic brain injury (TBI).
2. To evaluate how effective different rehabilitation therapies are by reviewing and combining results from multiple studies.
3. To provide evidence-based information that can guide researchers, doctors, and caregivers in improving rehabilitation for children with TBI.

Rationale

This study is needed to comprehend the efficacy of comprehensive rehabilitation strategies in improving the quality of life in children with TBI by taking into consideration the behavioural and cognitive outcomes after TBI. It can also help the clinician and rehabilitation worker in providing better outcomes by combining various approaches to enhance the day-to-day lifestyle of children with TBI.

Materials and methods

Secondary data analysis is used to study the research question. Various databases and registers like PubMed, Scopus, Google Scholar, Web of Science, Journals, and books were searched for articles relating to the rehabilitation of behavioural and psychosocial outcomes in pediatric patients experiencing Traumatic Brain Injury. Studies published between 2010-2024 in the English language were taken into account for this study.

Population

Patients diagnosed as pediatrics traumatic brain injury.

Study identification

For this study, data were considered from correlational studies, systematic reviews, case reports, cross-sectional studies, and interventional studies (Figure 1).

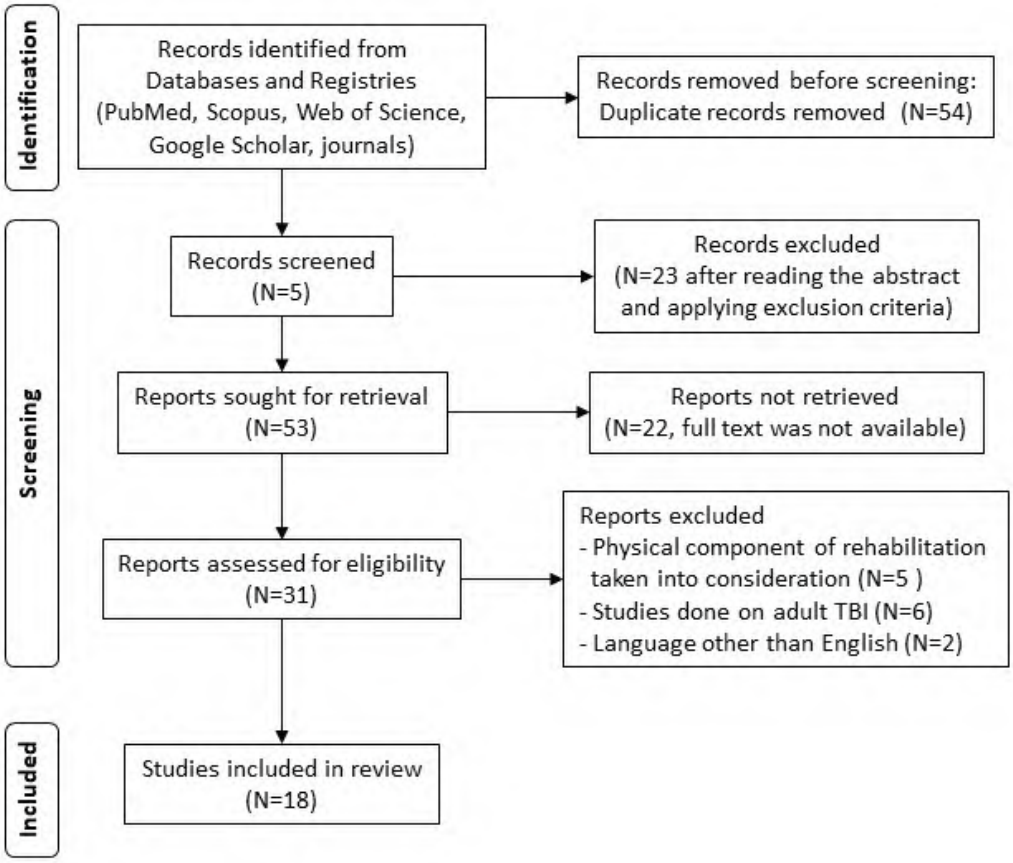


Figure 1. *Interventional studies.*

Eligibility criteria

- Studies published between 2010-2024.
- Studies in the English language.
- Studies involving pediatric TBI patients.
- Studies addressing the psychosocial aspect of pediatric TBI and its rehabilitation.

Exclusion criteria

- Studies addressing the physical domain of rehabilitation.
- Studies concerning the medical management of pediatric.

Results

By enhancing executive function, emotional regulation, social cognition, and adaptive behaviours, rehabilitation interventions—from cognitive-behavioural therapies and metacognitive training to family-centered and game-based interventions—significantly improve psychosocial outcomes in pediatric traumatic brain injury, according to the literature. Multiple trial analysis reveals that although many rehabilitation techniques (such as Goal Management Training, Attention Process Training, and Acceptance and Commitment Therapy) are successful in the short term, they are not always addressed in the long run, which indicates significant gaps in current practices. When taken as a whole, these findings provide important, empirically supported insights that can help researchers, clinicians, and caregivers choose and modify rehabilitation strategies that are sensitive to the environment and development of children with TBI (Table 1).

Table 1. Review of literature.

No.	Author	Sample size	Methodology	Result
1	Wang C *et al.*[19]	60 children aged 6-12 years	The experimental group received a metaphorical intervention based on children's games.	The intervention group scored higher on the Mini-Mental State Examination (MMSE) and Loewenstein Occupational Therapy Cognitive Assessment (LOTCA) and had lower scores on the Child Behaviour Checklist (CBCL), depression inventory, and emotional disorder screening tools compared to the control group.
2	Sood NT *et al*.[20]	31 children post-TBI	Participants were randomly assigned to either the active control group (Lexia Reading Core5, N=15) or the treatment group (Cogmed, N=16). For five weeks, both groups underwent computerized instruction with a physician's assistance via an online video platform for five weeks.	A significant improvement was observed in verbal working memory after therapy, but no lasting cognitive changes were seen six months later.
3	Sargénius HL[21]	38 pediatric clients and their parents	Goal Management Training (GMT) and Psychoeducational Control (Pediatric Brain Health Workshop, pBHW) were used.	Both GMT and pBHW (Pediatric Brain Health Workshop) enhanced executive functions based on parental reports (Behaviour Rating Inventory of Executive Function, BRIEF).
4	On ZX *et al*.[22]	11 articles included 482 children with TBI	Systematic Review and meta-analysis.	Children with traumatic brain injury display significant social cognition impairments, especially in theory of mind and emotion recognition.
5	Polinder S *et al*.[23]		Review article	Early educational intervention, cognitive behavioural therapy (CBT), and neuropsychological rehabilitation are effective for cognitive symptoms following mild TBI in children.
6	Séguin M *et al*.[24]	17 children post-TBI	Standardized tests for attention and executive function, along with behavioural checklists, assessed pre- and post-intervention in children with attention issues randomized to receive either Homework Assistance (HWA) or Attention Process Training (APT).	The study showed improvements in cognitive flexibility, working memory, and inhibition.
7	Ryan NP *et al*.[25]	78 children with TBI, 40 typically developing children, and their parents	Parents completed surveys after one or two years regarding their child's social skills and environmental factors.	Recovery over time correlates with injury severity; social outcomes depend on family environment and caregiver coping skills.
8	Barman A *et al*.[14]		Review article	Advocates for using restorative and compensatory methods in cognitive rehabilitation.
9	Brown FL *et al*.[26]	59 caregivers of children with ABI	Caregivers were split into two groups: one received Acceptance and Commitment Therapy (ACT) combined with Stepping Stones Triple P (SSTP), and the other received usual care (CAU).	Post-therapy, the treatment group showed small to moderate improvements in parental mental health, emotional flexibility, confidence in managing behaviours, family adaptation, and peer-parent relationships compared to the CAU (care-as-usual) group.

Table 1. Review of literature (*continued*).

No.	Author	Sample size	Methodology	Result
10	Coyne JH *et al*.[27]	Children aged 8-16 post-TBI with below-average memory	Three intervention types—massive restudy (MR), spaced restudy (SR), and retrieval practice (RP)—using two stimulus types: verbal paired associates (VPAs) and face-name pairs (FNPs).	The learning condition significantly affected delayed recall; retrieval practice yielded higher recall of verbal paired associates (VPAs) and face-name pairs (FNPs) than massive restudy (MR) and spaced restudy (SR).
11	Cook LG *et al*.[28]	20 adolescents post-TBI	Cognitive training comparing gist reasoning versus rote memory, with 8 sessions of 45 minutes each.	The gist reasoning group (N=10) showed notable improvements in recall, ability to abstract meaning, and generalization to working memory and inhibitory control.
12	Krasny-Pacini A *et al*.[29]	Five children with severe TBI and significant executive functioning difficulties	1 per week of metacognitive training treatment over 5 months	Both prospective memory and executive functioning improved.
13	Bower J *et al*.[30]	One child with severe TBI	Videotaped data from before, during, and after music therapy sessions using the Agitated Behaviour Scale and micro-analysis.	The child displayed behaviours such as neutral, acceptance, recruitment, and rejection during sessions.
14	Haslam C *et al*.[31]	30 participants: 15 with ABI and 15 non-injured controls	Participants underwent errorless learning, self-generated learning, and errorful learning.	No significant differences were observed between groups or methods.
15	Ho J *et al*.[32]	15 children with ABI	Memory rehabilitation using diary training, self-instruction, and case examples: six sessions per week.	Significant improvements were seen in retrieving information for Activities of Daily Living (ADLs).
16	Chan DY *et al.*[15]	32 children with ABI	Metacognitive problem-solving training.	Using education level, IQ scores, and baseline data as covariates, significant changes were observed in post-test scores between the experimental and control groups.
17	Feeney TJ[33]	Two children with severe TBI	Documented combined cognitive and executive function therapies using a variety of standard approaches.	Post-intervention reduced challenging behaviours and improved school performance were observed.
18	Yeates KO *et al*.[34]	19 children with severe TBI, 56 with mild/moderate TBI, 99 with orthopedic injury (OI)	Analyzed the relationship between family environment and psychosocial outcomes over time using mixed model analysis.	Major differences in social behaviour were identified between the orthopedic injury (OI) and TBI groups. Adaptive functioning improved.

Discussion

This integrative review explores various study designs to understand the role of rehabilitation in improving cognitive and psychosocial functions post-TBI in the pediatric population.

Wang C. found the use of metaphorical intervention relative to pediatrics games as a way to improve psychosocial abilities by reducing anxiety and depression and significantly improving the cognitive scores in children post-TBI.[19] Sood NT, *et al*. studied the effectiveness of Cogmed Working Memory Training during cognitive rehabilitation of children with TBI and supported the need for further studies to govern its efficacy for betterment of academic skills and quality of life of children with TBI.[20] Sargénius HL, *et al*. studied the effectiveness of Goal Management Training and pediatric psychoeducative control treatment in pediatric ABI using 2 2-year follow-up of RCT and extended continued executive functioning improvement from baseline.[21] On ZX, *et al*. conducted a systematic review and meta-analysis and found significant impairment in social cognition following TBI in children, mainly in 5 dimensions, including emotion recognition or perception, theory of Mind (ToM), pragmatic language, moral reasoning, and social problem solving.[22]

Polinder S, *et al*. presented a comprehensive approach for TBI using management of post-concussion symptoms. For psycho-cognitive management, they advocated the use of early educational treatment, CBT, to improve quality of life and neuropsychological rehabilitation to improve coping and cognitive regulation.[23] Seguin M, *et al*. studied the efficacy of an intensive attention training program for cognitive remediation following pediatric TBI. They supported the use of a combination of insight training and process-specific strategies for cognitive remediation.[24]

Ryan NP. found the correlation between social problems following pediatric TBI and family environment and emphasized improving family dynamics and coping strategies to improve socialization after TBI.[25]

Barman A. reported the use of various rehabilitation strategies for cognitive remediation following pediatric TBI, such techniques included Attention Process Training (APT) and compensatory techniques like memory notebook for attention, restorative techniques like word list, visual imagery, mnemonic strategies for memory, Virtual reality games for visuospatial perception and constraint induced aphasic treatment and Transcranial Direct Current Stimulation (tDCS) for language and communication, metacognitive training methods for executive functions. They also advocated the role of positive family environment on behaviour after pediatric TBI.[14]

Brown FL, *et al.* found parent skill training and acceptance and commitment therapy important to improve family bonding following childhood ABI.[26] Coyne JH. supported the use of retrieval practice to improve academic outcomes and learning difficulties following TBI.[27] Cook LG. compared the top-down gist training and bottom-up rote memory learning in adolescents with chronic stage TBI, finding the top-down approach better than the latter.[28] Pacini AG, *et al*. studied context-sensitive Goal Management Training to improve ADL functioning following severe TBI; this technique was found significant but needed more generalization.[29]

Bower J, *et al*. supported the use of familiar music therapy to evoke patient's interest towards intervention and further improving the cognitive abilities in the acute phase post childhood TBI.[30] Haslam C, *et al*. applied the principles of errorless learning techniques to pediatric TBI but found no significant results for the application of the technique; hence, further studies may be needed to apply this technique in practice.[31]

Ho J, *et al*. applied diary training, self-instruction training, and case examples on children with ABI having memory deficits. She found the technique effective in improving memory deficits and the mental well-being of children with ABI.[32] Chan DYK, *et al*. studied the effect of problem-solving skills training to improve ADL functioning in students with ABI attending mainstream schools. Intervention included paying attention, remembering and organising, defining the problem and goal setting, planning and monitoring.[15]

Feeney TJ. suggested the application of a support-oriented approach that reduces problematic behaviours through the integration of positive behaviour therapies and reinforcements with adaptable, dependent on context, cognitive/behavioural scripts.[33] Yeates KO, *et al*. in their study with TBI children and orthopedic injuries reported the dynamics of family in moderating psychological and social outcomes. They suggested social competence was significantly different between both groups, better home environments led to safer conditions following TBI, and behaviour problems were pronounced with severe TBI, irrespective of the parenting styles.[34]

This review of studies highlights that children with TBI experience various behavioural and cognitive deficits post-injury; a positive family environment, along with directed cognitive remediation therapy, is needed to provide better outcomes for the children as well as improve the well-being of caregivers.

Limitations

Variability in study designs, terminology, and measurement instruments among included studies may be one of the study's shortcomings, making comparisons difficult. Generalizability may be limited by the small sample sizes in certain studies and the variety of juvenile TBI populations (e.g., age, damage severity). Furthermore, many studies only provide short-term follow-ups and lack standardized rehabilitation techniques, which may limit the capacity to evaluate long-term efficacy. The review is made more difficult by reporting bias and the difficulty of assessing behavioural outcomes and social cognition. These elements emphasize the necessity for larger-scale, more regular studies in this field.

Conclusion

The substantial influence of rehabilitation on enhancing social cognition and behavioural outcomes in children and adolescents with traumatic brain injury (TBI) is highlighted

in this integrated review. The results imply that specific rehabilitation techniques can improve social functioning, emotional control, and adaptive behaviours, especially those that emphasize cognitive-behavioural therapy, social skills training, and neurocognitive therapies. Further study is required to refine therapeutic techniques and determine the most effective interventions for various age groups and severity levels of TBI, even if the evidence supports the effectiveness of rehabilitation in this setting. Improving long-term results in this population requires a multidisciplinary approach and the incorporation of tailored, evidence-based treatment strategies.

Further research into the neurobiological causes of TBI and how it affects social cognition is essential to developing rehabilitation strategies and raising the standard of living for young TBI patients.

Ethical approval

This self-study was conducted in line with the *Declaration of Helsinki* and ICMJE guidelines. As it involved only the researcher's personal reflections with no participant data, formal ethics 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 author declares no conflicts of interest related to this work.

Credit authorship contribution statement

Gupta Bandita: conception and design of the study, data collection, initial drafting, and data analysis of the manuscript; **Upadhyay Ishika**: interpretation and critical revision of the manuscript for important intellectual content. Both authors approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Acknowledgements

The author would like to express sincere gratitude to all those who contributed to the completion of this work. Special thanks are extended to colleagues, mentors, and peers for their valuable guidance, constructive feedback, and encouragement throughout the process. The author also acknowledges the support of family and Mr Tarun Joshi for their patience and motivation.

References

[1] Gardner MT, O'Meara AM, Miller Ferguson N. Pediatric traumatic brain injury: an update on management. Curr Pediatr Rep. 2017; 5(4): 213-9. doi: 10.1007/s40124-017-0144-5.

[2] Anderson V, Moore C. Age at injury as a predictor of outcome following pediatric head injury: a longitudinal perspective. Child Neuropsychol. 1995; 1(3): 187-202. doi: 10.1080/09297049508400224.

[3] Li L, Liu J. The effect of pediatric traumatic brain injury on behavioural outcomes: a systematic review. Dev Med Child Neurol. 2013; 55(1): 37-45. doi: 10.1111/j.1469-8749.2012.04414.x.

[4] Langlois JA, Rutland-Brown W, Thomas KE. Traumatic brain injury in the United States: emergency department visits, hospitalizations, and deaths. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control; 2006. Available from: https://stacks.cdc.gov/view/cdc/12294

[5] Bechtel K, Stoessel K, Leventhal JM, et al. Characteristics that distinguish accidental from abusive injury in hospitalized young children with head trauma. Pediatrics. 2004; 114(1): 165-8. doi: 10.1542/peds.114.1.165.

[6] Billmire ME, Myers PA. Serious head injury in infants–accident or abuse? Pediatrics. 1985; 75(2): 340-2. doi: 10.1007/BF00335074 .

[7] Laatsch L, Harrington D, Hotz G, Marcantuono J, Mozzoni MP, Walsh V, et al. An evidence-based review of cognitive and behavioural rehabilitation treatment studies in children with acquired brain injury. J Head Trauma Rehabil. 2007; 22(4): 248-56. doi: 10.1097/01.HTR.0000281841.92720.0a.

[8] Fuentes MM, Wang J, Haarbauer-Krupa J, Yeates KO, Durbin D, Zonfrillo MR, et al. Unmet rehabilitation needs after hospitalization for traumatic brain injury. Pediatrics. 2018; 141(5): e20172859. doi: 10.1542/peds.2017-2859.

[9] Laatsch L, Dodd J, Brown T, Ciccia A, Connor F, Davis K, et al. Evidence-based systematic review of cognitive rehabilitation, emotional, and family treatment studies for children with acquired brain injury literature: From 2006 to 2017. Neuropsychol Rehabil. 2020; 30(1): 130-61. doi: 10.1080/09602011.2019.1678490.

[10] DeLisa JA, Gans BM, Walsh NE. Physical medicine and rehabilitation: principles and practice. 4th Ed. Philadelphia: Lippincott Williams & Wilkins; 2005.

[11] Köhler R, Wilhelm EE, Shoulson I. Cognitive Rehabilitation Therapy for Traumatic Brain Injury: Evaluating the Evidence. Washington, DC: National Academies Press; 2012. doi: 10.17226/13220.

[12] Schretlen DJ, Shapiro AM. A quantitative review of the effects of traumatic brain injury on cognitive functioning. Int Rev Psychiatry. 2003; 15(4): 341-9. doi: 10.1080/09540260310001606728.

[13] Vaishnavi S, Rao V, Fann JR. Neuropsychiatric problems after traumatic brain injury: unraveling the silent epidemic. Psychosomatics. 2009; 50(3): 198-205. doi: 10.1176/appi.psy.50.3.198.

[14] Barman A, Chatterjee A, Bhide R. Cognitive impairment and rehabilitation strategies after traumatic brain injury. Indian J Psychol Med. 2016; 38(3): 172-81. doi: 10.4103/0253-7176.183086.

[15] Chan DY, Fong KN. The effects of problem-solving skills training based on metacognitive principles for children with acquired brain injury attending mainstream schools: a controlled clinical trial. Disability and Rehabilitation. 2011; 33(21-22): 2023-

32. doi: 10.3109/09638288.2011.556207.
[16] Hanten G, Bartha M, Levin HS. Metacognition following pediatric traumatic brain injury: a preliminary study. Dev Neuropsychol. 2000;18(3):383–98. doi: 10.1207/S1532694206Hanten
[17] Popernack ML, Gray N, Reuter-Rice K. Moderate-to-severe traumatic brain injury in children: complications and rehabilitation strategies. J Pediatr Health Care. 2015; 29(3): e1-7. doi: 10.1016/j.pedhc.2014.09.003.
[18] Resch C, Rosema S, Hurks P, de Kloet A, van Heugten C. Searching for effective components of cognitive rehabilitation for children and adolescents with acquired brain injury: a systematic review. Brain Inj. 2018; 32(6): 679-92. doi: 10.1080/02699052.2018.1458335.
[19] Wang C, Li X, Liu Y. Influence of play-based metaphors on negative emotion and behaviour in children recovering from brain injury. Chinese J Child Health Care. 2024; 32(5): 566-71. doi: 10.11852/zgetbjzz2023-0742.
[20] Sood NT, Godfrey C, Krasts D, Morrison E, Chavez Arana C, Hearps SJ, et al. Rehabilitation of Executive Function in Pediatric Traumatic Brain Injury (REPeaT): outcomes of a pilot randomized controlled trial. Neuropsychology. 2024; 38(5): 392-402. doi: 10.1037/neu0000951.
[21] Sargénius HL, Andersson S, Haugen I, Hypher R, Brandt AE, Finnanger TG, et al. Cognitive rehabilitation in paediatric acquired brain injury—a 2-year follow-up of a randomized controlled trial. Front Neurol. 2023; 14: 1173480. doi: 10.3389/fneur.2023.1173480.
[22] On ZX, Ryan NP, Konjarski M, Catroppa C, Stargatt R. Social cognition in paediatric traumatic brain injury: A systematic review and meta-analysis. Neuropsychol Rev. 2022;32(4):924-945. doi:10.1007/s11065-022-09553-1
[23] Polinder S, Cnossen MC, Real RGL, Covic A, Gorbunova A, Voormolen D, et al. A multidimensional approach to post-concussion symptoms in mild traumatic brain injury. Front Neurol. 2018; 9: 1113. doi: 10.3389/fneur.2018.01113.
[24] Séguin M, Lahaie A, Matte-Gagné C, Beauchamp MH. Ready! Set? Let's Train!: Feasibility of an intensive attention training program and its beneficial effect after childhood traumatic brain injury. Ann Phys Rehabil Med. 2018; 61(4): 189-96. doi: 10.1016/j.rehab.2017.05.001.
[25] Ryan NP, van Bijnen L, Catroppa C, Beauchamp MH, Crossley L, Hearps SJ, et al. Longitudinal outcome and recovery of social problems after pediatric traumatic brain injury (TBI): Contribution of brain insult and family environment. Int J Dev Neurosci. 2016; 49: 23-30. doi: 10.1016/j.ijdevneu.2015.12.004.
[26] Brown FL, Whittingham K, Boyd RN, McKinlay L, Sofronoff K. Does Stepping Stones Triple P plus Acceptance and Commitment Therapy improve parent, couple, and family adjustment following paediatric acquired brain injury? A randomised controlled trial. Behav Res Ther. 2015; 73: 58-66. doi: 10.1016/j.brat.2015.07.001.
[27] Coyne JH, Borg JM, DeLuca J, Glass L, Sumowski JF. Retrieval practice as an effective memory strategy in children and adolescents with traumatic brain injury. Arch Phys Med Rehabil. 2015; 96(4): 742-5. doi: 10.1016/j.apmr.2014.09.022.
[28] Cook LG, Chapman SB, Elliott AC, Evenson NN, Vinton K. Cognitive gains from gist reasoning training in adolescents with chronic-stage traumatic brain injury. Front Neurol. 2014; 5: 87. doi: 10.3389/fneur.2014.00087.
[29] Krasny-Pacini A, Limond J, Evans J, Hiebel J, Bendjelida K, Chevignard M. Context-sensitive goal management training for everyday executive dysfunction in children after severe traumatic brain injury. J Head Trauma Rehabil. 2014; 29(5): E49-64. doi: 10.1097/HTR.0000000000000015.
[30] Bower J, Catroppa C, Grocke D, Shoemark H. Music therapy for early cognitive rehabilitation post-childhood TBI: an intrinsic mixed methods case study. Dev Neurorehabil. 2014; 17(5): 339-46. doi: 10.3109/17518423.2013.778910.
[31] Haslam C, Bazen-Peters C, Wright I. Errorless learning improves memory performance in children with acquired brain injury: a controlled comparison of standard and self-generation techniques. Neuropsychol Rehabil. 2012; 22(5): 697-715. doi: 10.1080/09602011.2012.686820.
[32] Ho J, Epps A, Parry L, Poole M, Lah S. Rehabilitation of everyday memory deficits in paediatric brain injury: self-instruction and diary training. Neuropsychol Rehabil. 2011; 21(2): 183-207. doi: 10.1080/09602011.2010.547345.
[33] Feeney TJ. Structured flexibility: the use of context-sensitive self-regulatory scripts to support young persons with acquired brain injury and behavioural difficulties. J Head Trauma Rehabil. 2010; 25(6): 416-25. doi: 10.1097/HTR.0b013e3181fbc0a2.
[34] Yeates KO, Taylor HG, Walz NC, Stancin T, Wade SL. The family environment as a moderator of psychosocial outcomes following traumatic brain injury in young children. Neuropsychology. 2010; 24(3): 345-56. doi: 10.1037/a0018387.

Journal of Associated Medical Sciences 2025; 58(3): 329-334

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

Normative nasalance in Indians speaking Malayalam and Kannada: Influence of native language, gender, and vowels

Anjitha A Prabhu and Deepthi KJ*

Nitte (Deemed to be University), Nitte Institute of Speech and Hearing, Mangalore, India.

ARTICLE INFO

Article history:
Received 16 April 2025
Accepted as revised 18 August 2025
Available online 26 August 2025

Keywords:
Malayalam, Kannada, Nasalance Normative.

ABSTRACT

Background: Normative nasalance values need to be updated frequently to account for dialectal, regional, and linguistic variations across populations. Regular updates ensure the data remains relevant and accurate for clinical assessment and diagnosis.

Objectives: The primary objective of the current study was to establish normative nasalance for adults who are native speakers of Malayalam and Kannada using the NasalView system. Furthermore, this study aimed to investigate the influence of language, gender, and stimuli on the nasalance scores.

Materials and methods: The study recruited 80 native Malayalam-speaking adults (40 males and 40 females) and 80 native Kannada-speaking adults (40 males and 40 females) between 18 to 24 years of age. The nasalance score were obtained with the NasalView system while the participants phonated vowels /a/, /i/, and /u/. Three trials were conducted for each vowel. The obtained data were tabulated for further analysis.

Results: The study revealed significant differences in nasalance values for the vowels /i/ and /u/ between genders and across languages, but not for /a/. The results further revealed that /i/ had the highest nasalance, followed by by/u/ and /a/. Nasalance was significantly different between /i/ and /a/ and between /i/ and /u/ but not between /u/ and /a/.

Conclusion: The study has been able to establish nasalance normative for two major Dravidian languages of India that can be used for the clinical diagnosis and management of resonance disorders. Further, it sheds light on the impact of nasalance in geographically diverse multilingual situations.

Introduction

The velopharyngeal (VP) port controls nasality and supports clear speech. Its function is commonly evaluated perceptually using rating scales for hypernasality and hyponasality.[1] Perceptual assessment of VP function is often supported by objective measures, most commonly by Nasometer™ (Kay Elemetrics, Lincoln Park, NJ, USA)[2] and NasalView™ (Tiger Electronics, Seattle, WA, USA), developed by Awan.[3] Both instruments assess nasalance by measuring nasal and oral sound energy. The formula for nasalance is as follows: nasalance = (nasal energy / [(nasal energy + oral energy)] × 100. This equation provides the percentage of nasal acoustic energy in relation to the total acoustic energy radiating in the nasal and oral cavities.[4,5] Though both instruments measure nasalance, their values aren't directly comparable due to differences in hardware, software, and filter settings.[6]

* **Corresponding contributor.**
Author's Address: *Nitte (Deemed to be University), Nitte Institute of Speech and Hearing, Mangalore, India.*
E-mail address: deepthi@nitte.edu.in
doi: 10.12982/JAMS.2025.104
E-ISSN: 2539-6056

Nasalance estimations using either of these instruments have been documented for a few languages spoken in India, such as Kannada, Hindi, and Tamil.[7-13] However, only a limited number of studies have reported normative data for nasalance, with most of these publications being nearly a decade old.[9-12, 14] Most of these studies used a Nasometer for nasalance measurement. Normative data thus established vary, with some studies focussing on children and others on adults, making its interpretation and application challenging.

Although Nasometer and NasalView are the most commonly used instruments currently for nasalance measurement, however both differ in their signal processing algorithms, which may account for differences in nasalance scores. The Nasometer employs a bandpass filter centered around 500 Hz, analyzing a limited portion of the speech spectrum, whereas the NasalView system analyzes the full frequency range without such filtering.[15] These methodological differences, along with variations in software resolution, affect the sensitivity and comparability of measurements across devices, reinforcing the need for device-specific normative data.

The NasalView system is gaining recognition as a cost-effective, reliable tool for measuring nasalance, with unique filters setting it apart from the Nasometer. While its affordability supports wider clinical use in India, most existing norms are based on Nasometer data. The lack of normative values for Indian languages in NasalView limits accurate interpretation. There is a pressing need to establish nasalance norms for different Indian languages. Existing values, mainly derived using the Nasometer nearly a decade ago, cannot be directly applied to NasalView data due to system differences. Nasalance measures are vital for diagnosing velopharyngeal dysfunction in individuals with cleft, craniofacial anomalies, dysarthria, and hearing loss. Hence, developing normative scores using the NasalView system is essential to improve diagnostic and treatment outcomes. This study aimed to establish language- and gender-specific nasalance norms for native Malayalam and Kannada speakers aged 18-24, and to compare values across languages, genders, and vowels.

Materials and methods

The study was initiated after obtaining ethical approval from the Institutional Ethics Committee (Ref. No. EC/NEW/INST/2022/KA/0174), and all procedures were conducted in accordance with the principles outlined in the Declaration of Helsinki. A total of 160 participants were included; 80 native Malayalam speakers (40 males, 40 females) and 80 native Kannada speakers (40 males, 40 females), aged between 18 and 24 years. Participants were recruited through convenience sampling from a multi-specialty hospital in India between September 2023 and June 2024. All participants were neurotypical and anatomically typical individuals within the specified age range. Informed consent was obtained prior to participation. Individuals with a history of oral-pharyngeal or craniofacial surgeries, recent respiratory infections, or syndromic conditions such as hearing loss or cognitive deficits were excluded from the study.

Nasometry procedure

Nasalance was measured using NasalView system v1.2 (DRS Inc.) installed on a personal computer equipped with an Intel Pentium IV processor and a PCI-28 sound card. Recordings took place in a sound-treated room at a 22.05 kHz sampling rate and 44,100 Hz frequency. Participants were seated upright, and the NasalView headgear was fitted as per guidelines. Following hygiene protocols, participants sustained vowels /a/, /i/, and /u/ at habitual pitch and loudness for as long as possible, with three trials per vowel. The mid, stable 5-second segment from each trial was analyzed, and mean nasalance values were calculated. This process was repeated for all vowels. Final data were analyzed using SPSS version 25.[16]

Results

Descriptive statistics were computed for the data, as shown in Table 1. Across both genders and languages, the vowel /i/ exhibited the highest mean nasalance value. The nasalance values for /a/ were similar to those of /u/, a pattern consistent across both male and female participants and across native speakers of Malayalam and Kannada.

Table 1. Mean (and SD) of Nasalance scores (%) for vowels /a/, /i/, and /u/ by language and gender.

Languages	Gender	N	Age (years)	Nasalance /a/	Nasalance /i/	Nasalance /u/
			Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Malayalam	Female	40	21.77 (2.52)	26.00 (6.77)	42.10 (7.55)	28.51 (6.16)
	Male	40	22.63 (1.75)	26.36 (7.34)	35.16 (5.38)	25.75 (4.10)
Kannada	Female	40	22.78 (1.23)	26.45 (6.10)	38.81 (6.55)	25.88 (3.27)
	Male	40	23.54 (1.07)	24.02 (4.99)	31.72 (5.86)	24.63 (4.73)

Subsequently, the Shapiro-Wilk test was used to assess the normality of the data, which indicated that the data did not follow a normal distribution (p<0.05). Therefore, the Mann-Whitney U test was used for pairwise comparisons of nasalance values across languages and genders. The results are presented in Table 2. As shown in Table 2, the analysis revealed significant differences in the nasalance values for the vowels /i/ and /u/ between genders and across languages but not for vowel /a/.

Table 2. Results of Mann-Whitney U tests comparing nasalance scores across gender and language (Significant values are indicated with asterisks).

Comparison	Vowel	*p* value
Gender	/a/	0.141
	/i/	0.000**
	/u/	0.004*
Language	/a/	0.610
	/i/	0.005*
	/u/	0.038*

***Note**: Significant values are indicated with asterisks).*statistically significant at p<0.05; **statistically highly significant at p<0.001.*

To examine the rank order of nasalance values for the vowels, a Friedman test was conducted. The analysis revealed that the vowel /i/ had the highest mean rank (2.93), vowel /u/ had the second highest mean rank (1.58), and vowel /a/ had the lowest mean rank (1.49) as shown in Table 3. A significant difference in nasalance values was observed among the three vowels (/a/, /i/, and /u/), with $p<0.001$.

Table 3. Mean ranks of nasalance values for vowels /a/, /i/, and /u/ based on Friedman test.

	Mean rank	N	160
Nasalance	1.49	Chi-Square	208.838
Nasalance	2.93	Df	2
Nasalance	1.58	Asymp Sig.	0.000

***Note**: The Friedman test revealed a statistically significant difference in Nasalance values across vowels, $\chi^2(2)=208.838$, p<0.001. Higher mean rank indicates higher Nasalance.*

The Wilcoxon Signed-Rank test was used as a post hoc analysis following the Friedman test to maintain statistical rationale and clarity. The Wilcoxon Signed-Rank Test revealed a statistically significant difference in nasalance between the vowels /i/ and /a/ ($p<0.05$), as well as between /i/ and /u/ ($p<0.05$). In contrast, no significant difference was found between the nasalance values of /u/ and / a/ ($p > 0.05$).

Discussion

India is a highly multilingual nation, comprising 22 scheduled languages and approximately 100 non-scheduled languages, including English. These 22 scheduled languages collectively encompass 243 distinct languages, while languages spoken by fewer than 10,000 individuals are not included in this enumeration.[17] Although Hindi accounts for a significant proportion of speakers (41.03%), Dravidian languages such as Kannada and Malayalam—each spoken by around 3-4% of the population—exhibit unique phonological characteristics of both linguistic and clinical importance. Despite their relevance, these languages remain underrepresented in nasalance research using standardized tools such as the NasalView system. The inclusion of Kannada and Malayalam in the present study helps fill a critical gap by generating normative nasalance data for linguistically diverse, yet underrepresented, Indian languages. Many languages within India's multilingual framework possess rich oral traditions that warrant focused empirical inquiry. By contributing data from two distinct Dravidian languages, this study offers a stepping stone for future nasometric research to encompass a wider range of Indian languages, supported by evolving speech assessment technologies.

Influence of stimuli and instrument

The current study established language- and gender-specific nasalance norms for native Malayalam and Kannada speakers aged 18-24 using the NasalView system. Comparison across languages, genders, and vowels showed significant differences for /i/ and /u/, but not for /a/. These results differ from the previous studies, who found language- and gender-based differences using different stimuli and the Nasometer by Kay Elemetrics.[8,14] One study reported higher nasalance in Malayalam speakers, while the other found gender differences using word and sentence stimuli.[8,14] The variation noted between the current and these previous studies can be due to differences in both stimuli and instruments used by them. The two instruments i.e., Nasometer and NasalView, differ in their signal processing algorithms, filter settings, and calibration procedures, which can impact the magnitude and sensitivity of nasalance measurements. Hence, instrument-related variability is likely a major contributor when comparing results across studies using different systems.[15]

The absence of gender differences for the stimuli /a/ in the current study may also be attributed to instrumental variations. Unlike earlier studies that used the Nasometer, this study employed the NasalView, which differs in filter settings and analytical algorithms. These hardware and

software differences could influence the detection of gender-based variations in nasalance.

Moreover, physiological variables—such as shorter vocal tract length in females, differences in fundamental frequency (F_0), and nasal airflow patterns—as well as methodological factors like stimulus type, number of repetitions, and experimental context, have all been shown to influence the sensitivity to gender-related differences in nasalance.[18-20] Additionally, instrumental settings, particularly differences in microphone sensitivity, filter algorithms, and calibration between devices like the NasalView and Nasometer, may affect nasalance measurement outcomes.[1,15]

In the present study, although gender differences emerged clearly for high vowels /i/ and /u/, no significant gender effect was observed for the low vowel /a/. This null result may be attributed to the open articulation and lower inherent vowel height of /a/, which leads to less oral constriction and more stable resonance, making it potentially less sensitive to subtle variations in velopharyngeal activity.[21] Furthermore, the NasalView system may be less sensitive in detecting such small differences compared to the Nasometer, especially for acoustically simpler vowels like /a/.[1,15]

Interestingly, despite Malayalam being nasal-dominant, it did not significantly affect the nasalance of /a/. Further, no measurable difference was found between Kannada and Malayalam speakers for this vowel. This finding stands in contrast to that of Girish and colleagues who reported differing nasalance values using the Nasometer.[21] In the present study, we propose that the lack of significant variation for the vowel /a/ may be influenced by India's unique sociolinguistic context—particularly macro-level multilingualism and individual-level (micro) multilingual experiences, which are further discussed in the following paragraphs.

Influence of India's macro-level multilingualism

India's multilingual landscape is globally unique and unparalleled. The Constitution of India recognizes 22 Scheduled Languages, reflecting the nation's rich linguistic diversity. According to the Indian Census of 2001, Kannada is the native language of 3.69% of India's population, while Malayalam is spoken by 3.21%.[17,22] Among Kannada speakers, 28.44% are bilingual, and 12.15% are trilingual, whereas 28.75% of Malayalam speakers are bilingual, and 16.60% are trilingual. A common trend among multilingual speakers of Kannada and Malayalam is the preference for English and Hindi as additional languages. Studies show English is often the first choice, followed by Hindi.[17]

Participants were primarily university students from English-medium schools, and Hindi was typically acquired through formal education or media exposure. Additionally, India's multilingual environment likely facilitated informal acquisition of other regional languages through family, social interactions, and media. This broad language exposure may reflect macro-level multilingualism, where multiple languages subtly influence each other. In this context, it is possible that the native language exerted a limited effect on the nasalance of /a/. These findings may be consistent with earlier research suggesting that multilingualism can influence aspects of a speaker's native language, including phonetic and prosodic features such as nasalance.[23,24] Thus, we hypothesize that India's macro-level multilingualism could have influenced vowel nasalance in our participants, highlighting the fluid and dynamic nature of speech production in a multilingual society.

Influence of India's micro-level multilingualism

Beyond national-level multilingualism, India's micro-level multilingualism may also explain the lack of nasalance differences for /a/ between Kannada and Malayalam speakers. The study was conducted in Mangalore, a linguistically diverse border city where multiple languages coexist. Daily exposure to Kannada, Malayalam, Tulu, Konkani, Byari, Hindi, and English fosters constant language switching across contexts—work, home, and social settings.

Such multilingual immersion promotes linguistic adaptation, influencing diction, prosody, and resonance. This ongoing interplay shapes phonetic planning and execution, affecting vocal tract resonance and, consequently, nasalance. In this study, /a/ showed similar nasalance across groups, likely due to its simple articulation. Vowels like /i/ and /u/, which demand more articulatory precision, may be less sensitive to linguistic influence, accounting for their observed variability. Thus, the local multilingual environment significantly impacts speech production and nasalance.

Limitations

One limitation of the present study is the use of only three vowels as stimuli. This approach was chosen to obtain nasalance measures from acoustically stable and coarticulation-free speech samples. While vowels are commonly used in nasometric evaluations due to their sustained nature and ability to reflect oral-nasal resonance more clearly, this limits the generalizability of findings to more complex speech. The study did not include words, sentences, or paragraphs, as standardized and validated linguistic materials in Kannada and Malayalam are currently unavailable. As a result, the absence of speech segments containing voiced consonants or connected speech may have restricted the ability to fully capture nasalance characteristics in more naturalistic speaking contexts. Larger sample sizes and detailed language background profiling in future studies may help strengthen the validity and clarity of the results.

Conclusion

The study establishes normative nasalance values for vowels /a/, /i/, and /u/ in native Malayalam and Kannada speakers across genders. Unlike previous studies, /a/ showed no significant variation by language or gender, while /i/ and /u/ did. These results highlight that nasalance is shaped by more than anatomy or equipment—it's also shaped by linguistic context. Factors like multilingual

exposure, gender, instrument used, and stimulus type all play a role. Hence, clinicians should adopt an individualized approach that considers a speaker's linguistic and environmental background for accurate diagnosis.

Ethical approval

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Institution Ethics Committee (Date: 30th August, 2023; INST. EC/EC/229/2023; Reg. No: EC/NEW/INST/2022/KA/0174).

Consent

Informed consent was obtained from all individual participants included in the study.

Data

Data are available from the corresponding author upon reasonable request.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict to interest.

CRedit authorship contribution statement

All authors contributed to the study's conception and design. **Anjitha A Prabhu**: material preparation, data collection, data analysis, first draft of the manuscript preparation and comment; **Deepthi KJ**: conceptualized the study, planned the methodology, supervised the research, contributed to result interpretation, and provided critical revisions to the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors sincerely appreciate the invaluable guidance and support provided by the senior faculty members of our organization throughout the preparation of this manuscript. We also extend our heartfelt gratitude to the institution's administration for their unwavering support.

References

[1] Henningsson G, Kuehn DP, Sell D, Sweeney T, Trost-Cardamone JE, Whitehill TL. Universal parameters for reporting speech outcomes in individuals with cleft palate. Cleft Palate Craniofac J. 2008; 45(1) :1-17. doi: 10.1597/06-086.1

[2] Fletcher SG, Frost M, Druecke P. Measurement of nasality with the Nasometer. Cleft Palate J. 1989; 26: 1-5.

[3] Awan SN, Ziegler W, Deger K. Analysis of nasalance: NasalView. In: Clinical Phonetics and Linguistics. London, Whurr: 1997; pp 518-25.

[4] Fletcher SG. Theory and instrumentation for quantitative measurement for nasality. Cleft Palate J. 1970; 7(2): 601-9.

[5] Fletcher SG. "Nasalance" vs. listener judgements of nasality. Cleft Palate J. 1976; 13(1): 31-44.

[6] Sudarshan Reddy BM, Sheela S, Kishore Pebbili G. Comparison of nasalance values obtained from nasality visualization system and Nasometer II. J All India Inst Speech Hear. 2012; 31: 30-3.

[7] Deepthi KJ, Pushpavathi M. The effect of vowels on the nasalance values in Kannada-speaking preschoolers with repaired cleft of lip and palate. Int J Interdiscip Res Innov. 2018; 6(3): 383-8. Available from: www.researchpublish.com

[8] Girish KS, Pushpavathi M, Satish HV. Influence of native language on nasalance values in Kannada and Malayalam speakers. J Cleft Lip Palate Craniofac Anomal. 2021; 8(2): 111-8.

[9] Jayakumar T, Pushpavathi M. Normative score for Nasometer in Kannada. Student Res AIISH. 2005; 7: 44-53.

[10] Arya P, Pushpavathi M. Normative nasalance value in Hindi language [Master's dissertation]. Mysore: University of Mysore; 2009. (Unpublished, personal communication).

[11] Sunitha, Roopa N, Prakash B. Proceedings of 4th Annual Congress of Indian Society of Cleft Lip, Palate & Craniofacial Anomalies (Abstract). Proc 4th Annu Congr Indian Soc Cleft Lip Palate Craniofac Anomal. 1994.

[12] Pokharel A, Naina P, Sebastian S, Syed K, John M, Varghese A. Normative nasalance scores in Tamil-speaking Indian children. Logoped Phoniatr Vocol. 2020; 47: 1-6. doi: 10.1080/14015439.2020.1849391

[13] Jayakumar T, Soonan VV, Thankamany VR, Benoy JJ. Nasalance, nasal airflow and perceived nasality in carnatic singers and non-singers. Indian J Otolaryngol Head Neck Surg. 2024; 76(1): 5-18. doi: 10.1007/s12070-023-03994-w.

[14] Kuppusamy G, Sankar G, Mariswamy P. Normative nasalance values across stimuli and gender in Malayalam-speaking individuals. Otolaryngol Online J. 2013; 3: 85-98.

[15] Bressmann T, Sader R, Whitehill TL, Samman N, Marciniak C. Comparison of nasalance scores obtained with the Nasometer, the NasalView, and the OroNasal System. Cleft Palate Craniofac J. 2005; 42(4): 423-9. doi: 10.1597/03-159.1

[16] IBM Corp. IBM SPSS Statistics for Windows, Version 25.0. Armonk (NY): IBM Corp; 2017.

[17] Mallikarjun B. Patterns of Indian multilingualism. Lang India [Internet]. 2010 Jun;10(6). Available from: http://www.languageinindia.com

[18] Mishima K, Mori Y, Sugahara T, Minami K. Gender and dialectal differences in nasalance scores in a Japanese population. J Craniomaxillofac Surg. 2008; 36(3): 146-9. doi: 10.1016/j.jcms.2007.09.005

[19] Kim HK, Park IS, Yoon HY. Dialectal and gender differences in nasalance for a Mandarin population. Clin Linguist Phon. 2016; 30(6): 451-64. doi: 10.3109/02699206.2016.1142591

[20] Ibrahim HM, Hamzah HS, Nordin NAM, Ahmad R, Abdul Wahab N. Speech stimuli and nasalance scores among Mandarin-speaking Malaysian children. Clin Linguist Phon. 2020; 34(12): 1063-74. doi: 10.1080/02699206.2020.1793815

[21] Zajac DJ, Mayo R. Aerodynamic and acoustic measures of velopharyngeal function in normal speakers. Cleft Palate Craniofac J. 1996; 33(3): 235-41. doi: 10.1597/1545-1569_1996_033_0235_aaamov_2.3. co_2

[22] Office of the Registrar General & Census Commissioner, India. Census 2001 data online [Internet]. India.Gov.in National Portal of India; 2001 [cited 2025 Aug 20]. Available from: https://www.india.gov.in/census-2001-data-online

[23] Kroll JF, Dussias PE. The benefits of multilingualism to the personal and professional development of residents of the US. Foreign Lang Ann. 2017; 50(2): 248-59.

[24] Han Q, Tian JX, Chen HC. L3 prosody: Cross-linguistic influence of prosodic features in Mandarin and English by Cantonese multilinguals. In: L3 Development After the Initial State. Amsterdam: John Benjamins Publishing Company; 2023. pp 96-120.

Journal of Associated Medical Sciences 2025; 58(3): 335-344

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 and validity of the Thai version of the 2007 revised Developmental Coordination Disorder Questionnaire

Natewimon Nantiwat, Sureelak Sutchritpongsa, Amornrat Penphattrarakul, Prakasit Wannapaschaiyong, Nattha Ketudat, Pat Rojmahamongkol*

Division of Developmental and Behavioral Pediatrics, Department of Pediatrics, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 22 February 2025
Accepted as revised 18 August 2025
Available online 26 August 2025

Keywords:
Developmental coordination disorder, DCD, developmental coordination disorder questionnaire, DCDQ.

ABSTRACT

Background: Developmental coordination disorder (DCD) manifests with clumsiness and affects children's daily activities, academic achievement, peer acceptance, and self-esteem. Early identification and intervention are crucial for improving quality of life. There is no screening test for DCD in Thailand. The 2007 revised version of the Developmental Coordination Disorder Questionnaire (DCDQ'07) is a caregiver-report screening test for 5 to 15-year-old children globally used with good sensitivity and specificity.

Objectives: To translate and validate the Thai version of the DCDQ'07 (DCDQ-T), assessing its content validity, internal consistency, test-retest reliability, and inter-rater reliability.

Materials and methods: After receiving permission from the original questionnaire developer, forward and backward translation processes were performed. The content validity index (CVI), consisting of item-level CVI (I-CVI) and scale-level CVI, universal agreement method (S-CVI/UA), was evaluated by five experts. Reliability was examined by internal consistency in 88 caregivers of 6 to 12-year-old students without the diagnosis of visual impairment, intellectual disability, and neurologic problems that influence movement, in a primary school in Bangkok, Thailand. Test-retest reliability was investigated three months apart in 85 caregivers. Inter-rater reliability was investigated in 20 pairs of both parents-teachers and parents-parents.

Results: Both the I-CVI and S-CVI/UA of DCDQ-T were 1.0. The internal consistency showed a Cronbach's alpha of 0.888. The intraclass correlation coefficient of test-retest reliability was 0.573. Inter-rater reliability between parents and teachers, and parents and parents were 0.199, and 0.149, respectively.

Conclusion: DCDQ-T demonstrated an excellent CVI, good internal consistency, moderate test-retest reliability, and poor inter-rater reliability. Therefore, DCDQ-T could be a useful DCD screening tool for 6 to 12-year-old Thai children. However, a study on its repeatability and the degree of agreement among raters is further required.

Introduction

Developmental coordination disorder (DCD) is a neurodevelopmental disorder related to movement and coordination presented as clumsiness. The overall prevalence of DCD in school-aged children is 5-6%.[1,2] However, there are different results in studies from different countries. In the UK, the prevalence of DCD in 7-year-olds was 3%.[3] In Sweden, a study found a 4.9% and 8.6% prevalence of severe and moderate DCD in 7-year-old children, respectively.[4] In Thailand, a study found that 10.83 % of 3- to 6-year-old children were suspected of

* **Corresponding contributor.**
Author's Address: *Division of Developmental and Behavioral Pediatrics, Department of Pediatrics, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand.*
E-mail address: patojung@yahoo.com
doi: 10.12982/JAMS.2025.105
E-ISSN: 2539-6056

having DCD.[5] In contrast, a study in India showed only 0.8% prevalence of DCD.[6] DCD is often found to be a coexisting condition with other neurodevelopmental disorders.[7] International studies demonstrated that DCD was found in 30-50% of children and adolescents with attention-deficit/hyperactivity disorder (ADHD),[8] up to 70% of those with language problems,[9] 17.8% of those with learning disabilities,[10] and 79% of children with autism spectrum disorder.[11] A study from Thailand found a 35.3 % comorbidity of DCD in 6- to 17- year- old children and adolescents with ADHD.[12]

Poor coordination and clumsiness in DCD negatively affect the child's academic function, participation in activities of the same age group, and acceptance by friends.[13] Its impacts can be far beyond the motor domains. Research found that children with DCD demonstrated poorer self-regulation and working memory than typical children.[14,15] These disadvantages can aggravate the child's academic underachievement and relationship problems with peers; and contribute to their psychological and behavioral complications including low self competence, social isolation, anxiety, depression, and aggressive behavior.[13] Early diagnosis and prompt intervention for DCD are crucial to improve the child's quality of life since it has positive effect on motor function and skill performance, which lead to restore self-esteem and mental health status.[11]

Currently, DCD is diagnosed by the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria (Figure 1) using information from developmental history, medical history, physical examination, school report together with standard test assessment.[1,11]

To be diagnosed with developmental coordination disorder, the child must present all of the following criteria;

A.) The child's coordinated motor skill development and achievement is lower than expectation for his/her chronological age and chance for the skill implementation. This can be presented with clumsiness (e.g. dropping or hitting against things), and retard and imprecision in motor skill performance (e.g. catching a ball, writing, or riding a bicycle);

B.) The motor skill problems in criteria A obviously disturb the child's daily activities, academic productivity and leisure time activities;

C.) The motor skill difficulties began at early age; and

D.) The motor skill difficulties are not due to other conditions which are intellectual disability, visual impairment, and neurologic problems (e.g. degenerative disorder, cerebral palsy, or muscular dystrophy)

***Figure 1**. DSM-5 diagnostic criteria for developmental coordination disorder.*

The standard tests widely used for the diagnosis of DCD are the Movement Assessment Battery for Children, 3rd Edition (MABC-3)[16] which takes 30 minutes to complete, and the Bruininks Oseretsky Test of Motor Proficiency, 3rd Edition (BOT-3),[17] which takes 50-90 minutes to administer. Both tests are limited to certified medical personnel such as psychologists, occupational or speech therapists.

The Developmental Coordination Disorder Questionnaire (DCDQ) is a free online questionnaire developed by Wilson BN, et al. in 2000 and was revised in 2007 (DCDQ'07) for DCD screening in 5 to 15 years old children.[18] This 10- to 15-minute parent questionnaire is time-saving and accessible. So, it is widely used and was translated into many languages such as Turkish,[19] French European,[20] French Canadian,[21] German,[22] and Korean.[23]

As previously mentioned, DCD is prevalent in Thai children; however, the unique cultural context in Thailand—characterized by lower participation in structured physical activities and team sports[24]—may mask the early signs of DCD, thereby delaying necessary interventions. This underscores the importance of a culturally adapted screening tool that can early and accurately identify DCD in Thai children. In addition, the implementation of a valid screening instrument tailored to the Thai population would facilitate better allocation of healthcare resources, promote targeted early interventions, and ultimately contribute to improved long-term developmental outcomes.

The DCDQ'07 contains 15 items in a paper-based format that takes caregivers minimal time to complete. It is a convenient screening tool which is practical for clinical use in Thailand. Therefore, this study aimed to translate the DCDQ'07 into Thai language and to test the validity and reliability of the DCDQ-Thai version (DCDQ-T). We specifically focused on the elementary school children (6 to 12 years old) because the participants were recruited from an elementary school, which comprised students within this age range. Although the original DCDQ'07 is designed for 5- to 15-year-old children, our sample reflects the practical age range in the Thai school system where early identification of the DCD is most feasible. This study could also serve as a foundation for future research to further refine DCD screening and intervention strategies in Thai context.

Materials and methods

Study design and participants

This study was approved by the Ethics Committee of the Human Research Protection Unit at the Faculty of Medicine of Siriraj Hospital, Mahidol University, COI number Si158/2023. Informed consent was signed before collecting research data.

This questionnaire-based, cross-sectional research included caregivers of the students who studied at an elementary school (grades 1-6) in Bangkoknoi District in Bangkok, Thailand. As this elementary school included exclusively grades 1 to 6, the students' age range in this study was 6 to 12 years old. The recruitment period was from May to August 2023. Children who were diagnosed with intellectual disability, visual impairment that cannot be corrected with eyeglasses, neurologic problems such as cerebral palsy, muscular dystrophy, and degenerative disorders; and caregivers who cannot read Thai were excluded. The sample size of 87 was calculated based on the Cronbach's alpha coefficient of the original 2007 revised version of Developmental Coordination Disorder Questionnaire (DCDQ'07), which was 0.94,[18] and the confidence level of 0.95 with an error of 0.05. To minimize bias, students were included by stratified randomization from every grade of the school.

Instrument

The DCDQ'07 is developed for caregivers of 5- to 15-year-old children. It consists of 15 Likert scale questions (score of 1 = *not at all like your child* to score of 5 = *extremely like your child*) comparing children's motor abilities with their equivalent age peers. The questionnaire is divided into three parts which are motor control during movement such as ball catching (6 question items), fine motor/handwriting such as paper cutting (4 question items), and general condition such as clumsiness (5 question items). The range of the score is 15 to 75. Lower scores suggest the presence of developmental coordination disorder (DCD), while higher scores do not. The cutoff points are different between three age groups, which are 5 years to 7 years and 11 months, 8 years to 9 years and 11 months, and 10 years to 15 years. The cutoff score increases with increasing age as shown in Table 1. The DCDQ'07 shows excellent psychometric properties[18] including strong internal consistency with the Cronbach's alpha of 0.94. Favorable concurrent validity demonstrated by its statistically significant correlation with standard motor ability test (Movement Assessment Battery for Children; MABC, and Test of Visual-Motor Integration) which are r=-0.55, and r=0.42, respectively. The overall sensitivity is 84.6% and the specificity is 70.8%.[18]

In this study, the original cut-off scores (Table 1) were used as a provisional reference to provide basic guidance to caregivers whose children scored within the at-risk range. These scores were not used for diagnostic purposes but rather as an initial indicator for recommending motor and coordination skills support.

Table 1. Interpretation of the cut-off scores of the 2007 revised version of Developmental Coordination Disorder Questionnaire.

Age range (years)	Indicative or suspicious of DCD	DCD is unlikely
5 years to 7 years 11 months	15-46	47-75
8 years 0 months to 9 years 11 months	15-55	56-75
10 years 0 months to 15 years	15-57	58-75

***Note**: DCD: developmental coordination disorder.*

Procedures

Translation process and validity testing

The DCDQ'07 was translated into Thai language after obtaining permission from the original developer. The translation process was adapted from standard guidelines for cross-cultural adaptation.[25] Initially, forward translation was independently performed by two bilingual translators who are fluent in both Thai and English, one being a developmental and behavioral pediatrician who expertizes on developmental disorders and experienced in medical and psychological instruments; and the other being a professional translator experienced in working in both Thailand and other countries. Afterward, these translators collaboratively reviewed their translations in a consensus meeting to synthesize into a single preliminary Thai version. An independent bilingual translator, who is fluent in both Thai and English and certified for working as an international schoolteacher, and who had no prior knowledge of the original DCDQ'07, conducted backward translation into English. Next, an expert committee including pediatricians and occupational therapists reviewed and made the consensus to all translated versions. Finally, the backward-translated version was approved by the original developer to ensure equivalence and adherence to the original questionnaire.

The content validity was measured by three developmental and behavioral pediatricians and two occupational therapists. The content validity index (CVI) was tested for the item-level content validity index (I-CVI) and the scale-level content validity index, universal agreement method (S-CVI/UA). The I-CVI was calculated from the number of the experts who rated the item as "quite relevant" and "highly relevant" for each question item divided by the number of all experts. The range of the value of I-CVI is 0 to 1. The S-CVI/UA was computed from dividing the sum of all question items that I-CVI value is 1 by the number of all items. The S-CVI/ UA values also range from 0 to 1.

Reliability testing process

The internal consistency of the questionnaire was evaluated by Cronbach's alpha coefficient. Test-retest reliability was performed in the same parent of the entire sample at the beginning of the study and three months apart. Parents and homeroom teachers of twenty students were stratified randomized from three age groups (6 - <8, 8 - <10, and 10-12 years old) by quota which were 8, 6 and 6 students, respectively. They were evaluated for inter-rater reliability between parent-teacher and parent-parent pairs. The weight and height of the children were collected to calculate BMI and converted to the BMI z score, and then categorized as normal (BMI z score between 1SD and -2SD), overweight (BMI z score more than +1SD), and thin (BMI z score less than -2SD) according to WHO BMI-for-ages cut-offs.[26] Information was also collected about physician diagnoses of underlying medical conditions, learning disorder (LD), and attention-deficit/hyperactivity disorder (ADHD), as well as parent/teacher reports regarding the child's Thai and English literacy, and behavioral problems, including aggression and anxiety. At the end of the study, the caregivers of the children identified as potentially having DCD, based on the original questionnaire's cutoff point, received guidance in motor and coordination skills practice.

Data analysis

The data was analyzed using IBM SPSS Statistics 22 (SPSS Inc., Chicago, IL, USA). The participants' characteristics were described by descriptive statistics. The numbers and percentage were used to describe the categorial variables. Mean and standard deviation were used to present normal distributed data. Cronbach's alpha coefficient was analyzed to indicate the internal consistency of the questionnaire, which 0.9 and above reflects excellent, 0.8 - <0.9 reflects good, 0.7 - <0.8 reflects acceptable, 0.6 - <0.7 reflects questionable, 0.5 - <0.6 reflects poor, and <0.5 reflects unacceptable internal consistency, consecutively.[27] The intraclass correlation coefficient was calculated to indicate the intra-rater (test-retest) and inter-rater reliability, which 0.9-1 describes as excellent, 0.75 - <0.9 describes as good, 0.5 - <0.75 describes as moderate, and less than 0.5 describes as poor reliability, consecutively.[28] Difference of the mean of the DCDQ-T scores between boys and girls were analyzed by independent T test.

Results

Ninety caregivers completed the Thai version of the Developmental Coordination Disorder Questionnaire (DCDQ-T). Two children with the diagnosis of intellectual disability were excluded. Information from 88 participants was analyzed for internal consistency. Demographic data from these 88 participants are exhibited in Table 2. Of the 88 participants, three were withdrawn because of the refusal to fill out the questionnaire 3 months later. As a result, a total of 85 were examined for 3-month test-retest reliability. Twenty pairs of parents and teachers were investigated for inter-rater reliability between two parents and between parent and teacher.

Table 2. Demographic data of children (N=88).

Characteristics	Number (%)
Child's sex	
- Boys	43 (49)
- Girls	45 (51)
Child's age	
- 6 years 0 months to 7 years 11 months	32 (36.4)
- 8 years 0 months to 9 years 11 months	28 (31.8)
- 10 years 0 months to 12 years 11 months	28 (31.8)
Term birth	
Preterm	3 (3.4)
Birth weight (N=71)	
<2,500 gm	8 (11.3)
2,500-4,000 gm	62 (87.3)
>4,000 gm	1 (1.4)
BMI (N=61)	
- Overweight	20 (32.8)
- Normal	37 (60.7)
- Thin	4 (6.5)
Underlying conditions	
- None	74 (84.1)
- ADHD	1 (1.1)
- LD	5 (5.7)
- Medical illnesses	8 (9.1)
Behavioral problems	
- None	83 (94.3)
- Aggression	4 (4.6)
- Anxiety	1 (1.1)
Learning problems	
- None	84 (95.5)
- Thai literacy	3 (3.4)
- English literacy	1 (1.1)
Caregivers	
- Parents	67 (76.1)
- Grandparents	19 (21.6)
- Others	2 (2.3)

***Note**: ADHD: attention-deficit/hyperactivity disorder, LD: learning disorder.*

The content validity index (CVI) of DCDQ-T was reported in Table 3. At the first round, all experts rated 10 question items as "highly relevant", and 5 as "quite relevant" and suggested some minor wording adjustments. As a result, the I-CVI at the first round was 1.0. After these 5 quite relevant items were changed according to experts' suggestion, all of them were rated as "highly relevant". Some items were adapted for Thai culture. For example, the example of the activity "building a cardboard or cushion fort" in item No.6 was removed because these activities are uncommon for Thai children. For item No.14, the idiom "a bull in a China shop", which is not used in Thailand, was replaced with "clumsy movement and trend to break things" which is more straightforward for Thai parents. The final validity indexes of the content were 1.0, which indicated acceptable validity, for both I-CVI and S-CVI/UA (Table 3).

Table 3 Content validity index, Cronbach's alpha and intraclass correlation coefficients of the Developmental Coordination Disorder Questionnaire, Thai version.

DCDQ-T items	I-CVI by 5 experts		Each item score (N=88) Mean±SD	Reliability			
	1st round	2nd round		Cronbach's alpha if item deleted (N=88)	Test-retest ($ICC_{Two\text{-}way\ mixed\text{-}effects\ model}$) (N=85)	Inter-rater ($ICC_{Two\text{-}way\ random\text{-}effects\ model}$)	
						Parent-Parent (N=20)	Parent-teacher (N=20)
1	1.00	1.00	3.64 ± 0.90	0.880	0.594 (0.435-0.717)	0.290 (-0.103-0.626)	-0.126 (-0.500-0.309)
2	1.00	1.00	3.35 ± 0.89	0.879	0.355 (0.151-0.530)	-0.302 (-0.673-0.165)	-0.094 (-0.469-0.335)
3	1.00	1.00	3.14 ± 0.97	0.879	0.454 (0.264-0.610)	-0.363 (-0.728-0.112)	-0.179 (-0.584-0.284)
4	1.00	1.00	4.03 ± 0.90	0.882	0.395 (0.196-0.563)	-0.096 (-0.524-0.359)	0.078 (-0.395-0.503)
5	1.00	1.00	4.28 ± 0.93	0.883	0.412 (0.218-0.575)	0.115 (-0.289-0.506)	-0.103 (-0.535-0.355)
6	1.00	1.00	3.84 ± 1.03	0.878	0.379 (0.178-0.549)	0.074 (-0.344-0.481)	-0.004 (-0.450-0.435)
7	1.00	1.00	3.91 ± 0.98	0.876	0.371 (0.169-0.543)	0.092 (-0.346-0.502)	0.130 (-0.239-0.503)
8	1.00	1.00	3.98 ± 0.97	0.880	0.547 (0.377-0.681)	0.485 (0.057-0.760)	0.054 (-0.371-0.469)
9	1.00	1.00	3.72 ± 0.90	0.881	0.339 (0.133-0.517)	0.342 (-0.089-0.671)	-0.042 (-0.326-0.323)
10	1.00	1.00	3.72 ± 0.97	0.879	0.496 (0.314-0.642)	0.057 (-0.307-0.446)	0.230 (-0.230-0.604)
11	1.00	1.00	4.31 ± 0.75	0.885	0.241 (0.029-0.432)	-0.024 (-0.486-0.427)	0.309 (-0.137-0.653)
12	1.00	1.00	3.56 ± 1.00	0.883	0.348 (0.143-0.524)	0.347 (-0.120-0.682)	-0.297 (-0.684-0.179)
13	1.00	1.00	3.81 ± 0.96	0.879	0.416 (0.221-0.579)	0.081 (-0.357-0.493)	0.215 (-0.253-0.596)
14	1.00	1.00	3.52 ± 1.13	0.884	0.248 (0.033-0.440)	0.052 (-0.366-0.466)	0.019 (-0.445-0.458)
15	1.00	1.00	3.65 ± 1.40	0.889	0.239 (0.027-0.431)	-0.106 (-0.547-0.357)	0.116 (-0.235-0.486)
Overall	S-CVI/UA = 1.0	S-CVI/UA = 1.0	-	0.888	0.573 (0.409-0.702)	0.149 (-0.330-0.554)	0.199 (-0.001-0.389)

Note: *I-CVI: item-level content validity index, S-CVI/UA: scale-level content validity index, universal agreement method. ICC values present with 95% confidence interval in the parentheses.*

Cronbach's alpha coefficient; Cronbach's alpha if item deleted; the intraclass correlation coefficient for test-retest reliability; and for inter-rater reliability between two parents, and between parent and teacher were demonstrated in Table 3. The Cronbach's alpha if each item deleted, which was slightly varies from 0.888, indicates that no item was needed to be deleted. The mean and standard deviation of the DCDQ-T subscale and total scores categorized by age group and gender are demonstrated in Table 4. Although girls gained significantly higher mean fine motor/handwriting subscale scores than boys, there were no significant difference of the mean total DCDQ-T scores between genders (Table 4).

Table 4. Mean subscale and total scores of the Developmental Coordination Disorder Questionnaire, Thai version of each age group, classified by gender.

Age group and DCDQ-T scores	All participants (N=88)	Boys (N=43)	Girls (N=45)	*p* value
6 years 0 months to 7 years 11 months (N=32)				
Control during movement	21.7±3.6	20.6±3.9	22.6±3.2	0.122
Fine motor/ handwriting	15.1±3.3	13.7±3.4	16.2±2.8	0.029*
General coordination	18.1±3.7	17.9±3.3	18.3±4.1	0.798
Total score	**54.8±9.9**	**52.1±9.9**	**56.8±9.6**	**0.193**
8 years 0 months to 9 years 11 months (N=28)				
Control during movement	22.3±4.0	22.1±3.4	22.6±4.8	0.704
Fine motor/ handwriting	14.9±2.5	13.9±2.2	16.3±2.7	0.012*
General coordination	18.8±3.3	17.8±3.3	20.1±2.9	0.054
Total score	**55.9±8.7**	**53.6±7.6**	**59.0±9.5**	**0.105**
10 years 0 months to 12 years 11 months (N=28)				
Control during movement	23.0±4.6	24.5±3.8	21.6±4.9	0.090
Fine motor/ handwriting	15.9±3.1	15.5±2.3	16.3±3.4	0.541
General coordination	19.7±3.6	19.2±3.9	20.1±3.4	0.517
Total score	**58.3±9.3**	**59.1±7.5**	**57.6±10.8**	**0.667**

Note: *Data were presented as mean±SD, *p<0.05.*

Discussion

This study is the first to translate and validate the 2007 revised version of Developmental Coordination Disorder Questionnaire (DCDQ'07) into Thai, providing the first standardized screening tool for developmental coordination disorder (DCD) in Thailand. The availability of a validated Thai version can facilitate early detection and intervention for children at risk, which was previously limited due to the lack of culturally adapted tools.

To the best of our knowledge, although there were plenty of information about construct and concurrent validity analyses from the original DCDQ'07,[18] and previous translation studies of the DCDQ'07,[20-22] there was no data regarding content validity index (CVI) in prior cultural adaptations. The DCDQ-T demonstrated excellent content validity (I-CVI and S-CVI/UA=1.0), suggesting strong agreement among experts. The perfect I-CVI and S-CVI/UA values indicate that all five experts consistently judged each item as highly or quite relevant to the construct of developmental coordination disorder. This suggests that the DCDQ-T comprehensively reflects the targeted domains of motor control, fine motor skills/handwriting, and general coordination, as defined by the DSM-5 criteria for DCD and international guidelines. In particular, the expert panel—comprising developmental and behavioral pediatricians and occupational therapists—ensured that the items were culturally and clinically relevant within the Thai context. For example, modifications made to item examples (e.g., replacing "a bull in a China shop" with a culturally understandable phrase) enhanced item clarity without altering the underlying construct. This process reinforces the content validity by ensuring both conceptual and cultural alignment between the original and translated versions. Therefore, the DCDQ-T appears to be an appropriate instrument for assessing motor coordination difficulties among Thai children.

The internal consistency was good (Cronbach's alpha=0.888), comparable not only to the original DCDQ'07 (α=0.94),[18] but also to previous adaptations such as the Turkish (α=0.890),[19] German (α=0.891),[22] and French Canadian (α=0.949)[21] versions. The test-retest reliability was moderate (ICC=0.573), indicating that caregivers provided relatively stable responses over time. The test-retest reliability ICC of the DCDQ-T is less than those of the French Canadian, French European, and Turkish version (0.8-1.00).[19-21] However, the interval for test-retest reliability of the DCDQ-T was 3 months apart, while the one of other versions was 5-30 days. The inter-rater reliability of the DCDQ-T was poor between two parents (ICC=0.149) and between parents and teachers (ICC=0.199), which can

reflect the difference of Thai parents' and Thai teachers' perception and interpretation of their children's motor abilities, however, there were no previous studies investigated inter-rater ICC of the original DCDQ'07 and other adaptations. The low inter-rater reliability of the DCDQ-T may be explained by Thai cultural context that the children, especially in urban areas, tend to participate less in team sports and outdoor physical activities than their Western counterparts.[24] This may limit exposure to tasks assessed by the DCDQ-T, such as throwing or catching a ball. Moreover, the noticeability of these tasks may vary across settings (e.g., home vs. school) and raters (e.g., parents vs. teachers). This underscores the need for population-specific cutoff scores and further investigation into strategies to improve inter-rater agreement, such as rater training or structured observation guidelines.

Consistent with the original DCDQ'07,[18] the DCDQ-T scores showed an increasing trend with age, reflecting the natural progression of motor development in children. This similar trend suggested that the tool effectively captures age-related improvements in motor coordination. There was no gender effect on total DCDQ-T score which is consistent with the original DCDQ'07,[18] French European,[20] and French Canadian[21] versions. The finding that the girls, whose age group were 6-<8, and 8-<10 years old, gained significantly more mean DCDQ-T scores than boys in fine motor/ handwriting subscales is compatible with one previous study which demonstrated that pre-adolescent girls demonstrated better manual performance than their same-age boy peers.[29] However, there were no information regarding gender difference in DCDQ subscale scores in the either original English version of DCDQ'07 or other translated languages.

Limitations

The strengths of this study include an adequate sample size for validity testing, which aligns with prior DCDQ'07 validation studies. The use of stratified randomization reduced selection bias, ensuring balanced age representation throughout all grades in the school. In addition, the low dropout rate minimized potential bias from missing data.

Several limitations were acknowledged. Since the constraints of our study context, this study was adapted from the Beaton's guideline for cross-cultural adaptation.[25] This led to two main deviations from the guideline. The backward translation used only one professional translator, and we did not pilot the Thai version with a small group of parents before the whole data collection. However, the approval of the back-translated version from the original developer of the DCDQ'07 questionnaire and excellent content validity could imply the quality of the translation. This study conducted in an elementary school whose students' age range is 6-12 years, which did not fully cover the 5-15 years targeted by the original DCDQ'07, and could lead to limiting generalizability. Additionally, the sample was drawn from an urban school, potentially excluding rural children with different activity patterns. Finally, socioeconomic status, which may influence parental observations and responses, was not assessed. Future studies should expand the sample to include younger and older children and consider the influence of socioeconomic factors on DCDQ-T responses. Moreover, more psychometric properties including concurrent and construct validity, sensitivity, specificity; as well as culturally appropriate cut-off scores for the DCDQ-T should be established.

Conclusion

The Developmental Coordination Disorder Questionnaire, Thai version could be beneficial to screen for developmental coordination disorder in 6- to 12-year-old Thai children due to its good validity, good internal consistency, moderate test-retest reliability, and poor inter-rater reliability. Further studies are required to investigate its repeatability and inter-rater agreement.

Ethical approval

This study was approved by the Ethics Committee of the Human Research Protection Unit at the Faculty of Medicine of Siriraj Hospital, Mahidol University, COI number Si158/2023.

Funding

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

Conflict of interest

The authors declare no conflict of interest

CRedit authorship contribution statement

Natewimon Nantiwat: methodology, investigation, formal analysis, data curation, writing: original draft; **Sureelak Sutchritpongsa**: validation, project administration; **Amornrat Penphattrarakul**: visualization, writing: review and editing; **Prakasit Wannapaschaiyong**: resources, writing: review and editing; **Nattha Ketudat**: resources, validation; **Pat Rojmahamongkol**: conceptualization, supervision, formal analysis, writing: review and editing

Acknowledgements

The authors gratefully acknowledged Professor Brenda N. Wilson for the permission to translate the 2007 revised Developmental Coordination Disorder Questionnaire (DCDQ) into Thai language; Mr. Chotchana Photinil from the Department of Pediatrics, Faculty of Medicine, Siriraj Hospital, Mahidol University and Dr. Sasima Tongsai, PhD from Siriraj Research Data Management Unit for their helping with statistical analysis; and Dr. Mark Simmerman for English language editing.

References

[1] American Psychiatric Association. Diagnostic and statistical manual of mental disorders : DSM-5. 5th Ed. Washington DC: American Psychiatric Publishing; 2013: pp 74-7.

[2] Blank R, Barnett AL, Cairney J, Green D, Kirby A,

Polatajko H, et al. International clinical practice recommendations on the definition, diagnosis, assessment, intervention, and psychosocial aspects of developmental coordination disorder. Dev Med Child Neurol. 2019; 61(3): 242-85. doi: 10.1111/dmcn.14132.

[3] Lingam R, Hunt L, Golding J, Jongmans M, Emond A. Prevalence of developmental coordination disorder using the DSM-IV at 7 years of age: a UK population-based study. Pediatrics. 2009; 123(4): e693-700. doi: 10.1542/peds.2008-1770.

[4] Kadesjo B, Gillberg C. Developmental coordination disorder in Swedish 7-year-old children. J Am Acad Child Adolesc Psychiatry. 1999; 38(7): 820-8. doi: 10.1097/00004583-199907000-00011.

[5] Paepetch-Suato B, Luangpon N, Kiatkulanusorn S. Development Coordination of Pre-Kindergarten and Kindergarten. Technical Education Journal: King Mongkut's University of Technology North Bangkok [Internet]. 2019; 10(2): 175-84. Available from: https://ojs.kmutnb.ac.th/index.php/jote/article/view/2889.

[6] Girish S, Raja K, Kamath A. Prevalence of developmental coordination disorder among mainstream school children in India. J Pediatr Rehabil Med. 2016; 9(2): 107-16. doi: 10.3233/PRM-160371.

[7] Harris SR, Mickelson ECR, Zwicker JG. Diagnosis and management of developmental coordination disorder. CMAJ. 2015; 187(9) : 659-65. doi: 10.1503/cmaj.140994.

[8] Fliers E, Rommelse N, Vermeulen SH, Altink M, Buschgens CJ, Faraone SV, et al. Motor coordination problems in children and adolescents with ADHD rated by parents and teachers: effects of age and gender. J Neural Transm (Vienna). 2008; 115: 211-20. doi: 10.1007/s00702-007-0827-0.

[9] Flapper BCT, Schoemaker MM. Developmental coordination disorder in children with specific language impairment: co-morbidity and impact on quality of life. Res Dev Disabil. 2013; 34(2): 756-63. doi: 10.1016/j.ridd.2012.10.014.

[10] Margari L, Buttiglione M, Craig F, Cristella A, de Giambattista C, Matera E, et al. Neuropsychopathological comorbidities in learning disorders. BMC Neurol. 2013; 13: 198. doi: 10.1186/1471-2377-13-198.

[11] Blank R, Smits-Engelsman B, Polatajko H, Wilson P; European Academy for Childhood Disability. European Academy for Childhood Disability (EACD): recommendations on the definition, diagnosis and intervention of developmental coordination disorder (long version). Dev Med Child Neurol. 2012; 54(1): 54-93. doi: 10.1111/j.1469-8749.2011.04171.x.

[12] Chaingam S, Narkpongphun A. Prevalence and associated factors of developmental coordination disorder among patients with deficit hyperactivity disorder at Maharaj Nakorn Chiang Mai Hospital. J Psychiatr Assoc Thailand [Internet]. 2019; 64(4): 309-16. Available from: https://he01.tci-thaijo.org/index.php/JPAT/article/view/231886.

[13] Zwicker JG, Harris SR, Klassen AF. Quality of life domains affected in children with developmental coordination disorder: a systematic review. Child Care Health Dev. 2013; 39(4): 562-80. doi: 10.1111/j.1365-2214.2012.01379.x.

[14] Rahimi-Golkhandan S, Steenbergen B, Piek JP, Wilson PH. Deficits of hot executive function in developmental coordination disorder: sensitivity to positive social cues. Hum Mov Sci. 2014; 38: 209-24. doi: 10.1016/j.humov.2014.09.008.

[15] Leonard HC, Bernardi M, Hill EL, Henry LA. Executive functioning, motor difficulties, and developmental coordination disorder. Dev Neuropsychol. 2015; 40(4): 201-15. doi: 10.1080/87565641.2014.997933.

[16] Henderson SE, Barnett A. Movement Assessment Battery for Children, Thiird edition (MABC-3). Pearson Education, Inc [Internet]. 2023. [cited 2025 Jan. 22]. Available from: https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/mabc-3/mabc-2-to-3-comparison.pdf.

[17] Bruininks BD, Bruininks RH. Bruininks-Oseretsky Test of Motor Proficiency, Third edition (BOT-3). Pearson Education, Inc. [Internet] 2024. [cited 2025 Jan. 22]. Available from: https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Motor-Sensory/Bruininks-Oseretsky-Test-of-Motor-Proficiency-%7C-Third-Edition/p/P100071001.html.

[18] Wilson BN, Crawford SG, Green D, Roberts G, Aylott A, Kaplan BJ. Psychometric properties of the revised developmental coordination disorder questionnaire. Phys Occup Ther Pediatr. 2009; 29(2): 182-202. doi: 10.1080/01942630902784761.

[19] Yildirim CK, Altunalan T, Acar G, Elbasan B, Gucuyener K. Cross-cultural adaptation of the Developmental Coordination Disorder Questionnaire in Turkish children. Percept Mot Skills. 2019; 126(1): 40-9. doi: 10.1177/0031512518809161.

[20] Ray-Kaeser S, Thommen E, Martini R, Jover M, Gurtner B, Bertrand AM. Psychometric assessment of the French European Developmental Coordination Disorder Questionnaire (DCDQ-FE). PLoS One. 2019; 14(5): e0217280. doi: 10.1371/journal.pone.0217280.

[21] Martini R, St-Pierre MF, Wilson BN. French Canadian cross-cultural adaptation of the Developmental Coordination Disorder Questionnaire'07: DCDQ-FC. Can J Occup Ther. 2011; 78(5): 318-27. doi: 10.2182/cjot.2011.78.5.7.

[22] Kennedy-Behr A, Wilson BN, Rodger S, Mickan S. Cross-cultural adaptation of the developmental coordination disorder questionnaire 2007 for German-speaking countries: DCDQ-G. Neuropediatrics. 2013; 44(5): 245-51. doi: 10.1055/s-0033-1347936.

[23] Ko JY, Lee WH, Woon JJ, Kim YA. Development of Korean Version of Developmental Coordination Disorder Questionnaire (DCDQ-K). J Kor Phys Ther [Internet]. 2020; 32(1): 44-51. Available from: https://www.researchgate.net/publication/340052197_

Development_of_Korean_the_Version_of_the_Developmental_Coordination_Disorder_Questionnaire_DCDQ-K.

[24] Widyastari DA, Saonuam P, Rasri N, Pongpradit K, Katewongsa P. Prevalence and trends of physical activity in Thai children and young people: pooled panel data analysis from Thailand's surveillance on physical activity 2012–2020. Int J Environ Res Public Health. 2021; 18(19): 10088. doi: 10.3390/ijerph181910088.

[25] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adpatation of self-report measures. Spine (Phila Pa 1976). 2000; 25(24): 3816-91. doi: 10.1097/00007632-200012150-00014.

[26] World Health Organization. BMI- for –ages (5-19 years). 2025 [cited 2025 Apr. 15]. Avaliable from: https://www.who.int/tools/growth-reference-data-for-5to19-years/indicators/bmi-for-age.

[27] Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ. 2018; 48: 1273-96. doi: 10.1007/s11165-016-9602-2.

[28] 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: 10.1016/j.jcm.2016.02.012. Epub 2016 Mar 31.

[29] Flatters I, Hill LJ, Williams JH, Barber SE, Mon-Williams M. Manual control age and sex differences in 4- to 11-year-old children. PLoS One. 2014; 9(2): e88692. doi: 10.1371/journal.pone.0088692. eCollection 2014.

Journal of Associated Medical Sciences 2025; 58(3): 345-353

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

Serotype distribution of dengue virus in Trang Province, Southern Thailand, 2024

Suwandee Sapcharoen[1*], Ativet Sawetadul[1], Tipattaraporn Panich[2], Khatayut Nigapruke[1]

[1]Regional Medical Sciences Center 12/1 Trang, Department of Medical Sciences, Trang Province, Thailand.
[2]National Institute of Health, Department of Medical Sciences, Nonthaburi Province, Thailand.

ARTICLE INFO

Article history:
Received 19 May 2025
Accepted as revised 20 August 2025
Available online 28 August 2025

Keywords:
Dengue fever, dengue virus serotype, Trang, Southern Thailand.

ABSTRACT

Background: Dengue fever is a significant public health concern in Thailand, which consists of *Aedes* mosquitoes as a vector. The incidence of dengue fever has increased, positioning it among the top 10 infectious diseases affecting health in the Southern region of Thailand.

Objectives: This study aimed to examine serotypes and genotypes of the dengue virus causing dengue fever in Trang province in 2024.

Materials and methods: From April to August 2024, a total of 125 acute serum or plasma samples were collected from patients with dengue NS1 antigen screening test positive results and/or were clinically diagnosed with dengue fever. The samples were tested for dengue viral RNA detection and serotyping using real-time RT-PCR. Genotype identification was performed using whole-genome sequencing.

Results: Of these, 73 samples (58.4%) were detected for dengue viral RNA. The predominant serotype was DENV-3 (N=31, 42.5%), followed by DENV-2 (N=27, 37.0%), DENV-4 (N=13, 17.8%), and DENV-1 (N=2, 2.7%). By analyzing and comparing the whole genome sequence of 3 samples (TR-22, TR-37, and TR-38), all results were DENV-4 Genotype I. The percent of identity for genotype classification ranged from 95.1% to 99.3%.

Conclusion: The findings would be useful for establishing effective control and prevention measures for the outbreaks and could contribute to the development of diagnostic tests and vaccines in the future.

Introduction

Dengue fever is a major public health issue in many countries worldwide, with the *Aedes aegypti* mosquito serving as the primary vector for transmitting the disease to humans.[1,2] It is commonly found in tropical and subtropical regions of the Americas, the eastern Mediterranean, Africa, the western Pacific, and Southeast Asia.[3] Globally, the World Health Organization (WHO) estimates that more than 390 million people are infected with the dengue virus annually, with approximately 20,000 deaths attributed to the infection. Seventy percent of all cases have occurred in countries within Asia.[3,4,5,6]

Dengue virus (DENV) is a single-stranded RNA virus with a genome size of approximately 10.7 kilobases.[5] The viral structure includes a lipid envelope and various structural proteins, such as premembrane/membrane, capsid and envelope proteins, as well as non-structural proteins, including NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5.[6] The Dengue virus belongs to genus Flavivirus and family Flaviviridae, with four distinct serotypes: DENV-1, DENV-2, DENV-3 and DENV-4.[5,6] These serotypes

∗ **Corresponding contributor.**
Author's Address: *Regional Medical Sciences Center 12/1 Trang, Department of Medical Sciences, Trang Province, Thailand.*
E-mail address: suwandee.s@dmsc.mail.go.th
doi: 10.12982/JAMS.2025.106
E-ISSN: 2539-6056

are different based on epitopes, which are small units or positions on the viral surface that interact with antibodies, and genetic differences. Each serotype can be further subdivided into multiple genotypes based on nucleotide sequences, with no more than 6% variation within a single genotype.[6] DENV-1 includes six genotypes (genotype I, II, III, IV, V and VI), DENV-2 includes five genotypes (Asian I, Asian II, Asian/American, American, Cosmopolitan, and sylvatic genotypes), DENV-3 includes five genotypes (genotype I, II, III, IV and V), and DENV-4 includes five genotypes (genotype I, IIA, IIB, III and sylvatic genotypes).[7] All four serotypes are in circulation in Thailand.[3] Clinical symptoms of dengue fever may range from mild dengue fever (DF) to severe dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS). DF presents with fever, including headache, bone or joint and muscular pains, nausea, vomiting, and small petechial hemorrhages on the skin. In severe cases, it can lead to severe fever (DHF and DSS), which is a result of circulatory failure, shock, and death.[8] Individuals who have been infected once develop lifelong immunity to the serotype that they were exposed to, and short-term immunity (approximately 3 to 12 months) to other serotypes, which means that they can be infected multiple times by different serotypes. Currently, rapid diagnostic tests that detect the NS1 antigen protein are used for screening dengue virus infections, and real-time RT-PCR assays are used for detecting viral RNA with high sensitivity and specificity, allowing confirmation of infection and serotype identification. Additionally, antibody testing for IgM and IgG specific to the dengue virus using the ELISA method can also confirm the presence of infection.[9]

The global incidence of dengue fever has risen, making it one of the top 10 global health concerns in 2019.[10.] In Thailand, continuous outbreaks of dengue fever have occurred, with a major outbreak occurring in 2013, when 154,773 cases were reported.[11] The high mortality rate was recorded in 80 cases, accounting for 0.13% in 2018.[12] According to the Epidemiology Surveillance Report 506 from the Department of Disease Control, the dengue situation from fiscal years 2018 to 2023 showed a significant outbreak in 2019, followed by a downward trend in 2020. However, in 2022, the number of infected cases and deaths started to rise again. In 2023, the highest number of deaths was reported with an outbreak comparable to 2019.[13] In 2023, there were 158,620 reported cases and 181 deaths (0.11%), while these cases were 3.5 times higher than in 2022.[14] DENV-2 serotype was the most predominant, accounting for 31.3% of cases, followed by DENV-1 at 29.9%. However, DENV-3 had a 5 to 16-fold increase compared to the last five years.[15] The shifting predominant serotypes over the years could be a significant factor contributing to the dengue outbreak, as patients lack immunity to the changing serotypes. Dengue fever exhibits seasonal variation with an increase in cases starting in April, peaking between June and August during the rainy season. The number of cases tends to decrease in September, but if the number of cases remains high toward the end of the year, it may lead to continued outbreaks in the following year.

The Department of Disease Control reported the situation for the year 2023 from January 1 to December 20, in regional health area 12, which includes seven provinces: Songkhla, Satun, Trang, Phatthalung, Pattani, Yala, and Narathiwat. The cumulative number of cases was 15,096 with an incidence rate of 301.64 per 100,000 people and resulted in 23 deaths.[16] In Trang, there were eight districts affected by the outbreak: Mueang Trang, Kantang, Yan Ta Khao, Palian, Huai Yot, Wang Wiset, Na Yong, and Ratsada.[16] According to the 2023 surveillance data, the total reported number of dengue fever cases in Trang was 1,089 cases. However, there is currently lack of data on dengue virus serotypes and genotypes of infected patients in Trang. Therefore, this study aimed at investigating serotypes and genotypes of dengue virus in Trang in 2024, to utilize this data for establishing effective control and prevention measures for the outbreaks. Additionally, the findings could contribute to the development of diagnostic tools and vaccines in the future.

Materials and methods

Sample Size

Sample size was calculated using Cochran's formula:

$n_o = (Z^2 \times p \times q) / e^2$

$n_o = (1.96^2 \times 0.5 \times 0.5) / 0.05^2$

Finite Population Correction

$n = n_o / (1 + ((n_o - 1) / N))$

$n = n_o / (1 + ((n_o - 1) / 1{,}089))$

$= 284.6 \approx 285$

Where:

n_0 = required sample size

Z = Z-score corresponding to the desired confidence level

p = estimated proportion of the population

q = 1 − p

e = acceptable margin of error

n = adjusted sample size

N = total population size

According to the 2023 surveillance data, the total reported number of dengue fever cases in Trang was 1,089. Based on this formula, assuming a confidence level of 95% (Z=1.96), an estimated population proportion (p) of 0.5, and a margin of error (e) of 0.05, the calculated sample size for a one-year study period was 285 participants. Since the current study spans only five months, the sample size was proportionally adjusted to 119 participants.

Specimen collection

From April to August 2024, acute serum or plasma was collected from 125 patients in Trang who tested positive using the NS1 antigen screening test using the careUS Dengue Combo NS1 & IgM/IgG kit (Wells Bio Inc., Republic of Korea) and/or those diagnosed with dengue fever by physicians. Inclusion criteria were an acute illness for 5 days post-symptom, including two or more of the following: fever, headache, nausea, vomiting, red

eyes, rash, joint pain, muscle pain. The age of patients ranged from 8 months to 75 years old. All specimens were confirmed to have dengue virus infection by quantitative reverse transcription polymerase chain reaction (qRT-PCR) at the Regional Medical Sciences Center 12/1 Trang and kept at -20°C until further use.

Nucleic acid extraction

Following the manufacturer's instructions, total nucleic acids were extracted from 200 µL of each serum or plasma using Zybio nucleic acid extraction kit (Magnetic bead method) with automated Zybio EXM3000 nucleic acid isolation system (Zybio Inc., China) with an elution volume of 80 µL. For whole genome sequencing, nucleic acids were extracted from 150 µL of each cell culture or serum or plasma using QIAamp Viral RNA Mini Kit (QIAGEN, Germany) with an elution volume of 30 µL.

Dengue virus serotyping

For the detection of dengue virus serotypes, one-step real-time RT-PCR was performed using a VIASURE Real-Time PCR Detection kit (Certest Biotec S.L., Spain) that targets the conserved region of the NS5 gene (DENV-1), envelope gene (DENV-2), prM gene (DENV-3) and NS2A gene (DENV-4) with qTOWER^{3}G Real-Time PCR (Analytik Jena, Germany) in 20 µL PCR reaction (5 µL of each extracted sample). A result is considered positive when the cycle threshold (Ct) value is less than 40, according to the manufacturer's instructions. The remaining serums were stored at -20 °C and transported under cold chain conditions to the National Institute of Health, Department of Medical Sciences.

Dengue virus isolation and amplification

Samples that tested positive for dengue virus serotyping via real-time RT-PCR with Ct values below 25 were cultured to propagate the virus for whole genome sequencing using next-generation sequencing. Dengue virus was cultured from 5 µL of serum or plasma samples in C6/36 cells (derived from *Aedes albopictus* mosquitoes) using Leibovitz's 15 cell culture medium (Life Technologies, USA) supplemented with Fetal Bovine Serum (FBS) (Sigma, Paraguay), Tryptose Phosphate Broth (TPB) (Himedia, USA) and Penicillin-Streptomycin 10,000 units/mL (Life Technologies, USA). The culture was incubated at 28±2 °C for 7 days. Cytopathogenic effects (CPE) were observed every 1-2 days. The culture supernatant containing the virus was collected, and the dengue virus serotype was confirmed by infecting the cells and performing the in-house Indirect Immunofluorescent Antibody (IFA) test. Positive results were indicated by the presence of green fluorescence in the cells under a fluorescence microscope (Nikon, Japan).

Library preparation and whole genome sequencing

Dengue virus whole genome nucleotide sequencing using Next Generation Sequencing (NGS). Five microliters of RNA were prepared for cDNA synthesis using the TURBO DNA-free™ Kit (Thermo Fisher Scientific Inc., USA). The RNA was then combined with one µL of Random Hexamers and used for first-strand cDNA synthesis with SuperScript™ IV Reverse Transcriptase and dNTP Mix (10 mM each) (Thermo Fisher Scientific Inc., USA). Second-strand cDNA synthesis was subsequently performed using Sequenase Version 2.0 DNA Polymerase (Thermo Fisher Scientific Inc., USA), followed by cDNA cleanup using Agencourt AMPure XP (Beckman Coulter Inc., USA). Twelve microliters of the eluted sample were used to prepare a DNA library (DNA preparation) with the Illumina DNA/RNA Prep Tagmentation Kit (Illumina, USA) and Illumina DNA/RNA UD Indexes Set D. Afterward, the library concentration was measured using a Qubit 2.0 fluorometer (Invitrogen, USA), and the samples were pooled with 0.2N NaOH and diluted with Hyb buffer (Illumina, USA). Finally, 600 µL of sample volume was loaded onto MiSeq Reagent Kit v3 (150 Cycles) and sequenced using the MiSeq sequencer platform (Illumina, USA), with single-end 150 bp read configuration. The sequenced data were in FASTQ format.

Data analysis of whole genome sequencing

The data obtained were analyzed for dengue virus genotypes using the BLAST (Basic Local Alignment Search Tool) program. The nucleotide sequence results from each sample were compared with reported nucleotide sequences of dengue virus available in the GenBank database, which can be accessed at https://blast.ncbi.nlm.nih.gov/Blast.cgi. The comparison of the results was based on the percent of identity. Additionally, the nucleotide sequence relationships based on the phylogeny of the dengue virus were constructed using MEGA 12 software (https://www.megasoftware.net). The TN+G+I model with the Maximum Likelihood method was applied, and the robustness of the analysis was evaluated using 1,000 bootstrap replicates. The nucleotide sequence of Zika virus (NC_012532) was designated as the outgroup.

Results

Patients who passed the screening using the NS1 antigen test and/or were diagnosed with dengue fever by physicians in Trang were enrolled. The number of suspected dengue cases during the study period by month (Figure 1). The highest number of patients was recorded in July 2024. A total of 125 samples were subsequently tested for real-time RT-PCR with serotyping. Dengue viral RNA was detected in 73 samples using real-time RT-PCR, accounting for 58.4%. The most common serotype was DENV-3 (N=31, 42.5%), followed by DENV-2 (N=27, 37.0%), DENV-4 (N=13, 17.8%), and DENV-1 (N=2, 2.7%), respectively (Figure 2).

It can be observed that the average rainfall has been steadily increasing since April, reaching a peak in August 2024. The average rainfall from April to August 2024 in Trang, reported by Trang's meteorological station, which was shown as 61.6, 227.6, 222.5, 238.8, and 514.8 millimeters, respectively, as illustrated in Figure 1. Due to limitations in the available data, statistical analysis of the relationship between rainfall and patient numbers was not conducted.

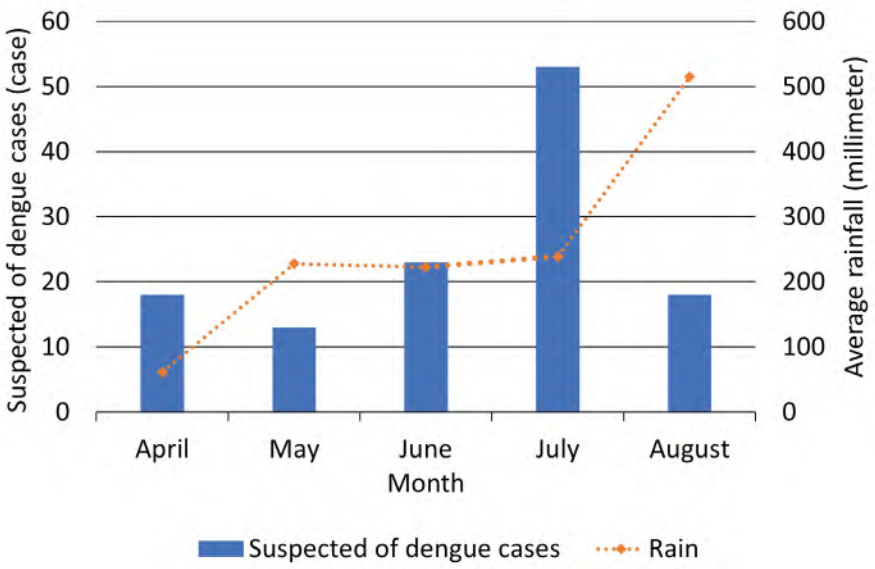
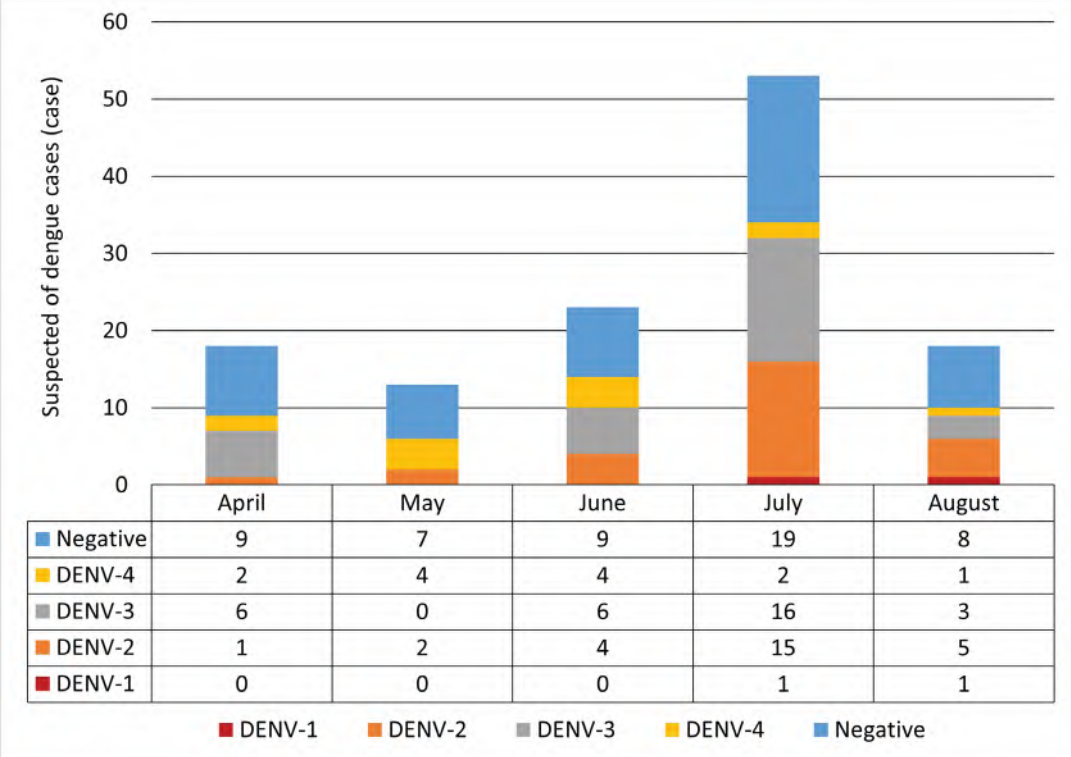


Figure 1. *Correlation between the number of suspected dengue cases and average rainfall during the study period in Trang.*



	April	May	June	July	August
Negative	9	7	9	19	8
DENV-4	2	4	4	2	1
DENV-3	6	0	6	16	3
DENV-2	1	2	4	15	5
DENV-1	0	0	0	1	1

Figure 2. *Monthly distribution of dengue cases from April to August 2024.*

Among the patients, the number of females (N=38) was slightly higher than males (N=35), accounting for a 1.1:1 ratio. Most infected individuals were in the age group 15-24 years (N=31), followed by the age group 10-14 years (N=21) and the age group 25-34 years (N=20), as shown in Table 1. The ages of the 73 dengue-infected patients ranged from 3 to 72 years old (mean 23.2 years old).

A total of 98 samples were screened with the NS1 antigen test, but only 69 samples showed concordant results with the detection of dengue viral RNA and serotype classification using real-time RT-PCR, including 55 samples with NS1 antigen positive and 14 samples with NS1 antigen negative. Additionally, 23 patients who were diagnosed with dengue fever tested positive for NS1 antigen but negative by real-time RT-PCR, while 6 patients were tested negative for NS1 antigen but positive by real-time RT-PCR, as shown in Table 2. Dengue NS1 antigen detection revealed a sensitivity of 90.2% (55/61) and specificity of 37.8% (14/37) when compared to the qRT-PCR results.

Table 1. Demographic characteristics of dengue patients, categorized by real-time RT-PCR positive and negative results.

Characteristics	Dengue virus positive, N (%)				Dengue virus negative N (%)	Total N (%)
	DENV-1	DENV-2	DENV-3	DENV-4		
Gender						
Male	0 (0.0)	12 (20.3)	18 (30.5)	5 (8.5)	24 (40.7)	59 (47.2)
Female	2 (3.0)	15 (22.7)	13 (19.7)	8 (12.1)	28 (42.4)	66 (52.8)
Total	2 (1.6)	27 (21.6)	31 (24.8)	13 (10.4)	52 (41.6)	125 (100.0)
Age (years)						
<1	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)	1 (0.8)
1-4	0 (0.0)	0 (0.0)	4 (80.0)	0 (0.0)	1 (20.0)	5 (4.0)
5-9	0 (0.0)	4 (22.2)	6 (33.3)	1 (5.6)	7 (38.9)	18 (14.4)
10-14	0 (0.0)	7 (33.3)	6 (28.6)	4 (19.1)	4 (19.1)	21 (16.8)
15-24	0 (0.0)	6 (19.4)	7 (22.6)	3 (9.7)	15 (48.4)	31 (24.8)
25-34	1 (5.0)	3 (15.0)	2 (10.0)	2 (10.0)	12 (60.0)	20 (16.0)
35-44	0 (0.0)	4 (40.0)	1 (10.0)	3 (30.0)	2 (20.0)	10 (8.0)
45-54	0 (0.0)	1 (11.1)	2 (22.2)	0 (0.0)	6 (66.7)	9 (7.2)
55-64	1 (25.0)	0 (0.0)	3 (75.0)	0 (0.0)	0 (0.0)	4 (3.2)
>64	0 (0.0)	2 (33.3)	0 (0.0)	0 (0.0)	4 (66.7)	6 (4.8)
Total	2 (1.6)	27 (21.6)	31 (24.8)	13 (10.4)	52 (41.6)	125 (100.0)

Table 2. Correlation between the NS1 antigen test and real-time PCR detection of dengue virus.

Parameter	Real-time RT-PCR positive, N (%)				Real-time RT-PCR negative, N (%)	Total, N (%)
	DENV-1	DENV-2	DENV-3	DENV-4		
NS1 antigen positive	2 (2.6)	16 (20.5)	27 (34.6)	10 (12.8)	23 (29.5)	78 (79.6)
NS1 antigen negative	0 (0.0)	4 (20.0)	2 (10.0)	0 (0.0)	14 (70.0)	20 (20.4)
Total	2 (2.0)	20 (20.4)	29 (29.6)	10 (10.2)	37 (37.8)	98 (100.0)

***Note**: Among the enrolled samples, there was no initial NS1 antigen test data for 27 samples. Of these, 12 samples were real-time RT-PCR positive (7 samples for DENV-2, 2 samples for DENV-3, and 3 samples for DENV-4), while 15 samples were real-time RT-PCR negative.*

Samples that tested positive for dengue virus serotyping using real-time RT-PCR with Ct values of less than 25 were cultured to obtain the virus for whole genome sequencing using NGS. The data was then analyzed and compared to determine the dengue virus genotype. A total of 12 samples were processed for culturing, but only 3 samples were cultured positive for further testing: TR-22, TR-37, and TR-38, all of which were DENV-4 serotype. The genotype analysis revealed that all 3 samples were identified as DENV-4 Genotype I, with the percent of identity ranging from 95.11% to 99.29%, as shown in Table 3.

By analyzing and comparing the nucleotide sequence data from samples with more than 80% genome coverage of the dengue virus using a phylogenetic tree, one sample (TR-38) was identified as DENV-4 of Genotype I, supported by a bootstrap value of 98%, as shown in Figure 3.

Table 3. Characterization of dengue virus using the BLAST program.

Sample number	Serotype	Genotype	% Identity	Similar GenBank Access Number
TR-22	DENV-4	Genotype I	95.11	KY451945
TR-37	DENV-4	Genotype I	98.85	LC410197
TR-38	DENV-4	Genotype I	99.29	LC410196

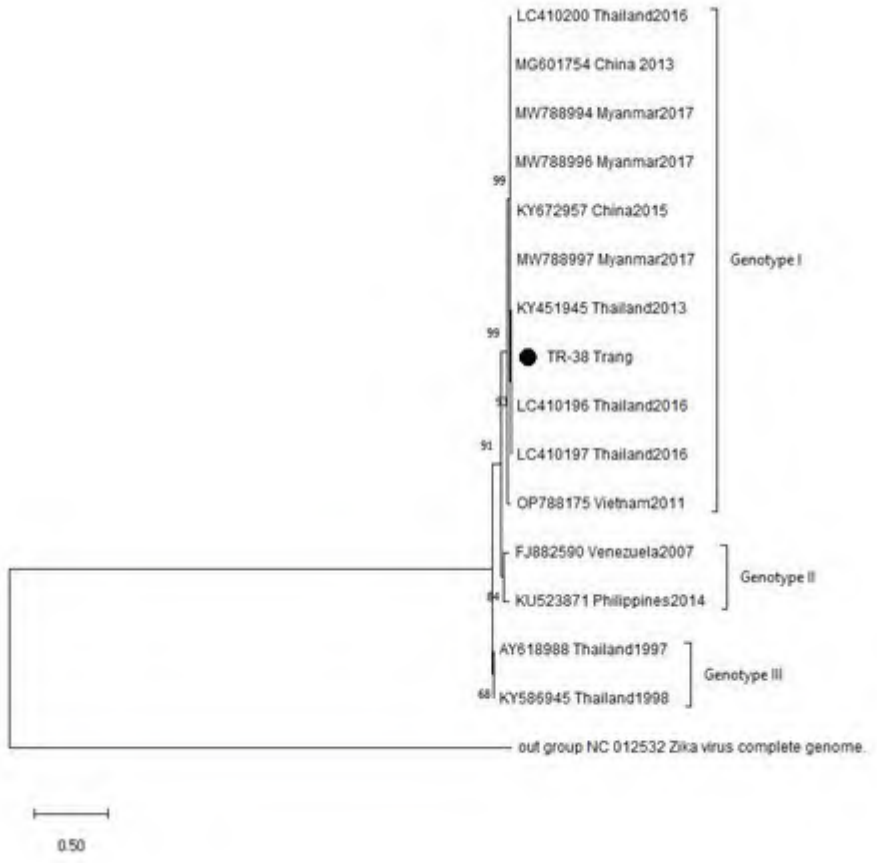


Figure 3. *Phylogenetic tree showing genotypic classification of DENV-4 dengue virus.*

Discussion

The outbreak of dengue fever often occurs in the rainy season. According to an epidemiological survey of dengue fever in Thailand between 2000 and 2011, the number of cases typically increases from June to August. Dengue prevention and control programs in Thailand are mainly based on hospital case reports within 24 hours in order to prevent transmission by spraying insecticides within 100 meters of the patient's house.[17] This study was limited to laboratory-confirmed dengue cases, and the incompleteness of patient data may have affected the accuracy of correlations drawn with provincial average rainfall data, particularly in Trang. All serotypes of dengue virus were found in the Southern region between 2005 and 2010, with DENV-1 being the most prevalent from 2005 to 2009. However, the trend in the epidemiological data of dengue serotypes shifted in 2010, and DENV-2 became the most common.[3] This study also found that

DENV-2 was the most prevalent, like the findings in 2010. The study in other regions of Thailand between 2005 and 2010 by Kriengsak Limkittikul *et al*., Central, Northern and Northeastern regions reported DENV-1 as the most prevalent serotype from 2005 to 2009, with DENV-2 being predominant in 2010.[3] Only few data of dengue serotype distribution in the upper Southern region of Thailand during from 2014 to 2021 revealed that 1,220 samples were positive accounting for 50.4%. The highest prevalent dengue serotype was DENV-1, accounting for 40.3%.[18] A study by Aranya Pinyorattanachot *et al*. revealed that epidemiological data on dengue fever in Thailand from 2014 to 2021 showed that the highest incidence rate occurred in 2018, followed by 2019, with a subsequent decline in 2020 and 2021, potentially due to the impact of the COVID-19 pandemic. Data from the Department of Medical Sciences on dengue virus serotype surveillance between 2005 and 2019 indicated that all four dengue virus serotypes (DENV-1 to DENV-4) have been co-circulating in Thailand. Notably, during the outbreak years, a distinct shift in the predominant circulating serotype was observed: DENV-3 was predominant in 2013, DENV-3 again in 2015, and DENV-1 became dominant in 2018.[18] Jacqueline Kyungah Lim *et al*. showed that dengue epidemics occur every 2-4 years in Ratchaburi province, Thailand.[19]

The study from other countries in Asia, Jean Claude Balingit *et al*. found an outbreak in Bali, Indonesia in 2022 where the ratio of female to male patients was 5:3, indicating a higher infection rate among females, like the findings in this study. Analysis of 48 patient samples revealed the following distribution: 28 samples of DENV-3, 14 samples of DENV-2, 5 samples of DENV-1, and 1 sample of co-infection between DENV-1 and DENV-3, with no cases of DENV-4 reported in this outbreak.[20] The study of Yu Kie Chem *et al*. on the molecular epidemiology of dengue virus in Malaysia from 2015 to 2021 analyzed 42,763 patient samples for serotype detection. The study found that DENV-2 and DENV-1 were the predominant serotypes, with detection rates of 35.9% and 33.4%, respectively, while DENV-3 and DENV-4 were also found in Malaysia.[21] A molecular characterization from dengue patients from 2019-2020 in northern Vietnam resulted in predomination of DENV-2.[22] In the Lao People's Democratic Republic, DENV-1 was predominant in 2010-2011 and in 2015, DENV-3 in 2012-2013 and DENV-4 in 2014 and 2016.[23] As previously described, the distribution of dengue virus serotypes appears to be influenced by the geographic localization of affected populations, correlates with variations in clinical disease severity, and is supported by laboratory-based diagnostic data.

Clinical symptoms of dengue fever range from mild flu-like symptoms to severe conditions, such as bleeding or shock, which can lead to death. Laboratory diagnosis during the acute phase is crucial for timely identification, and molecular techniques for detecting the viral genetic materials, such as qRT-PCR and RT-PCR, are the standard diagnostic methods. In developing countries with limited molecular diagnostic tools, serological tests are often used as alternatives, such as methods detecting NS1 antigen and antibodies (IgM and IgG) against the virus, including ELISA. For diagnosing dengue fever using rapid diagnostic tests, detecting the NS1 antigen is also a preliminary screening method that is fast and able to identify the NS1 antigen in patients with both primary and secondary infections. The NS1 protein can be detected up to 9 days after the onset of symptoms. However, the study by Mya Myat Ngwe Tun *et al*. in 2024 evaluated the newly developed Fujifilm Dengue non-structural antigen diagnostic kit and compared it to the SD Bioline NS1 antigen test kit with 140 samples. The rapid diagnostic test kit showed sensitivity, ranging from 88.6% to 94.3% and specificity of 100%. Nonetheless, real-time PCR remains the gold standard for diagnosis.[24] In this study, the careUS Dengue Combo NS1 & IgM/IgG kit showed 90.2% sensitivity and 37.8% specificity. The timing of sample collection may contribute to discrepancies between NS1 antigen and PCR results. These results may represent false-positive NS1 antigen detections since the samples were negative for DENV RNA. However, longitudinal studies show that NS1 antigen can persist longer than viral RNA, likely due to its extended half-life.[25] The study by Danielle Gyurech *et al*. reported a false-positive dengue NS1 antigen test in a traveller with an acute Zika virus infection.[26] However, this study did not include the detection of Chikungunya and Zika viruses, which are other co-circulating pathogens; there may have been cases of these infections.

This study could analyze the full genome nucleotide sequence of dengue virus using NGS for only 3 samples, probably due to the poor storage and/or transportation conditions of the samples. One potential limitation of this study is the possibility of viral inactivation during specimen transport from the Medical Science Center 12/1 Trang to the National Institute of Health, Nonthaburi, as the considerable distance and transit duration could have affected sample integrity. The study by Caio Santos de Souza *et al*. used samples with high viral genomic concentration for whole genome sequencing, whereas samples with insufficient viral genomic concentration were analyzed for the nucleotide sequence of the envelope (E) gene.[6] Patcharaporn Nonyong *et al*. suggested that the shorter fragments of PCR products have been more suitable for sequencing.[17] This method is beneficial for determining genotypes in future studies. Three samples from Trang were found to be DENV-4 Genotype I, while DENV-4 Genotype II has been reported in Indonesia.[27] In the Americas, DENV-4 Genotype IIb has been identified in Panama.[28] Alejandra Rojas *et al*. also revealed a DENV-4 Genotype II outbreak in Paraguay.[29] Thus, dengue genotype varies by geographical regions.

Limitation

None

Conclusion

Dengue fever is considered a significant public health issue, particularly in the Southern region of Thailand, which experiences substantial rainfall throughout the

entire year. It is important to annually monitor trends in the changing epidemiological patterns of dengue serotypes. By performing the viral RNA detection and serotype classification of dengue virus using real-time RT-PCR of 73 patient samples in Trang in 2024, all four dengue virus serotypes were identified, while the proportions of DENV-3, DENV-2, DENV-4, and DENV-1 were 42.5%, 37.0%, 17.8% and 2.7%, respectively. Furthermore, the whole genome sequencing of dengue virus from 3 samples using NGS revealed DENV-4 Genotype I. However, this study has limitations in analyzing genotypes due to the small sample size, and further studies are still needed. The findings from this study would be useful for establishing effective control and prevention measures for the outbreaks, and the data could also contribute to the development of diagnostic tools and vaccines in the future.

Ethical approval

The study was approved by the Institutional Review Board of the Department of Medical Sciences, the Ministry of Public Health, Thailand (DMSc-EC011).

Funding

The project is financially supported by the National Institute of Health for reagents for dengue virus serotyping and genotyping, and the Regional Medical Sciences Center 12/1 Trang for extraction reagents.

Conflict of interest

The authors declare no conflicts of interest.

CRedit authorship contribution statement

Suwandee Sapcharoen: conceptualization, investigation, data curation, writing original draft, review and editing; **Ativet Sawetadul**: investigation; **Tipattaraporn Panich**: conceptualization, investigation, data curation; **Khatayut Nigapruke**: supervision.

Acknowledgements

We would like to express our sincere gratitude to the Department of Medical Sciences for providing the reagents for dengue viral genetic detection and genotyping. We are sincerely thankful for the laboratory network in Trang province for providing samples for testing, which contributed to the successful completion of this study.

References

[1] Tsegaye MM, Mekonnen AT, Gebretsion DT, Gelanew T, Alemayehu DH, Tefera DA, et al. Predominance of dengue virus serotype-1/genotype-I in Eastern and Southeastern Ethiopia. Viruses. 2024; 16(8): 1334. doi: 10.3390/v16081334.

[2] Sirikut P, Santarattiwong P. Dengue serotype. Pediatric Infectious Disease Society of Thailand [Internet]. 2015 [cited 2024 Oct 11]. Available from: https://www.pidst.or.th/A456.html (in Thai).

[3] Limkittikul K, Brett J, L'Azou M. Epidemiological trends of dengue disease in Thailand (2000-2011): A systematic literature review. PLoS Negl Trop Dis. 2014; 8(11): e3241. doi: 10.1371/journal.pntd.0003241. eCollection 2014.

[4] World Health Organization. Dengue: prevention and control. 2014; EB136/24. [cited 2025 Jan 29]. Available from: https://apps.who.int/gb/ebwha/pdf_files/eb136/b136_24-en.pdf.

[5] Mo L, Shi J, Guo X, Zeng Z, Hu N, Sun J, et al. Molecular characterization of an imported dengue virus serotype 4 isolate from Thailand. Arch Virol. 2018; 163(10): 2903-6. doi: 10.1007/s00705-018-3906-7.

[6] de Souza CS, Caleiro GS, Claro IM, de Jesus JG, Coletti TM, da Silva CAM, et al. Phylogenetics, epidemiology and temporal patterns of dengue virus in Araraquara, São Paulo State. Viruses. 2024; 16(2): 274. doi: 10.3390/v16020274.

[7] Phadungsombat J, Lin MY, Srimark N, Yamanaka A, Nakayama EE, Moolasart V, et al. Emergence of genotype cosmopolitan of dengue virus type 2 and genotype III of dengue virus type 3 in Thailand. PLoS One. 2018; 13(11): e0207220. doi: 10.1371/journal.pone.0207220.

[8] Wang WH, Urbina AN, Chang MR, Assavalapsakul W, Lu PL, Chen YH, et al. Dengue hemorrhagic fever – A systemic literature review of current perspectives on pathogenesis, prevention and control. J Microbiol Immunol Infect. 2020; 53(6): 963-78. doi: 10.1016/j.jmii.2020.03.007.

[9] Department of Medical Services. Guidelines for diagnosis and management of dengue patients (concise edition). Nonthaburi: Ministry of Public Health; 2023. [in Thai].

[10] Leng XY, Zhao LZ, Liao L, Jin KH, Feng JM, Zhang FC. Genotype of dengue virus serotype 1 in relation to severe dengue in Guangzhou, China. J Med Virol. 2024; 96(5): e29635. doi: 10.1002/jmv.29635.

[11] Suangtho P, Buathong R. Dengue virus infection (dengue fever, dengue hemorrhagic fever, dengue shock syndrome). Annual Epidemiological Surveillance Report; 2013: 30-2. [in Thai].

[12] Division of Vector-Borne Diseases, Department of Disease Control. Annual report 2023. 2018 [cited 2025 Jan 29]. Available from: https://drive.google.com/drive/folders/1tV85bvXvX6FZp6EJRvdPJ9J0yu6xFPx9 [in Thai].

[13] Division of Vector-Borne Diseases, Department of Disease Control. Annual report 2023. 2023 [cited 2025 Jan 29]. Available from: https://drive.google.com/drive/folders/1tV85bvXvX6FZp6EJRvdPJ9J0yu6xFPx9 [in Thai].

[14] Division of Vector-Borne Diseases, Department of Disease Control. Forecast of dengue situation, 2023. 2024 [cited 2024 Oct 10]. Available from: https://lookerstudio.google.com/reporting/43e588b9-9773-4918-a56b-76d446496e61/page/IjbnD [in Thai].

[15] Division of Vector-Borne Diseases, Department of Disease Control. Risk assessment report on vector-borne disease. 2023. [cited 2025 Jan 29]. Available from: http://www.ppb.moi.go.th/midev08/upload/moi257.pdf [in Thai].

[16] Division of Vector-Borne Diseases, Department of

Disease Control. Disease Report Form 506: Division of Vector-Borne Diseases. 2023. [cited 2025 Jan 29]. Available from: https://www.tessabantambonbor.go.th/contents/post/000000385-8a63602b02c-d7a544d733b729cf 7108e.pdf [in Thai].

[17] Nonyong P, Ekalaksananan T, Phanthanawiboon S, Aromseree S, Phadungsombat J, Nakayama EE, et al. Dengue virus in humans and mosquitoes and their molecular characteristics in northeastern Thailand 2016-2018. PLoS One. 2021; 16(9): e0257460. doi: 10.1371/journal.pone.0257460.

[18] Pinyorattanachot A, Tempachana M, Senapong N. An analysis of dengue serotypes distribution for surveillance and control in upper southern area of Thailand, 2014-2021. Lanna Public Health J. 2024; 20(1): 99-105 (in Thai).

[19] Lim JK, Chanthavanich P, Limkittikul K, Lee JS, Sirivichayakul C, Lee KS, et al. Clinical and epidemiologic characteristics associated with dengue fever in 2011-2016 in Bang Phae district, Ratchaburi province, Thailand. PLoS Negl Trop Dis. 2021; 15(6): e0009513. doi: 10.1371/journal.pntd.0009513.

[20] Balingit JC, Denis D, Suzuki R, Hayati RF, Ngwe Tun MM, Takamatsu Y, et al. Impact of pre-existing cross-reactive antibodies on cyclic dengue outbreaks in the hyperendemic region of Bali, Indonesia. Virus Res. 2024: 348: 199445. doi: 10.1016/j.virusres.2024.199445.

[21] Chem YK, Yenamandra SP, Chong CK, Mudin RN, Wan MK, Tajudin N, et al. Molecular epidemiology of dengue in Malaysia: 2015-2021. Front Genet. 2024; 15: 1368843. doi: 10.3389/fgene.2024.1368843.

[22] Phadungsombat J, Vu HTT, Nguyen QT, Nguyen HTV, Nguyen HTN, Dang BT, et al. Molecular characterization of dengue virus strains from the 2019-2020 epidemic in Hanoi, Vietnam. Microorganisms. 2023; 11(5): 1267. doi: 10.3390/microorganisms11051267.

[23] Balière C, Calvez E, Thiberge JM, Somlor S, Vandenbogaert M, Grandadam M, et al. A six years (2010-2016) longitudinal survey of the four serotypes of dengue viruses in Lao PDR. Microorganisms. 2023; 11(2): 243. doi: 10.3390/microorganisms11020243.

[24] Ngwe Tun MM, Kapandji M, Wada A, Yamamoto K, Dumre SP, Nwe KM, et al. Performance of Fujifilm dengue NS1 antigen rapid diagnosis kit compared to quantitative real-time polymerase chain reaction. Pathogens. 2024; 13(9): 818. doi: 10.3390/pathogens13090818.

[25] Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, et al. Evaluation of commercially available diagnostic tests for the detection of dengue virus NS1 antigen and anti-dengue virus IgM antibody. PLoS Negl Trop Dis. 2014; 8(10): e3171. doi: 10.1371/journal.pntd.0003171.

[26] Gyurech D, Schilling J, Schmidt-Chanasit J, Cassinotti P, Kaeppeli F, Dobec M. False positive dengue NS1 antigen test in a traveller with an acute Zika virus infection imported into Switzerland. Case Reports Swiss Med Wkly. 2016; 146: w14296. doi: 10.4414/smw.2016.14296.

[27] Aryati A, Wrahatnala BJ, Yohan B, Fanny M, Hakim FKN, Sunari EP, et al. Dengue virus serotype 4 is responsible for the outbreak of dengue in East Java City of Jember, Indonesia. Viruses. 2020; 12(9): 913. doi: 10.3390/v12090913.

[28] Chen-Germán M, Araúz D, Aguilar C, Vega M, Gonzalez C, Gondola J, et al. Detection of dengue virus serotype 4 in Panama after 23 years without circulation. Front Cell Infect Microbiol. 2024; 14: 1467465. doi: 10.3389/fcimb.2024.1467465.

[29] Rojas A, Shen J, Cardozo F, Bernal C, Cabellero O, Ping S, et al. Characterization of dengue virus 4 cases in Paraguay, 2019-2020. Viruses. 2024; 16(2): 181. doi: 10.3390/v16020181.

Journal of Associated Medical Sciences 2025; 58(3): 354-360

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

Wearable inertial sensors for clinical gait analysis in Parkinson's disease: A test-retest reliability study

Thanakorn Angkurasiripaiboon[1,2], Chernkhuan Stonsaovapak[3], Maria Justine[1,4], Akkradate Siriphorn[1*]

[1]Department of Physical Therapy, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand.
[2]Department of Rehabilitation Medicine, King Chulalongkorn Memorial Hospital, Thai Red Cross Society, Bangkok, Thailand.
[3]Department of Rehabilitation Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand.
[4]Centre for Physiotherapy Studies, Faculty of Health Sciences, Universiti Teknologi MARA, Puncak Alam Campus, Selangor, Malaysia.

ARTICLE INFO

Article history:
Received 12 February 2025
Accepted as revised 26 August 2025
Available online 29 August 2025

Keywords:
Gait analysis, inertial measurement units, Parkinson disease, reproducibility of results, spatiotemporal analysis.

ABSTRACT

Background: Parkinson's disease (PD) is a neurodegenerative disorder that significantly affects gait parameters. Accurate and reliable gait assessment tools are essential for monitoring disease progression and treatment efficacy. The G-Walk®, a wearable inertial measurement device, offers a convenient and portable solution for evaluating spatiotemporal gait parameters, but its reliability in individuals with moderate-stage PD remains underexplored.

Objectives: This study aimed to evaluate the test-retest reliability of the G-Walk® sensor system in assessing gait parameters in individuals with moderate-stage PD.

Materials and methods: Thirteen participants with idiopathic PD completed two walking trials on a 14-meter pathway, conducted 3 to 7 days apart. Gait parameters, including stride length, cadence, gait velocity, and limb support times, were recorded using the G-Walk® system. Intraclass correlation coefficients (ICCs), standard error of measurement (SEM), and minimal detectable change (MDC) were calculated.

Results: The G-Walk® system demonstrated excellent reliability for stride length (ICC=0.979), cadence (ICC=0.958), and gait velocity (ICC=0.997). Moderate reliability was observed for single limb support (ICC=0.699) and double limb support times (ICC=0.656). SEM and MDC values further supported the precision of the system for clinical applications.

Conclusion: The G-Walk® sensor system is a reliable tool for assessing gait parameters in individuals with moderate-stage PD, particularly for stride length, cadence, and gait velocity. Its portability and ease of use make it suitable for clinical settings to track disease progression and evaluate interventions.

Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disorder primarily affecting individuals over the age of 60, with a prevalence estimated at 2%-3% in this population.[1] The increasing prevalence with age is attributed to cumulative exposure to neurodegenerative factors, reduced neuronal resilience, and aging-related biological changes that exacerbate dopaminergic neuron vulnerability.[2] The disorder is characterized by the degeneration of dopaminergic neurons and α-synuclein aggregation in the substantia nigra pars compacta within the basal ganglia, leading to dopamine deficiency.[1] This neurotransmitter plays a critical role in motor control, and its depletion results in hallmark motor symptoms such as bradykinesia, resting tremor, rigidity, and postural instability.[3] Among these symptoms, gait disturbances are

∗ **Corresponding contributor.**
Author's Address: *Department of Physical Therapy, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand.*
E-mail address: Akkradate.s@chula.ac.th
doi: 10.12982/JAMS.2025.107
E-ISSN: 2539-6056

particularly significant, profoundly impacting the quality of life in individuals with PD.[4]

Gait abnormalities in individuals with PD typically manifest as reduced stride length, increased cadence, decreased gait velocity, reduced single limb support time, and increased double limb support time.[5] These spatiotemporal gait deficits contribute to a heightened risk of falls and injuries due to impaired balance and reduced stability during walking.[6] They also lead to reduced physical activity levels as individuals may avoid movement to minimize the risk of falling, which in turn exacerbates deconditioning and mobility limitations.[7] Additionally, the cumulative impact of these deficits diminishes the quality of life,[4] as individuals face increasing challenges in performing daily activities and maintaining social engagement. Moreover, these impairments are associated with increased morbidity, mortality, and economic burdens due to higher healthcare costs and social dependencies.[8] Consequently, the ability to accurately assess and monitor gait disturbances is crucial for managing the progression of PD and evaluating the efficacy of therapeutic interventions.

The G-Walk® sensor system, a wearable inertial measurement unit, has emerged as a promising tool for gait analysis. Unlike traditional optoelectronic systems that require extensive setups, fixed laboratory conditions,[9] and are limited by their dependence on walkway-based measurements, the G-Walk® offers portability, ease of use, and affordability, making it highly suitable for clinical applications and enabling gait assessments in more diverse and practical settings.[10, 11] The device measures key spatiotemporal gait parameters, such as stride length, cadence, and limb support times, using advanced sensor fusion technology.[10] This portable device is widely used in clinical and research settings to assess spatiotemporal gait parameters during walking tests, which are essential for assessing gait disturbances in various populations,[10, 12] including those with neurodegenerative diseases like PD.[13] Comparative studies have demonstrated its accuracy in assessing gait, showing comparable results to gold-standard tools like optoelectronic systems and the GAITRite system.[9-11] Despite its widespread use, however, there is limited research on its reliability in assessing gait parameters in individuals with moderate-stage PD, a group characterized by more pronounced gait abnormalities.

Reliability is a critical aspect of any measurement tool, particularly in clinical settings where consistent and accurate data are necessary for decision-making. A previous study on the G-Walk® primarily focused on populations with mild PD or healthy individuals, where the device has demonstrated high test-retest reliability.[14] In moderate-stage PD, clinical symptoms involve more severe gait impairments such as increased postural instability, motor variability, higher fall risk, and fluctuating motor performance due to Levodopa wearing-off effects.[15] These factors pose challenges for reliable assessment, making accurate monitoring in this subgroup essential and highlighting the need for further investigation into the device's consistency in this population.

This study aims to evaluate the test-retest reliability of the G-Walk® sensor system in individuals with moderate-stage PD, focusing on key gait parameters such as stride length, cadence, gait velocity, and limb support times. The findings are expected to inform clinical practice by identifying reliable parameters for tracking disease progression and tailoring therapeutic interventions. Although clinical assessments such as the MDS-UPDRS III and H&Y staging are commonly used, they are limited in their sensitivity to gait variability.[14] While the GAITRite® system, a pressure sensitive walkway, is considered the gold standard in gait parameter measurement and analysis, it presents several limitations: the examination mat requires a large space, is heavy to handle, and has a high initial cost.[10] These factors limit its practicality for routine clinical use. Additionally, the study incorporates the calculation of the standard error of measurement (SEM) and minimal detectable change (MDC) to provide a comprehensive understanding of the system's precision and clinical utility.

This research seeks to establish the G-Walk® as a reliable and accessible tool for monitoring gait impairments in moderate-stage PD, facilitating improved management and intervention strategies.

Materials and methods

Study design

This prospective study assessed the test-retest reliability of the G-Walk® sensor system during walking tasks in individuals with moderate-stage PD which defined as patients with a modified Hoehn and Yahr (H&Y) stage of 2.5 to 3. Stage 2.5 represents a transitional phase with bilateral symptoms and early postural instability that does not yet require assistance, while stage 3 involves bilateral involvement with mild to moderate postural instability, with patients remaining physically independent.[16]

Participants

Thirteen participants with idiopathic PD were recruited from King Chulalongkorn Memorial Hospital, The Thai Red Cross Society, Bangkok, Thailand. Neurologists confirmed the diagnosis based on established clinical criteria. Participants were aged 60–75 years and in the "ON" phase of their PD medication cycle. Inclusion criteria required participants to have a modified H&Y stage of 2.5 to 3, the ability to walk 10 meters independently without assistive devices, and normal cognitive function, defined as a Thai Mental State Examination (TMSE) score >23.[17] Exclusion criteria included individuals with severe visual impairments that were not correctable with devices, as these may interfere with gait assessment. Participants with significant musculoskeletal disorders (e.g., arthritis, joint deformities), cardiovascular conditions (e.g., heart failure, arrhythmias), or neurological conditions (e.g., stroke, multiple sclerosis) unrelated to PD that impair gait patterns were also excluded. Additionally, individuals participating in other clinical trials that might affect gait parameters or interfere with study outcomes were not eligible.

Sample size calculation

The required sample size was calculated using an online sample size calculator for intraclass correlation coefficient (ICC) hypothesis testing.[18] The minimum acceptable reliability (ICC_0) was set at 0.50, and the expected reliability (ICC_1) was set at 0.90. A significance level (α) of 0.05 and power (1-β) of 80% were used, with two repetitions per subject (k=2). Based on these parameters, the sample size was calculated to be 11 participants. To allow for a possible 10% dropout rate, this number was increased, resulting in a final sample size of 13 participants.

Procedure

Relevant demographic and clinical data, including age, sex, weight, height, medical history, duration of disease, timing of medication intake, and any previous injuries or conditions affecting gait, were collected from participants with PD.

Test-retest assessments were conducted over two sessions, separated by an interval of 3 to 7 days, to ensure consistent testing conditions.[19] Practice trials aimed to reduce variability and ensure participants were comfortable with the walking task. Participants were instructed to walk at a self-selected, comfortable pace along a 14-meter straight path. Excluding the initial and final 2 meters minimized errors due to acceleration and deceleration, focusing on stable walking performance (Figure 1).[10] Each participant completed two trials during each session. Participants were given a 2-minute rest or additional recovery time between trials to ensure accuracy and prevent fatigue.

The G-Walk® sensor (BTS G-Sensor 2, BTS Bioengineering Company, Milan, Italy) was used to collect gait data. This portable device features a 4-sensor inertial platform, transmitting data via Bluetooth® 3.0 technology with a range of up to 60 meters. The accelerometer operates at frequencies of 4-1000 Hz, the gyroscope at 4-8000 Hz, and the magnetometer up to 100 Hz, with sensor fusion at a rate of 200 Hz. The portability of the device is enhanced by its compact dimension (70×40×18 mm) and lightweight design (37 grams). For optimal performance, the device was placed in a designated belt pocket aligned with the participant's S1 vertebra in a standing position.[20, 21]

Key gait parameters were recorded during the trials, including gait velocity (meters per second), cadence (steps per minute), stride length (meters), single limb support time (% gait cycle; %GC), and double-limb support time (% gait cycle; %GC). Averages were calculated from two trials for each parameter.

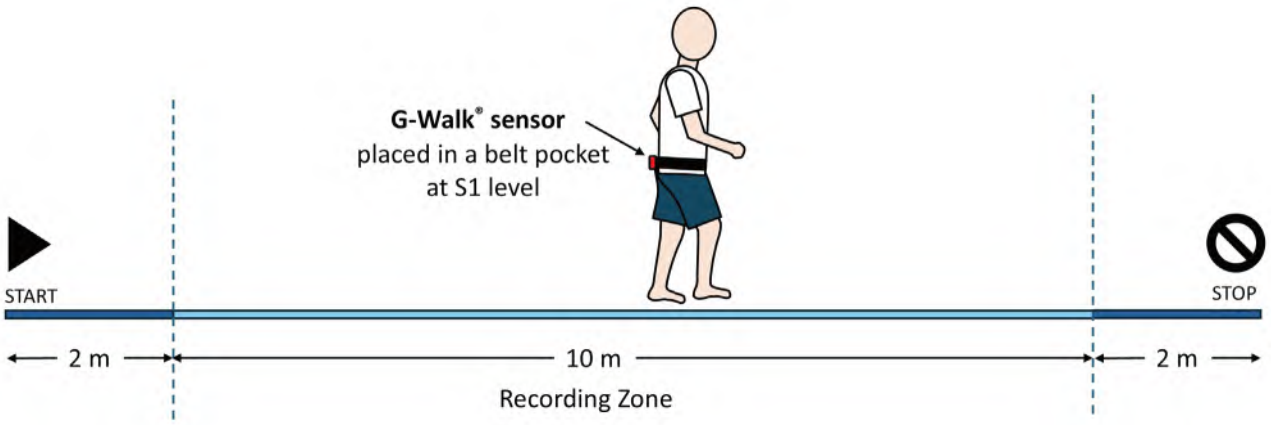


Figure 1*. Study setup for gait assessment using the G-Walk®, a wearable inertial sensor system. Participants walked along a 14-meter pathway, with gait parameters recorded over the middle 10 meters to minimize acceleration and deceleration effects.*

Statistical analysis

Data analysis was conducted using SPSS version 29 (IBM Corp., Armonk, NY, USA). The Shapiro-Wilk test was used to assess the normality of the data, ensuring appropriate statistical tests were applied based on the distribution characteristics. Descriptive statistics summarized the demographic and clinical characteristics of participants, including age, sex, weight, height, medical history, duration of disease, modified H&Y stage, and TMSE scores.

To evaluate test-retest reliability, ICCs were calculated for gait parameters, including stride length, cadence, gait velocity, single limb support time, and double limb support time. The $ICC_{[3,1]}$ model, a two-way mixed effects model assessing absolute agreement for a single rater, was selected, as measurements were obtained by the same rater across multiple sessions. ICC values were interpreted according to established guidelines: values less than 0.49 indicated poor reliability, 0.50-0.74 indicated moderate reliability, 0.75–0.89 indicated good reliability, and values greater than 0.90 indicated excellent reliability.[22]

Additionally, SEM was calculated to estimate the level of error in individual scores, using the formula: $SEM = SD \times \sqrt{(1-ICC)}$,[23] where SD represents the standard deviation of the test scores, and ICC is the intraclass correlation coefficient. To determine the smallest change in a parameter that exceeds measurement error and reflects a meaningful change, the MDC was computed using the formula: $MDC = 1.96 \times \sqrt{2} \times SEM$.[23]

Results

Participant Characteristics

A total of 13 participants with PD were included in the study. The demographic and clinical characteristics are presented in Table 1. Regarding disease severity, six participants were classified as modified H&Y stage 2.5, while seven were classified as stage 3. Cognitive function, assessed using the TMSE, revealed a mean score of 28.62 ±1.61, indicating preserved cognitive function across the cohort.

Regarding medical history, 13 participants reported comorbid conditions that could potentially affect gait, including hypertension (N=2), orthostatic hypotension (N=1), and tachycardia (N=1). However, none of these conditions were active during the tests. No participants reported any previous injuries to the lower limbs or spine. Medication timing varied slightly among participants; however, all participants took their dopaminergic medication within one hour prior to the gait assessment.

Table 1. Demographic and clinical characteristics of participants with Parkinson's disease (N=13).

Characteristics	PD (N=13)
Sex, N (%)	
Male	6 (46.2%)
Female	7 (53.8%)
Age (year)	
Range	60-75
Mean±SD	69.69±5.71
Weight (kg)	59.04±14.20
Height (cm)	160.23±9.13
Duration of disease (year)	6.38±3.07
Modified H&Y stage, N (%)	
Stage 2.5	6 (46.2%)
Stage *3*	7 (53.8%)
TMSE scores	28.62±1.61

Note: *H&Y: Hoehn and Yahr, TMSE, Thai mental state examination.*

Test-retest reliability

The test-retest reliability of the gait parameters assessed using the G-Walk® sensor system is presented in Table 2. ICCs indicated excellent reliability for stride length (ICC=0.979; 95% CI [0.933-0.993]), cadence (ICC=0.958; 95% CI [0.871-0.988]), and gait velocity (ICC=0.997; 95% CI [0.991-0.999]). These results demonstrate a high level of consistency for these parameters across repeated measures.

Moderate reliability was observed for single limb support time (ICC=0.699; 95% CI [0.052-0.907]) and double limb support time (ICC=0.656; 95% CI [0.082-0.894]). The wider confidence intervals for these parameters reflect greater variability and suggest the need for cautious interpretation when assessing temporal support times.

Measurement precision and minimal detectable change

The SEM and MDC, as shown in Table 2, provide additional insights into the precision and clinical relevance of the gait parameter measurements. Stride length exhibited the lowest SEM (0.02 m) and MDC (0.06 m), followed by gait velocity (SEM=0.01 m/s; MDC=0.02 m/s).

In contrast, single limb support time and double limb support time demonstrated higher SEM (0.83% and 0.89%, respectively) and MDC (2.31% and 2.45%, respectively), as detailed in Table 2.

Table 2. Test-retest reliability results for gait parameters assessed using the G-Walk® sensor system.

Gait parameter	Test Mean±SD	Retest Mean±SD	ICC [3,1] [95% CI]	SEM	MDC
Stride length (m)	1.12±0.16	1.13±0.16	0.979 [0.933-0.993]	0.02	0.06
Cadence (steps/min)	113.04±11.30	112.26±10.67	0.958 [0.871-0.988]	2.25	6.24
Gait velocity (m/sec)	1.04±0.16	1.04±0.16	0.997 [0.991-0.999]	0.01	0.02
Single limb support time (%GC)	38.79±1.82	38.36±1.22	0.699 [0.052-0.907]	0.83	2.31
Double limb support time (%GC)	11.21±1.79	11.66±1.23	0.656 [0.082-0.894]	0.89	2.45

Note: *SEM: standard error of measurement, MDC: minimal detectable change, ICC: intraclass correlation coefficient, CI: confidence interval, %GC: percentage of the Gait cycle.*

Discussion

This study aimed to evaluate the reliability of the G-Walk® sensor system in measuring gait parameters in individuals with moderate-stage of PD. The findings demonstrated excellent test-retest reliability for gait parameters such as stride length, cadence, and gait velocity. These results align with a previous study that reported the reliability of the G-Walk® system in individuals with early-stage PD.[10] Additionally, a study assessing test-retest reliability using markerless motion capture systems and traditional motion capture systems in individuals with moderate-stage PD further supports the G-Walk® system's utility in this population.[24]

The high ICC values observed for stride length (0.979), cadence (0.958), and gait velocity (0.997) in this study indicate excellent reliability, consistent with prior research in the early stages of PD.[14] This can be attributed to the fact that the measurements were taken while the patients were in an "ON" phase of their PD medication, allowing them to maintain a consistent walking rhythm. Consequently, variations in stride length, cadence, and gait velocity were minimal [25], resulting in highly stable gait parameters across repeated assessments. However, moderate reliability was observed for single limb support time (ICC=0.699) and double limb support time (ICC= 0.656), which is lower than findings from an earlier study conducted with individuals in the early stages of PD.[14] A possible explanation for this difference is the progression-related postural instability frequently observed in moderate-stage PD.[5] Postural instability may introduce greater variability in limb support times due to increased reliance on compensatory mechanisms or fluctuations in motor control. This limitation may contribute to gait fluctuations and temporal alterations during walking,[26] thereby affecting the consistency of measurements during retesting. Another explanation involves the selective effect of dopamine on gait variability. A previous study found that double limb support time was not significantly influenced by dopaminergic medication, suggesting that dynamic postural control may be more closely associated with cholinergic pathways, particularly involving the pedunculopontine nucleus.[27]

Although the minimum acceptable reliability (ICC_o) for sample size calculation in this study was set at 0.50, it is widely recognized that for key clinical gait parameters such as stride length, cadence, and gait velocity, a higher threshold of ICC (≥ 0.75) is generally recommended to ensure clinical utility.[28] The rationale for selecting ICC_o at 0.50 was to meet statistical requirements for hypothesis testing and to ensure sufficient sample size for detecting reliability across a range of gait parameters, including those expected to have moderate reliability. However, the actual ICC values observed for stride length (0.979), cadence (0.958), and gait velocity (0.997) in this study far exceeded the clinically relevant threshold, indicating that the G-Walk® system provides excellent reliability for these parameters in individuals with moderate-stage PD. This supports the clinical applicability of the device for monitoring disease progression and treatment response.

The calculated SEM and MDC allowed for a more detailed interpretation of measurement variability and responsiveness.[23] Stride length and gait velocity exhibited the smallest SEM (0.02 m and 0.01 m/s, respectively) and MDC (0.06 m and 0.02 m/s, respectively), suggesting their high sensitivity for detecting clinically meaningful changes. These findings reinforce their use as primary indicators in clinical settings for tracking disease progression and evaluating treatment efficacy in individuals with moderate-stage PD. Conversely, the higher SEM and MDC values for single and double limb support times indicate greater variability and a lack of stability in these parameters, potentially limiting their utility for short-term monitoring in this population. The SEM and MDC values observed in this study tend to be lower than those reported in previous research involving children with cerebral palsy using the G-Walk® system, which is the only other population with published MDC data for this device.[12] This comparison suggests that the G-Walk® system may have good sensitivity to detect meaningful changes in populations with different motor control impairments. However, further studies are needed to establish MDC values in other populations such as older adults or stroke patients.

A notable strength of this study is the use of a standardized protocol for test-retest assessments, ensuring consistency and minimizing measurement bias. The inclusion of individuals with moderate-stage PD provides critical insights into the application of the G-Walk® system in this specific population, addressing a gap in the existing literature. Additionally, the portability and ease of use of the G-Walk® system make it highly practical for clinical and community-based settings. The availability of detailed spatiotemporal gait data supports precise monitoring of disease progression and enables timely intervention. From a clinical application perspective, the findings highlight the potential of the G-Walk® system as a compact device for accurately and reliably assessing gait parameters in individuals with PD. It can serve as a convenient and appropriate tool in clinical settings. The obtained data can serve as benchmarks for evaluating and monitoring gait changes across different stages of the disease. Moreover, it aids physicians and physical therapists in using reliable information to plan appropriate treatment and care strategies for patients.

Limitation

This study has certain limitations, including a sample size of 13 participants, which was calculated to meet statistical requirements for ICC hypothesis testing with adequate power. However, despite being statistically adequate, the sample size may still limit the generalizability of the findings to broader populations. Future studies with larger, more diverse cohorts are recommended to further validate these results and improve generalizability. Additionally, this study lacked a healthy control group for comparison. The inclusion of such a group in future research would provide a valuable baseline, allowing for direct comparison of gait parameter

reliability between individuals with PD and the general population. This addition would enhance the validity and clinical relevance of the findings by distinguishing disease-specific measurement variability from normal variability in gait parameters. Furthermore, the selection of patients who could walk without assistive devices makes the results specific to those with moderate-stage PD. Variations in medication timing and individual responses to Levodopa may also have influenced gait performance, potentially introducing variability into the results. Future research should aim to expand the sample size and include participants with more advanced stages of PD or severe gait impairments to provide more comprehensive and generalizable data. Longitudinal studies could also help to assess the responsiveness of the G-Walk® system to changes in disease severity or therapeutic interventions over time.

Conclusions

In conclusion, this study highlights the excellent reliability of the G-Walk® sensor system for measuring key gait parameters such as stride length, cadence, and gait velocity in individuals with PD. In contrast, single and double limb support times showed only moderate reliability, indicating that these parameters should be interpreted with caution in clinical practice. The analysis of SEM and MDC highlights the accuracy and stability of repeated measurements, recommending the use of stable parameters like stride length and gait velocity as primary indicators for tracking changes in clinical settings. Future research is crucial for enhancing the utility of G-Walk® in evaluating gait parameters in PD.

Ethical approval

The study was approved by the Institutional Review Board (COA No.1512/2024). Written informed consent was obtained from all participants prior to their participation.

Funding

This work received no specific grant from any funding agency.

Conflicts of Interest Statement

The authors declare no conflict of interest.

CRedit authorship contribution statement

Thanakorn Angkurasiripaiboon: conceptualization, methodology, validation, formal analysis, investigation, writing: original draft, review and editing; **Chernkhuan Stonsaovapak:** writing: original draft, review and editing **Maria Justine:** writing: original draft, review and editing; **Akkradate Siriphorn:** conceptualization, methodology, formal analysis, writing: original draft, review and editing

Acknowledgements

The author gratefully acknowledges King Chulalongkorn Memorial Hospital, The Thai Red Cross Society, for providing access to its facilities for data collection. Sincere appreciation is also extended to the Department of Rehabilitation Medicine, Chulalongkorn Hospital, for providing the instruments and equipment necessary for this study.

References

[1] Poewe W, Seppi K, Tanner CM, Halliday GM, Brundin P, Volkmann J, et al. Parkinson disease. Nat Rev Dis Primers. 2017; 3: 17013. doi: 10.1038/nrdp.2017.13.

[2] Reeve A, Simcox E, Turnbull D. Ageing and parkinson's disease: Why is advancing age the biggest risk factor? Ageing Res Rev. 2014;14(100):19-30. doi: 10.1016/j.arr.2014.01.004.

[3] Bloem BR, Okun MS, Klein C. Parkinson's disease. Lancet. 2021; 397(10291): 2284-303. doi: 10.1016/S0140-6736(21)00218-X.

[4] Zhao N, Yang Y, Zhang L, Zhang Q, Balbuena L, Ungvari GS, et al. Quality of life in parkinson's disease: A systematic review and meta-analysis of comparative studies. CNS Neurosci Ther. 2021; 27(3): 270-9. doi: 10.1111/cns.13549.

[5] Zanardi APJ, da Silva ES, Costa RR, Passos-Monteiro E, Dos Santos IO, Kruel LFM, et al. Gait parameters of parkinson's disease compared with healthy controls: A systematic review and meta-analysis. Sci Rep. 2021; 11(1): 752. doi: 10.1038/s41598-020-80768-2.

[6] Dibble LE, Addison O, Papa E. The effects of exercise on balance in persons with parkinson's disease: A systematic review across the disability spectrum. J Neurol Phys Ther. 2009; 33(1): 14-26. doi: 10.1097/NPT.0b013e3181990fcc.

[7] Pickering RM, Grimbergen YA, Rigney U, Ashburn A, Mazibrada G, Wood B, et al. A meta-analysis of six prospective studies of falling in parkinson's disease. Mov Disord. 2007; 22(13): 1892-900. doi: 10.1002/mds.21598.

[8] Albarmawi H, Zhou S, Shulman LM, Gandhi AB, Johnson A, Myers DE, et al. The economic burden of parkinson disease among medicare beneficiaries. J Manag Care Spec Pharm. 2022; 28(4): 405-14. doi: 10.18553/jmcp.2022.28.4.405.

[9] Kleiner AFR, Pacifici I, Vagnini A, Camerota F, Celletti C, Stocchi F, et al. Timed up and go evaluation with wearable devices: Validation in parkinson's disease. J Bodyw Mov Ther. 2018; 22(2): 390-5. doi: 10.1016/j.jbmt.2017.07.006.

[10] Vítečková S, Horáková H, Poláková K, Krupička R, Růžička E, Brožová H. Agreement between the gaitrite(®) system and the wearable sensor bts g-walk(®) for measurement of gait parameters in healthy adults and parkinson's disease patients. PeerJ. 2020; 8: e8835. doi: 10.7717/peerj.8835.

[11] De Ridder R, Lebleu J, Willems T, De Blaiser C, Detrembleur C, Roosen P. Concurrent validity of a commercial wireless trunk triaxial accelerometer system for gait analysis. J Sport Rehabil. 2019; 28(6). doi: 10.1123/jsr.2018-0295.

[12] Yazıcı MV, Çobanoğlu G, Yazıcı G. Test-retest reliability and minimal detectable change for measures of wearable gait analysis system (g-walk) in children

with cerebral palsy. Turk J Med Sci. 2022; 52(3): 658-66. doi: 10.55730/1300-0144.5358.

[13] Harro CC, Shoemaker MJ, Coatney CM, Lentine VE, Lieffers LR, Quigley JJ, et al. Effects of nordic walking exercise on gait, motor/non-motor symptoms, and serum brain-derived neurotrophic factor in individuals with parkinson's disease. Front Rehabil Sci. 2022; 3: 1010097. doi: 10.3389/fresc.2022.1010097.

[14] Bailo G, Saibene FL, Bandini V, Arcuri P, Salvatore A, Meloni M, et al. Characterization of walking in mild parkinson's disease: Reliability, validity and discriminant ability of the six-minute walk test instrumented with a single inertial sensor. Sensors (Basel). 2024; 24(2): 662. doi: 10.3390/s24020662.

[15] Smulders K, Dale ML, Carlson-Kuhta P, Nutt JG, Horak FB. Pharmacological treatment in parkinson's disease: Effects on gait. Parkinsonism Relat Disord. 2016; 31: 3-13. doi: 10.1016/j.parkreldis.2016.07.006.

[16] Goetz CG, Poewe W, Rascol O, Sampaio C, Stebbins GT, Counsell C, et al. Movement disorder society task force report on the hoehn and yahr staging scale: Status and recommendations. Mov Disord. 2004; 19(9): 1020-8. doi: 10.1002/mds.20213.

[17] Muangpaisan W, Assantachai P, Sitthichai K, Richardson K, Brayne C. The distribution of thai mental state examination scores among non-demented elderly in suburban bangkok metropolitan and associated factors. J Med Assoc Thai. 2015; 98(9): 916-24.

[18] Arifin WN. Sample size calculator (web) 2024 [cited 2024 Oct 31]. Available from: http://wnarifin.github.io.

[19] Hutin E, Ghédira M, Mardale V, Boutou M, Santiago T, Joudoux S, et al. Test-retest and inter-rater reliability of the 20-meter ambulation test in patients with parkinson's disease. J Rehabil Med. 2023; 55: jrm00378. doi: 10.2340/jrm.v55.4381.

[20] BTS. G-sensor2 hardware manual version 1.2.2. Milan, Italy: BTS Bioengineering Corp.; 2016.

[21] BTS. G-walk user manual version 8.1.0. Milan, Italy: BTS Bioengineering Corp.; 2017.

[22] 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: 10.1016/j.jcm.2016.02.012.

[23] Beninato M, Portney LG. Applying concepts of responsiveness to patient management in neurologic physical therapy. J Neurol Phys Ther. 2011; 35(2): 75-81. doi: 10.1097/NPT.0b013e318219308c.

[24] Kluge F, Gaßner H, Hannink J, Pasluosta C, Klucken J, Eskofier BM. Towards mobile gait analysis: Concurrent validity and test-retest reliability of an inertial measurement system for the assessment of spatio-temporal gait parameters. Sensors (Basel). 2017; 17(7). doi: 10.3390/s17071522.

[25] Ambrus M, Sanchez JA, Sanchez Miguel JA, Del-Olmo F. Test-retest reliability of stride length-cadence gait relationship in parkinson's disease. Gait Posture. 2019; 71: 177-80. doi: 10.1016/j.gaitpost.2019.05.009.

[26] Peterson DS, Horak FB. Neural control of walking in people with parkinsonism. Physiology (Bethesda). 2016; 31(2): 95-107. doi: 10.1152/physiol.00034.2015.

[27] Lord S, Baker K, Nieuwboer A, Burn D, Rochester L. Gait variability in parkinson's disease: An indicator of non-dopaminergic contributors to gait dysfunction? J Neurol. 2011; 258(4): 566-72. doi: 10.1007/s00415-010-5789-8.

[28] Parati M, Ambrosini E, B DEM, Gallotta M, Dalla Vecchia LA, Ferriero G, et al. The reliability of gait parameters captured via instrumented walkways: A systematic review and meta-analysis. Eur J Phys Rehabil Med. 2022; 58(3): 363-77. doi: 10.23736/s1973-9087.22.07037-x.

Journal of Associated Medical Sciences 2025; 58(3): 361-375

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 feasibility of fabricated 3D-printed immobilization masks for radiotherapy: mechanical and dosimetric analysis

Kochakorn Phantawong[1*], Nongnapat Singthuan[1], Kanjana Pinichsai[1], Somchat Taertulakarn[2]

[1]Department of Radiological Technology, Faculty of Allied Health Sciences, Thammasat University, Pathum Thani Province, Thailand.
[2]Department of Medical Technology, Faculty of Allied Health Sciences, Thammasat University, Pathum Thani Province, Thailand.

ARTICLE INFO

Article history:
Received 30 May 2025
Accepted as revised 26 August 2025
Available online 29 August 2025

Keywords:
Radiotherapy, thermoplastic mask, 3D printing, infill pattern.

ABSTRACT

Background: 3D printing technology is becoming widely popular in the medical field, particularly in radiotherapy, where it is used to fabricate immobilization masks from thermoplastic materials because it can generate geometric figures and produce sophisticated templates eligible for individual patients.

Objectives: This study aims to evaluate the mechanical and dosimetric properties of three types of plastic: polylactic acid (PLA), polyethylene terephthalate glycol-modified (PETG), and acrylonitrile styrene acrylate (ASA) to identify an appropriate material for use as a 3D-printed thermoplastic mask in radiotherapy.

Materials and methods: The 3D-printed square plastic samples were designed with four infill patterns: triangle, honeycomb, gyroid, and honeycomb-gyroid, and were manufactured using fused deposition modeling (FDM) 3D printing technology. This study analyzed the characteristics regarding mechanics, uniformity, and radiation transmission.

Results: Mechanical testing demonstrated that PETG provided high strength (10.88 MPa with gyroid pattern) and flexibility (7.36% with honeycomb pattern). The mean uniformity across tests ranged from 84.80 to 85.03, with a standard deviation between 5.93 and 7.53. All plastic types have Hounsfield units (HU) ranging from -774.88 to -620.21. The 3D-printed plate demonstrated radiation transmission like the efficiency of a commercial thermoplastic mask, in the range of 99.30% to 99.91%.

Conclusion: All three types of plastic have proven to be appropriate alternative materials for the integration of an immobilization mask. PLA plates displayed advantageous results in terms of radiation transmission, while PETG emerged as the most balanced performer in terms of mechanical properties, showing high strength and flexibility. ASA, a polymer confirmed by FTIR analysis that contains a benzene ring, demonstrated that it is more resistant to radiation. The 3D-printed masks provide a viable alternative to traditional thermoplastic mask systems. Nevertheless, further investigation is required.

Introduction

Radiation therapy for cancer is typically divided into multiple doses known as fractionation. In daily irradiation treatment, the patient should constantly be positioned in the same way. Immobilization devices are used to help set up patient placement. The immobilization device aids in positioning the patient in a position that remains constant during treatment simulation. This allows the tumor to get precisely the dose of radiation required. It also decreases radiation exposure to adjacent normal tissue areas.

The fabrication of products by 3D printing is currently

* **Corresponding contributor.**
Author's Address: *Department of Radiological Technology, Faculty of Allied Health Sciences, Thammasat University, Pathum Thani Province, Thailand.*
E-mail address: kochakorn.pha@allied.tu.ac.th
doi: 10.12982/JAMS.2025.108
E-ISSN: 2539-6056

prevalent. This method can generate geometric figures and produce sophisticated templates. It is utilized as a material for fabricating thermoplastic masks to immobilize patients during radiation therapy. Carbon fibers are commonly used as the material for these masks. Nonetheless, carbon fibers are slightly costly.[1] In the last two decades, 3D printing has been used for medical purposes, such as surgical planning, dosimetry devices, and brachytherapy applications.[2] 3D printing of immobilizers has created a lot of interest in radiotherapy. The fabrication method was revealed to be essentially fused deposition modeling (FDM) printing, which continues to develop rapidly and seems to hold unlimited possibilities across numerous applications. Medical applications have applied the layer-by-layer approach of FDM to produce prosthetics, medical devices, and organ models. The capabilities of this technology are consistently growing due to ongoing advancements and findings, especially in the areas of materials and printing devices.[3] The literature highlights several benefits of 3D-printed immobilizers, such as enhanced patient experience and comfort compared to traditional methods. Additionally, they illustrate high accuracy in fitting patients, allow for repeatable setups, and exhibit beam attenuation properties comparable to those of thermoformed immobilizers. The drawbacks included the slow production rate of 3D printing and the potential for inaccuracies in digitizing patient geometry.[4]

This study examined three types of plastics: 1) polylactic acid (PLA), 2) polyethylene terephthalate glycol-modified (PETG), and 3) acrylonitrile styrene acrylate (ASA). All three sorts of plastics are classified as thermoplastics, yet they have different qualities. PLA is a highly potential biopolymer because its monomers may be derived from non-toxic resources, and it is a naturally produced organic acid. PLA is utilized for several applications in the medical industry. PLA as an appropriate material for 3D printing in external beam radiotherapy.[5,6] PETG is a material that is very reliable and flexible, resistant to chemicals, and frequently utilized in the medical industry since it can be sterilized without deforming the product. ASA material has properties like ABS plastic. However, ASA is more robust, resistant to UV radiation, and less expensive than ABS plastic.[7,8]

The objective of this study is to identify a suitable material for a 3D-printed (additive manufacturing) thermoplastic mask in radiotherapy. This was assessed by looking at the strength and radiation properties of three types of plastic: polylactic acid (PLA), polyethylene terephthalate glycol-modified (PETG), and acrylonitrile styrene acrylate (ASA), using different internal designs like triangular, honeycomb, gyroid, and a mix of honeycomb and gyroid patterns.

Material and methods

Manufacturing of the samples

The sample was designed using Autodesk Fusion 360 software. The design involved creating samples measuring 10 x 10 cm^2 with a thickness of 2 mm using three types of plastics (Ultramo, Thailand) : PLA, PETG and ASA (Table 1). All the materials are available in the form of filaments with diameters of 1.75 mm. Four internal structures were designed for each plastic: triangle, honeycomb (hexagon), gyroid, and honeycomb-gyroid, as shown in Figure 1. Models were exported to Stereolithography (STL) file format and converted to G-code in HP SmartStream 3D Build Manager software (HP, Palo Alto, USA). G-code is the low-level machine language that controls the printer.[9] A 3D model is created layer by layer using the 3D printer (Flashforge, Model: Creator 3, China) based on the Fused Deposition Modeling (FDM) technique.

Table 1. Properties of plastic material and print setting.

Physical property	PLA	PETG	ASA
Chemical formula	$(C_3H_4O_2)_n$	$(C_{10}H_8O_4)x(C_8H_{16}O_2)_m$	$C_{18}H_{23}NO_2$
Density (gm/cm^3)	1.25	1.27	1.08
Flexibility	Low	High	Medium
Durability	Medium	High	High
Printing temperature (°C)	190-210	190-210	240-260
Print speed (mm/sec)	50-80	40-60	30-50
Nozzle temperature (°C)	200	240	245
Bed temperature (°C)	50-60	50-60	75-95

Note: *PLA: polylactic acid, PETG: polyethylene terephthalate glycol-modified, ASA: acrylonitrile styrene acrylate.*

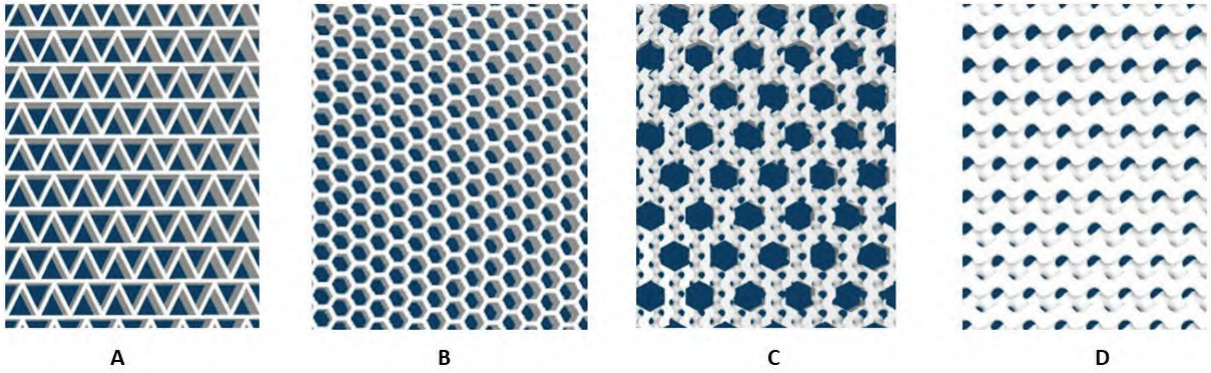


Figure 1. *Infill pattern. A: triangle, B: honeycomb or hexagonal, C: honeycomb-gyroid, D: gyroid.*

Testing chemical properties

Fourier transform infrared (FTIR) analysis

Specimens of the 3D-printed plastic samples were investigated by Fourier-transform infrared (FT-IR) spectroscopy. The FTIR uses interferometry to encode information from a couple of materials placed within the IR beam. The profile was generated by the Nicolet iS50 FTIR Spectrometer (Thermo Scientific, USA) equipped with a rapid recovery deuterated triglycine sulfate (DTGS) detector. FTIR analysis could be conducted in Attenuated Total Reflection (ATR) mode, with a diamond as the crystal used. The obtained spectra were analyzed with the integrated OMNIC software (Thermo Scientific, USA). Spectra were collected between 4000 and 400 cm^{-1}, with a resolution of 4.0 cm^{-1} and 64 co-added scans.

Testing of the mechanical properties

Testing with simulation mode

Autodesk Fusion 360 software is used for the mechanical simulation step. Stress, tension, and flexibility tests were applied to the plastic plates. A force of 10 Newtons or 1.02 kilograms was applied, based on the average force applied to thermoplastic masks for head and neck.[10] The software requires the specification of the sample material to mimic its mechanical properties. An attempt was made to reverse the g-code of each plate and utilize the STL format to acquire the various infills as an initial modeling approach for the plates. This approach, however, failed because the STL files lacked infill information about gyroid and honeycomb-gyroid.[1]

Testing with universal testing machine (UTM)

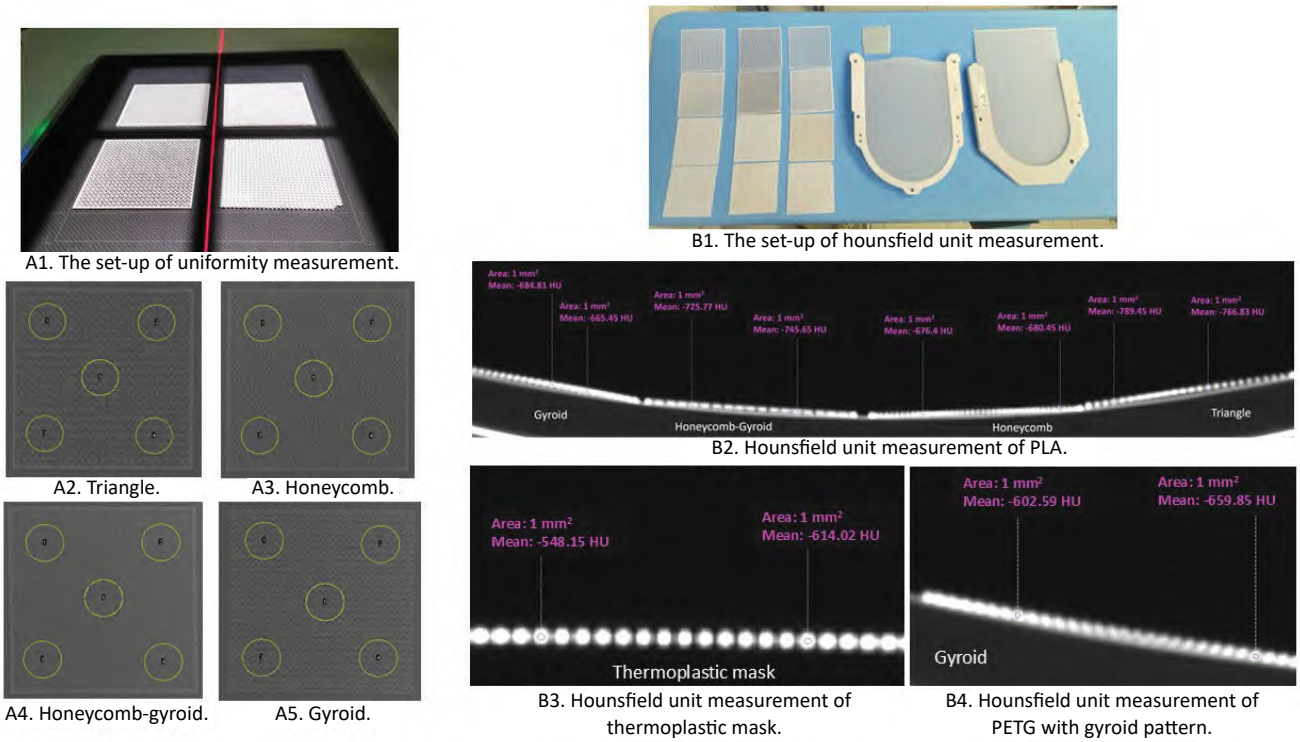
Tensile testing was conducted with universal test machines (Lloyd Instruments LS 2.5, United Kingdom). All materials and infill patterns were subjected to tension testing at a speed of 20 mm/min, with a gauge length of 30 mm, a thickness ranging from 1.76 to 2.38 mm, and a cross-sectional area between 18.3 and 23.8 mm^2. The tensile mechanical qualities, including maximum load and machine extension at maximum load, were examined. Following testing, the obtained data underwent analysis, and the results are provided through ultimate tensile strength, strain at maximum load, Young's modulus, and elongation.

Testing of the uniformity

The sample was x-rayed by a general x-ray machine (FDR Smart X, Fujifilm Medical System, Japan) with an exposure technique of 50 kVp, 3.2 mAs, and a source-to-image distance of 100 cm. The resulting images were examined for uniformity using ImageJ software, which included providing five same-size regions of interest (ROIs) and calculating both the mean and standard deviation. All regions of interest were clearly defined with an area of 5035-$pixel^2$.The measurement of uniformity for 3D-printed plastic sheets was illustrated in Figure 2 (A1-A5).

Testing of the density of samples using Hounsfield unit (HU) values

Samples of 10×10 cm^2 for Hounsfield unit (HU) assessments were scanned applying a Siemens SOMATOM Definition AS CT scanner (Global Siemens Healthcare, Erlangen, Germany) with 120 kVp tube potential, 270 mAs tube current and scan time, 500 mm field of view (FOV), 1.00 mm slice thickness, 0.4 pitch were used. The circular regions designated for measuring HU were precisely located at 1 mm^2 on middle axial slice CT images, with average HU values derived from five ROIs for each 3D-printed plate. The resulting images were evaluated for density by comparing the HU values of the sample plastic plates with those of a thermoplastic mask. Figure 2 (B1-B4).



A1. The set-up of uniformity measurement.

B1. The set-up of hounsfield unit measurement.

A2. Triangle.

A3. Honeycomb.

B2. Hounsfield unit measurement of PLA.

A4. Honeycomb-gyroid.

A5. Gyroid.

B3. Hounsfield unit measurement of thermoplastic mask.

B4. Hounsfield unit measurement of PETG with gyroid pattern.

Figure 2. *The measurement of uniformity and density for 3D-printed plastic sheets. A1-A5: measurement of uniformity for 3D-printed plastic sheets, B1-B4: testing the density of the samples using Hounsfield unit (HU) values*

Testing of the radiation transmission

The sample irradiation was performed in the Radiation Oncology Center at Thammasat University hospital, Pathum Thani, Thailand. TRS-398 guidelines were followed in the calibration output, as recommended by the International Atomic Energy Agency (IAEA). The radiation dose was determined by inserting a 0.6 cm^3 Farmer-type ionization chamber (PTW 30013, Freiburg, Germany) into a plastic slab phantom (Virtual Water™, Med-Cal, Tennessee, USA). The dose of radiation was measured at a depth of 5 cm from the surface of a 30 cm tissue-equivalent material (density was 1.04 gm/cm^3, and a 30x30 cm size) or the surface-to-ionization chamber distance was 5 cm. The field size was 10x10 cm^2, and the source-to-skin distance (SSD) was 95 cm. The radiation was delivered at a dose of 2 Gy (600 MU/min) using a 6 MV photon beam on the Elekta Versa HD linac (Elekta Oncology Systems, Crawley, UK). The radiation passing through the object was measured using a Farmer-type ionization device with a volume of 0.6 cm^3. The object was placed on the tissue-equivalent material. The dose was quantified using an ionization chamber, followed by a comparison with a conventional immobilization thermoplastic mask of approximately 2 mm thick. The setup for testing radiation transmission was shown in Figure 3. The transmission of radiation emerges as it passes each plate, calculated as follows:

$$\% \text{ of transmitted dose} = \frac{\text{plastics sample measurement}}{\text{dose from control reading}} \times 100$$

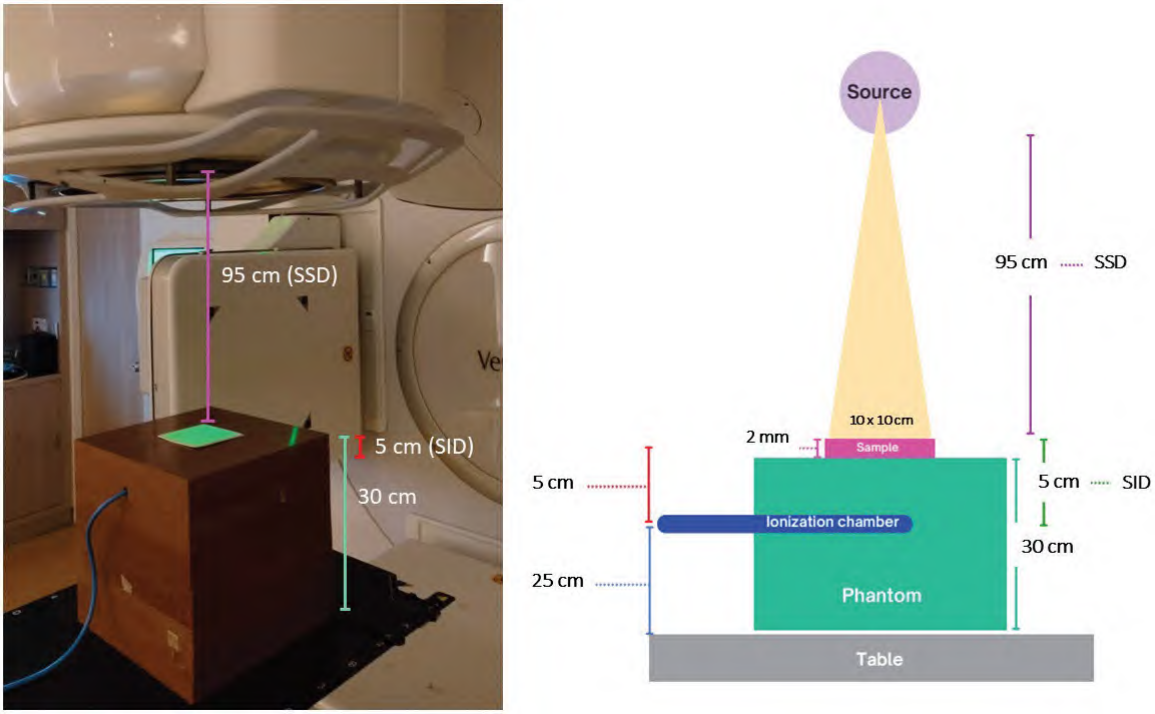


Figure 3. *Equipment setup for measuring the radiation transmission.*

Result

Fourier transform infrared (FTIR) analysis

FTIR analysis for PLA, PETG, and ASA filaments is given in Figure 4 (A). The FTIR spectra of the 3D printed plastics, recorded in the 4000-400 cm^{-1} wavenumber range, are graphically described.

The characteristic stretching peaks of PLA are the C-O stretching at 1081.83 cm^{-1}, the C-O-C stretching at 1180.49 cm^{-1}, the C-H bending at 1358.90 cm^{-1} and 1452.15 cm^{-1}, the C=O stretching at 1747.47 cm^{-1}. The CH3 asymmetric and CH3 symmetric bending frequencies were observed at 2995.13 and 2944.51 cm^{-1}, respectively.

In case of PETG, the characteristic bands of -CH were observed at 724.42 cm^{-1}. The peaks at 1016.47 and 1093.45 cm^{-1} are related with C-H bond vibrations. The peaks at 1407.97 and 1239.95 cm^{-1} are associated with the -CH2 deformation band and C-O-O stretching of ester groups, respectively. The peak at 1712.98 cm^{-1} frequency shows C=O ester groups. Asymmetric and symmetrical aliphatic C-H stretching are identified at frequencies of 2855.37 and 2927.41 cm^{-1}, respectively.[11,12]

For investigation of ASA, the apparent peak at 1732.36 cm^{-1} was the C=O stretching vibration. The appearance of C≡N absorption band at 2236.85 cm^{-1} gave strong supporting evidence that the ASA. The bands with a peak value at around 1601.90, 1493.67, and 1452.67 cm^{-1} were assigned to stretching vibration peak of the benzene ring skeleton. Besides, peaks at 759.16 and 698.42 cm^{-1} represented C-H bending.[12] Figure 4 (B-D) showed the FTIR significant peaks for PLA, PETG, and ASA along with the associated assignments.

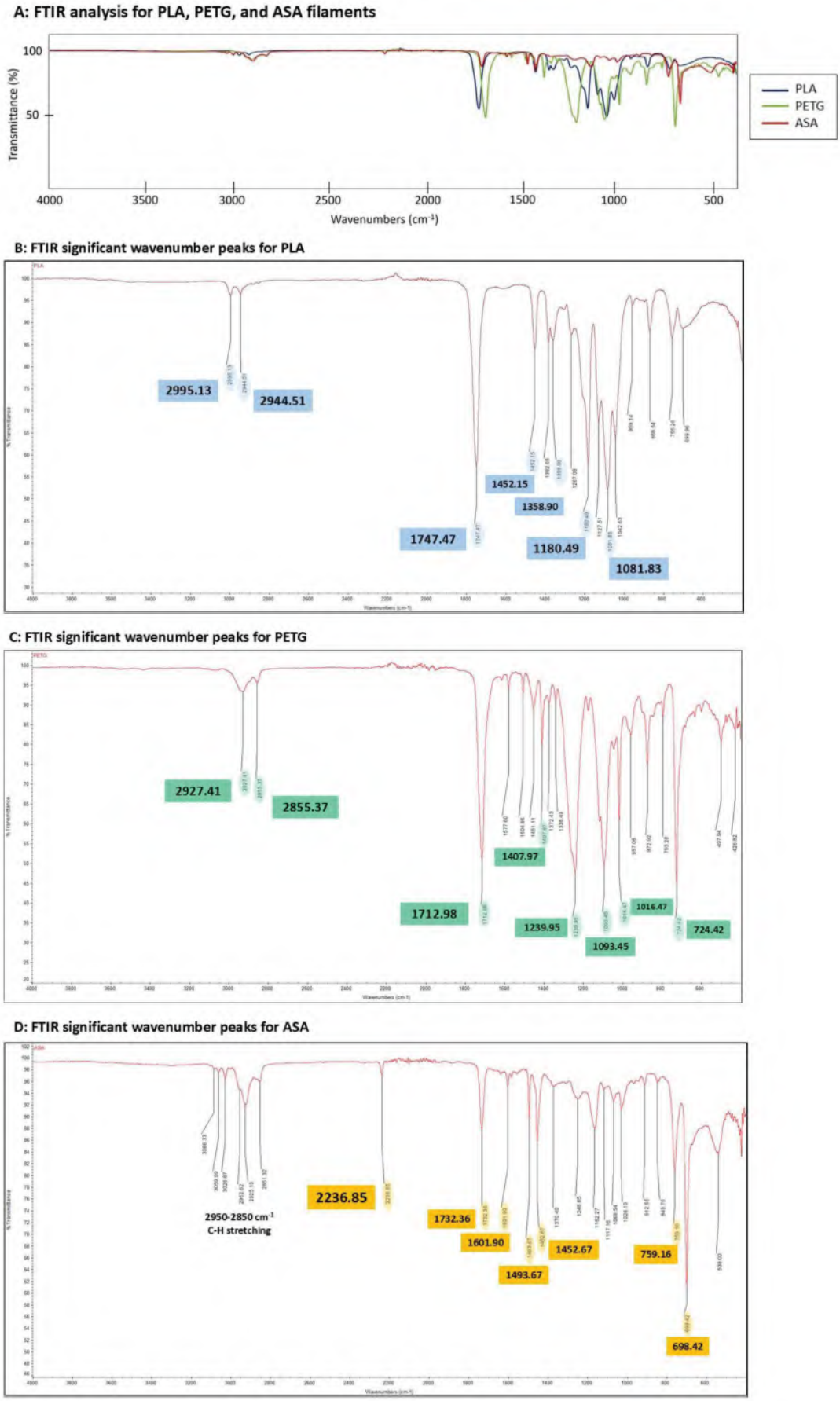


Figure 4. *FTIR analysis of 3D printed specimens. A: FTIR analysis for PLA, PETG, and ASA filaments, B: FTIR significant wavenumber peaks for PLA, C: FTIR significant wavenumber peaks for PETG, D: FTIR significant wavenumber peaks for ASA.*

Mechanical testing

Simulation-based testing using Autodesk Fusion 360

The mechanical characteristics were evaluated using numerical simulation. The outcomes of the mechanical test are presented in Table 2. All testing was performed utilizing Autodesk Fusion 360 in simulation mode. The stress testing, which evaluates the object's capacity to withstand external forces and its flexibility, revealed that the PLA plastic sample featuring a honeycomb design had the highest strength at 41.46 MPa and a flexibility of 1.46 mm in the triangular infill pattern. In strain testing, which measures the force inducing deformation in an item, the ASA-honeycomb sample demonstrated a maximum value of 0.037 MPa. Figure 5 demonstrates the normalized physical properties of the material manufactured by the 3D printer.

Table 2. Mechanical properties test

Plastic type	Infill pattern	Stress	Strain	Flexibility
		Maximum (MPa)	Maximum (MPa)	Maximum (mm)
PLA	Triangle	20.82	0.02	1.46
	Honeycomb	41.46	0.04	0.39
	Honeycomb-gyroid	0.33	0.00*	0.01
	Gyroid	0.44	0.00*	0.07
PETG	Triangle	16.84	0.02	0.54
	Honeycomb	30.07	0.04	0.54
	Honeycomb-gyroid	0.33	0.00*	0.02
	Gyroid	0.40	0.00*	0.10
ASA	Triangle	20.81	0.02	0.42
	Honeycomb	39.56	0.04	0.42
	Honeycomb-gyroid	0.32	0.00*	0.01
	Gyroid	0.43	0.00*	0.01

Note: **values less than 0.01.*

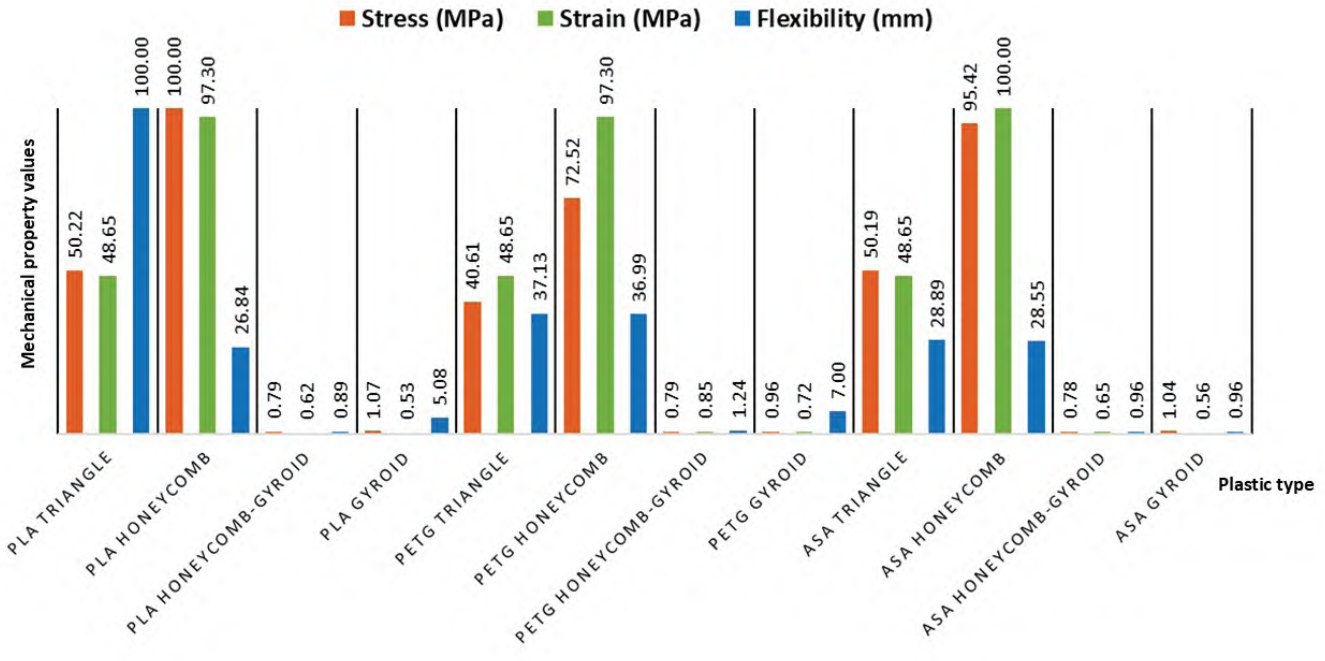


Figure 5. *The mechanical properties, including stress, strain, and flexibility, of each plastic material and infill pattern.*

Strength and flexibility testing using universal testing machine (UTM)

Testing was performed utilizing the Lloyd Instruments LS 2.5 Universal Testing Machine. PLA exhibited the highest ultimate tensile strength with the gyroid pattern (8.59±1.67 MPa) and Young's modulus (461±19 MPa); however, the triangle pattern displayed superior flexibility, achieving maximum strain at failure (3.36±0.67%). The gyroid pattern for PETG had the highest ultimate tensile strength (10.88±1.55 MPa), but the honeycomb pattern displayed enhanced flexibility with a maximum strain at failure of 7.36±1.87%. Furthermore, the gyroid pattern demonstrated superior stiffness with a Young's modulus of 258±41 MPa, significantly outperforming other patterns. For ASA, The gyroid pattern demonstrated superior stiffness (415±15 MPa Young's modulus) and high tensile strength (9.38±0.40 MPa), whereas the triangle pattern showed exceptional flexibility (5.09±0.17% strain at maximum load). The values of strength and flexibility obtained using the universal testing machine (UTM) are presented in Table 3 and Figure 6.

Table 3. The mechanical performance of each 3D-printed specimen with four different infill patterns.

Plastic type	Infill pattern	Ultimate Tensile Strength (MPa)±SD	Strain at Max load (%)±SD	Young's Modulus (MPa)±SD	Elongation (mm)
PLA	Triangle	7.79±1.35	3.36±0.67	292±24	1.16
	Honeycomb	8.26±0.21	2.96±0.53	336±47	1.07
	Honeycomb-gyroid	2.63±0.53	2.48±0.95	160±11	1.05
	Gyroid	8.59±1.67	2.48±0.27	461± 9	0.85
PETG	Triangle	6.67±0.51	4.13±0.27	192±8	1.34
	Honeycomb	9.40±1.49	7.36±1.87	214±7	2.81
	Honeycomb-gyroid	5.37±0.45	6.76±1.50	175±10	2.35
	Gyroid	10.88±1.55	5.86±0.23	258±41	1.81
ASA	Triangle	4.62±0.26	5.09±0.17	151±8	1.58
	Honeycomb	8.93±0.75	3.76±0.56	278±10	1.24
	Honeycomb-gyroid	4.00±0.14	4.10±0.41	164±7	1.32
	Gyroid	9.38±0.40	2.89±0.20	415±15	0.94

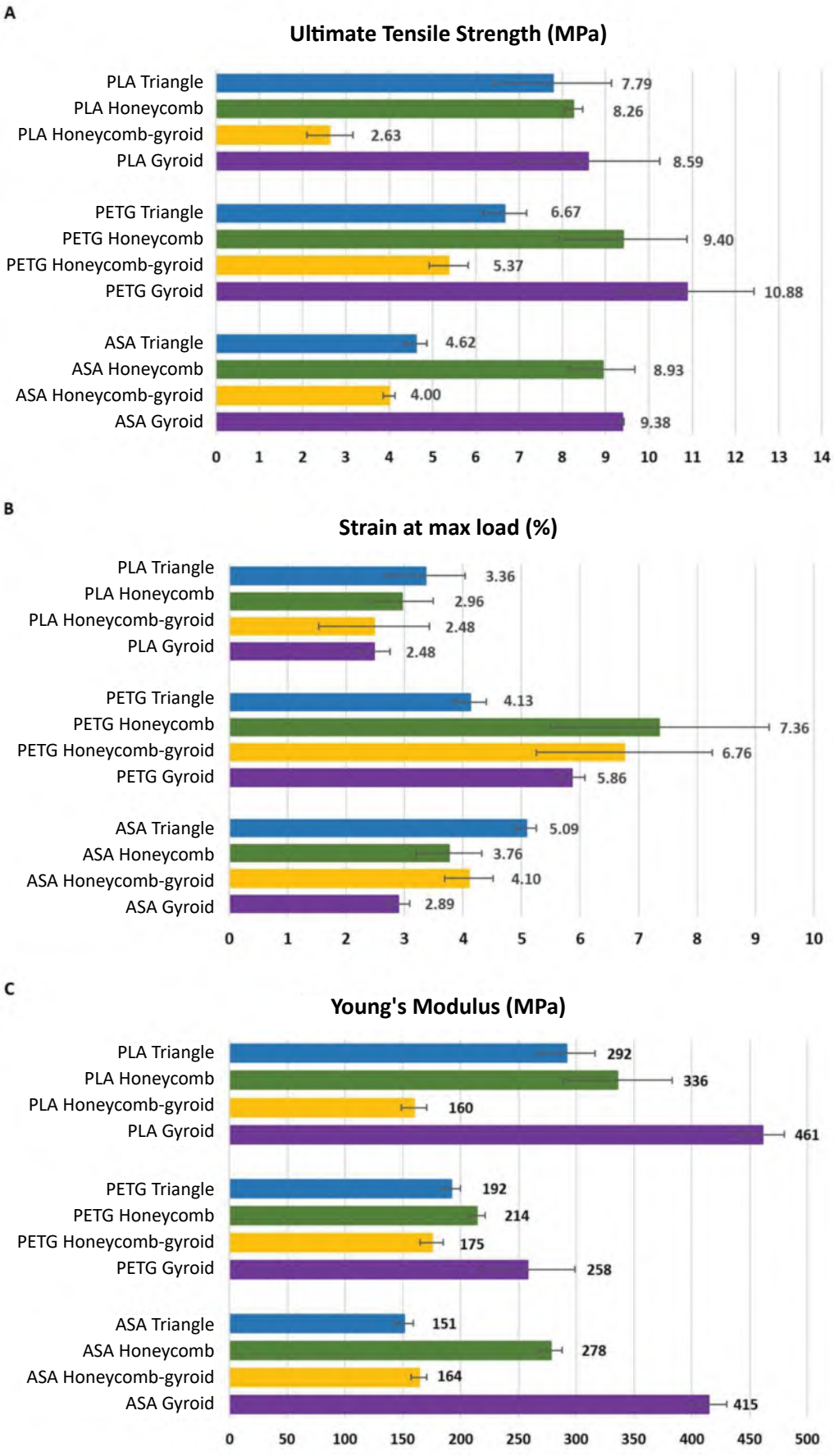


Figure 6. *The mechanical properties of UTM of each plastic material and infill pattern. A: ultimate tensile strength, B: strain at max Load, C: Young's Modulus (stiffness).*

Uniformity and Hounsfield unit (HU) Testing

The uniformity assessment of the objects was performed by radiography each sample and evaluating the resultant images for uniformity utilizing the Image J software. The region of interest (ROI) was delineated at five spots on the pictures, and the homogeneity of each sample together with the thermoplastic mask was evaluated. The homogeneity of each plastic type—PLA, PETG, and ASA—exhibits the following ranges of mean and standard deviation: 80.31-89.13 (5.27-10.33), 79.81-89.76 (5.44-8.98), and 79.99-88.66 (4.05-8.18), respectively. The previous values are presented in Table 4. The density of the 3D-printed plastic plates, measured in Hounsfield units (HU), was evaluated using the Synapse® PACS (Fujifilm Medical System, Japan), as shown in Table 4. The HU ranges for PLA, PETG, and ASA are -774.88 to -677.6, -730.27 to -649.24, and -757.7 to -620.21, respectively.

Table 4. Uniformity and Hounsfield unit of 3D-printed samples and standard thermoplastic mask

Plastic type	Infill pattern	Uniformity (gray values) Mean±SD	Hounsfield unit (HU) Mean±SD
Uni-frame® Thermoplastic mask		79.00±7.37	-567.53±40.09
Type-S™ Thermoplastic mask		79.90±7.44	-559.17±42.0
PLA	Triangle	80.31±10.33	-774.88±12.09
	Honeycomb	84.09±6.68	-699.43±22.23
	Honeycomb-gyroid	89.13±8.23	-722.48±25.29
	Gyroid	85.67±5.27	-677.6±23.27
PETG	Triangle	79.81±6.59	-730.27±36.36
	Honeycomb	83.98±6.47	-724.29± 20.61
	Honeycomb-gyroid	89.76±8.98	-711.25±34.87
	Gyroid	86.55±5.44	-649.24±29.10
ASA	Triangle	79.99±8.18	-757.7±34.02
	Honeycomb	84.22±6.76	-699.5±28.32
	Honeycomb-gyroid	88.66±4.05	-702.59±26.68
	Gyroid	86.58±4.73	-620.21±27.70

Dosimetric properties

As illustrated in Table 5, the percentages of dose differences and transmitted radiation are presented. The dose difference was determined by reducing the initial radiation dose from the dose that passed through the sample. A standard immobilization thermoplastic mask (CIVCO Medical Solutions, Iowa, USA) approximately 2.4 mm thick, was irradiated to determine a term of comparison.

Table 5. Radiation transmission in different plastic type plates and infill pattern.

Plastic type	Infill pattern	% Dose difference	% Transmitted dose
No mask (control)		-	-
Uni-frame® Thermoplastic mask		0.4391	99.56
Type-S™ Thermoplastic mask		0.3585	99.64
PLA	Triangle	0.0896	99.91
	Honeycomb	0.1344	99.87
	Honeycomb-gyroid	0.4584	99.54
	Gyroid	0.1165	99.88
PETG	Triangle	0.5992	99.40
	Honeycomb	0.5834	99.42
	Honeycomb-gyroid	0.5318	99.47
	Gyroid	0.6994	99.30
ASA	Triangle	0.4693	99.53
	Honeycomb	0.5229	99.48
	Honeycomb-gyroid	0.5229	99.48
	Gyroid	0.5854	99.41

Comparable results were seen when the differences in transmission between each plate and the conventional mask were below 0.7%. PLA with triangular infill demonstrated superior performance compared to other materials for 6 MV photon energy, exhibiting a lower attenuation of 0.09% versus 0.4% for the typical thermoplastic mask. The analysis of radiation dose transmission through 3D-printed masks made from PLA plastic (triangle, gyroid, honeycomb) reveals a transmission percentage ranging from 99.87% to 99.91%. This percentage is notably higher than the standard thermoplastic mask, which transmits between 99.56% and 99.64%, as illustrated in Figure 7.

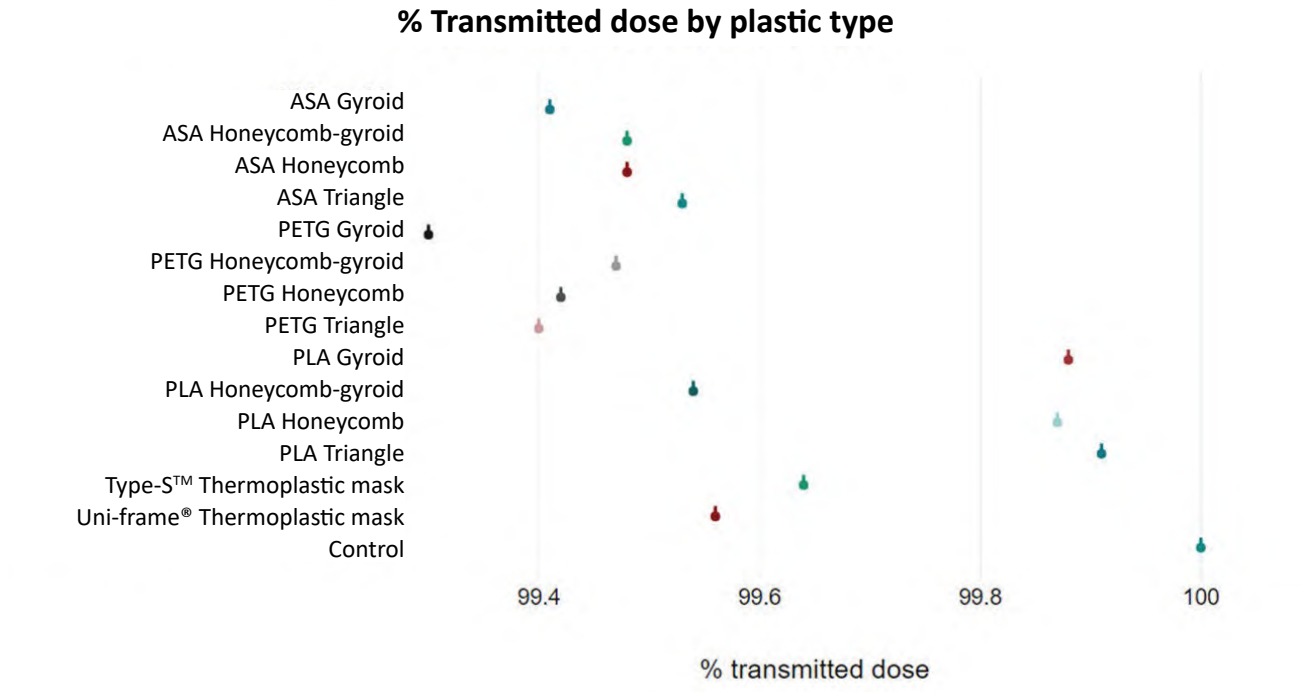


Figure 7. *Variation of the transmitted radiation in percentage for different plastic type plates and infill pattern. (Graph plotted by DATAtab.net, Graz, Austria)*

Discussion

Surgery, radiotherapy, and chemotherapy are widely utilized approaches in the treatment of cancer. A significant majority of patients with cancer receive treatment through radiotherapy. The primary objective in radiotherapy is to achieve optimal dose delivery to the tumor while reducing the side effects associated with the treatment. Therefore, it is crucial to carefully execute the radiation treatment design. While the introduction of advanced radiotherapy technologies has minimized errors associated with reconstructing the patient's geometry, certain inaccuracies remain in the production of accessories. The application of medical devices is not appropriate for every patient. Currently, 3D printing technology is regarded as a promising approach that can efficiently adapt to the treatment of individuals on a personalized basis. Consequently, there is significant interest in the utilization of 3D printing technology within the radiotherapy process.[13]

A variety of materials have been utilized to create the intended 3D object using the 3D printer, such as polylactic acid (PLA), acrylonitrile butadiene styrene (ABS), polyethylene terephthalate glycol (PETG), thermoplastic elastomers (TPE), polyamide (PA, commonly referred to as nylon), thermoplastic polyurethane (TPU), and polyvinyl acetate, all of which have been assessed for their application in radiation therapy.[14,15]

FTIR analysis was used to examine the chemical characteristics of three different types of filaments that were produced for 3D-printed sheets. The transmittance bands from 2996–2944 cm^{-1} correspond to the symmetric and asymmetric stretching of the CH3 group. The band approximately of 1750 cm^{-1} (range of 1745-1758 cm^{-1}) corresponds to the vibration of the C=O bond of PLA because of the presence of the ester. This peak was intense due to the concentration of PLA and ester in the formulation. The band at 1267 cm^{-1} is associated with the C=O bending vibration, while the bands at 1180 and 1130 cm^{-1} correspond to the C-O-C stretching vibration of the ester groups in PLA polymer.[16]

The % transmittance at around 1717 cm^{-1} indicates a significant peak of PETG associated with the C=O stretching vibration in the ester group, contrasting with PLA, which exhibits a band at 1750 cm^{-1}. The PETG spectrum has peaks at around 1505, 875, and 730 cm^{-1}, indicative of aromatic bond vibrations that are absent in the PLA spectra.[17]

The C≡N stretching band (2236-2238 cm^{-1}) is the most significant and characteristic peak of ASA, indicating the vibration of the C≡N bond in the nitrile group of the acrylonitrile unit. The C=O stretching seen at around 1733-1735 cm^{-1} indicates the vibration of the C=O bond within the ester group of the acrylate component in ASA. The aromatic C=C stretching peaks are observed at roughly 1602 and 1493 cm^{-1}, signifying the vibration of the

C=C bond within the benzene ring, a hallmark of the aromatic structure in the styrene component. This verifies that ASA is a terpolymer consisting of three primary constituents: acrylonitrile, styrene, and acrylate.[12]

FTIR analysis in this experiment revealed that PLA corresponded with polylactide at 95.65%; PETG aligned with poly (ethylene terephthalate) at 95.11%; and ASA matched with copier toner (ether extract) styrene acrylate copolymer at 87.40% and poly (styrene: acrylonitrile: methyl acrylate) at 83.91%. The matching was conducted with the reference spectrum from the OMNIC software library.

Different materials and filling patterns will yield varying outcomes based on mechanical test analysis. In Autodesk Fusion 360 simulation testing, the strength testing showed that honeycomb patterns are the strongest, supporting the results of previous study[3], which found that hexagons have better stress and strain scores than other shapes. Flexural testing indicates that PLA with a triangle infill pattern exhibits the greatest flexibility, followed by PETG and ASA materials. Most types of thermoplastic masks allow deviations of 2 to 5 mm.[18] In all plastic plates, the displacement (in mm) was less than 1.46 mm. This was below the previously reported range of 2 to 5 mm and proved to be an excellent result.

While, in universal tensile machine (UTM) testing, the gyroid infill pattern revealed the highest ultimate tensile strength, closely followed by honeycomb across all plastic types. These findings correspond with recent studies indicating that complex 3D infill patterns, such as gyroid, offer enhanced strength-to-weight ratios owing to their isotropic load distribution properties.[19,20]

Our results demonstrate that material-infill pattern interactions significantly influence mechanical performance. The gyroid infill pattern emerged as the optimal choice across all three materials for applications requiring high stiffness and strength[21,22], which demonstrate gyroid's superior mechanical resistance compared to conventional patterns.

For Flexibility, the triangle infill pattern consistently provided superior flexibility across PLA and ASA materials, achieving strain values of 3.36% (PLA) and 5.09% (ASA) at maximum load. This aligns with recent literature indicating triangular patterns' exceptional energy absorption and impact resistance capabilities.[23]

While honeycomb patterns showed high theoretical strength potential, our analysis revealed significant manufacturing consistency challenges, 3D-printed plastics sheet with honeycomb has strength less than gyroid pattern. This contrasts with some literature suggesting honeycomb as the universally strongest pattern[24], highlighting the importance of material-specific optimization and quality control in FDM processing.

On material-specific performance characteristics, PLA demonstrated exceptional stiffness characteristics, achieving the highest Young's modulus values (up to 461 MPa with gyroid infill pattern), consistent with literature reporting PLA's superior mechanical properties in controlled environments. However, our findings confirm PLA's brittleness limitations, with relatively low strain values and reduced performance under environmental stress conditions, supporting recent weathering studies.[25,26]

PETG emerged as the most balanced performer, achieving the highest ultimate tensile strength (10.88 MPa with gyroid) while maintaining excellent flexibility (7.36% strain with honeycomb).[25] The material's consistent performance across all infill patterns supports its growing adoption for functional applications requiring both strength and durability.[27]

ASA showed competitive mechanical properties with exceptional environmental resistance capabilities, achieving high stiffness (415 MPa) comparable to PLA while maintaining better flexibility characteristics. However, manufacturing inconsistencies, particularly with complex infill patterns, highlight processing optimization needs.[27]

Nonetheless, various factors can affect the analysis of the mechanical properties of plastic plates: infill density, extrusion temperature, infill orientation, layer thickness, and even filament color. Consequently, it is essential for the printed mask to undergo mechanical testing with standard material dimension size to ensure a more accurate assessment of its properties and to determine the actual displacement.[28]

The plastic plates produced by the 3D printer demonstrate consistent quality, as analyzed through the ImageJ program, with an average standard deviation between 5.93 and 7.63. The thermoplastic mask illustrates a value of 7.4, suggesting that the 3D printer has the capability of consistently producing each piece.

The analysis of sample density through Hounsfield unit (HU) revealed that the ASA featuring a gyroid structure has a HU of -620.21, which is nearest to the Uni-frame® and Type-S™ Thermoplastic mask. The HU values for all 3D-printed samples range from -774.88 to -620.21, indicating a very low radiological density—comparable to lung density (-950 to -550)[29], which is advantageous for radiotherapy applications as it minimizes beam attenuation.[1,30] The filling pattern may influence the size of the air gap and the value of HU. The dimensions of the air gap and the HU value were significant factors in getting equivalent tissue.

The radiological properties of 3D-printed materials, particularly their Hounsfield unit (HU) values, have been the subject of extensive research by numerous research groups. These studies have revealed substantial variations that are dependent on the material composition, infill density, and printing parameters. Across various research groups, PLA material exhibits the most extensively investigated HU characteristics. Fundamental baseline values were reported in the previous study, which demonstrated a strong infill density dependence with HU values ranging from -904 at 10% infill to +61 at 100% infill. At the 60% infill density that is frequently employed, PLA demonstrated HU values of -530±25, which places it within the lung-to-soft tissue equivalent range.[30] In the fabrication of the anthropomorphic phantom or thermoplastic mask of radiation therapy, the infill density of 45-55% is equivalent to the HU of lung tissue when

constructed from other materials, such as PETG, ASA, and ABS.[31,32] This is in accordance with this research that was printed with a 40% infill density, and the Hounsfield units (HU) of the three varieties of plastic are near the density of lung tissue.

The triangle structure in PLA allows the highest radiation transmission at 99.91%, displaying a minimal dose difference of 0.09%. By contrast, the PLA gyroid structure attains a 99.88% transmission rate with a 0.12% dose difference. The honeycomb-gyroid structure made from PETG demonstrates the lowest radiation transmission at 99.30%, showing a dose difference of 0.70%. The results show that all three types of plastic plates radiation similarly to the thermoplastic mask samples, which have transmission rates of 99.56% and 99.64%, with dose differences of 0.44% and 0.36%. The increased thickness in plate manufacture may affect radiation attenuation.

Limitations

The strength and flexibility values tested through Autodesk Fusion 360 in simulation mode using a compressive force of 10 Newtons (equivalent to 1.02 kilograms) are only predicted values from the program. In simulation mode does not fully support the analysis of the strength and flexibility of internal structures like gyroid and honeycomb-gyroid as comprehensively as other internal structures. Therefore, the strength and flexibility values for gyroid and honeycomb-gyroid internal structures may deviate from their actual values.

The mechanical testing conducted through simulation in Autodesk Fusion 360 provides valuable preliminary information. The test plastics utilized were flat sheets, inadequate to resemble the curvature required for actual anatomical patient application. Additional investigations will be conducted using 3D models to align with realistic patients.

The mechanical testing conducted by UTM has been limited by factors such as the dimensions of plastic sheets, the size of specimens for cross-sectional area assessments, the restriction to a single environmental testing setting, and the limited number of samples for certain material-pattern combinations.

Recommendations

Strength and flexibility tests of the plastic sheets should be conducted using a compressive testing machine alongside Autodesk Fusion 360 to increase the accuracy of the strength and flexibility values obtained.

Perform experiments focused on the production of larger components and their fabrication using 3D patient models, which are obtained from techniques like computed tomography (CT) or magnetic resonance imaging (MRI), combined with sophisticated computer-aided design (CAD), 3D scanning, and virtual planning software.

Subsequent research should address these shortcomings by implementing expanded testing matrices and consistent specimen preparation techniques for tensile testing and comparative investigation in different material filament brands.

Conclusion

This pilot study examines Fused Deposition Modeling (FDM) technology in 3D printing using the four different infill patterns of Polylactic Acid (PLA), Polyethylene Terephthalate Glycol-Modified (PETG), and Acrylonitrile Styrene Acrylate (ASA) for the fabrication of radiation immobilization devices. Based on comprehensive property testing, all three types of plastic have proven to be appropriate alternative materials for the integration of an immobilization mask. As a result, PLA plates displayed advantageous results in terms of radiation transmission, while PETG emerged as the most balanced performer in terms of mechanical properties, showing high strength and flexibility. ASA, a polymer containing a benzene ring, demonstrated that it is more resistant to radiation. Additionally, the Hounsfield unit (HU) value that is related to each soft tissue could be varied according to the characteristics and % infill pattern density.

Funding

The authors gratefully acknowledge the financial support provided by the Faculty of Allied Health Sciences Research Fund, Contract No. AHSNS1/2565

Conflict of interest

The authors declare that they have no conflicts of interest.

CRedit authorship contribution statement

Kochakorn Phantawong: conceptualization, methodology, validation, formal analysis, data curation, writing: original draft, review and editing, project administration and funding acquisition; **Nongnapat Singthuan:** data collection, analysis, interpretation of results and writing: original draft; **Kanjana Pinichsai:** data collection, analysis, interpretation of results and writing: original draft; **Somchat Taertulakarn:** supervision and funding acquisition

Acknowledgements

The authors would like to thank the staff of the Radiation Oncology Center at Thammasat University hospital for their assistance and use of their facilities during this research.

References

[1] Duarte J, Loja MAR, Portal R, Vieira L. 3D printing of abdominal immobilization masks for therapeutics: dosimetric, mechanical and financial analysis. Bioengineering (Basel). 2022; 9(2): 55. doi: 10.3390/bioengineering9020055.

[2] Mattke M, Rath D, Häfner M, Unterhinninghofen R, Sterzing F, Debus J, et al. Individual 3D-printed fixation masks for radiotherapy: first clinical experiences. Int J Comput Assist Radiol Surg. 2021; 16(6): 1043-9. doi: 10.1007/s11548-021-02393-2.

[3] Sa'ude N, Ab Latib NAN. A study on the mechanical properties of PLA, ABS and PETG filament printed by various type of infill design using 3D printing

machine. RPMME. 2023; 4(2): 162-7. [cited 2025 May 11]. Available from: https://publisher.uthm.edu.my/periodicals/index.php/rpmme/article/view/11492.

[4] Asfia A, Novak JI, Mohammed MI, Rolfe B, Kron T. A review of 3D printed patient specific immobilisation devices in radiotherapy. Phys Imaging Radiat Oncol. 2020; 13: 30-5. doi: 10.1016/j.phro.2020.03.003.

[5] Farah S, Anderson DG, Langer R. Physical and mechanical properties of PLA, and their functions in widespread applications—A comprehensive review. Adv Drug Deliv Rev. 2016; 107: 367-92. doi: 10.1016/j.addr.2016.06.012.

[6] Van der Walt M, Crabtree T, Albantow C. PLA as a suitable 3D printing thermoplastic for use in external beam radiotherapy. Australas Phys Eng Sci Med. 2019; 42: 1165-76. doi: 10.1007/s13246-019-00818-6.

[7] Ramírez-Revilla S, Camacho-Valencia D, Gonzales-Condori EG, Márquez G. Evaluation and comparison of the degradability and compressive and tensile properties of 3D printing polymeric materials: PLA, PETG, PC, and ASA. MRS Commun. 2023; 13(1): 55-62. doi: 10.1557/s43579-022-00311-4.

[8] Robar JL, Kammerzell B, Hulick K, Kaiser P, Young C, Verzwyvelt V, et al. Novel multi jet fusion 3D-printed patient immobilization for radiation therapy. J Appl Clin Med Phys. 2022; 23(11): e13773. doi: 10.1002/acm2.13773.

[9] Novak JI, Liu MZ-E, Loy J. Designing thin 2.5 D parts optimized for fused deposition modeling. Additive Manufacturing Technologies From an Optimization Perspective. Hershey, PA: IGI Global; 2019: pp 134-64. doi: 10.4018/978-1-5225-9167-2.

[10] Kim T-H, Cho M-S, Shin D-S, Shin DH, Kim S. Development of a real-time thermoplastic mask compression force monitoring system using capacitive force sensor. Front Robot AI. 2022; 9: 778594. doi: 10.3389/frobt.2022.778594.

[11] Demir E, Duygun İK, Bedeloğlu A. The mechanical properties of 3D-printed polylactic acid/polyethylene terephthalate glycol multi-material structures manufactured by material extrusion. 3D Print Addit Manuf. 2024; 11(1): 197-206. doi: 10.1089/3dp.2021.0321.

[12] Zhang Y, Zhang X, Cao Y, Feng J, Yang W. Acrylonitrile-styrene-acrylate particles with different microstructure for improving the toughness of poly (styrene-co-acrylonitrile) resin. Adv Polym Technol. 2021; 2021(1): 3004824. doi: 10.1155/2021/3004824.

[13] Moghaddam SHZ. Applications of Three-Dimensional Printing Technology in Radiotherapy: Advances in 3D Printing. London: IntechOpen; 2023. doi: 10.5772/intechopen.109398.

[14] Alssabbagh M, Abdulmanap M, Zainon R. Evaluation of 3D printing materials for fabrication of a novel multi-functional 3D thyroid phantom for medical dosimetry and image quality. Radiat Phys Chem. 2017; 135: 106-12. doi: 10.1016/j.radphyschem.2017.02.009.

[15] Zhao Y, Moran K, Yewondwossen M, Allan J, Clarke S, Rajaraman M, et al. Clinical applications of 3-dimensional printing in radiation therapy. Med Dosim. 2017; 42(2): 150-5. doi: 10.1016/j.meddos.2017.03.001.

[16] Moldovan A, Cuc S, Prodan D, Rusu M, Popa D, Taut AC, et al. Development and characterization of polylactic acid (PLA)-based nanocomposites used for food packaging. Polymers (Basel). 2023; 15(13): 2855. doi: 10.3390/polym15132855.

[17] Talataisong W, Ismaeel R, Marques TH, Abokhamis Mousavi S, Beresna M, Gouveia M, et al. Mid-IR Hollow-core microstructured fiber drawn from a 3D printed PETG preform. Sci Rep. 2018; 8(1): 8113. doi: 10.1038/s41598-018-26561-8.

[18] Haefner M, Giesel F, Mattke M. 3D-Printed masks as a new approach for immobilization in radiotherapy—A study of positioning accuracy. Oncotarget. 2018; 9(5): 6490-8. doi: 10.18632/oncotarget.24032.

[19] Aboelella MG, Ebeid SJ, Sayed MM. Layer combination of similar infill patterns on the tensile and compression behavior of 3D printed PLA. Sci Rep. 2025; 15(1): 11759. doi: 10.1038/s41598-025-94446-8.

[20] Dawood LL, AlAmeen ES. Influence of infill patterns and densities on the fatigue performance and fracture behavior of 3D-printed carbon fiber-reinforced PLA composites. AIMS Mater Sci. 2024; 11(5): 833–57. doi: 10.3934/matersci.2024041.

[21] Pandžić A, Hodžić D, Milovanović A, editors. Effect of infill type and density on tensile properties of PLA material for FDM process. Proceedings of the 30th DAAAM international symposium; 2019: 0545-0554. doi: 10.2507/30th.daaam.proceedings.074.

[22] Pandzic A, Hodzic D, editors. Mechanical properties comparison of PLA, tough PLA and PC 3D printed materials with infill structure–Influence of infill pattern on tensile mechanical properties. IOP Conf Ser Mater Sci Eng. 2021; 1208 (1): 012019. doi: 10.1088/1757-899X/1208/1/012019.

[23] Amador AMG, Avendano RAV, González AQ, Fernández LP. Mechanical characterization and testing of multi-polymer combinations in 3D printing. Heliyon. 2025; 11(3): e42420. doi: 10.1016/j.heliyon.2025.e42420.

[24] Eryildiz M. The effects of infill patterns on the mechanical properties of 3D printed PLA parts fabricated by FDM. Ukr. J Mech Eng Mater Sci. 2021; 7(1-2): 1-8. doi: 10.23939/ujmems2021.01-02.001.

[25] Sedlak J, Joska Z, Jansky J, Zouhar J, Kolomy S, Slany M, et al. Analysis of the mechanical properties of 3D-printed plastic samples subjected to selected degradation effects. Materials (Basel). 2023; 16(8): 3268. doi: 10.3390/ma16083268.

[26] Cressall S, Phillips CO, Al-Shatty W, Deganello D. The effect of high-intensity gamma radiation on PETG and ASA polymer-based fused deposition modelled 3D printed parts. J Mater Sci. 2024; 59(4): 1768-82. doi: 10.1007/s10853-023-09309-2.

[27] Iacob DV, Zisopol DG, Minescu M. Technical-economical study on the optimization of FDM

parameters for the manufacture of PETG and ASA parts. Polymers (Basel). 2024; 16(16): 2260. doi: 10.3390/polym16162260.

[28] Asfia A, Deepak B, Novak JI, Rolfe B, Kron T. Multi-jet fusion for additive manufacturing of radiotherapy immobilization devices: Effects of color, thickness, and orientation on surface dose and tensile strength. J Appl Clin Med Phys. 2022; 23(4): e13548. doi: 10.1002/acm2.13548.

[29] Kalender WA. Computed tomography: fundamentals, system technology, image quality, applications: 3rd Edition. Germany: John Wiley & Sons; 2011.

[30] Park S-Y, Choi N, Choi BG, Lee DM, Jang NY. Radiological characteristics of materials used in 3-dimensional printing with various infill densities. Prog Med Phys. 2019; 30(4): 155-9. doi: 10.14316/pmp.2019.30.4.155.

[31] Hariyanto AP, Christianti KH, Rubiyanto A, Nasori N, Haekal M, Endarko E. The effect of pattern and infill percentage in 3D Printer for phantom radiation applications. JID. 2022; 23(2): 87-92. doi: 10.19184/jid.v23i2.27256.

[32] Bustillo JPO, Mata JL, Posadas JRD, Inocencio ET, Rosenfeld AB, Lerch ML. Characterization and evaluation methods of fused deposition modeling and stereolithography additive manufacturing for clinical linear accelerator photon and electron radiotherapy applications. Phys Med. 2025; 130: 104904. doi: 10.1016/j.ejmp.2025.104904.

Journal of Associated Medical Sciences 2025; 58(3): 376-383

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 home-based isometric handgrip training on vascular function in middle-aged women with elevated pulse pressure

Worrawut Usupharach[1,2], Hataichanok Boonpim[3], Sawitri Wanpen[2], Raoyrin Chanavirut[2], Ponlapat Yonglitthipagon[2], Saowanee Nakmareong[2*]

[1]MSc student in Physical Therapy program, School of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen Province, Thailand.
[2]School of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen Province, Thailand.
[3]Department of Physical Therapy, Faculty of Allied Health Sciences, Nakhonratchasima College, Nakhonratchasima Province, Thailand.

ARTICLE INFO

Article history:
Received 18 June 2025
Accepted as revised 26 August 2025
Available online 31 August 2025

Keywords:
Arterial stiffness, blood pressure, isometric handgrip exercise, resistance exercise.

ABSTRACT

Background: Wide pulse pressure (PP) is a significant predictor of cardiovascular events. Isometric handgrip exercise (IHG) is a form of exercise used to manage blood pressure (BP). However, research on its effects in individuals with above-normal pulse pressure remains limited.

Objectives: This study aimed to examine the impact of a home-IHG program on arterial stiffness and blood pressure in middle-aged individuals with prehypertension or stage 1 hypertension and elevated PP.

Materials and methods: Twenty-eight participants were randomly assigned to either the IHG group (IHG, N=14) or the control group (CON, N=14). The IHG group performed IHG at 30% of their maximum voluntary contraction, using a modified mercury sphygmomanometer as the exercise device, 3 times/week, for 8 weeks. Brachial-ankle pulse wave velocity (baPWV) and BP were assessed at baseline and post-intervention.

Results: The results demonstrated an improvement in vascular function following IHG implementation, indicated by a significant decrease in baPWV in the IHG group compared with the control group (IHG: 1240±121 cm/s vs CON: 1365±159 cm/s; $p<0.05$). IHG also significantly reduced systolic blood pressure (SBP) compared with the control group (IHG: 119.86±5.26 mmHg vs CON: 131.21±3.64 mmHg; $p<0.05$). Moreover, changes in baPWV were positively correlated with changes in SBP ($r=0.65$, $p<0.001$) and PP ($r=0.52$, $p<0.01$).

Conclusion: An 8-week home-based IHG program using a mercury sphygmomanometer for training reduced BP and arterial stiffness in middle-aged women with elevated PP.

Introduction

Elevated BP, even at a marginal level, is a significant clinical concern and increased cardiovascular morbidity and mortality.[1] Therefore, appropriate BP control is crucial. Age-related changes in BP are characterized by an increase in SBP and a stable or mildly decreasing diastolic pressure (DBP), resulting in higher PP. Notably, elevated PP is related to vascular elasticity and serves as an index of arterial stiffness.[2] The European Society of Hypertension also acknowledges that a widened PP is a distinct risk factor, independent of elevated SBP.

Guidelines from health organizations recommend at least 150 minutes/week of moderate-intensity aerobic exercise as a first-line non-pharmacological strategy for the prevention, management, and lowering of hypertension (HTN).[3,4] A recent systematic review and meta-analysis

∗ **Corresponding contributor.**
Author's Address: *School of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen Province, Thailand.*
E-mail address: saowna@kku.ac.th
doi: 10.12982/JAMS.2025.109
E-ISSN: 2539-6056

demonstrated that aerobic training reduced SBP and DBP by approximately 4 and 3 mmHg, respectively.[5] Despite the established benefits of regular exercise, in Thailand, 43% of adults do not meet recommended exercise levels.[6] Lack of time is frequently cited as a barrier to engaging in aerobic exercise,[7] highlighting the need for time-efficient alternative approaches to BP management.

IHG has gained attention for its simplicity, short exercise duration, and efficacy in reducing BP.[8-10] Previous studies have demonstrated the anti-hypertensive effects of IHG interventions, with durations ranging from 3 to 12 weeks across diverse populations, with average BP reductions comparable to those achieved with standard exercise recommendations.[8,10] Nevertheless, only a limited number of trials have explored the effect of IHG on vascular adaptation, yielding inconsistent results.[11-13] Prior IHG implementation typically required specialized equipment, potentially impeding exercise accessibility. Hence, to overcome this limitation, this research employed a modified mercury sphygmomanometer as an accessible tool for grip strength exercises.[14-16]

Therefore, the aims of this study were to investigate the effect of an 8-week IHG intervention on vascular function BP in middle-aged individuals with pre-HTN and stage 1 HTN who exhibited elevated PP. We hypothesized that IHG would lead to improved vascular function, with a concomitant decrease in BP.

Materials and methods

Participants

Women aged 40-59 years with elevated BP or stage 1 HTN (resting SBP ≥120-139 mmHg and/or DBP ≥80-89 mmHg) were recruited from Samliam sub-district Health Promoting Hospital and Nong Kung sub-district Health Promoting Hospital, Khon Kaen province, Thailand. The inclusion criteria were: 1) PP of 50-79 mmHg, 2) body mass index (BMI) of 18.5-24.9 kg/m², and 3) sedentary behavior, defined as a low level of physical activity assessed using the Thai version of the short-form International Physical Activity Questionnaire (IPAQ-SF). Exclusion criteria included: 1) a diagnosis of heart disease, diabetes mellitus, kidney disease, pulmonary disease, or neurological conditions;,2) use of regular medications, supplements, or hormone replacement therapy within the 12 months prior to study enrollment, 3) upper extremity fracture within the past 6 months, 4) upper extremity deformity that could impair IHG, 5) moderate to severe pain, 6) current regular smoking or alcohol consumption, and 7) participation in any other exercise programs within 12 months prior to trial enrollment.

The sample size was calculated based on a previous study[17] using the formula for the comparison of two independent means. A sample size of 14 participants per group was required to achieve a statistical power of 0.80 with an alpha (α) level of 0.05. Consequently, a total of 28 participants were randomly assigned to the study groups.

Modified mercury sphygmomanometer

This study employed a modified mercury sphygmomanometer (SPIRIT, CK-101, Germany) for both measurement and training device (Figure 1). Previous research has demonstrated a correlation between grip strength values obtained using mercury sphygmomanometers and standard dynamometers in various populations. This evidence supports the use of modified sphygmomanometers as a reliable and alternative method for assessing handgrip strength.[14-16] As previously described, the sphygmomanometer was adapted using the bag method, which involved folding the inflatable bladder cuff into a cylindrical form with an approximate circumference of 7 inches and encasing it in a soft, non-slip, non-stretch fabric.[15,16] The modified mercury sphygmomanometer was calibrated and validation prior to data collection to ensure accuracy. The calibration results confirmed the validity for research application, with high measurement reliability at 95% confidence, complying with international standards (EA-4/02 and GUM).

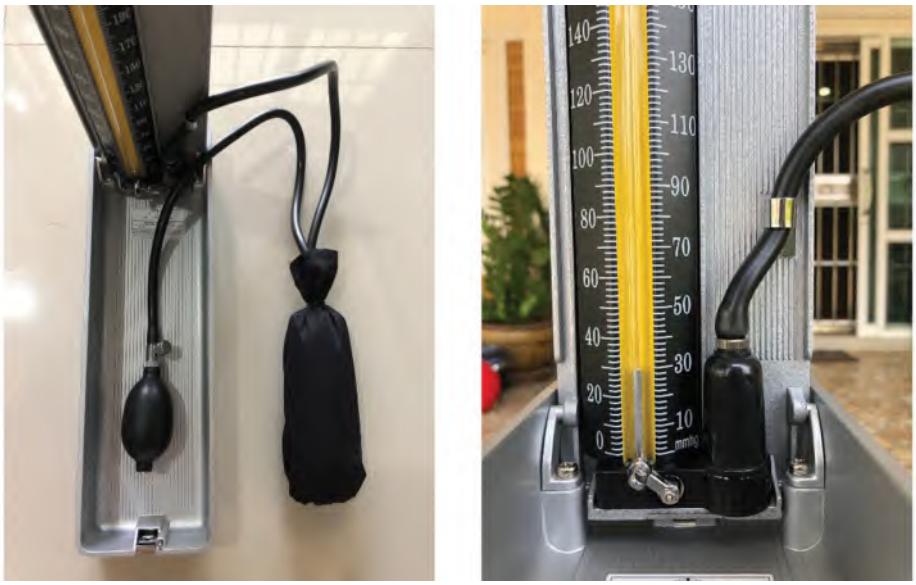


Figure 1. *Modified mercury sphygmomanometer.*

Procedure

The maximal voluntary contraction (MVC) of the dominant hand was assessed using a modified mercury sphygmomanometer, consistent with the recommendations of the American Society of Hand Therapists.[18] All measurements were performed by a trained physiotherapist. Following a familiarization session, the MVC test was conducted with participants in a seated position, maintaining the shoulder in a neutral position and elbow flexion at 90°. Each participant performed two trials with a minimum rest period of 1 minute between attempts. The higher value of the two trials, recorded in mmHg, was used as the participant's MVC.

Isometric handgrip training protocol

In the initial 2 weeks, participants performed exercises under the closed supervision of a physical therapist and received an instructional exercise booklet for continued practice. Furthermore, the participants were informed of the precautions and the indications for stopping exercise. For the subsequent 6 weeks, participants independently conducted home-based training. The IHG program, performed 3 times/week for 8 weeks, consisted of four sets of two-minute contractions using a modified mercury sphygmomanometer with the dominant hand, maintaining a target pressure at 30% of the participant's MVC, with one-minute rest periods between each set.[8] The MVC was reassessed every two weeks to adjust the exercise load. Throughout the study, participants were asked to maintain their routine activities and record each home exercise session in an exercise diary. Adherence to the home-based exercise program was monitored through weekly phone calls and researcher home visits every two weeks. All of participants in the IHG group completed all 24 scheduled exercise sessions.

Control group

Participants in the control group maintained their usual daily activities during the study period and received no additional interventions.

Measurements

Before testing, subjects were instructed to avoid strenuous exercise for 24 hours and to abstain from smoking, alcohol, and caffeine for 4 hours. A minimum 10-minute rest period was required prior to data collection. All assessments were conducted in a quiet, temperature-controlled environment. Outcome measurements were collected at baseline and after the 8-week program. Post-intervention assessments were conducted 48 hours after the final exercise session to minimize the influence of immediate exercise responses. All data were collected by the same investigator who blinded to group allocation. After a 10-minute rest, brachial SBP and DBP were measured using a calibrated, automated device (Masimo, WelchAllyn Vonnex Spot Monitor, USA) following a standardized protocol.[19] The mean of two readings was recorded for analysis, PP was then calculated from SBP and DBP.

Vascular function

baPWV was determined using a data acquisition system (Arterial Compliance Monitor, Bart's and The London School of Medicine and Dentistry, UK) with Doppler probes (Dopplex MDII, Huntleigh Healthcare, Cardiff, UK). Two consecutive waveform tracings were obtained for each participant, and the mean value was used for subsequent analysis. The arterial path length was estimated based on participant height.[20]

Statistical analysis

All data are expressed as mean±SD. Statistical analyses were performed using Statistical software (SPSS version 28.0.1.0 (142)). The Shapiro-Wilk test was used to assess the normality of data distribution. For within-group comparisons, paired samples T-tests were applied to evaluate differences between baseline and post-intervention values. Between-group differences in outcome variables were analyzed using independent samples t-tests. Pearson's correlation coefficient was calculated to assess the relationship between changes in vascular function and BP parameters. A <0.05 was considered statistically significant for all analyses.

Results

Baseline characteristics

Twenty-eight middle-aged women (age range: 43-57 years, mean age: 50.57±3.90 years) participated in this study. At baseline, there were no significant differences in demographic characteristics between the IHG and control groups (Table 1).

Table 1. Participant's characteristics.

Variables	Control group (N=14)	IHG group (N=14)
Age (year)	49.23±4.14	51.24±3.68
Weight (kg)	56.77±5.27	56.55±6.67
Height (cm)	161.14±4.53	158.50±6.37
BMI (kg/m^2)	21.84±1.50	22.45±1.52

Note: *IHG group: isometric handgrip exercise group, kg: kilogram, cm: centimeter, BMI: body mass index, m: meter. Data are expressed as mean±SD.*

Blood pressure

The 8-week IHG program significantly reduced SBP by approximately 10 mmHg and PP by approximately 8 mmHg from baseline. (Table 2). Following the intervention, SBP and PP were significantly lower in the exercise group compared to the control group (p<0.05). However, DBP did not change significantly over the 8-week period.

Initial baPWV did not differ significantly between the IHG and control groups (p>0.05; Figure 2). After 8 weeks, baPWV in the IHG group decreased significantly by 117 cm/s from baseline values (p<0.05). Furthermore, arterial function significantly improved in the IHG group, as evidenced by a lower baPWV compared to the control group at the post-intervention assessment (p<0.05).

Table 2. Changes in blood pressure before and after 8-weeks of intervention.

Variables	Control group (N=14)		IHG group (N=14)	
	Baseline	8 weeks	Baseline	8 weeks
SBP (mmHg)	130.79±4.89	131.21±3.64	130.75±4.51	119.86±5.26*,#
DBP (mmHg)	76.43±3.13	77.50±5.50	77.57±3.37	75.21±4.90
PP (mmHg)	54.36±3.82	53.71±7.05	53.18±3.24	44.64±4.57*,#

Note: *IHG group: isometric handgrip exercise group, SBP: systolic blood pressure, mmHg: millimeter of mercury, DBP: diastolic blood pressure, PP: pulse pressure, *significant different from baseline (p<0.05), #significant different from control group (p<0.05). Data are express as mean±SD*

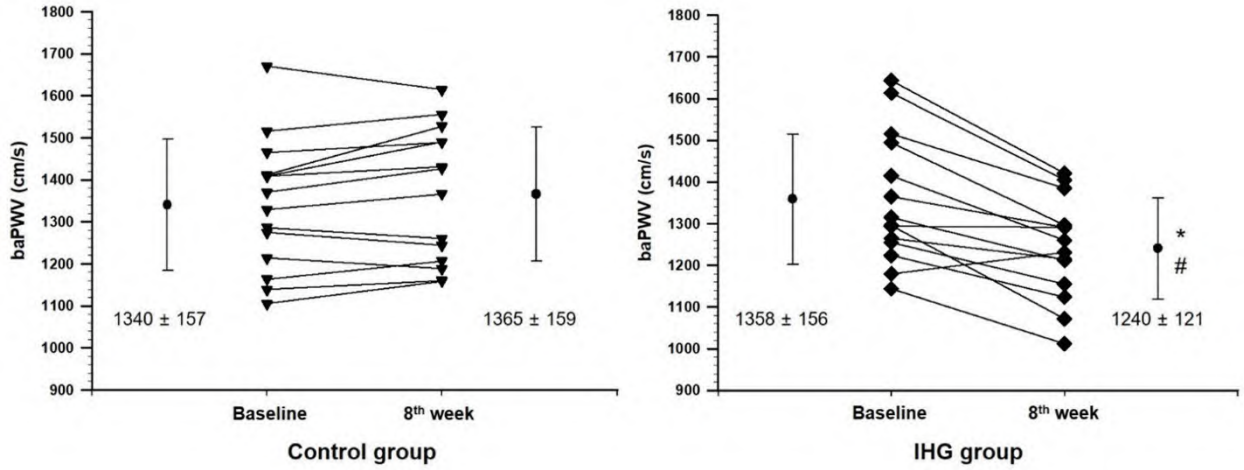


Figure 2. *Individual changes and mean baPWV before and after 8-weeks of intervention.*
*IHG group: isometric handgrip exercise group, baPWV: brachial ankle pulse wave velocity, cm: centimeter, s: second, * significant different from baseline (p<0.05), # significant different from control group (p<0.05).*

Pearson correlation analysis was conducted to examine the relationship between changes in arterial stiffness and BP parameters. There was a significant positive association between changes in baPWV and changes in SBP (r=0.65, p<0.001; Figure 3A). Similarly, alterations in baPWV were significantly correlated with changes in PP (r=0.52, p<0.01; Figure 3B).

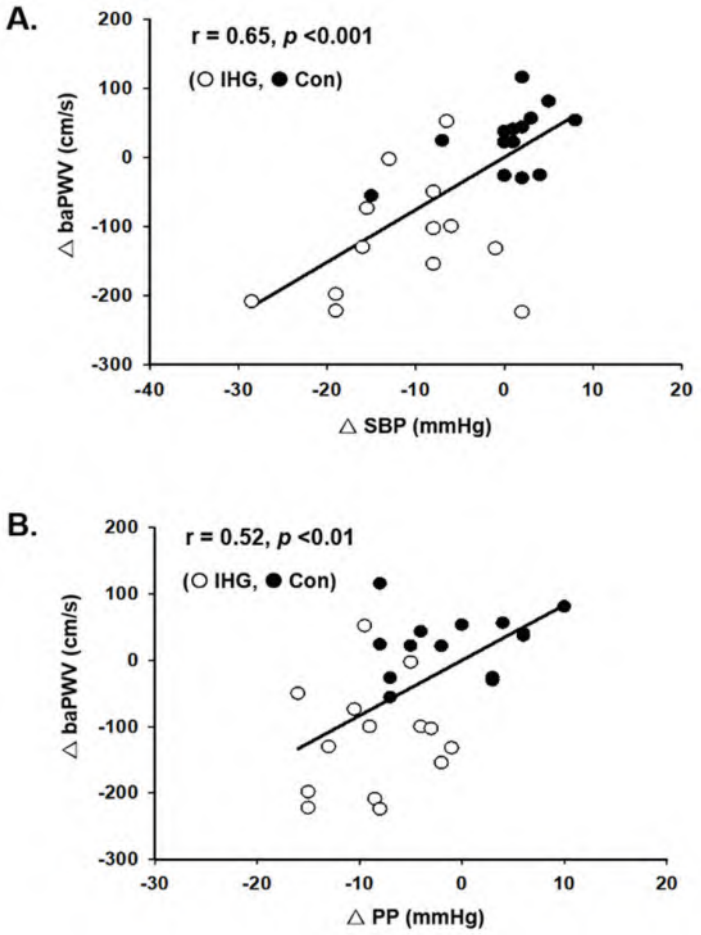


Figure 3. *Relationship between changes in brachial-ankle pulse wave velocity (baPWV) and changes in systolic blood pressure (A) and PP (B) all participants.*
IHG: isometric handgrip exercise group, Con: control group, baPWV: brachial ankle pulse wave velocity, cm: centimeter, s: second, SBP: systolic blood pressure, mmHg: millimeter of mercury, PP: pulse pressure.

Discussion

In this cohort of middle-aged women with elevated PP an 8-week home-based IHG program demonstrated beneficial effects on vascular function and reduced BP. Increased arterial stiffness is a recognized characteristic in HTN patients. In the present study, baPWV was used to assess arterial stiffness. A significant reduction in baPWV (approximately 100 cm/s) was observed following the implementation of IHG exercise. This magnitude of PWV reduction is considered clinically meaningful.[21] This finding aligns with a study by Patil *et al*.,[22] which demonstrated an alleviation of arterial stiffness (both carotid-femoral pulse wave velocity: cfPWV and baPWV) following a yoga intervention in elderly subjects with elevated PP. Similarly, previous research has shown a significant decrease in cfPWV after 12 weeks of IHG in elderly hypertensive patients. However, they reported no significant change in peripheral arterial stiffness, assessed via leg PWV.[23] IHG-induced vascular adaptations may involve several mechanisms, including enhanced local endothelium-dependent vasodilation, vascular remodeling, and/or modulation of autonomic vasomotor control.[8]

In the current study, a 8-week IHG training resulted in a clinically significant reduction in resting SBP by approximately 10 mmHg.[24] Following the intervention, 35% of participants in the handgrip group achieved optimal SBP levels (<120 mmHg). Additionally, 85% of participants achieved a PP less than 50 mmHg, a level associated with reduced cardiovascular risk.[25] Our findings align with a previous systematic review of low- to moderate-intensity IHG training programs which reported a mean reduction in SBP of 5 mmHg.[26] These findings support the efficacy of IHG for lowering BP in pre-HTN and stage 1 HTN middle-aged adults with above-normal PP. Indeed, the magnitude of BP reduction following IHG exercise appears to be related to the participants' baseline clinical characteristics and the IHG exercise protocol. Millar et al.[27] reported a positive correlation between baseline SBP and the magnitude of SBP reduction after IHG training, indicating greater benefits in individuals with higher initial SBP levels. Regarding the

IHG prescription, prior meta-analyses suggest that IHG training durations exceeding 8 weeks may yield larger BP reductions.[28]

The present study found that handgrip training had no significant effect on DBP. Our findings are consistent with prior research and likely attributed to the baseline characteristics of our participants, the majority of whom presented with DBP values within the normal range prior to the intervention. Therefore, any additional decrease in DBP was limited by a physiological floor effect.[28,29] This lack of a significant DBP change aligns with findings from a systematic review and meta-analysis by Almeida et al., which also reported no significant diastolic reductions following 8-10 weeks of IHG training in individuals with normotensive baseline DBP. The researchers proposed that IHG training mainly influences SBP by decreasing vascular resistance, leading to a greater impact on systolic blood pressure.[30] Conversely, DBP is primarily determined by cardiac output, which appear to be less responsive to this form of exercise training.[31]

Elevated BP accelerates structural and functional vascular changes, leading to reducing arterial distensibility. Conversely, increased arterial stiffness elevates BP by enhancing vascular resistance, establishing a bidirectional relationship.[27,28] In the present study, a positive association between reductions in BP and PWV was observed, suggesting that BP lowering following IHG training was proportional to improvements in vascular function. SBP has been demonstrated to be associated with arterial stiffness. Supporting this relationship, a large cross-sectional study (N=12,517) by Tomiyama *et al*. revealed that systolic blood pressure was correlated with baPWV among both male (r=0.61, p<0.01) and female participants (r=0.75, p<0.01).[34] This association has also been observed in intervention studies. Prior research in individuals with chronic obstructive pulmonary disease reported a significant positive association between changes in baPWV and changes in SBP (r=0.79, p=0.004) after a 4-week aerobic exercise program.[35] Similarly, Figueroa et.al., found correlations between decrease in baPWV and reductions in both radial systolic blood pressure (r=0.57, p<0.01) and aortic systolic blood pressure (r=0.64, p<0.01) following 6 weeks of watermelon supplementation in post-menopausal women.[36] These findings suggest that the intervention-induced reductions in systolic blood pressure may be accompanied by concurrent decreases in arterial stiffness.

The potential mechanisms underlying the BP-lowering effects of handgrip exercise might be involved an attenuation of vascular resistance by enhanced nitric oxide (NO)-induced vasodilation.[8] Although our study did not directly measure NO levels, the assessed flow-mediated dilation (FMD) as an indicator of endothelium-dependent, NO-mediated vasodilation in response to IHG. Consistent with this, Javidi et al. reported an enhancement in brachial artery %FMD in IHG training group, implying increased NO bioavailability after the IHG intervention.[37] Further supporting this mechanism, improvements in endothelial function after IHG have been demonstrated by a reduction in shear rate area under the curve after occlusion in hypertensive patient.[23]

Limitations

This study has several limitations. The absence of measurements for nitric oxide, oxidative stress, or inflammatory markers limits our ability to fully elucidate the mechanisms underlying the BP reductions observed following IHG. Further research is needed to explore the mechanistic pathways. Additionally, studies with longer durations and larger sample sizes are warranted to confirm the long-term benefits of IHG training and only specific for middle aged women.

Conclusion

Moderate-intensity IHG using a modified sphygmomanometer significantly reduced BP and arterial stiffness, potentially lowering cardiovascular risk. These findings suggest that IHG exercise may serve as a therapeutic adjunct for improving BP control and vascular function in middle-aged women with pre-HTN or stage 1 HTN and elevated PP.

Ethical approval

The research procedures were approved by the Khon Kaen University Ethics Committee for Human Research (HE661093), in accordance with the Declaration of Helsinki and the Good Clinical Practice Guidelines. Written informed consent was obtained from all participants before their enrollment.

Funding

This study was supported by research and graduate studies, Khon Kaen University, Thailand.

Conflict of interest

The authors have no conflicts of interest associated with the material presented in this paper.

CRedit authorship contribution statement

Worrawut Usupharach: data collection, formal analysis, data curation, initial drafting; **Hataichanok Boonpim:** formal analysis, data curation, initial drafting; **Sawitri Wanpen:** formal analysis, data curation; **Raoyrin Chanavirut:** conceptualized and designed, formal analysis; **Ponlapat Yonglitthipagon:** conceptualized and designed, formal analysis; **Saowanee Nakmareong:** conceptualized and designed, writing – review and editing, project administration and funding acquisition. All authors approved the final version of the manuscript.

Acknowledgements

The authors would like to sincerely thank all participants for their valuable contributions to this study and to Samliam and Nong Kung sub-district Health Promoting Hospital for the dedicated support and cooperation of their staff in facilitating participant recruitment.

References

[1] Lewington S, Clarke R, Qizilbash N, Peto R, Collins R, Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. Lancet Lond Engl. 2002; 360(9349): 1903-13. doi: 10.1016/s0140-6736(02)11911-8.

[2] Steppan J, Barodka V, Berkowitz DE, Nyhan D. Vascular stiffness and increased pulse pressure in the aging cardiovascular system. Cardiol Res Pract. 2011; 2011: 263585. doi: 10.4061/2011/263585.

[3] Riebe DK, Ehrman J, Liguori G, Magal M. ACSM's guidelines for exercise testing and prescription. Wolters Kluwer; 2022.

[4] Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2018; 71(19): e127-248. doi:10.1161/HYP.0000000000000066.

[5.] Edwards JJ, Deenmamode AHP, Griffiths M, Arnold O, Cooper NJ, Wiles JD, et al. Exercise training and resting blood pressure: a large-scale pairwise and network meta-analysis of randomised controlled trials. Br J Sports Med. 2023; 57(20): 1317-26. doi: 10.1136/bjsports-2022-106503.

[6] Widyastari DA, Saonuam P, Pongpradit K, Wongsingha N, Choolers P, Kesaro S, et al. Results from the Thailand 2022 report card on physical activity for children and youth. J Exerc Sci Fit. 2022; 20(4): 276-82. doi:10.1016/j.jesf.2022.06.002.

[7] Justine M, Azizan A, Hassan V, Salleh Z, Manaf H. Barriers to participation in physical activity and exercise among middle-aged and elderly individuals. Singapore Med J. 2013; 54(10): 581-6. doi: 10.11622/smedj.2013203.

[8] Edwards JJ, Coleman DA, Ritti-Dias RM, Farah BQ, Stensel DJ, Lucas SJE, et al. Isometric exercise training and arterial hypertension: An updated review. Sports Med Auckl NZ. 2024 May 19. doi: 10.1007/s40279-024-02036-x.

[9] Hansford HJ, Parmenter BJ, McLeod KA, Wewege MA, Smart NA, Schutte AE, et al. The effectiveness and safety of isometric resistance training for adults with high blood pressure: a systematic review and meta-analysis. Hypertens Res Off J Jpn Soc Hypertens. 2021; 44(11): 1373-84. doi: 10.1038/s41440-021-00720-3.

[10] Oliveira PC, Silva MR, Lehnen AM, Waclawovsky G. Isometric handgrip training, but not a single session, reduces blood pressure in individuals with hypertension: a systematic review and meta-analysis. J Hum Hypertens. 2023; 37(9): 844-53. doi: 10.1038/s41371-022-00778-7.

[11] Rodrigues SLC, Farah BQ, Silva G, Correia M, Pedrosa R, Vianna L, et al. Vascular effects of isometric handgrip training in hypertensives. Clin Exp Hypertens N Y N 1993. 2020; 42(1): 24-30. doi: 10.1080/10641963.2018.1557683.

[12] Edwards JJ, Jalaludeen N, Beqiri A, Wiles JD, Sharma R, O'Driscoll JM. The effect of isometric exercise training on arterial stiffness: A randomized crossover-controlled study. Physiol Rep. 2023; 11(10): e15690. doi: 10.14814/phy2.15690.

[13] Lopes S, Afreixo V, Teixeira M, Garcia C, Leitão C, Gouveia M, et al. Exercise training reduces arterial stiffness in adults with hypertension: a systematic review and meta-analysis. J Hypertens. 2021; 39(2): 214-22. doi: 10.1097/HJH.0000000000002619.

[14] Brito SAF de, Santana M de M, Benfica P do A, Aguiar LT, Gomes G de C, Faria CDC de M. The modified sphygmomanometer test for assessment of muscle strength of community-dwelling older adults in clinical practice: reliability and validity. Disabil Rehabil. 2022; 44(1): 131-8. doi: 10.1080/09638288.2020.1758804.

[15] Seephim B, Nualnetr N. Intra-rater reliability of grip strength measurement by using a manual sphygmomanometer in patients with rheumatoid arthritis. Phys Ther J. 2021; 43(2): 59-66.

[16] Souza LAC, Martins JC, Moura JB, Teixeira-Salmela LF, De Paula FVR, Faria CDCM. Assessment of muscular strength with the modified sphygmomanometer test: what is the best method and source of outcome values? Braz J Phys Ther. 2014; 18(2): 191-200. doi: 10.1590/S1413-35552012005000149.

[17] Punia S, Kulandaivelan S. Home-based isometric handgrip training on RBP in hypertensive adults—Partial preliminary findings from RCT. Physiother Res Int. 2020; 25(1): e1806. doi: 10.1002/pri.1806.

[18] Külkamp W, Ache-Dias J, Kons RL, Detanico D, Dal Pupo J. The ratio standard is not adequate for scaling handgrip strength in judo athletes and nonathletes. J Exerc Rehabil. 2020 Apr;16(2):175–82. doi:10.12965/jer.2040108.054.

[19] Unger T, Borghi C, Charchar F, Khan NA, Poulter NR, Prabhakaran D, et al. 2020 International Society of Hypertension Global Hypertension Practice Guidelines. Hypertens Dallas Tex 1979. 2020; 75(6): 1334-57. doi:10.1161/HYPERTENSIONAHA.120.15026.

[20] Munakata M. Brachial-Ankle Pulse Wave Velocity: Background, Method, and Clinical Evidence. Pulse Basel Switz. 2016; 3(3-4): 195-204. doi: 10.1159/000443740.

[21] Vlachopoulos C, Aznaouridis K, Terentes-Printzios D, Ioakeimidis N, Stefanadis C. Prediction of cardiovascular events and all-cause mortality with brachial-ankle elasticity index: a systematic review and meta-analysis. Hypertens Dallas Tex 1979. 2012; 60(2): 556-62. doi: 10.1161/HYPERTENSIONAHA.112.194779.

[22] Patil SG, Aithala MR, Das KK. Effect of yoga on arterial stiffness in elderly subjects with increased

pulse pressure: A randomized controlled study. Complement Ther Med. 2015; 23(4): 562-9. doi: 10.1016/j.ctim.2015.06.002.

[23] Rodrigues SLC, Farah BQ, Silva G, Correia M, Pedrosa R, Vianna L, et al. Vascular effects of isometric handgrip training in hypertensives. Clin Exp Hypertens N Y N 1993. 2020; 42(1): 24-30. doi: 10.1080/10641963.2018.1557683.

[24] Ettehad D, Emdin CA, Kiran A, Anderson SG, Callender T, Emberson J, et al. Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. The Lancet. 2016; 387(10022): 957-67. doi:10.1016/S0140-6736(15)01225-8.

[25] Asmar R, Vol S, Brisac AM, Tichet J, Topouchian J. Reference values for clinic pulse pressure in a nonselected population. Am J Hypertens. 2001; 14(5Pt1): 415-8. doi: 10.1016/s0895-7061(01)01284-5.

[26] Betancur AFL, Medrano IC. Is low-intensity isometric handgrip exercise an efficient alternative in lifestyle blood pressure management? A systematic review. Sports Health. 2020; 12(5): 470-7. doi: 10.1177/1941738120943882.

[27] Millar PJ, Bray SR, McGowan CL, MacDonald MJ, McCartney N. Effects of isometric handgrip training among people medicated for hypertension: a multilevel analysis. Blood Press Monit. 2007; 12(5): 307-14. doi:10.1097/MBP.0b013e3282cb05db.

[28] Inder JD, Carlson DJ, Dieberg G, McFarlane JR, Hess NC, Smart NA. Isometric exercise training for blood pressure management: a systematic review and meta-analysis to optimize benefit. Hypertens Res Off J Jpn Soc Hypertens. 2016; 39(2): 88-94. doi:10.1038/hr.2015.111.

[29] Cornelissen VA, Smart NA. Exercise training for blood pressure: a systematic review and meta-analysis. J Am Heart Assoc. 2013; 2(1): e004473. doi:10.1161/JAHA.112.004473.

[30] Almeida JPA de S, Bessa M, Lopes LTP, Gonçalves A, Roever L, Zanetti HR. Isometric handgrip exercise training reduces resting systolic blood pressure but does not interfere with diastolic blood pressure and heart rate variability in hypertensive subjects: a systematic review and meta-analysis of randomized clinical trials. Hypertens Res Off J Jpn Soc Hypertens. 2022; 45(5): 930-1. doi:10.1038/s41440-021-00681-7.

[31] Millar PJ, McGowan CL, Cornelissen VA, Araujo CG, Swaine IL. Evidence for the role of isometric exercise training in reducing blood pressure: potential mechanisms and future directions. Sports Med Auckl NZ. 2014; 44(3): 345-56. doi:10.1007/s40279-013-0118-x.

[32] Safar ME, Asmar R, Benetos A, Blacher J, Boutouyrie P, Lacolley P, et al. Interaction Between Hypertension and Arterial Stiffness. Hypertens Dallas Tex 1979. 2018; 72(4): 796-805. doi:10.1161/HYPERTENSIONAHA.118.11212.

[33] Wilson J, Webb AJS. Systolic blood pressure and longitudinal progression of arterial stiffness: A quantitative meta-analysis. J Am Heart Assoc. 2020; 9(17): e017804. doi:10.1161/JAHA.120.017804.

[34] Tomiyama H, Yamashina A, Arai T, Hirose K, Koji Y, Chikamori T, et al. Influences of age and gender on results of noninvasive brachial-ankle pulse wave velocity measurement--a survey of 12517 subjects. Atherosclerosis. 2003; 166(2): 303-9. doi:10.1016/s0021-9150(02)00332-5.

[35] Vivodtzev I, Minet C, Wuyam B, Borel JC, Vottero G, Monneret D, et al. Significant improvement in arterial stiffness after endurance training in patients with COPD. Chest. 2010; 137(3): 585-92. doi:10.1378/chest.09-1437.

[36] Figueroa A, Wong A, Hooshmand S, Sanchez-Gonzalez MA. Effects of watermelon supplementation on arterial stiffness and wave reflection amplitude in postmenopausal women. Menopause. 2013; 20(5): 573. doi:10.1097/GME.0b013e3182733794.

[37] Javidi M, Ahmadizad S, Argani H, Najafi A, Ebrahim K, Salehi N, et al. Effect of Lower- versus Higher-Intensity Isometric Handgrip Training in Adults with Hypertension: A Randomized Controlled Trial. J Cardiovasc Dev Dis. 2022; 9(9): 287. doi:10.3390/jcdd9090287.

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/bulletinAMS/

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 Sciences' copyright.

Generative AI and AI-Assisted Technologies for the Authors

JAMS's policy refers only to manuscript preparation, not to using AI and AI-assisted tools to analyze and draw insight from data as a part of the research.

It is recommended that the authors use AI and AI-assisted tools to improve English writing for the readability and language of the manuscript. The authors should carefully review and edit the result with human oversight and control since AI may generate authoritative-sounding output incorrectly or bias.

Before submission, The authors must ensure the work is original and does not infringe third-party rights.

Authorship implies responsibilities and tasks that can only be attributed to and performed by humans. By this, AI and AI-assisted technologies must not be listed as authors, co-authors, or cited AI.

The authors must be responsible and accountable for the content of the manuscript. Moreover, the authors should disclose in the manuscript the use of AI and AI-assisted technologies, and a statement must appear in the published article since it will support transparency and trust among the authors, readers, reviewers, editors, and contributors.

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 not permitted.

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 on to the Editor for review.

Examples of References list

Multiple Authors: List up to the first 6 authors/editors, and use "et al." for any additional authors.

Journal Articles (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 8]. Available from: http://www.who.int/mental_health/neurology/dementia/dementia_thematicbrief_epidemiology.pdf.

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