JOURNAL OF ASSOCIATED MEDICAL SCIENCES

JAMS

JOURNAL OF ASSOCIATED MEDICAL SCIENCES

THE OFFICIAL PEER-REVIEWED ONLINE JOURNAL

Volume 57 Number 3 September - December 2024 E-ISSN: 2539-6056

Journal of Associated Medical Sciences

Aims and scope

The Journal of Associated Medical Sciences belongs to Faculty of Associated Medical Sciences (AMS), Chiang Mai University, Thailand. The journal specifically aims to provide the platform for medical technologists, radiologic technologists, occupational therapists, physical therapists, speech-language pathologists and other related professionals to distribute, share, 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 scope of the journal will be processed strictly following the double-blinded peer review process of the journal. Therefore, the final decision can be completed in 1-3 months average, depending on the number of rounds of revision.

Objectives

The Journal of Associated Medical Sciences aims to publish integrating research papers in areas of Medical Technology, Physical Therapy, Occupational Therapy, Radiologic Technology, and related under peer-reviewed via double-blinded process by at least two internal and external reviewers.

Types of manuscript

Manuscripts may be submitted in the form of review articles, original articles, 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).
- **Original articles** must not exceed 15 journal pages (not more than 3,500 words), including 6 tables/figures, and 40 reference (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).

Peer review process

By submitting a manuscripts to 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 manuscripts is received, it is assigned by a specific Associate Editor. The Associate Editor prepares a list of expert reviewers, which may include some suggested by the Editor-in-Chief. Authors can indicate specific 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. Manuscripts 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 are appended.

Publication frequency

Journal of Associated Medical Sciences publishes 3 issues a year

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

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. 57 No. 3 : September - December 2024

Content

Content

Content

Journal of Associated Medical Sciences 2024; 57(3): 1-7

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Radiation dose and image quality optimization in lumbar spine digital radiography for overweight and obese patients: Phantom study

Ausanai Prapan[1] Kanthicha Chuprempri[1] Phattharawadee Fong-in[1] Panida Kheawtubtim[1] Natch Rattanarungruangchai[2] Thanyawee Pengpan[1*]

[1]Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.
[2]Department of Radiation Dose Measurement and Assessment, Nuclear Technology Service Center, Thailand Institute of Nuclear Technology (Public Organization), Nakornnayok Province, Thailand.

ARTICLE INFO

Article history:
Received 22 November 2023
Accepted as revised 12 May 2024
Available online 16 May 2024

Keywords:
Lumbar spine radiography, radiation dose, image quality, obesity, overweight

ABSTRACT

Background: Lumbar spine radiography plays an important role in routine use for clinical practice in overweight and obese patients. The radiographer is responsible for setting suitable exposure factors for the tradeoffs between radiation dose and image quality (IQ).

Objective: To investigate the effect of different kVp values combined with AEC used on radiation dose and IQ for routine lumbar spine radiography in overweight and obese patients.

Materials and methods: A 1.5 and 3 cm thickness of frozen pork lard was placed on a pelvis phantom to simulate an overweight and obese patient, respectively. The phantom was imaged at 10 kVp intervals in combination with AEC used. For IQ evaluation, CNR and SNR were calculated, and the observer study was determined using visual grading scores (VGS) with a 5-point Likert scale. The radiation dose was measured using a DAP meter, and then the figure of merit (FOM) was calculated.

Results: SNR and CNR for both AP and lateral projection showed a slightly decreasing trend when kVp increased in all groups. The DAP values decreased when the higher kVp was selected with AEC used in each group. The VGS by five radiographers showed good image quality in all groups, while the FOM at 100 and 109 kVp was the highest score for both AP and lateral projections.

Conclusion: The optimal kVp setting in this study ranged from 100 to 109 kVp in combination with AEC used, indicating minimal radiation dose, while maintaining diagnostic IQ.

Introduction

The number of overweight and obese people has increased dramatically in recent years. The World Health Organization (WHO) stated that 13% of the world's adult population was obese, with the global prevalence of obesity increasing 3-fold between 1975 and 2016.[1] Obesity is categorized as one of the most serious public health issues, impacting the increased risk of numerous diseases and reducing life expectancy.[2] Medical examinations and procedures, including radiography, play a significant role in the diagnosis, treatment, and care of obese patients.[3] Radiographic imaging procedures in overweight and obese patients present practical challenges because of the increased radiation dose required, as well as reduced image quality.[4] The attenuation of the X-ray beam, scatter radiation, and long exposure times result in motion artifacts

∗ **Corresponding contributor.**
Author's Address: *Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*
E-mail address: thanyaweep@nu.ac.th
doi: 10.12982/JAMS.2024.041
E-ISSN: 2539-6056

during radiographic examination.[5,6] More adipose tissue in overweight and obese patients leads to poor photon penetration, resulting in high quantum noise.[7] Therefore, it becomes critical to optimize exposure parameters within the framework to be as low as reasonably achievable (ALARA), while the compromise between radiation dose and image quality needs to remain consistent with the diagnostic purpose.

Lumbar spine radiography plays a vital role in routine clinical practice, and it can assist the doctor to assess damage to the lumbar bone. However, lumbar spine radiography is conducted at the highest collective dose, with higher radiation-induced cancer risk compared with other X-ray examinations. Using high-exposure parameters (i.e., tube voltage and tube current) delivers increased radiation doses in lumbar spine radiography, with many X-ray photons penetrating the human body.[8] Therefore, it is necessary to understand and optimize the exposure parameters for lumbar spine radiography by keeping the radiation dose as low as possible while not diminishing the image quality.

In this study, we aimed to optimize the exposure parameter of lumbar spine radiography in the DR imaging system for overweight and obese patients by increasing the X-ray tube voltage in combination with the use of automatic exposure control (AEC).

Materials and methods

Experimental Setup, Imaging Acquisition, and Phantom

A general radiographic unit (SIEMENS Multix TOP, München, Germany) and a gadolinium oxysulfide (GADOX) based DR detector (VIVIX-S, Vieworks Co., Ltd., Gyeonggi-do, Korea) were used for image processing. Quality assurance tests for the X-ray units and the detector were conducted routinely including the consistency of AEC.

In this experimental study, we set up the phantom into 3 groups: normal, overweight, and obese group. In the normal group, an anthropomorphic pelvis phantom (RS-113T, Radiology Support Devices, Inc., CA, USA) was used to image radiographs. The phantom was positioned in supine and lateral positions with 20x43 cm for the AP view and 25x43 cm for the lateral view of beam collimation sizes with 12:1 grid ratio. The central ray was centered at a point at the level of the 3rd lumbar vertebral body and the SID was set at 100 cm for all images. For the lateral projections, the phantom was turned to its left side until the midcoronal plane of the phantom was at 90 degrees to the image receptor. Experimental images were taken across a range of kVp settings in combination with the use of a central AEC chamber, starting with the 70 kVp and increasing to 109 kVp at approximately 10 kVp intervals.

Simulation of overweight and obese patients was achieved by placing 1.5 cm and 3 cm thick layers of frozen pork lard on the frontal and lateral aspects of the phantom, respectively (Figure 1). The CT scan (Philips Medical Systems, Cleveland, USA) was performed to measure CT density measurements for validation of the frozen pork lard which uses as stimulating abdominal fat. Our result data of CT density of frozen pork lard was -112 HU, indicating a similar tissue density of human fat.[9]

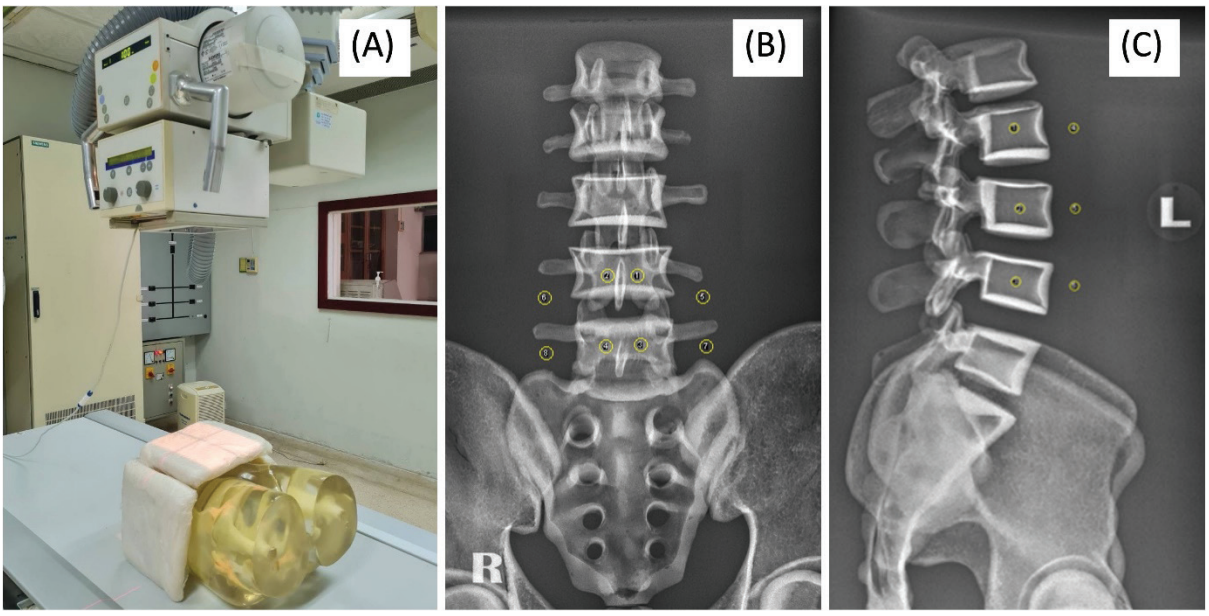


Figure 1. *The experimental set-up. A: with 3 cm slices of frozen pork lard placed on the pelvis phantom, B,C: The ROIs were drawn on the image of AP and Lateral lumbar spine, respectively*

Radiation Dose Measurement

The dose area product (DAP) was measured by attaching a DAP meter (VacuDAP™; VacuTec Meßtechnik GmbH, Dresden, Germany) to the center of the radiographic tube.

Image quality evaluation

The signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR)

The images were evaluated by calculating the SNR and CNR by using the ImageJ software.[10] All the ROIs must have the same size and be placed at the same location for all the images. An object and a background region of interest (ROI) were determined to obtain SNR and CNR. The object ROIs of the frontal view contained 4 individual areas, each of which covered the vertebral regions of L4 and L5 respectively. The background ROIs of the AP view contained 4 areas, covering soft tissue regions on either side (left and right) of the intervertebral joints L4/L5, and L5/S1 but excluded areas of the transverse processes (Figure 1B). For the lateral view, the object ROIs of the frontal view contained 3 individual areas, each of which covered the vertebral regions of L3, L4, and L5 and the background ROI contained 3 areas that covered soft tissue regions and anterior to the vertebral body (Figure 1C). The SNR and CNR were calculated using the following Equation 1 and 2.[11]

$$SNR = \frac{Average\ pixel\ value\ of\ object}{\sigma_{object}} \quad (1)$$

$$CNR = \frac{Pixel\ value\ object - Pixel\ value\ background}{\sigma} \quad (2)$$

furthermore, σ is calculated as $\sqrt{\frac{SD_1^2 + SD_2^2}{2}}$ where SD_1 is the standard deviation for the ROI_{Object} and SD_2 is the standard deviation of the $ROI_{Background.}$[11]

Observer study: Visual Grading Score (VGS)

The images were assessed by 5 radiographers, each which a range of 3-10 years of experience. Observers were blinded to the exposure parameter used for each image, which were displayed in randomized order. Thirty images of the lumbar spine in each group were created, including 15 radiographs of AP projection and 15 radiographs of lateral projection. The images were evaluated on the same diagnostic monitor using INFINITT software (distributed by THAI GL CO., LTD., Bangkok, Thailand). All radiographs were assessed on a five-point Likert scale according to the criteria listed in the European guidelines.[12] The image quality criteria and the rating on the scale were demonstrated in Table 1.

Table 1. Anatomical criteria of lumbar spine and scoring scale

Part and Projection	
AP (anteroposterior) Lumbar Spine	*Visually sharp reproduction of the:* • *Upper and lower-end plate surfaces.* • *Pedicles.* • *Cortex/trabecular patterns.* *Reproduction of the:* • *Intervertebral joints.* • *Spinous and transverse processes.* • *Sacro-iliac joints*
Lateral Lumbar Spine	*Visually sharp reproduction of the:* • *Upper and lower-end plate surfaces.* • *Cortex/trabecular patterns.* *Reproduction of the:* • *Pedicles and intervertebral foramina* • *Spinous processes* *Full superimposition of the posterior vertebral edges*
Visual grading scale	*1 = Not visible* *2 = Probably not visible* *3 = Acceptable visibility* *4 = Good visibility* *5 = Very good visibility*

Optimization: Figure of Merit (FOM)

In this study, the FOM was calculated to correlate our finding of DAP and CNR. With CNR and DAP, it can calculate the FOM, which is described as follows: [13]

$$FOM = \frac{CNR^2}{DAP} \quad (3)$$

Statistical Analysis

The results were compared using an ANOVA-like test by using the GraphPad Prism software (GraphPad, La Jolla, CA, USA) for statistical analyses. A *P* value less than 0.05 was considered to indicate a statistically significant difference. The degrees of the agreement were assessed using the intraclass correlation coefficient (ICC) and the results were considered significant at the 95% confidence level by using SPSS Software Version 17.00 (SPSS, Inc., Chicago, IL, USA). According to criteria to evaluate the ICC, the value of less than 0.4 represents poor interobserver agreement, 0.40-0.59 represents fair agreement, 0.6-0.74 represents good agreement and 0.75-1.00 indicates excellent agreement.[14]

Results

SNR and CNR

The resulting SNR and CNR for AP and lateral view of the lumbar spine are shown in Figure 2 with the corresponding values for the normal, overweight, and obese groups. The SNR and CNR, in both AP and lateral projection, were slightly decreased in all groups when higher kVp was applied. The SNR of the obese group was significantly lower when compared with the normal group *(P<0.001).* Furthermore, there were no significant differences in the CNR among the three groups at the same tube voltage in the AP projection. For comparisons of CNR between the obese and normal group in lateral projection, CNRs were significantly superior when increasing tube voltage in lateral projection (p<0.001).

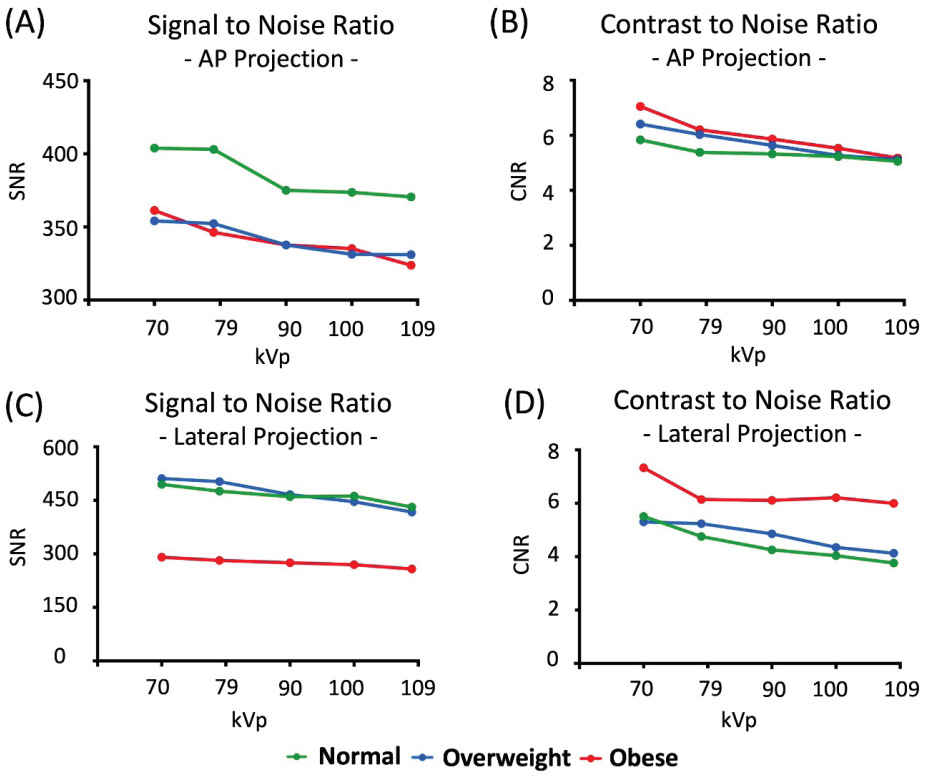


Figure 2*. SNR and CNR of lumbar spine radiography. A,B: AP projection, C,D: lateral projection.*

DAP and Image Quality Score

The results of DAP values are shown in Figure 3(A,C). For AP projection, the DAP ranged from 62 to 119 µGym^2 in the normal group, 75 to 146 µGym^2 in the overweight group, and 85 to 163 µGym^2 in the obese group. For lateral projection, the DAP ranged from 117 to 231 µGym^2 in the normal group, 144 to 279 µGym^2 in the overweight group, and 259 to 500 µGym^2 in the obese group. Overall, for both projections, the DAP values gradually decreased, indicating a decrease in radiation doses, when the higher kVp values were used.

The average of the VGS of the AP and lateral lumbar spine in various kVp used is shown in detail in Figure 3(B,D). The mean image score of the lumbar spine in each group and in various kVp ranged from 3.6 to 4.5 and no observer scored the images as less than 3 on the 5-point rating scale, indicating that all of them agreed to accept image visibility. Moreover, the result demonstrated there was no significant difference in an overall subjective image quality when using the higher kVp values (p>0.05).

Moreover, the ICC values were below 0.4 in both AP and lateral projections. It can be noted that there was no correlation in the results, indicating poor agreement.

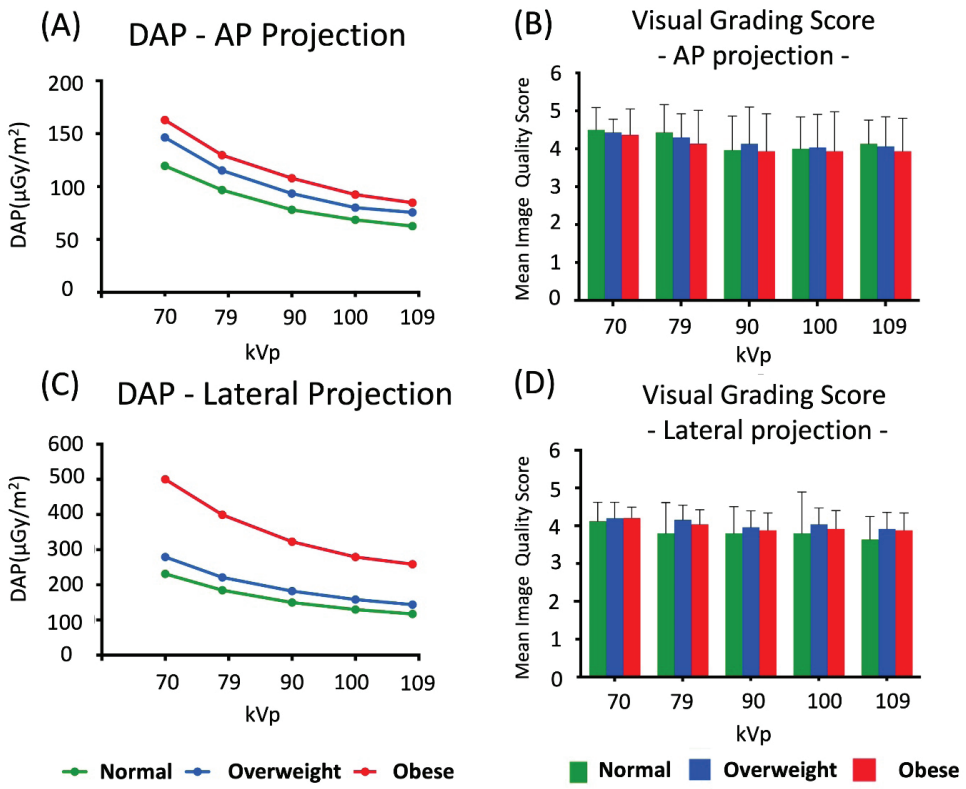


Figure 3. *DAP values and Image Quality Score of lumbar spine radiography. A,B: AP projection, C,D: lateral projection.*

The Figure of Merit (FOM)

For the AP projection, the highest FOM was at 109 kVp for the normal and overweight groups, and 100 kVp for the obese groups (Figure 4A). However, there was no statistical difference between 100 and 109 kVp for AP projection in all groups. For the lateral projection, 90 kVp provided the highest FOM in the overweight group (Figure 4B). Moreover, 100 and 109 kVp provided the highest FOM in the obese group, while 70 kVp provided the highest FOM for the normal group. In both projections of the lumbar spine, AP and lateral, FOM values increased as kVp increased, except in the lateral projection of the normal group. However, the FOM of the lateral projection in the normal group had no statistical difference in each kVp.

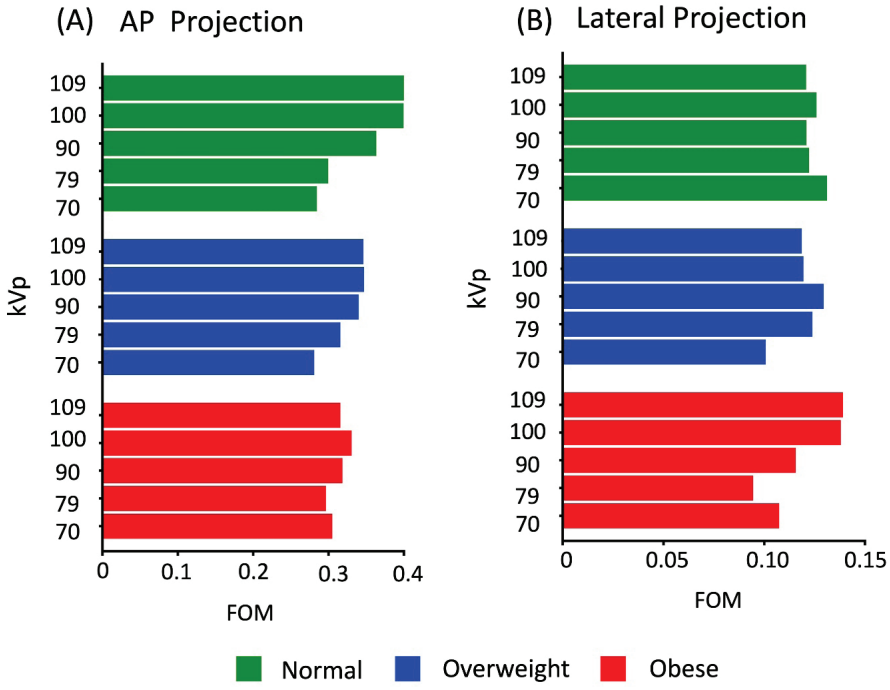


Figure 4. *The figure of merit (FOM)*

Discussion

This study has investigated the effect of increasing the kVp values in combination with AEC used for producing routine lumbar spine radiography for overweight and obese patients using a DR imaging system.

The overweight and obese groups, in this study, were categorized by waist circumference (WC). The 1.5 and 3 cm. thickness of frozen fat lard was placed on a phantom pelvis to simulate the difference in AP abdominal diameter for the overweight and obese groups, respectively. The WC is correlated with abdominal fat,[15] and represents a valuable, convenient, and simple measurement method to quickly identify the risk factors and morbidity of obesity-related diseases.[16,17] The WC is a more suitable and simple body measurement method for obese patients than body mass index (BMI) because high WC reveals that patients have high central obesity.[18]

For dose measurement, results showed that the radiation dose delivered to patients by DAP measurement was suitable, and provided a good estimate of the total radiation energy delivered during a procedure, being directly related to the amount of radiation produced by the X-ray tube.[19] Our results showed DAP reduction when using high kVp, concurring with previous research reporting that DAP decreased at higher kV and copper filtration.[20] Further studies should be conducted to determine other factors affecting DAP such as radiographers' experience and type of X-ray machine used. Moreover, in obese patients need higher kVp to increase photon energy for passing through the patient's body resulting in more back scatter radiation occurred. Further study should consider the back scatter factor (BSF) added to the patient dose.

Tube voltage (kV) and tube current (mA) control the quantity of radiation, while the amount of radiation delivered is the product of mA and exposure time or milliampere seconds (mAs) and affects the noise. Using AEC detected mAs rather than kVp and can controlled the noise in the DR image receptor at a minimal level. In this study, AEC was used for exposures and lead to the adjustment of the mAs to obtain a constant radiation dose level. Moreover , the AEC was used as a method for the termination of radiation exposure, while AEC is also commonly used for performing routine projections in lumbar spine radiography examinations in clinical practice.[21]

This study had several limitations. Firstly, only one type of DR system was investigated, therefore the result may not be representative of other systems. Secondly, fat was simulated using frozen pork lard, which does not fully represent the distribution of human fat in overweight and obese patients. This study investigated only one parameter (kVp setting), and other parameters such as SID or additional filters should also be considered. The radiation dose used in this study (DAP) does not account for the BSF which influences on patient dose. However, DAP is an uncomplicated radiation measurement and could be used for determination as well as the calculation of the entrance skin dose.[22] Lastly, the ICC was not satisfactory for the evaluation of inter-rater reliability and demonstrated poor interobserver agreement for both AP and lateral projections (ICC<0.4). The observers had varying skill levels (experience 3-10 years) and were not trained before performing image ratings. In future studies, practice and training should be conducted for the observers before image grading.

In conclusion, this study finding suggested the potential of using high kVp combined with AEC used for lumbar spine radiography imaging in overweight and obese patients using the DR system. Our data suggested that utilization of 100-109 kVp in combination with the use of AEC gave the optimal lumbar spine image in terms of minimal radiation dose and good image quality for overweight and obese patients. The DAP decreased in both AP and lateral projections, while VGS had good visibility at higher kV values.

Conflict of interest

The authors declare no conflict of interest.

Funding

This research study was supported by the Thailand Institute of Nuclear Technology (Public Organization), Thailand, through its program TINT to University. (2022)

Acknowledgements

The authors would like to acknowledge the Thailand Institute of Nuclear Technology (Public Organization) for their kind assistance for dose measurement. The authors would also like to thank the radiographers at Chomthong Hospital, Chiang Mai, Thailand for image quality assessment. Thanks are due to Pitsanuvej Hospital, Phisanulok Province, Thailand, for providing the CT scanner to obtain CT data. This research study was supported by the Thailand Institute of Nuclear Technology (Public Organization), Thailand, through its program of TINT to University 2022.

References

[1] World Health Organization. Obesity and overweight world health organization [Internet]. [cited 2021 Apr 1]. Available from: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

[2] De Lorenzo A, Gratteri S, Gualtieri P, *et al*. Why primary obesity is a disease? J Transl Med. 2019; 17: 1-13. doi.org/10.1186/s12967-019-1919-y.

[3] Le NTT, Robinson J, Lewis SJ. Obese patients and radiography literature: what do we know about a big issue? J Med Radiat Sci. 2015; 62: 132-41. doi.org/10.1002/jmrs.105

[4] Yanch JC, Behrman RH, Hendricks MJ, *et al*. Increased radiation dose to overweight and obese patients from radiographic examinations. Radiology. 2009; 252: 128-39. doi.org/10.1148/radiol.2521080141.

[5] Van den Heuvel J, Punch A, Aweidah L, *et al*. Optimizing projectional radiographic imaging of the abdomen of obese patients: an e-delphi study. J Med Imaging Radiat Sci. 2019; 50:289-96. doi.org/10.1016/j.jmir.2019.01.004.

[6] Carucci LR. Imaging obese patients: problems and solutions. Abdom Imaging. 2013; 38: 630-46. doi.org/10.1007/s00261-012-9959-2.

[7] Alqahtani SJ, Knapp KM. Imaging patients with obesity. J Med Radiat Sci. 2022; 69(1): 3-4. doi.org/10.1002/jmrs.560.

[8] Chan CTP, Fung KKL. Dose optimization in lumbar spine radiographic examination by air gap method at CR and DR systems: A phantom study. J Med Imaging Radiat Sci. 2015; 46: 65-77. doi.org/10.1016/j.jmir.2014.08.003.

[9] Yoshizumi T, Nakamura T, Yamane M, *et al*. Abdominal fat: standardized technique for measurement at CT. Radiology. 1999; 211: 283-86. doi.org/10.1148/radiology.211.1.r99ap15283.

[10] Ferreira T, Rasband W. ImageJ User Guide-Version 1. 44. 2012.

[11] Gatt S, Portelli JL, Zarb F. Optimisation of the AP abdomen projection for larger patient body thicknesses. Radiography. 2022; 28: 107-14. doi.org/10.1016/j.radi.2021.08.009.

[12] Office for Official Publications of the European Communities. European guidelines on quality criteria for diagnostic radiographic images. Luxemburg; 1996.

[13] Hauge IH, Aandahl IJ, Baranzelli JP, *et al*. Radiography: impact of lower tube voltages on image quality and radiation dose in chest phantom radiography for averaged sized and larger patients. In: OPTIMAX 2017: optimising image quality for medical image. 2017. p. 47-62.

[14] Cicchetti D V. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess. 1994; 6: 284-90. doi.org/10.1037/1040-3590.6.4.284.

[15] Birwadkar G, Ratta AK. Assessment of risk factors of obesity among residents of class iv employees' quarters of a tertiary care hospital. J Evol Med Dent Sci. 2019; 8: 2367-70. doi.10.14260/jemds/2019/518.

[16] Han TS, Lean ME. A clinical perspective of obesity, metabolic syndrome and cardiovascular disease. JRSM Cardiovasc Dis. 2016; 5. doi.org/10.1177/2048004016633371.

[17] Anuradha R, Hemachandran S, Dutta R. The waist circumference measurement: a simple method for assessing the abdominal obesity. J Clin Diagnostic Res. 2012; 6: 1510-3. doi.org/10.7860/JCDR/2012/4379.2545.

[18] Ismail AA, Othman NS, Mohamad M, *et al*. Evaluating the relationship of body mass index and waist circumference on the image quality of abdominal computed radiography. J Sains Kesihat Malaysia. 2020; 18: 11-8. https://ejournal.ukm.my/jskm/article/view/26118/10301.

[19] Allisy-Roberts P, Williams J. Farr's physics for medical imaging. 2nd Ed. Elsevier Health Sciences; 2007.

[20] Mifsud K, Portelli JL, Zarb F, *et al*. Evaluating the use of higher kVp and copper filtration as a dose optimisation tool in digital planar radiography. Radiography. 2022; 28: 586-92. doi.org/10.1016/j.radi.2022.04.002.

[21] Bontrager KL, Lampignano JP. Textbook of positioning and related anatomy. 8th Ed. 2014.

[22] Ghayour-Saffar N, Ehsanbakhsh A, Keshtkar M, Pandesh S. Evaluation of DAP Values Obtained from Chest X-rays in Children under 12 Years of Age Referred to Educational Hospitals of Birjand University of Medical Sciences in 2020. Frontiers Biomed Technol. 2022; 9(3): 185-90. doi.org/10.18502/fbt.v9i3.9644.

Journal of Associated Medical Sciences 2024; 57(3): 8-17

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Breast cancer characterization using region-based convolutional neural network with screening and diagnostic mammogram

Jaroonroj Wongnil[1] Anchali Krisanachinda[1] Rajalida Lipikorn[2*]

[1]Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand.
[2]Machine Intelligence and Multimedia Information Technology Laboratory, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 16 March 2024
Accepted as revised 17 May 2024
Available online 23 May 2024

Keywords:
Malignant calcification, artificial intelligence, two-stage detection model, one-stage detection model, region-based convolutional neural network.

ABSTRACT

Background: Detection and classification of microcalcifications in breast tissues is crucial for early breast cancer diagnosis and long-term treatment.

Objective: This paper aims to propose a robust model capable of detection and classification of breast cancer calcifications in digital mammogram images using Deep Convolutional Neural Networks (DCNN).

Materials and methods: An expert breast radiologist annotated the 3,265 clinical mammogram images to create a comprehensive ground truth dataset comprising 2,500 annotations for malignant and benign calcifications. This dataset was utilized to train our model, a two-stage detection system incorporating a Region-based Convolutional Neural Network (RCNN) with AlexNet and support vector machines to enhance the system's robustness. The proposed model was compared to the one-stage detection, utilizing YOLOv4 combined with the Cross-Stage Partial Darknet53 (CSPDarknet53) architecture. A separate dataset of 504 mammogram images was explicitly set aside for model testing. The efficacy of the proposed model was evaluated based on key performance metrics, including precision, recall, F1 score, and mean average precision (mAP).

Results: The results showed that the proposed RCNN-2 model could automatically identify and categorize calcifications as malignant or benign, outperforming the YOLOv4 models. The RCNN-2's overall effectiveness, as evaluated by precision, recall, F1 score, and mean average precision (mAP), achieved scores of 0.82, 0.85, 0.83, and 0.74, respectively.

Conclusion: The proposed RCNN-2 model demonstrates very effective detection and classification of calcification in mammogram images, especially in high-dense breast images. The performance of the proposed model was compared to that of YOLOv4, and it can be concluded that the proposed RCNN model yields outstanding performance. The model can be a helpful tool for radiologists.

Introduction

Breast cancer is one of the most common cancers and has the highest mortality rate among women worldwide. In 2022, the World Health Organization (WHO) reported that the first rank of new cancer cases in Thai women was breast cancer, with 21,628 patients (23.2%) and the second-highest cause of death in the same year.[1] Mammography is a crucial medical imaging technique and the gold standard for breast cancer detection.[2] Digital mammography, which replaces screen-film technology, provides superior images.[3-5] Its widespread use in breast cancer screening and diagnosis has significantly contributed

* **Corresponding contributor.**
Author's Address: *Machine Intelligence and Multimedia Information Technology Laboratory, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok, Thailand.*
E-mail address: rajalida.l@chula.ac.th
doi: 10.12982/JAMS.2024.042
E-ISSN: 2539-6056

to early detection, reducing breast cancer mortality by 40%.[6-8] Microcalcification is an early indicator of breast cancer, identified as bright, white spots or dots on the breast tissue in mammography images.[9,10] Malignant calcifications are usually smaller and typically range from <0.5-1 mm.

In contrast, benign calcifications are generally more extensive, with a more defined and coarse appearance.[11-13] However, breast density influences the diagnostic sensitivity and efficacy of mammography.[14,15] The Breast Imaging Reporting and Data System (BI-RADS), developed by the American College of Radiology (ACR), indicates that high-dense breast tissue significantly impacts mammography's sensitivity and diagnostic accuracy.[16-18]

Artificial intelligence powered by deep learning with the convolutional neural network (CNN) has been widely applied in medical imaging.[19] They enable automatic and adaptive feature learning across low to high-level complexity patterns across spatial levels.[20,21] In Thailand, numerous studies have utilized CNNs for the classification of breast cancer in various breast imaging. For instance, Aphinives *et al.* explored AI development using free-trial services to detect microcalcifications in mammography.[22] Additionally, Intasam *et al*. investigated deep learning models for classifying mammograms as benign or malignant.[23] Recently, Labcharoenwongs *et al*. developed an automated breast tumor detection and classification system using deep learning techniques based on the computerized analysis of breast ultrasound images.[24] Despite these advancements, targeted research on detecting and classifying microcalcifications remains relatively limited, especially in high-density breast tissues.

Therefore, this study aims to propose a robust model capable of detecting and characterizing breast cancer calcifications in digital mammogram images using Deep Convolutional Neural Networks (DCNN) by developing a two-stage network architecture utilizing authentic clinical breast images in various formats, coupled with critical parameter adjustments. The model was compared with a one-stage network to determine the most suitable method for detecting and classifying breast cancer calcifications in digital mammography for Thai women.

Materials and methods

This research is a retrospective diagnosis that received approval from the Institutional Review Boards (IRB) of two hospitals that provided digital mammogram images. Due to the retrospective nature, informed consents from patients were waived. Mammogram images from two hospitals were merged to strengthen the learning model's robustness. Patient demographics and pathological data were extracted from electronic medical records (EMR). The process consists of data preparation, model construction, and performance evaluation.

Data preparation

Three thousand two hundred sixty-five clinical mammogram images were collected from January 1, 2018, to December 31, 2019. The dataset comprises Mediolateral Oblique (MLO) and Craniocaudal (CC) views for each breast. Breast density categories, routinely assigned by radiologists in standard clinical workflows using the BI-RADS system, were retrieved from mammography reports.[25,26] Malignant calcifications on mammograms, identified and reviewed by radiologists with histopathological confirmation through biopsy, were also utilized as ground truth. Patients lacking histopathological data were excluded from the research.

The first step was manually removing artifacts from the images, such as location markers and views. Regions of Interest (ROIs) were defined and extracted from MLO and CC views. The ROIs of suspicious areas were manually cropped according to the distribution of calcifications, and adjustments were made to ensure that all relevant areas were included. ROI criteria were derived from radiological and pathological reports executed by expert breast radiologists. Each ROIs was saved as a new image with 227x227 pixels. These ROIs were labeled as one of three categories: malignant calcification, benign calcification, or normal breast tissue. The number of ROIs per image could range from one to four. A total of 5,000 ROI images were generated, consisting of 2,500 ROIs identified as malignant calcifications, 1,250 as benign calcifications, and another 1,250 as normal breast tissue. Examples of ROIs for each category are shown in Figure 1(a).

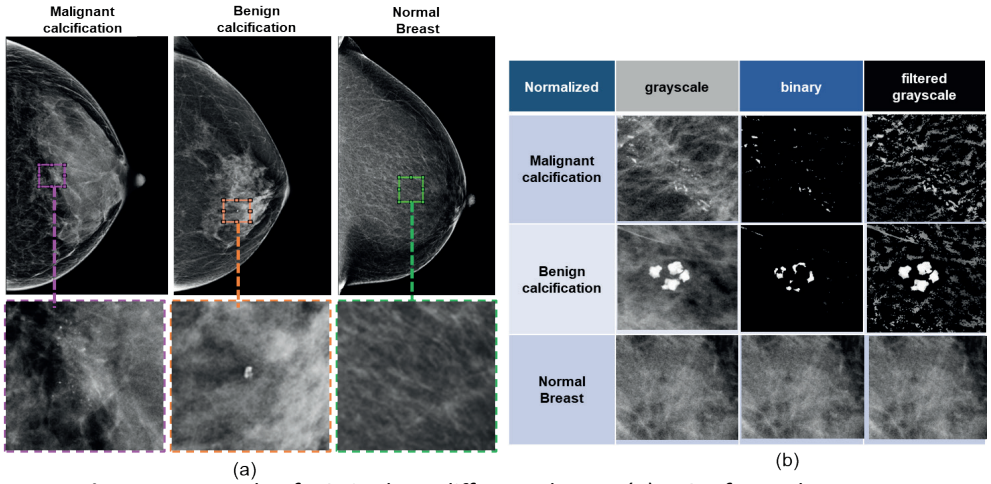


Figure 1. *Example of ROI in three different classes. (a): ROIs for each category, (b): Three sets of generated data from ROIs for training CNN.*

The next step was to generate data from ROIs for training CNN using grayscale normalization, binarization, and logical AND operation to obtain grayscale images, binary images, and filtered images that represent potential real-life scenarios in mammogram imaging, as shown in Figure 1(b). For grayscale normalization, ROI images were transformed and normalized into grayscale with an intensity range between 0 and 1 to ensure consistency and remove variations in intensity scaling, allowing us to focus on structural attributes and intensity fluctuations within the breast tissue that are crucial for detecting abnormalities. For binarization, ROI images were converted to binary images using adaptive thresholding. The foreground polarity was used to specify that the desired foreground (object) is brighter than the background. A threshold value was estimated by setting the sensitivity parameter to 0.5. Any pixels surpassed this threshold value were set to 1 (white), while others were set to 0 (black). These binary images facilitate feature extraction and region-based analysis. For logical operation, the AND operator was applied to each grayscale ROI image and a binary image to produce a filtered grayscale image that highlighted lesions and minimized background distractions. This method helped sharpen lesion visibility, reduces noise, and preserves essential intensity details for better analysis. Consequently, each dataset comprised 5,000 ROI images.

The final stage was to create a ground truth dataset for training and evaluating the proposed and comparative models. To establish the ground truth dataset, 2,500 bounding boxes were drawn on 883 mammogram images displaying malignant features and annotated on 552 images identified with benign features. Each mammogram image may contain up to four bounding boxes. Expert breast radiologists supervised this entire process.

Model construction

The proposed model is a two-stage detection system based on RCNN using AlexNet as the base network, as shown in Figure 2. All computations were performed by MATLAB version R2022a on a personal computer (CPU: Intel Core i7, RAM: 24 GB RAM, NVIDIA 64-bit operating system). A five-fold cross-validation was employed to evaluate and select the model.

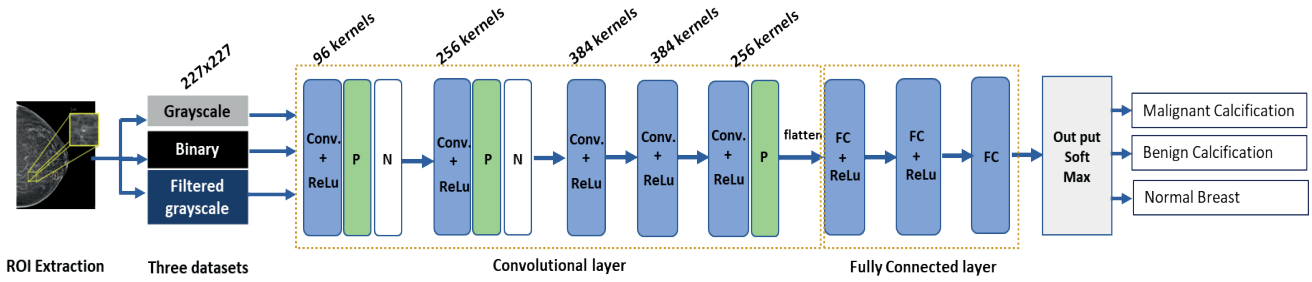


Figure 2. *CNN backbone architecture.*

AlexNet was first trained by 4000 cropped ROI images that were resized to 227x227 pixels for the input layer to improve the classification accuracy. These input data were fed to convolutional layers deploying 96 and 256 kernels for the initial feature map that were enhanced with max pooling and normalization techniques in the first and the second layers. Then, feature analysis was performed through successive convolutional layers, utilizing 384 kernels without pooling until the final convolutional layer. The final convolutional layer was constructed by utilizing 256 kernels. The flattening was used to convert the 2-dimensional arrays obtained from max pooling into a vector. The vector was fed to the fully connected layer. The network's final layer was customized to distinctively craft for three classes: malignant, benign, and normal tissue. The learning rates were increased for quicker adaptation to mammogram data.

A five-fold cross-validation process was employed. AlexNet models 1 through 5 represented each model built from each validation cycle, where the model performances were evaluated from randomly segmented data. The AlexNet model that demonstrated the highest diagnostic precision was chosen as the base network of the RCNN framework, ensuring comprehensive validation against various data patterns and potential anomalies.

Among the 5,000 ROI images in each dataset, 1,000 were used as test data, and the remaining 4,000 ROI images were divided into a training set (80%) and a validation set (20%). The model with the highest performance was selected as the base network of the RCNN model.

The proposed RCNN model for calcification detection and classification that consisted of two processes was constructed starting from using the edge box method[27] to find the region proposals that might contain calcification and using AlexNet as a base network to classify each region proposal whether it is malignant or benign as shown in Figure 3. This figure shows the proposed two-stage model for detecting and classifying breast calcifications. The model begins with an edge box method to locate potential regions of interest (ROIs) or region proposals containing calcifications. These identified region proposals are then processed using AlexNet, which serves as the base network of the proposed RCNN model. The process starts with feature extraction layers, and the sequence involves five convolutional layers (Conv) applying filters to regions of interest (ROIs) to extract detailed features. The output of each layer feeds into the next, while Local Response Normalization (N) is applied after the first and second pooling layers to normalize the responses. Max pooling (P) is applied after the first, second, and fifth convolutional

layers to reduce their dimensionality and processed by ReLU activation functions. The extracted features are flattened and passed through multiple fully connected layers to synthesize the learned information. Finally, a Support Vector Machine (SVM) is integrated to enhance the classification accuracy by optimally separating the identified classes with maximum margins, significantly boosting the precision and robustness of the model. The process culminates in a softmax output layer that classifies the regions into malignant or benign categories, clearly depicting the classification results. The proposed RCNN model was trained with 2,888 images of the whole breast randomly selected from the ground truth dataset. The remaining 377 images of the entire breast were used as test data.

For comparison, the one-stage network model utilizes You Only Look Once version 4 (YOLOv4), known for its efficient one-stage object detection was also implemented. It incorporated the Cross-Stage-Partial-connections Darknet-53 (CSPDarknet53), a 53-layer CNN that used residual connections and Leaky ReLU activation to improve training efficiency and accuracy. YOLOv4 was trained and tested using the same ground truth dataset, ensuring a consistent basis for comparison.

The proposed RCNN model and YOLOv4 were trained to utilize the same ground truth dataset. Critical hyperparameters such as learning rate, batch size, and epochs were meticulously optimized to enhance model performance. This optimization developed distinct configurations: RCNN Models 1 and 2 and YOLOv4 Models 1 and 2. Specifically, RCNN Model 1 and YOLOv4-1 shared hyperparameters with a learning rate 0.001, a batch size of 128, and 50 epochs. Conversely, RCNN Model 2 and YOLOv4-2 were configured with a learning rate 0.0001, maintaining the same batch size but extending to 100 epochs.

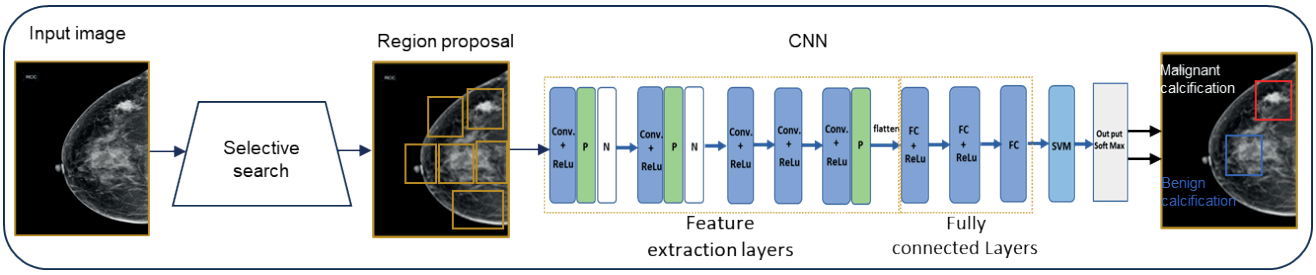


Figure 3. *The proposed two-stage model.*

Performance evaluation

This research utilized multi-statistical metrics to evaluate CNN backbone and the two-stage model and to compare the two-stage and one-stage networks. The equations used in this research are displayed in Table 1. The models' performances were determined using confusion matrices, employing a 3x3 matrix for the multi-class classification by AlexNet and a 2x2 matrix for evaluating the performance of the proposed RCNN model and YOLOv4. These detailed evaluations thoroughly analyzed the models' accuracies and potential utility in medical imaging diagnostics.

Table 1. Confusion matrix of three models.

Evaluation metrics	Equation	Description
True positive	TP	Predicted that a bounding box exists, object is was correct
False positive	FP	Predicted that a bounding box exists, but object is was wrong
False negative	FN	Did not predict a bounding box, even though an object is there
True negative	TN	Not typically defined in the context of object detection evaluation metrics.
Precision (Positive predictive value)	$\frac{TP}{TP+FP} = \frac{\text{Correct Predictions}}{\text{Total Predictions}}$	Probability of the predicted bounding boxes that matched the actual ground truth boxes
Recall (Sensitivity, True positive rate)	$\frac{TP}{TP+FN} = \frac{\text{Correct Predictions}}{\text{Total GroundTruth}}$	Probability of correctly detecting ground truth objects
F1 Score	$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$	A balanced performance measure of the model performance (harmonic mean precision and recall)
Micro average F1-Score	$\frac{\sum_{i=1}^{n} TP_i}{\sum_{i=1}^{n} TP_i + 1/2(\sum_{i=1}^{n} FP_i + \sum_{i=1}^{n} FN_i)}$	Sums result from all classes, including TP, FN, and FP, to compute an overall F1 score, making it suitable for evaluating models on imbalanced datasets.
Macro average F1 score	$\frac{\sum_{i=1}^{n} F1\_Score_i}{n}$	Averages the F1 scores for all classes, treating each equally, and is ideal for assessing model performance across varied class distributions.
Weighted average F1 score	$\sum_{i=1}^{n} Wi \times F1\_Score_i$	Multiplying each class's F1 score by its proportion in the dataset and then summing these values, providing a metric that accounts for class imbalance.
Mean average precision (mAP)	$\frac{1}{n} \sum_{k=1}^{k=n} APk$	Performance measurement across multiple classes AP_k is AP of class *k*, *n* is the number of classes

Results

The performances of AlexNet across three datasets of 5,000 ROI-cropped images using five-fold cross-validation are shown in Table 2. The results highlight that AlexNet4, AlexNet5, and AlexNet6 achieved the highest accuracies of 90.40%, 87.10%, and 85.00% for the grayscale, binary, and filtered grayscale datasets, respectively. The top-performing model for each dataset was further evaluated on a separate test set of 1,000 ROI images, with the classification outcomes depicted through a 3x3 confusion matrix for each dataset, as shown in Figure 4. Furthermore, the evaluation of the three AlexNet models includes micro average F1 score, macro average F1 score, and weighted average F1 score, as shown in Table 3. Notably, the AlexNet4 model using the grayscale dataset outperforms all other F1 scores, leading to its selection as the base network of the proposed RCNN model.

Table 2. Performances of AlexNet models on three datasets.

Model	Grayscale	Binary	Filtered grayscale
AlexNet1	89.60	85.30	85.00
AlexNet2	88.80	85.30	84.30
AlexNet3	88.80	86.20	83.80
AlexNet4	90.40	84.30	83.30
AlexNet5	89.33	87.10	85.00

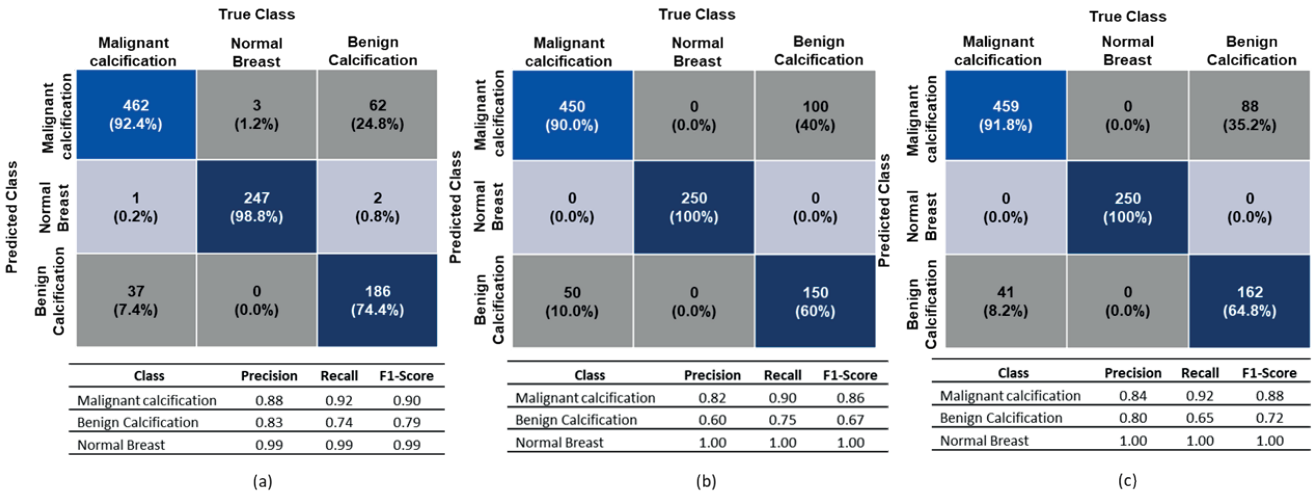


Figure 4. *The 3x3 confusion matrices represent the classification results. (a): AlexNet4 on grayscale ROIs, (b): AlexNet5 on binary ROIs, (c): AlexNet5 on filtered grayscale ROIs with precision, recall, and F1-score for each class.*

Table 3. Performances of the best AlexNet models across three variations of ROI images.

Dataset	Micro average F1	Macro average F1	Weighted F1
AlexNet4 on grayscale	0.90	0.89	0.89
AlexNet5 on binary	0.85	0.84	0.85
AlexNet5 on filtered grayscale	0.87	0.86	0.87

The proposed RCNN models employing AlexNet4 and YOLOv4 with CSPDarknet53 underwent training using the same ground truth dataset. The trainings were conducted under two different configurations to assess the model's effectiveness and the influence of hyperparameters on their performances compared to YOLOv4, as shown in Table 4.

Table 4. Hyperparameters and training time of the detection and classification models.

Detector network model	Learn rate (LR)	Batch size (BS)	Epochs	Training hours
R-CNN-1	0.001	128	50	34.31±3.50
R-CNN-2	0.0001	128	100	66.77±3.53
YOLOv4-1	0.001	128	50	1.87±0.14
YOLOv4-2	0.0001	128	100	3.46+0.14

RCNN-1, RCNN-2, YOLOv4-1, and YOLOv4-2 were performed using a test set of 377 mammogram images. Table 5 summarizes the performances of four models across various metrics, considering both scenarios with a confidence score at a threshold value equal to 0.5 (CF) and without a confidence score (No CF), which provides a holistic view of model performances. Furthermore, the use of five-fold cross-validation and varying confidence score threshold values enable detailed calculations of average precision and recall for each class, enhancing the robustness and clarity of the evaluation.

Table 5. Classification performances of RCNN-1, RCNN-2, YOLOv4-1, and YOLOv4-2 based on precision and recall.

Model	Class		Precision		Recall	
			No CF	CF 0.5	No CF	CF 0.5
R-CNN-1	Malignant calcification	Average	0.31	0.76	0.51	0.58
		SD	0.03	0.04	0.02	0.01
	Benign calcification	Average	0.34	0.75	0.54	0.61
		SD	0.03	0.03	0.01	0.03
R-CNN-2	Malignant calcification	Average	0.57	0.82	0.83	0.84
		SD	0.05	0.03	0.01	0.02
	Benign calcification	Average	0.60	0.83	0.83	0.85
		SD	0.05	0.02	0.02	0.02
YOLOv4-1	Malignant calcification	Average	0.32	0.79	0.53	0.55
		SD	0.03	0.01	0.03	0.01
	Benign calcification	Average	0.34	0.78	0.54	0.56
		SD	0.02	0.03	0.03	0.01
YOLOv4-2	Malignant calcification	Average	0.44	0.73	0.65	*0.75*
		SD	0.05	0.04	0.02	0.03
	Benign calcification	Average	0.43	0.78	0.66	*0.78*
		SD	0.04	0.07	0.02	0.02

Note: No CF: without a confidence factor, CF 0.5: confidence factor at threshold value equal to 0.5.

Table 6 shows the comparative performance metrics of the four models based on average precision, recall, F1-score, and mean Average Precision (mAP). These values are presented along with their ranges to account for variability in the five-fold cross-validation process. The results indicate that RCNN-2 achieves superior performance metrics compared to other models.

Table 6. Performance metrics for the four models.

Model	Precision	Recall	F1 Score	mAP
R-CNN-1	0.72 (0.66-0.78)	0.66 (0.58-0.74)	0.69 (0.65-0.73)	0.66 (0.65-0.67)
R-CNN-2	0.82 (0.80-0.84)	0.85 (0.83-0.87)	0.83 (0.82-0.84)	0.74 (0.73-0.75)
YOLOv4-1	0.72 (0.64-0.80)	0.57 (0.54-0.60)	0.64 (0.54-0.74)	0.70 (0.63-0.77)
YOLOv4-2	0.77 (0.73-0.81)	0.78 (0.74-0.80)	0.77 (0.76-0.78)	0.70 (0.66-0.74)

Discussion

The proposed RCNN model with a grayscale dataset achieved superior performance, as evidenced by the highest F1 scores. As presented in Figure 5, the comparison across different dataset forms clearly shows the advantages of using grayscale images. In the grayscale image (Figure 5a), essential intensity details are maintained, aiding in the more precise differentiation of classes. On the other hand, binary and filtered grayscale images (Figure 5b and c) might simplify the foreground but at the cost of losing fine details, which can be detrimental when analyzing tiny lesions.

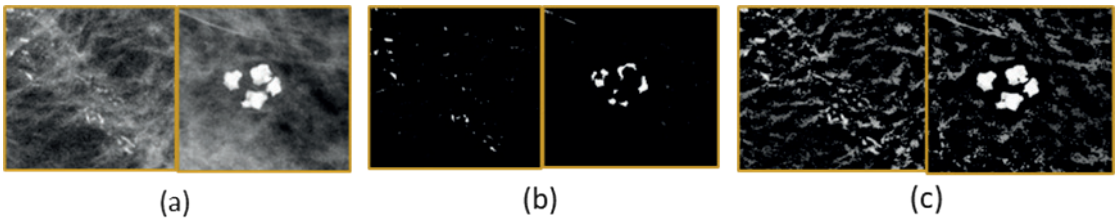


Figure 5. *Images of malignant calcification and benign calcification of breast tissue. (a): grayscale, (b): binary, (c): filtered grayscale.*

This paper proposes the RCNN model that underscores the importance of precise hyperparameter adjustments, notably in learning rates. Such fine-tuning was the key to the RCNN-2 model's exceptional performance, enabling it to analyze complex mammographic patterns accurately. Despite newer CNN models, RCNN was chosen for its proven efficacy in object-scale datasets like mammograms. The RCNN-2 model with a learning rate 0.0001 and 100 epochs yields the highest precision, recall, F1 score, and Map. A low learning rate facilitates a gradual understanding of complex patterns in mammogram images that could lead to better convergence and enhance performance, showcasing its ability to differentiate subtle details in dense breast images and, furthermore, setting the confidence score threshold at 0.5 enhanced detection accuracy across all four models. This threshold level could effectively reduce less reliable detections, particularly in the cases that contributed to partial false positives and false negatives, as depicted in Figure 6. Confidence scoring is pivotal in refining model performance.

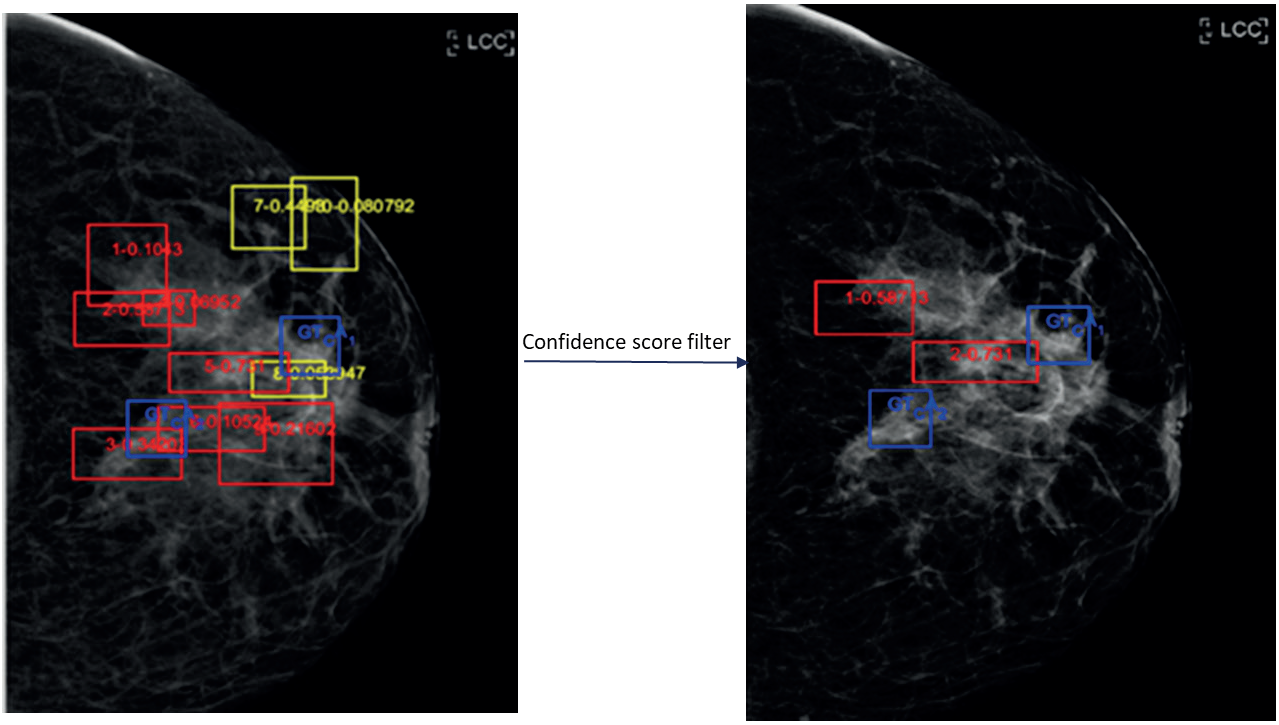
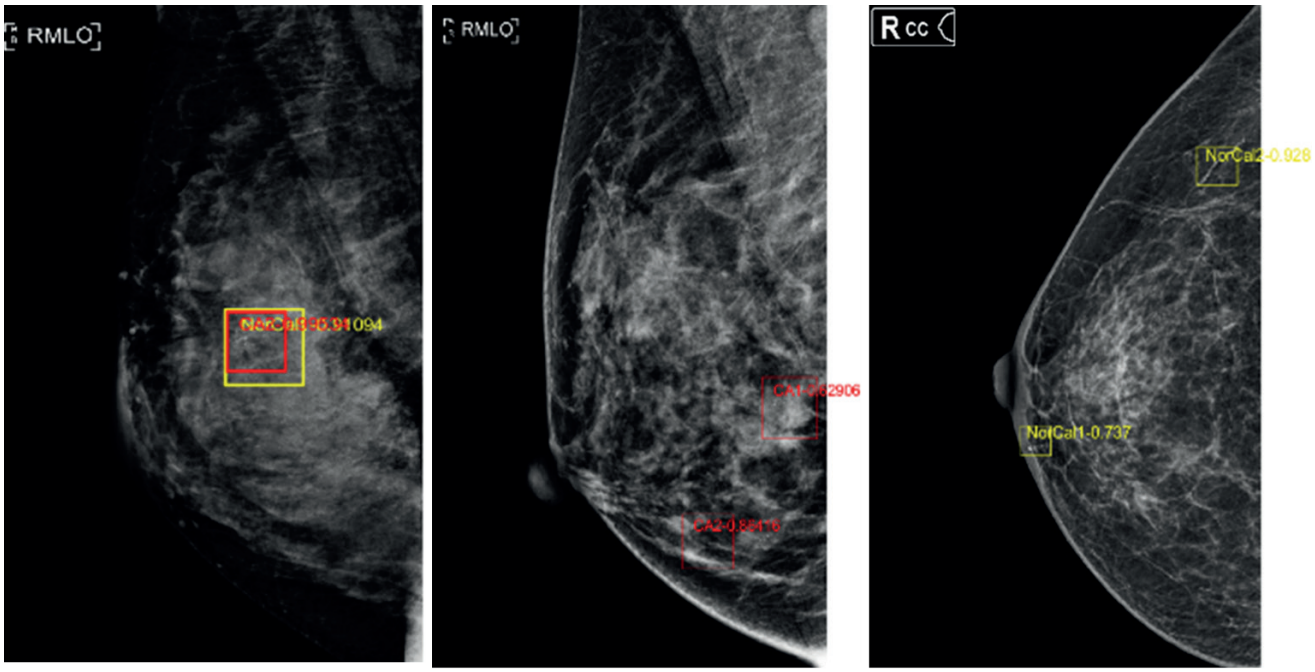


Figure 6. *Utilization of a confidence score helped reduce the impact of unreliable detections that resulted in false positives (FP) and false negatives (FN).*

The RCNN-2 model excels with precision (0.82) and recall (0.85), showcasing its strength in accurate classification and valid positive identification, resulting in an F1 score of 0.83 and mAP of 0.74. Meanwhile, YOLOv4-2 closely follows precision (0.77) and recall (0.78), with an F1 score of 0.77 and mAP of 0.70. Despite lower metrics, RCNN-1 and YOLOv4-1 still post F1 scores of 0.69 and 0.64 with a mAP of 0.70 each. RCNN-2 notably outperforms in detecting and classifying calcifications in dense breasts, a challenging task due to the overlapping characteristics of benign and malignant calcifications.

The proposed model demonstrates its capability to distinguish between benign and malignant microcalcifications in dense breast tissue, as shown in Figure 7(a). This ability indicates that the proposed model can extract and analyze distinctive patterns and characteristics hidden within the highly dense breasts. Figure 7(b) exemplifies highly dense breast tissue accuracy with two correctly classified region proposals. Figure 7(c) highlights the ease of detection in a non-dense breast due to its sharper background.



(a) (b) (c)

Figure 7. *Examples of mammogram images. (a): model's ability to differentiate between benign and malignant microcalcifications in dense breast tissue, (b): two correctly classified region proposals, (c): clear detection in a non-dense breast.*

In clinical validation, the assessment outcomes stem from comparing the model results with the diagnostic results from the radiologist. This section employed the latest mammogram images and preliminary tests to demonstrate the model's effectiveness with current clinical mammogram images. The expert breast radiologists agree with the detection and classification results obtained from the proposed model, demonstrating the model's proficiency in interpreting and classifying mammographic abnormalities. Furthermore, it signifies the model's potential as an auxiliary tool in the diagnostic process. However, it is essential to acknowledge that the images can sometimes lead to disparities in classification. These discrepancies typically arise when the contents of some images present subtle or ambiguous features that require subjective interpretation, especially in high-dense breast mammogram images.

In comparing our research with other studies on breast cancer using deep learning, it is evident that each approach offers unique insights. Aphinives *et al.* highlighted AI's capability in detecting microcalcifications with a precision of 80.0% and a recall of 12.5%, depending on training duration.[22] Intasam *et al*. reported an accuracy of 86.76% by evaluating various CNN architectures.[23] Labcharoenwongs *et al*. advanced a system for tumor detection and volume estimation in ultrasound images, achieving high accuracy and robust classification.[24] Our study employs a two-stage detection system using RCNN integrated with AlexNet, improving robustness and accuracy and addressing both detection and classification of microcalcifications in high-density breast tissues. This comprehensive approach enhances diagnostic tools for early breast cancer detection, especially in challenging dense breast tissues, and has been validated in clinical settings by expert breast radiologists.

Limitation

The limitation of this research is the lack of information on female patients who might be suspected of having breast cancer; it does not consistently offer complete care information, such as pathological reports that affect the collection of mammogram images and the model's ability to fully understand and predict based on localized demographic and clinical nuances. The computational constraints also impact the deep learning model's effectiveness and increase processing times.

Conclusion

In conclusion, this research successfully addresses its objective to develop a robust two-stage Deep Convolutional Neural Network (DCNN) model for detecting and classifying breast cancer calcifications in digital mammogram images. The proposed RCNN model, built upon the proposed CNN based on AlexNet, demonstrates significant potential in aiding radiologists in identifying microcalcifications, particularly within high-density breast tissues of Thai subjects. The comparative analysis with the one-stage YOLOv4 underscored the superior precision and recall scores achieved by our RCNN model, emphasizing the benefits of model fine-tuning and training with varied datasets. Future research will explore other alternative models and develop specifically address the distinct features of Thai female breasts, aiming to enhance the precision and dependability of breast cancer diagnostics. This study also highlights that the identification of breast cancer calcifications in highly dense breasts by expert

radiologists, combined with a model that, while not the latest version, still maintains high capability and is tuned with appropriate parameters, can result in artificial intelligence aiding healthcare professionals in early breast cancer detection. This is particularly crucial in dense breast tissues, where traditional mammography may falter, aiming to improve patient outcomes and screening efficiency.

Conflict of interest

The authors declare no conflict of interest.

Funding

No funding was received to conduct this study.

Ethical Approval

The Ethical Committee of the Faculty of Medicine, Chulalongkorn University, Thailand (IRB no. 503/63), and Bumrungrad International Hospital, Bangkok, Thailand (Grant No. 20:205/CRC:ai) approved the research.

Acknowledgements

The authors would like to thank expert breast radiologists, Department of Radiology, Faculty of Medicine, Chulalongkorn University, and Bumrungrad International Hospital, Bangkok, Thailand, for their guidance and analysis.

References

[1] Ferlay J, Ervik M, Lam F, Laversanne M, Colombet M, Mery L, *et al*. Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. 2024. https://gco.iarc.fr/today [accessed 1 February 2024].

[2] Padia K, Douglas TS, Cairncross LL, Baasch RV, Vaughan CL. Detecting Breast Cancer with a Dual-Modality Device. Diagnostics. 2017; 7(1): 17. doi: 10.3390/diagnostics7010017

[3] Haus AG, Yaffe MJ. Screen-Film and Digital Mammography: Image Quality and Radiation Dose Considerations. Radiol Clin North Am. 2000; 38(4): 871-98. doi: 10.1016/s0033-8389(05)70207-4

[4] Sechopoulos I. A review of breast tomosynthesis. Part I. The image acquisition process. Med Phys. 2013; 40(1): 014301. doi: 10.1118/1.4770279

[5] International Atomic Energy Agency, Quality Assurance Programme for Digital Mammography, IAEA Human Health Series No. 17, IAEA, Vienna, 2011.

[6] Yilmaz R, Aydiner A, Igci A, Soran A. Breast Imaging: Breast Cancer A Guide to Clinical Practice. 1st ed. Springer Nature, Switzerland, AG, 2019.

[7] American College of Radiology, Mammography saves lives. https://www.acr.org/Practice-Management-Quality-Informatics/Practice-Toolkit/Patient-Resources/Mammography-Saves-Lives. Accessed on February 20, 2023.

[8] Monticciolo DL, Newell MS, Hendrick RE, Helvie MA, Moy L, Monsees B, *et al*. Breast Cancer Screening for Average-Risk Women: Recommendations From the ACR Commission on Breast Imaging. J Am Coll Radiol. 2017; 14: 1137-43. doi: 10.1016/j.jacr.2017.06.001

[9] Nyante SJ, Lee SS, Benefield T, Hoots TN, Henderson LM. The association between mammographic calcifications and breast cancer prognostic factors in a population-based registry cohort. Cancer. 2017; 123: 219-27. doi: 10.1002/cncr.30281

[10] Logullo AF, Prigenzi KCK, Nimir CCBA, Franco AFV, Campos MSDA. Breast microcalcifications: Past, present and future (Review). Mol Clin Oncol. 2022; 16(4): 81. doi: 10.3892/mco.2022.2514

[11] Cai H, Huang Q, Rong W, Song Y, Li J, Wang J. *et al*. Breast Microcalcification Diagnosis Using Deep Convolutional Neural Network from Digital Mammograms. Comput Math Methods Med. 2019; 1-10. doi: 10.1155/2019/2717454

[12] Chen Z, Strange H, Oliver A, Denton ER, Boggis C, Zwiggelaar R, et al. Topological Modeling and Classification of Mammographic Microcalcification Clusters. IEEE Trans Biomed Eng. 2015; 62: 1203-14. doi: 10.1109/TBME.2014.2385102

[13] Singh N, Marak J, Joshi P, Singh DK. Morphological and distribution pattern of calcifications on full field digital mammography versus digital breast tomosynthesis and comparison of diagnostic abilities of the two modalities: A retrospective study. J Clin of Diagn Res. 2023; 17(3): TC36-TC41. doi: 10.7860/JCDR/2023/55632.17675

[14] Mann RM, Athanasiou A, Baltzer PAT, *et al*. Breast cancer screening in women with extremely dense breasts recommendations of the European Society of Breast Imaging (EUSOBI). Eur Radiol. 2022; 32(6): 4036-45. doi: 10.1007/s00330-022-08617-6

[15] Edmonds CE, O'Brien SR, Conant EF, *et al*. Mammographic Breast Density: Current Assessment Methods, Clinical Implications, and Future Directions. Seminars in Ultrasound, CT, and MRI. Epub. 2022; 44: 35-45. doi: 10.1053/j.sult.2022.11.001

[16] Sickles EA, D'Orsi CJ, Bassett LW. American College of Radiology. ACR BI-RADS® Atlas, 5th ed. Reston, VA, USA, 2013.

[17] Gordon PB. The Impact of Dense Breasts on the Stage of Breast Cancer at Diagnosis: A Review and Options for Supplemental Screening. Curr Oncol. 2022; 17; 29(5): 3595-636. doi: 10.3390/curroncol29050291

[18] Bodewes FTH, van Asselt AA, Dorrius MD, Greuter MJW, de Bock GH. Mammographic breast density and the risk of breast cancer: A systematic review and meta-analysis. Breast. 2022; 66: 62-8. doi: 10.1016/j.breast.2022.09.007

[19] Sarvamangala DR, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. Evol Intell. 2022; 15(1): 1-22. doi: 10.1007/s12065-020-00540-3

[20] Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights Imaging. 2018; 9; 611-29. doi: 10.1007/s13244-018-0639-9

[21] Ahmed SF, Alam MSB, Hassan M, Rozbu MR, Ishtiak

T, Rafa N, *et al*. Deep learning modelling techniques: current progress, applications, advantages, and challenges. Artif Intell Rev. 2023; 56: 13521-617. doi: 10.1007/s10462-023-10466-8

[22] Aphinives C, Aphinives P, Nawapan S. Artificial Intelligence Development for Detecting Micro-calcification within Mammography. J Med Assoc Thai. 2021; 104(4): 560-4. doi: 10.1002/jmat.2021.104.560

[23] Intasam A, Promworn Y, Thanasitthichai S, Piyawattanametha W. A comparative study of convolutional neural networks for mammogram diagnosis. Proceedings of 14th the Biomedical Engineering International Conference (BMEiCON-2022); 2022 Nov 10-13; Songkhla, Thailand. p.1-4. doi: 10.1109/BMEiCON56653.202210012074. Available from: http://ieeexplore.ieee.org/document/10012074.

[24] Labcharoenwongs P, Vonganansup S, Chunhapran O, Noolek D, Yampaka T. An Automatic Breast Tumor Detection and Classification including Automatic Tumor Volume Estimation Using Deep Learning Technique. Asian Pac J Cancer Prev. 2023; 24(3): 1081-8. doi: 10.31557/APJCP.2023.24.3.1081

[25] Lehman CD, Yala A, Schuster T, Dontchos B, Bahl M, Swanson K, *et al*. Mammographic Breast Density Assessment Using Deep Learning: Clinical Implementation. Radiol. 2019; 290(1): 52-8. doi.org/10.1148/radiol.2018180694

[26] Sabani A, Landsmann A, Hejduk P, Schmidt C, Marcon M, Borkowski K, *et al*. BI-RADS-Based Classification of Mammographic Soft Tissue Opacities Using a Deep Convolutional Neural Network. Diagnostics. 2022; 12(7): 1564. doi: 10.3390/diagnostics12071564

[27] Zitnick CL, Dollár P. Edge Boxes: Locating Object Proposals from Edges. European Conference on Computer Vision. 2014: Part V, LNCS 8693 p.391-405.

Journal of Associated Medical Sciences 2024; 57(3): 18-29

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

School performance readiness of elementary students with disabilities before starting the occupational therapy program in special education schools

Suchitporn Lersilp Tep-aksorn Tipsut Phattariya Maboontham Rattanaporn Sonngai Kewalin Panyo*

**Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*

ARTICLE INFO

Article history:
Received 30 March 2024
Accepted as revised 28 May 2024
Available online 7 June 2024

Keywords:
Performance, readiness, elementary student, disabilities, special education school.

ABSTRACT

Background: School-based occupational therapists (SBOTs) work as health professionals in educational settings. School performance readiness is within the scope of SBOTs in providing a service for students who might have experienced decreasing performance during their school life.

Objective: This study aimed to explore the school performance readiness of elementary students with disabilities before starting the occupational therapy program in special education schools.

Materials and methods: The school performance readiness checklist for students with disabilities in special education schools was the research instrument. It comprised four areas: physical, social and emotional, pre-academic, and self-care readiness. Seventy-five elementary students with disabilities participated in this study. They consisted of 41 students with intellectual disability, 21 students with physical disability, and 13 students with sensory disability.

Results: Results from the initial semester in special education schools showed that most of the students with disabilities (85.33%) needed support in promoting their school performance readiness, particularly in pre-academic readiness. This included most of those with intellectual disability (33.33%) and all of those with sensory disabilities, while all of the students with physical disabilities needed support in promoting physical readiness.

Conclusion: Most students with disabilities needed support in promoting their school performance readiness according to their type of disability. Although the special education schools had enrolment criteria, SBOTs and school professionals should be concerned with providing related intervention programs to promote school readiness, particularly pre-academic readiness.

Introduction

In 2021, there were 139,640 school-aged children with disabilities in Thailand, which amounted to 8.94% of people or 3.84% of school-aged children with disabilities in the country.[1] Although the recent special education model drives towards inclusive education to encourage students with disabilities to gain a place in regular schools, most of them in Thailand enroll in special education schools. Educators and related professionals realize the benefits of inclusive education for students with disabilities. However, the inclusive system requires changes at all levels of society, including the school, community, and national levels, and more time is needed to prepare for these alterations, particularly in a developing country.[2] During a move to inclusive education, special

∗ **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: kewalin.panyo@cmu.ac.th
doi: 10.12982/JAMS.2024.043
E-ISSN: 2539-6056

education schools are necessary as a learning setting for students with disabilities in a country that faces barriers to change. Most students with disabilities who can participate in regular schools have high function or independent performances or invisible disabilities such as attention deficit and hyperactive disorder (ADHD), learning disorder (LD), high function autism spectrum disorder (ASD), and so forth.

On the other hand, students with obvious or severe disabilities are rarely accepted in regular schools. Therefore, special education schools are the leading choice for these students in Thailand, which includes those with intellectual, physical, visual, and hearing disabilities. Those who study in special education learn from a modified or parallel curriculum, providing more accessibility to participation in school learning activities. These learning activities are designed to consider the strengths and limitations of the students. Generally, the enrolment criteria in each special education school are based on the student's disability and not on their performance and skills. Before starting school, most students receive an early intervention program to promote development and prepare for the necessary performance in the school context.

Occupational therapy (OT) is a related service that supports the function and performance of students in their daily life activities and their engagement in human occupation.[3] School-based occupational therapists (SBOTs) are health professionals who work collaboratively with education professionals in providing services to students with disabilities in the school context. In Thailand, four types of special education schools follow the types of disabilities: intellectual, physical, visual, and hearing. Regarding elementary students with intellectual disability, the SBOTs focus on developmental stimulation, academic performance preparation, and the school curriculum.[4] Regarding elementary students with a physical disability, the SBOTs focus on the activities of daily living (ADL) training, rehabilitation, academic performance preparation, and the school curriculum.[5] Elementary students with visual and hearing disability are in the group of sensory disability that usually refers to impairment of the senses such as sight, hearing, taste, touch, smell, and spatial awareness. However, sensory disability mainly covers visual or blindness and hearing impairment or deafness.[6] The SBOTs focus on sensory stimulation and training for students with sensory disability by compensation to retain sensory functions for independent living in all everyday activities.[7,8]

In general, the SBOTs provide an intervention program that promotes the functional abilities and participation of the students in their daily routines by following the role of the student.[9] However, in cases of students needing to meet the minimal school criteria, the SBOTs are expected to evaluate and improve the required performances of the students as they progress through school life. Those students were provided with OT intervention programs for establishing new routines, developing new skills needed for independence in school, engaging in academic tasks, and participating in appropriate social interactions with others. These intervention programs for school performance readiness are for students who need extra support in meeting the minimal criteria to become elementary students. Moreover, these programs fulfil the required performance of students who experienced decreased performance during school.

School performance readiness is a crucial indicator that predicts the potential and achievement of children when they start as students. This includes academic performance and all developmental performances in physical, cognitive, social, and emotional areas.[10,11] Most students with disabilities receive the early intervention program before entering school. However, some of them need related services at the beginning and during their school life. In Thailand, occupational therapists are related to health professionals who provide services in early intervention programs and school settings. Unfortunately, by the time students with disabilities move to schools, the outcome of the early intervention program in school performance readiness has been hardly investigated.

Therefore, this study aimed to explore the school performance readiness of elementary students with disabilities before starting the OT program in special education schools. The results could provide valuable information that reflects the plans of occupational therapists in the early intervention program for promoting school performance readiness of students with disabilities in each type of special education school.

Materials and methods

This descriptive study explored the school performance readiness of elementary students with disabilities before they started the OT program in special education schools in Chiang Mai province, Thailand. These schools consisted of students with physical, visual, hearing, and intellectual disabilities, and they provided education and related services for children living in upper northern Thailand. Each school had only one occupational therapist. A total of four SBOTs evaluated their students by using the school performance readiness checklist, which was a research instrument in this study.

Participants

Seventy-five participants comprised 41 students with intellectual disability, 21 students with physical disability, 13 students with sensory disability, 5 students with visual disability, and 8 students with hearing disability. Before entering the schools, these participants were screened informally by the school professional team, who showed the need for OT services. Most participants were male (66.67%, N=50) and aged 8.0-8.11 years (38.67%, N=29). The demographics of the participants in this study are shown in Table 1.

Table 1. Demographics of the participants (N=75)

Characteristics	Intellectual disability N (%)	Physical disability N (%)	Sensory disability N (%)	Total N (%)
Gender				
Male	31 (41.34)	13 (17.33)	6 (8.00)	50 (66.67)
Female	10 (13.34)	8 (10.66)	7 (9.33)	25 (33.33)
Age (years old)				
6.0-6.11	0 (0.00)	0 (0.00)	3 (4.00)	2 (2.67)
7.0-7.11	7 (9.33)	5 (6.67)	4 (5.33)	15 (20.00)
8.0-8.11	21 (28.01)	4 (5.33)	2 (2.67)	29 (38.67)
9.0-9.11	9 (12.00)	10 (13.34)	2 (2.67)	21 (28.01)
10.0-10.11	4 (5.33)	2 (2.67)	2 (2.67)	8 (10.66)

Instrument

The school performance readiness checklist of students with disabilities in special education schools was a research instrument in this study. It was applied by the normal development of children and the School Readiness Checklist,[12] which comprised four areas of school performance readiness such as physical, social-emotional, pre-academic, and self-care readiness. It was examined for content validity and reliability. Content validity was examined by five related specialists, occupational therapists, and education professionals in the early intervention program and school contexts. The first draft of the checklist consisted of 71 items. After consideration by the specialists, the item-object congruence (IOC) index was calculated to be between 0.40 and 1.00. Suggestions from the specialists were used to correct the checklist before resending it to the specialists. After that, 62 items remained on the checklist with the IOC=1.00. Regarding reliability, the checklist was administered to 12 elementary students with disabilities. Cronbach's alpha coefficient indicated α=0.96, which presented excellent reliability.[13] The checklist items were on a nominal scale, with scores of 1 and 0, in which 1 (Yes) meant the present performance of the students in those items, and 0 (No) meant no present performance. If a student had no opportunity to perform an item, those items were recorded as NA and not included in the scoring. Total scores of items performed and total scores of all items, without the NA ones, were calculated as the performance percentage, as shown in the formula below.

$$\text{Performance percentage} = \frac{\text{Total scores of performed items}}{\text{Total scores of all items without the NA items}} \times 100$$

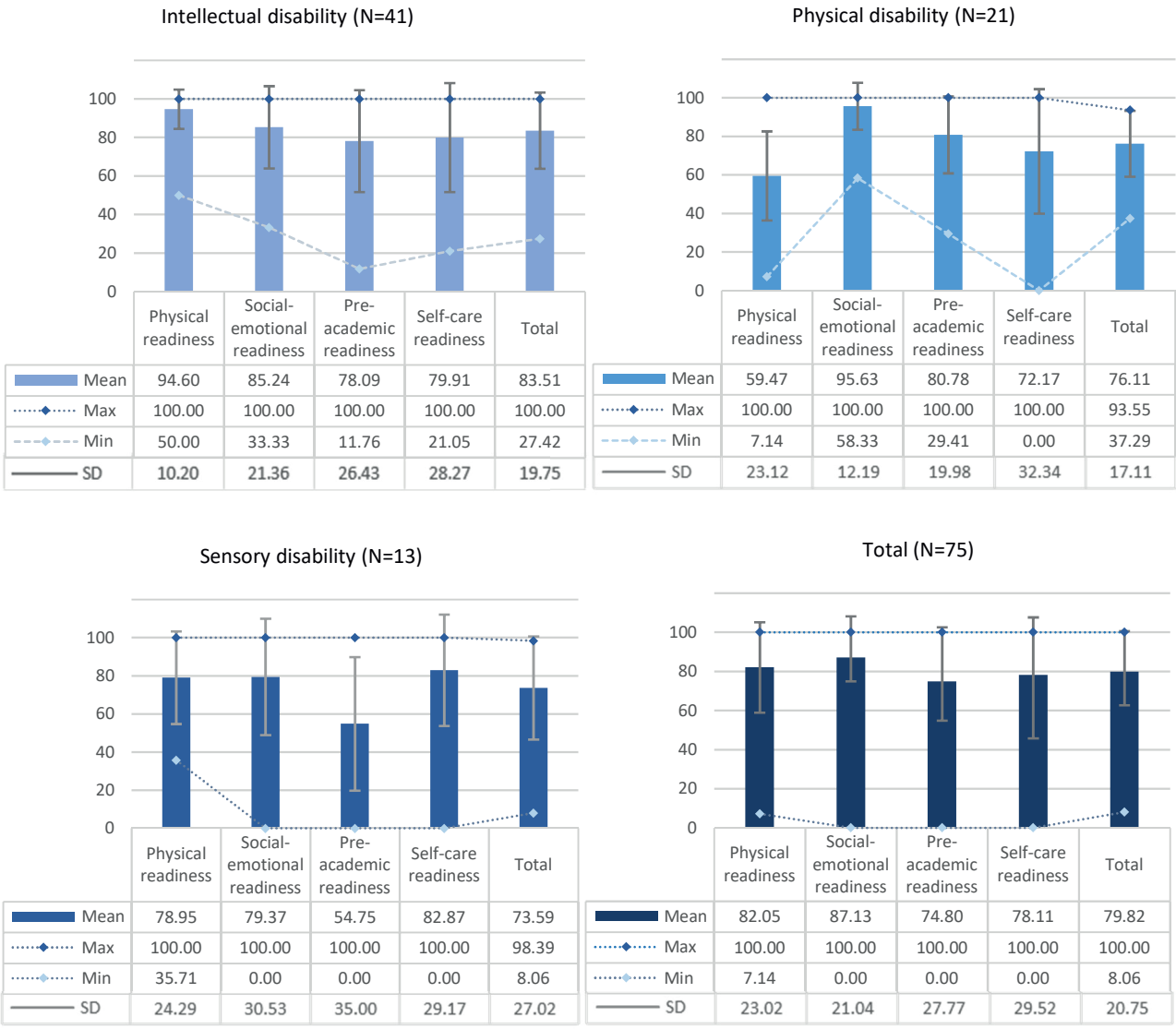
The interpretation of school performance readiness was based on the performance percentage of each item and overall items, but it did not include the NA items. All performance items were considered as minimally required performance for elementary school students. Thus, if any items were checked "no," the child needed support. The students who were able to perform in all of the items were marked for reaching school performance readiness. On the other hand, those unable to get the total performance percentage of each item and overall items were seen to need support.

Statistical analysis

Demographic data, school performance readiness, and the number of students with disabilities were analyzed for each item using descriptive statistics, including frequency and percentage. The percentage of school performance readiness was analyzed using descriptive statistics, including maximum, minimum, mean, and standard deviation.

Results

Results indicated that the minimum, maximum, and average performance percentage of school performance readiness in students with disabilities was 8.06, 100.00, and 79.82$\pm$20.75, respectively. Regarding the type of school performance readiness, social-emotional readiness showed the highest performance percentage, followed by physical, self-care, and pre-academic readiness in that order. When considering each type of student with disabilities, the results showed that those with intellectual disability presented the highest performance percentage in physical readiness. The students with physical disability and those with sensory disability presented the highest performance percentages in social-emotional readiness and self-care readiness, respectively. Details of the school performance readiness percentage in each type of student with disabilities are shown in Figure 1.



Intellectual disability (N=41)	Physical readiness	Social-emotional readiness	Pre-academic readiness	Self-care readiness	Total
Mean	94.60	85.24	78.09	79.91	83.51
Max	100.00	100.00	100.00	100.00	100.00
Min	50.00	33.33	11.76	21.05	27.42
SD	10.20	21.36	26.43	28.27	19.75

Physical disability (N=21)	Physical readiness	Social-emotional readiness	Pre-academic readiness	Self-care readiness	Total
Mean	59.47	95.63	80.78	72.17	76.11
Max	100.00	100.00	100.00	100.00	93.55
Min	7.14	58.33	29.41	0.00	37.29
SD	23.12	12.19	19.98	32.34	17.11

Sensory disability (N=13)	Physical readiness	Social-emotional readiness	Pre-academic readiness	Self-care readiness	Total
Mean	78.95	79.37	54.75	82.87	73.59
Max	100.00	100.00	100.00	100.00	98.39
Min	35.71	0.00	0.00	0.00	8.06
SD	24.29	30.53	35.00	29.17	27.02

Total (N=75)	Physical readiness	Social-emotional readiness	Pre-academic readiness	Self-care readiness	Total
Mean	82.05	87.13	74.80	78.11	79.82
Max	100.00	100.00	100.00	100.00	100.00
Min	7.14	0.00	0.00	0.00	8.06
SD	23.02	21.04	27.77	29.52	20.75

Figure 1. *Performance percentage of school performance readiness.*

Regarding interpreting school performance readiness, results showed that during the initial semester in special education schools, most students with disabilities (85.33%) needed support in promoting their school performance readiness, especially pre-academic readiness. On the other hand, 14.67% of the students did not need the OT program to facilitate their school performance readiness. When considering each type of student with disabilities, the results showed that most students with intellectual disability (33.33%) needed support in promoting their pre-academic readiness. In comparison, most of them (37.33%) had completed physical readiness. All students with physical disability needed support in promoting their school performance readiness, particularly physical readiness. All students with sensory disability needed support in promoting their school performance readiness, especially pre-academic readiness. Details of school performance readiness of the students with disabilities are shown in Figure 2.

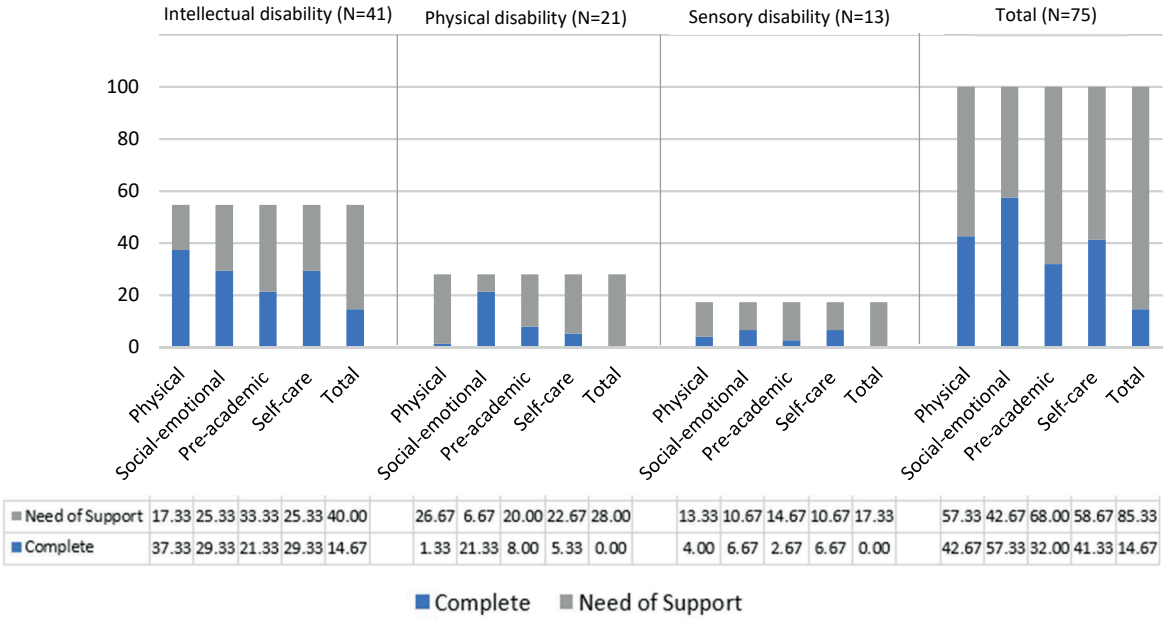


Figure 2. *Percentage of students with disabilities together with school performance readiness.*

When considering the performance items, in terms of physical readiness, most of the students (44.00%) could not perform "use scissors correctly", including those with intellectual disability (31.71%), while most of the students with sensory disabilities (30.00%) could not perform "Operate eye-hand activities coordinately". In addition, most of the students with a physical disability (80.95%) could not perform "grasp objects dexterously." In terms of social-emotional readiness, most of the students (22.67%) could not perform "control aggressive behaviors when facing unsatisfactory situations", including those with intellectual (31.71%) and sensory disabilities (23.08%). In comparison, most of the students with a physical disability (9.52%) could not perform "keep toys after play", "be patient by waiting", and "adapt to an unfamiliar environment easily". In terms of pre-academic readiness, most of the students (36.00%) could not perform "solve problems with age appropriately" and "understand the concept of numbers with age appropriately," including those with a physical disability (52.38%). In addition, most of the students with intellectual disability (34.15%) were unable to perform "solve problems with age appropriately", "perceive person, time, and place orientation", and "give directions and locations of places". Meanwhile, most of the students with sensory disability (46.15%) could not perform "give directions and locations of places", and "inform about the usefulness of objects in daily life". In terms of self-care readiness, most of the students (38.67%) could not perform "button up independently," including those with a physical disability (61.90%). Moreover, most of the students with intellectual disability (36.58%) could not perform "maintain personal devices", "being aware of the danger in daily activities", and "making the bed after getting up". In comparison, most of those with sensory disability (38.46%) were unable to perform "being aware of the danger in daily activities." Details of the number of participants in each item of school performance readiness are shown in Table 2.

Table 2. Number and Percentage of participants in each item of school performance readiness

School performance readiness	Intellectual disability (N=41)			Physical disability (N=21)			Sensory disability (N=13)			**Total** (N=75)		
	Yes	No	NA	Yes	No	NA	Yes	No	NA	Yes	No	NA
Physical Readiness												
1. Control upright head and neck while doing activities	41 (100.00)	0 (0.00)	0 (0.00)	21 (100.00)	0 (0.00)	0 (0.00)	13 (100.00)	0 (0.00)	0 (0.00)	75 (100.00)	0 (0.00)	0 (0.00)
2. Sit steadily on floor and chair	41 (100.00)	0 (0.00)	0 (0.00)	17 (80.95)	4 (19.05)	0 (0.00)	13 (100.00)	0 (0.00)	0 (0.00)	71 (94.67)	4 (5.33)	0 (0.00)
3. Keep balance when facing an external force	39 (95.12)	2 (4.88)	0 (0.00)	7 (33.33)	14 (66.67)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	58 (77.33)	17 (22.67)	0 (0.00)
4. Move actively without fatigue	41 (100.00)	0 (0.00)	0 (0.00)	6 (28.57)	15 (71.43)	0 (0.00)	10 (76.92)	3 (23.08)	0 (0.00)	57 (76.00)	18 (24.00)	0 (0.00)
5. Be mobile independently	41 (100.00)	0 (0.00)	0 (0.00)	17 (80.95)	4 (19.05)	0 (0.00)	13 (100.00)	0 (0.00)	0 (0.00)	71 (94.67)	4 (5.33)	0 (0.00)
6. Reach out, grasp, carry and release voluntarily	40 (97.56)	1 (2.44)	0 (0.00)	20 (95.24)	1 (4.76)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	72 (96.00)	3 (4.00)	0 (0.00)
7. Grasp objects steadily	40 (97.56)	1 (2.44)	0 (0.00)	15 (71.43)	6 (28.57)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	67 (89.33)	8 (10.67)	0 (0.00)
8. Grasp objects dexterously	37 (90.24)	4 (9.76)	0 (0.00)	4 (19.05)	**17 (80.95)**	0 (0.00)	11 (84.62)	2 (15.38)	0 (0.00)	52 (69.33)	23 (30.67)	0 (0.00)
9. Use bilateral hand coordinately	41 (100.00)	0 (0.00)	0 (0.00)	16 (76.19)	5 (23.81)	0 (0.00)	11 (84.62)	2 (15.38)	0 (0.00)	68 (90.67)	7 (9.33)	0 (0.00)
10. Grasp stationery with age appropriately	37 (90.24)	4 (9.76)	0 (0.00)	11 (52.38)	10 (47.62)	0 (0.00)	8 (61.54)	3 (23.08)	2 (15.38)	56 (74.67)	17 (22.66)	2 (2.67)
11. Transfer objects from one hand to the other	41 (100.00)	0 (0.00)	0 (0.00)	17 (80.95)	4 (19.05)	0 (0.00)	10 (76.92)	3 (23.08)	0 (0.00)	68 (90.67)	7 (9.33)	0 (0.00)
12. Grasp and rotate objects with one hand	37 (90.24)	4 (9.76)	0 (0.00)	9 (42.86)	12 (57.14)	0 (0.00)	10 (76.92)	3 (23.08)	0 (0.00)	56 (74.67)	19 (25.33)	0 (0.00)
13. Use scissors correctly	28 (68.29)	**13 (31.71)**	0 (0.00)	5 (23.81)	16 (76.19)	3 (10.00)	8 (61.54)	4 (30.77)	1 (7.69)	41 (54.67)	**33 (44.00)**	4 (5.33)
14. Operate eye-hand activities coordinately	39 (95.12)	2 (4.88)	0 (0.00)	15 (71.43)	6 (28.57)	0 (0.00)	8 (61.54)	**5 (38.46)**	0 (0.00)	62 (82.67)	13 (17.33)	0 (0.00)

Table 2. Number and Percentage of participants in each item of school performance readiness (cont.)

School performance readiness	Intellectual disability (N=41)			Physical disability (N=21)			Sensory disability (N=13)			**Total** (N=75)		
	Yes	No	NA	Yes	No	NA	Yes	No	NA	Yes	No	NA
Social-emotional Readiness												
1. Share things with peers	35 (85.37)	6 (14.63)	0 (0.00)	21 (100.00)	0 (0.00)	0 (0.00)	11 (84.62)	2 (15.38)	0 (0.00)	67 (89.33)	8 (10.67)	0 (0.00)
2. Control emotion appropriately	29 (70.73)	12 (29.27)	0 (0.00)	20 (95.24)	1 (4.76)	0 (0.00)	10 (76.92)	**3 (23.08)**	0 (0.00)	59 (78.67)	16 (21.33)	0 (0.00)
3. Keep toys after play	34 (82.93)	7 (17.07)	0 (0.00)	19 (90.48)	**2 (9.52)**	0 (0.00)	10 (76.92)	**3 (23.08)**	0 (0.00)	63 (84.00)	12 (16.00)	0 (0.00)
4. Be patient by waiting	29 (70.73)	12 (29.27)	0 (0.00)	19 (90.48)	**2 (9.52)**	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	60 (80.00)	15 (20.00)	0 (0.00)
5. Follow class rules	34 (82.93)	7 (17.07)	0 (0.00)	20 (95.24)	1 (4.76)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	66 (88.00)	9 (12.00)	0 (0.00)
6. Greet others	38 (92.68)	2 (4.88)	1 (2.44)	21 (100.00)	0 (0.00)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	71 (94.67)	3 (4.00)	1 (1.33)
7. Make and maintain relationships with friends	39 (95.12)	2 (4.88)	0 (0.00)	21 (100.00)	0 (0.00)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	72 (96.00)	3 (4.00)	0 (0.00)
8. Change or move from one activity to another smoothly	40 (97.56)	1 (2.44)	0 (0.00)	20 (95.24)	1 (4.76)	0 (0.00)	11 (84.62)	2 (15.38)	0 (0.00)	71 (94.67)	4 (5.33)	0 (0.00)
9. Adapt to an unfamiliar environment easily	40 (97.56)	1 (2.44)	0 (0.00)	19 (90.48)	**2 (9.52)**	0 (0.00)	11 (84.62)	1 (7.69)	1 (7.69)	70 (93.34)	4 (5.33)	1 (1.33)
10. Repent for mistakes	41 (100.00)	0 (0.00)	0 (0.00)	21 (100.00)	0 (0.00)	0 (0.00)	9 (69.23)	**3 (23.08)**	1 7.69)	71 (94.67)	3 (4.00)	1 (1.33)
11. Collaborate with group activities	41 (100.00)	0 (0.00)	0 (0.00)	20 (95.24)	1 (4.76)	0 (0.00)	11 (84.62)	2 (15.38)	0 (0.00)	72 (96.00)	3 (4.00)	0 (0.00)
12. Control aggressive behaviors when facing unsatisfactory situations	27 (65.85)	**13 (31.71)**	1 (2.44)	20 (95.24)	1 (4.76)	0 (0.00)	10 (76.92)	**3 (23.08)**	0 (0.00)	57 (76.00)	**17 (22.67)**	1 (1.33)

Table 2. Number and Percentage of participants in each item of school performance readiness (cont.)

School performance readiness	Intellectual disability (N=41)			Physical disability (N=21)			Sensory disability (N=13)			**Total** (N=75)		
	Yes	No	NA	Yes	No	NA	Yes	No	NA	Yes	No	NA
Pre-academic Readiness												
1. Understand and express needs by natural gestures	39 (95.12)	2 (4.88)	0 (0.00)	21 (100.00)	0 (0.00)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	72 (96.00)	3 (4.00)	0 (0.00)
2. Understand verbal, non-verbal or sign language	39 (95.12)	2 (4.88)	0 (0.00)	21 (100.00)	0 (0.00)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	72 (96.00)	3 (4.00)	0 (0.00)
3. Use two-way communication with verbal, non-verbal or sign language	38 (92.68)	3 (7.32)	0 (0.00)	21 (100.00)	0 (0.00)	0 (0.00)	9 (69.23)	4 (30.77)	0 (0.00)	68 (90.67)	7 (9.33)	0 (0.00)
4. Understand age-appropriated vocabularies	36 (87.80)	5 (12.20)	0 (0.00)	17 (80.95)	4 (19.05)	0 (0.00)	8 (61.54)	5 (38.46)	0 (0.00)	61 (81.33)	14 (18.67)	0 (0.00)
5. Tell stories continuously	31 (75.61)	10 (24.39)	0 (0.00)	14 (66.67)	6 (28.57)	1 (4.76)	8 (61.54)	5 (38.46)	0 (0.00)	53 (70.67)	21 (28.00)	1 (1.33)
6. Optimal level of arousal	38 (92.68)	3 (7.32)	0 (0.00)	20 (95.24)	1 (4.76)	0 (0.00)	10 (76.92)	3 (23.08)	0 (0.00)	68 (90.67)	7 (9.33)	0 (0.00)
7. Optimal attention and concentration	36 (87.80)	5 (12.20)	0 (0.00)	14 (66.67)	7 (33.33)	0 (0.00)	9 (69.23)	4 (30.77)	0 (0.00)	59 (78.67)	16 (21.33)	0 (0.00)
8. Able to finish all assignments	35 (85.37)	6 (14.63)	0 (0.00)	13 (61.90)	8 (38.10)	0 (0.00)	9 (69.23)	4 (30.77)	0 (0.00)	57 (76.00)	18 (24.00)	0 (0.00)
9. Tell his/her name correctly	36 (87.80)	5 (12.20)	0 (0.00)	20 (95.24)	1 (4.76)	0 (0.00)	10 (76.92)	3 (23.08)	0 (0.00)	66 (88.00)	9 (12.00)	0 (0.00)
10. Be flexible in uncertain situations	34 (82.93)	7 (17.07)	0 (0.00)	20 (95.24)	1 (4.76)	0 (0.00)	10 (76.92)	3 (23.08)	0 (0.00)	64 (85.33)	11 (14.67)	0 (0.00)
11. Solve problems with age appropriately	27 (65.85)	**14 (34.15)**	0 (0.00)	10 (47.62)	**11 (52.38)**	0 (0.00)	11 (84.62)	2 (15.38)	0 (0.00)	48 (64.00)	**27 (36.00)**	0 (0.00)
12. Understand concept of numbers with age appropriately	29 (70.73)	11 (26.83)	1 (2.44)	10 (47.62)	**11 (52.38)**	0 (0.00)	8 (61.54)	5 (38.46)	0 (0.00)	47 (62.67)	**27 (36.00)**	1 (1.33)
13. Compare and analyze similarity and difference of tangible objects	28 (68.29)	12 (29.27)	1 (2.44)	20 (95.24)	1 (4.76)	0 (0.00)	9 (69.23)	4 (30.77)	0 (0.00)	57 (76.00)	17 (22.67)	1 (1.33)
14. Compare and analyze similarity and difference of intangible objects	29 (70.73)	11 (26.83)	1 (2.44)	15 (71.43)	6 (28.57)	0 (0.00)	10 (76.92)	3 (23.08)	0 (0.00)	54 (72.00)	20 (26.67)	1 (1.33)
15. Perceive person, time, and place orientation	26 (63.41)	**14 (34.15)**	1 (2.44)	20 (95.24)	1 (4.76)	0 (0.00)	9 (69.23)	4 (30.77)	0 (0.00)	55 (73.33)	19 (25.33)	1 (1.33)
16. Give directions and locations of places	26 (63.41)	**14 (34.15)**	1 (2.44)	20 (95.24)	1 (4.76)	0 (0.00)	7 (53.85)	**6 (46.15)**	0 (0.00)	53 (70.67)	21 (28.00)	1 (1.33)
17. Inform about the usefulness of objects in daily life	27 (65.85)	13 (31.71)	1 (2.44)	16 (76.19)	5 (23.81)	0 (0.00)	7 (53.85)	**6 (46.15)**	0 (0.00)	50 (66.67)	24 (32.00)	1 (1.33)

Table 2. Number and Percentage of participants in each item of school performance readiness (cont.)

School performance readiness	Intellectual disability (N=41)			Physical disability (N=21)			Sensory disability (N=13)			**Total** (N=75)		
	Yes	No	NA	Yes	No	NA	Yes	No	NA	Yes	No	NA
Self-care Readiness												
1. Face washing independently	41 (100.00)	0 (0.00)	0 (0.00)	20 (95.24)	1 (4.76)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	73 (97.33)	2 (2.67)	0 (0.00)
2. Tooth brushing independently	39 (95.12)	2 (4.88)	0 (0.00)	18 (85.71)	3 (14.29)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	69 (92.00)	6 (8.00)	0 (0.00)
3. Combing hair independently	40 (97.56)	1 (2.44)	0 (0.00)	19 (90.48)	2 (9.52)	0 (0.00)	11 (84.62)	1 (7.69)	1 (7.69)	70 (93.33)	4 (5.33)	1 (1.33)
4. Eating independently	39 (95.12)	2 (4.88)	0 (0.00)	20 (95.24)	1 (4.76)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	71 (94.67)	4 (5.33)	0 (0.00)
5. Toileting independently	31 (75.61)	10 (24.39)	0 (0.00)	13 (61.90)	8 (38.10)	1 (4.76)	11 (84.62)	2 (15.38)	0 (0.00)	55 (73.33)	20 (26.67)	0 (0.00)
6. Cleaning after toileting independently	31 (75.61)	10 (24.39)	0 (0.00)	19 (90.48)	2 (9.52)	0 (0.00)	11 (84.62)	2 (15.38)	0 (0.00)	61 (81.33)	14 (18.67)	0 (0.00)
7. Putting on clothes independently	35 (85.37)	6 (14.63)	0 (0.00)	15 (71.43)	6 (28.57)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	62 (82.67)	13 (17.33)	0 (0.00)
8. Taking-off clothes independently	35 (85.37)	6 (14.63)	0 (0.00)	18 (85.71)	3 (14.29)	0 (0.00)	11 (84.62)	2 (15.38)	0 (0.00)	64 (85.33)	11 (14.67)	0 (0.00)
9. Zipping clothes independently	31 (75.61)	10 (24.39)	0 (0.00)	14 (66.67)	7 (33.33)	0 (0.00)	11 (84.62)	2 (15.38)	0 (0.00)	56 (74.67)	19 (25.33)	0 (0.00)
10. Button up independently	28 (68.29)	13 (31.71)	0 (0.00)	8 (38.10)	**13 (61.90)**	0 (0.00)	10 (76.92)	3 (23.08)	0 (0.00)	46 (61.33)	**29 (38.67)**	0 (0.00)
11. Unbutton independently	29 (70.73)	12 (29.77)	0 (0.00)	18 (88.71)	3 (14.29)	0 (0.00)	11 (84.62)	2 (15.38)	0 (0.00)	58 (77.33)	17 (22.67)	0 (0.00)
12. Putting on socks independently	31 (75.61)	10 (24.39)	0 (0.00)	15 (71.43)	6 (28.57)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	58 (77.33)	17 (22.67)	0 (0.00)
13. Taking-off socks independently	37 (90.24)	4 (9.76)	0 (0.00)	18 (85.71)	3 (14.29)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	67 (89.33)	8 (10.67)	0 (0.00)
14. Putting on shoes independently	34 (82.93)	7 (17.07)	0 (0.00)	15 (71.43)	6 (28.57)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	61 (81.33)	14 (18.67)	0 (0.00)
15. Taking-off shoes independently	37 (90.24)	4 (9.76)	0 (0.00)	18 (85.71)	3 (14.29)	0 (0.00)	12 (92.31)	1 (7.69)	0 (0.00)	67 (89.33)	8 (10.67)	0 (0.00)
16. Maintaining personal devices	25 (60.98)	**15 (36.58)**	1 (2.44)	12 (57.14)	9 (42.86)	0 (0.00)	10 (76.92)	3 (23.08)	0 (0.00)	47 (62.67)	27 (36.00)	1 (1.33)
17. Being aware of danger in daily activities	23 (56.10)	**15 (36.58)**	3 (7.32)	19 (90.48)	2 (9.52)	0 (0.00)	8 (61.54)	**5 (38.46)**	0 (0.00)	50 (66.67)	22 (29.33)	3 (4.00)
18. Being able to do basic chores	26 (63.41)	13 (31.71)	2 (4.88)	9 (42.86)	11 (52.38)	1 (4.76)	11 (84.62)	2 (15.38)	0 (0.00)	46 (61.33)	26 (34.67)	3 (4.00)
19. Making the bed after getting up	24 (58.54)	**15 (36.58)**	2 (4.88)	16 (76.19)	5 (23.81)	0 (0.00)	9 (69.23)	3 (23.08)	1 (7.69)	49 (65.33)	23 (30.67)	3 (4.00)

Discussion

Generally, related federal laws address the rights of children with disabilities to access educational participation in the school system. However, due to the limitation of their disability status, they need support in achieving school performance readiness. As the students with disabilities in this study were screened informally by the school professional team before entering school, it was indicated that the students needed OT services and that some of them showed essential readiness to study in the school system as elementary students.

Although the results showed that students with intellectual disability had the highest performance percentage in physical readiness, they also had the lowest in pre-academic readiness. Those students had a neurodevelopmental condition that brings about cognitive and functional performance issues, including language learning, imitation, symbolic activities, reasoning, problem-solving, planning, abstract thinking, judgment, generalization, and adaptive functioning.[14] These performances might be essential capacities that contribute to developing students' academic potential with intellectual disability.[14-16]

In terms of students with physical disability, although they had high performance in social-emotional readiness, they faced barriers when performing related physical activities due to health conditions and severity of physical impairment. These findings are noteworthy because students with physical disability spend a significant portion of their lives in the school environment. Previous studies indicated that environmental issues in the school context could affect students with physical disability in active participation.[17,18] For this reason, the environment for such students in special education schools was designed universally so that they could participate in school activities independently. Moreover, they could participate in group activities such as camping, dancing, field trips, etc. This support could encourage their social-emotional readiness, even though they had physical limitations. This finding related to a previous study in South Africa, in which the government released policies that provided equal opportunity for students with disabilities to participate in sports and recreational activities as much as their non-disabled peers.[19] Social-emotional development took a complicated path for students learning their occupation, including studying in an educational setting and developing skills in daily life.[20] Promoting social-emotional skills for students with disabilities should begin in the preschool and elementary years, particularly for initiatives and social problem-solving. These skills are associated with favorable long-term outcomes that are important in encouraging the development of self-determination skills for the students.[21,22] Therefore, health professionals might be concerned about addressing appropriate support in the school environment; for instance, environmental modification should focus on improving specific areas of risk or enhancing areas of competence in the educational setting.

Regarding students with sensory disability, the results indicated they had the highest performance percentage in self-care readiness and the lowest in pre-academic readiness. The severity of disability might impact the learning ability of these students. This is because the awareness of sensory input was crucial for interpretation and response to surrounding information.[23] The students with sensory impairments such as hearing and visual disabilities often had difficulty in perceiving and learning lessons in the classroom due to their surrounding sensory inputs. They would have limitations in learning that might be caused by delayed language development when communicating with other people, which contributed to learning problems and poor academic achievement.[24] Indeed, they might need to achieve successful pre-academic readiness.[25] Therefore, they needed OT intervention to promote their academic performances and related skills. In addition, related professionals in school settings should focus on providing services in early intervention programs and promoting continual pre-academic readiness.[26-27] In other words, the students with hearing and visual disabilities had unique challenges. However, appropriate preparation of school readiness skills for the kindergarten or early intervention period could bring about their future academic success.[28]

Although special education schools for students with disabilities had enrolment criteria, most students needed support in reaching school performance readiness before starting elementary school. This finding related to a previous study, which indicated that school readiness of children with disabilities has significantly less likely foundational reading and numeracy skills when compared to children without disabilities, particularly in low- and middle-income countries.[29] Moreover, this finding could act as valuable information for SBOTs and school professionals for planning and providing related intervention programs that promote school performance readiness according to the student's disabilities. This is because each type of student with disabilities has different preparation needs that relate to a previous study, which explained why the effects of treatment varied by the diagnosis and context of the children.[30] This finding significantly impacts OT and related service provision programs before entering the school context for children with disabilities. This transitional period of school performance readiness needs to be promoted to children with disabilities to improve their pre-academic and fundamental skills in achieving student life in a school setting.[31]

Limitation

This study used an exploratory research methodology to explain the school performance readiness of elementary students with disabilities before starting the occupational therapy program in special education schools. These findings revealed a broad perspective that could not be explained in deep perspectives. Therefore, these findings might be used in further study to develop potential research questions and generate hypotheses using the inferential statistical method. Another limitation of this study was the sample size. This study was performed in special education schools in upper northern Thailand. Thus, the number of

schools and students needed to be increased. The sample size was small, and it was difficult to analyze each type of student, especially those with sensory disabilities. For future research, more students with sensory disabilities could be recruited, which might reveal their similarities to and differences from students with visual and hearing disabilities and clarify more understanding. In addition, further research might expand to other areas of Thailand. Also, the results of this study were in the context of special education schools. Therefore, future research could develop a specific research instrument for students in the context of mainstream or regular schools. Furthermore, before elementary school, comparing school performance readiness between students in special education schools and those in mainstream or regular schools would be interesting in future research.

Conclusion

The students with disabilities in special education schools were perceived as potential learners. They could perform their age-appropriated occupations when they received opportunities and related services. However, due to their disabilities, they had activity limitations and participation restrictions in their role as students, especially during the transition period from the early intervention setting to an elementary school. For this reason, they needed OT services to enhance school performance and readiness in the school system and their school life.

Conflict of interest

The author(s) declared no potential conflicts of interest concerning this article's research, authorship, and publication.

Ethical approval

This study was approved by the Research Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University, Thailand (AMSEC-63EX-044).

References

[1] Office of the National Economic and Social Development Council. National report of persons with disabilities in Thailand. 2021 [cited 2023 Mar 21]. Available from: https://opendata.nesdc.go.th/en/dataset/situation-report-on-dep.

[2] UNICEF. Inclusive education. 2023 [cited 2023 August 9]. Available from: https://www.unicef.org/education/inclusive-education.

[3] American Occupational Therapy Association. Occupational therapy practice framework: domain and process. Am J Occup Ther. 2002; 56(6): 609-39. doi: 10.5014/ajot.56.6.609.

[4] Kawila Anukul School. Mission and target. 2023 [cited 2024 April 29]. Available from: http://www.kawila-anukul.ac.th/datashow_11483. [in Thai]

[5] Sri Sangwan School. Mission and target. 2024 [cited 2024 April 29]. Available from: http://www.swn.ac.th/datashow_67892. [in Thai]

[6] Abdullah N, Low KEY, Feng Q. Sensory Disability. In: Gu D, Dupre ME, editors. Encyclopedia of gerontology and population aging. Springer; 2021. [cited 2024 April 29]. Available from: doi.org/10.1007/978-3-030-22009-9_480.

[7] Kesaraksha W. Occupational therapy guideline for children with visual impairment. 2019 [cited 2024 April 29]. Available from: https://anyflip.com/ugfmx/mfmf. [in Thai]

[8] Phetchabun School for the Deaf. Information 2021. 2021 [cited 2024 April 29]. Available from: http://sotphetchabun.com/pdf/Information.pdf. [in Thai]

[9] Panyo K, Lersilp S, Putthinoi S, Hsu HY. Occupational therapy service during transitional periods in special education schools. J Occup Ther Sch Early Interv. 2023; 16(1): 91-105. doi: 10.1080/19411243.2021.2009082.

[10] Dangol R, Shrestha M. Learning readiness and educational achievement among school students. Int J Indian Psychol. 2019; 7(12): 467-76. doi: 10.25215/0702.056.

[11] University of California. School readiness and health. Childcare Health Program; 2006.

[12] Your Kids OT. School readiness checklist. 2023 [cited 2023 Oct 20]. Available from: https://www.yourkidsot.com/store/p80/FREE_School_Readiness_Checklist_v2_2023.html.

[13] Konting MM, Kamaruddin N, Man NA. Quality assurance in higher education institutions: exit survey among Universiti Putra Malaysia graduating students. Int Educ Stud. 2009; 2(1): 25-31. doi: 10.5539/IES.V2N1P25.

[14] Sajewicz-Radtke U, Jurek P, Olech M, Łada-Maśko AB, Jankowska AM, Radtke BM. Heterogeneity of cognitive profiles in children and adolescents with mild intellectual disability (MID). Int J Environ Res Public Health. 2022; 19(12): 1-12. doi: 10.3390/ijerph19127230.

[15] Adrien JL, Taupiac E, Thiébaut E, Paulais MA, Van-Gils J, Kaye K, Blanc R, Gattegno MP, Contejean Y, Michel G, Dean A, Barthélémy C, Lacombe D. A comparative study of cognitive and socio-emotional development in children with Rubinstein-Taybi syndrome and children with autism spectrum disorder associated with a severe intellectual disability, and in young typically developing children with matched developmental ages. Res Dev Disabil. 2021; 116: 1-12. doi: 10.1016/j.ridd.2021.104029.

[16] Echavarría-Ramírez LM, Tirapu-Ustárroz J. Neuropsychological examination in children with intellectual disabilities. Rev Neurol. 2021; 73(2): 66-76. doi: 10.33588/rn.7302.202102.

[17] Lersilp S, Putthinoi S, Lersilp T. Facilitators and barriers of assistive technology and learning environment for children with special needs. Occup Ther Int. 2018; 2018(3705946): 1-9. doi: 10.33588/rn.7302.202102.

[18] Egilson ST, Traustadottir R. Participation of students with physical disabilities in the school environment. Am J Occup Ther. 2009; 63(3): 2664-72. doi: 10.5014/ajot.63.3.264.

[19] Chiwandire D. Students with disabilities' lack of opportunity for sport and recreational activities: the case of South African universities. In: Falola T, Hamel N, editors. Disability in Africa: inclusion, care, and the ethics of humanity. Boydell & Brewer; 2018: 361-88.

[20] Besi M, Sakellariou M. Transition to primary school the importance of social skills. Int J Humanit Soc Sci. 2019; 6(1): 33-6. doi: 10.14445/23942703/IJHSS-V6I1P107.

[21] Coster WJ, Haltiwanger JT. Social-behavioral skills of elementary students with physical disabilities included in general education classrooms. Remedial Spec Educ. 2004; 25(2): 95-103. doi: 10.1177/07419325040250020401.

[22] Larcombe TJ, Joosten AV, Cordier R, Vaz S. Preparing children with autism for transition to mainstream school and perspectives on supporting positive school experiences. J Autism Dev Disord. 2019; 49(8): 3073-88. doi: 10.1007/s10803-019-04022-z.

[23] Benson, JD, Breisinger E, Roach M. Sensory-based intervention in the schools: a survey of occupational therapy practitioners. J Occup Ther Sch Early Interv. 2019;12(1):115-28.doi:10.1080/19411243.2018.1496872

[24] Lang-Roth R. Hearing impairment and language delay in infants: diagnostics and genetics. GMS Curr Top Otorhinolaryngol Head Neck Surg. 2014; 13: 1-31. doi: 10.3205/cto000108.

[25] Shaver DM, Marschark M, Newman L, Marder C. Who is where? characteristics of deaf and hard-of-hearing students in regular and special schools. J Deaf Stud Deaf Educ. 2014; 19(2): 204-19. doi: 10.1093/deafed/ent056.

[26] Britto PR. School readiness: a conceptual framework. United Nations Children's Fund; 2012.

[27] Maluleke NP, Khoza-Shangase K, Kanji A. School readiness and academic achievement of children with hearing impairment: a South African exploratory study. S Afr J Child Educ. 2021; 11(1): 1-7. doi: 10.4102/sajce.v11i1.898.

[28] McConachie L. School readiness and kindergarten transitions: children with vision impairment and blindness. In: Mashburn A, LoCasale-Crouch J, Pears K, editors. Kindergarten transition and readiness. Springer; 2018: 205-223.

[29] Niblocka J, Clarkb GF, Vosc TC, Liebermand D, Hunter EG. Systematic review of occupational therapy interventions to enhance cognitive development in children 0–5 years: Part 2, at-risk due to environmental factors and promoting cognitive development. J Occup Ther Sch Early Interv. 2021; 14(4): 486-504. doi: 10.1080/19411243.2021.1941493.

[30] Nair MKC, Radhakrishnan R, Olusanya BO. Promoting school readiness in children with developmental disabilities in LMICs. Front Public Health. 2023; 11(993642): 1-10. doi: 10.3389/fpubh.2023.993642

[31] Fang Z, Liu X, Zhang C, Qiao D. Early childhood interventions in educational settings that promote school readiness for children with autism and other developmental disabilities: systematic review Res Autism Spectr Disord. 2023; 108(102257). doi: 10.1016/j.rasd.2023.102257.

Journal of Associated Medical Sciences 2024; 57(3): 30-35

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

Interpretability and appropriate cut-off score of occiput-bed distance to indicate risk of hyperkyphosis in older adults

Roongnapa Intaruk[1,2] Sirirut Multakorn[1] Wasunan Sornsamran[1] Metinee Yamtawech[1] Sugalya Amatachaya[1,2] Pipatana Amatachaya[2] Thiwabhorn Thaweewannakij[1,2*]

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

ARTICLE INFO

Article history:
Received 30 April 2024
Accepted as revised 11 June 2024
Available online 14 June 2024

Keywords:
Spinal curvature, Dowager's hump, Cobb's method, test-retest reliability, measurement error.

ABSTRACT

Background: Age-related hyperkyphosis has been described as a new geriatric syndrome. Therefore, early screening is critical. The occiput-bed distance (OBD) was developed as a new tool; however, there needed to be clear evidence supporting its clinical utility compared to the data from standard measurement.

Objective: To investigate the test-retest reliability and minimal detectable change (MDC) with standard measurement error (SEM). Moreover, the practical cut-off scores to determine the risk of hyperkyphosis are crucial compared to a standard Cobb's method.

Materials and methods: This study was designed as a cross-sectional study conducted in ninety-six hyperkyphosis older adults aged at least 60 years. All participants were assessed for their hyperkyphosis using the OBD. Within seven days, they were at a hospital to complete a radiographic examination to determine the appropriate cut-off score of hyperkyphosis. Additionally, the first 30 participants were assessed for their hyperkyphosis again using the OBD to determine the interpretability.

Results: The outcomes of the OBD had excellent test-retest reliability ($ICC_{3,3}$ =0.887, p<0.001). The SEM and MDC_{95} values were 0.75 and 2.08 cm, respectively. In addition, the cut-off score of OBD was reported to be at least 7.40 cm (sensitivity 71.80%, specificity 73.70%, and AUC =0.734) to indicate the risk of hyperkyphosis.

Conclusion: The OBD is consistent, reliable, and has good diagnostic properties for hyperkyphosis. The findings confirm the use of OBD as a practical alternative method for early detection of hyperkyphosis in older individuals, particularly those who cannot stand for long and cannot access radiology.

Introduction

Hyperkyphosis has been described as a new geriatric syndrome.[1] Its characteristic is an anterior curvature in the thoracic spine, and tends to increase with age (up to 40% in older adults).[2] Previous reports have found that hyperkyphosis can induce adverse health consequences and superimpose substantial effects on levels of independence and mortality rate in older adults.[3,4] Although hyperkyphosis does not always lead to severe injuries, this condition can induce undesirable cosmetic consequence that causes low self-esteem.[5] To reduce the risk of adverse consequences of hyperkyphosis, an early assessment with a standard practical screening tool should be conducted, followed by monitoring and providing an appropriate intervention.[4,6]

Previous studies appraised that there were several hyperkyphosis measurements commonly used in community

* **Corresponding contributor.**
Author's Address: *School of Physical Therapy, Faculty of Associated Medical Science, Khon Kaen University, Khon Kaen Province, Thailand.*
E-mail address: thiwth@kku.ac.th
doi: 10.12982/JAMS.2024.044
E-ISSN: 2539-6056

settings, including the flexicurve ruler, inclinometer, goniometer, distance from the wall, and block method.[7,8] However, most of them needed a specialist to identify bony landmarks of the spine, except the block method, which is an easy measurement in lying position using wooden blocks by experts and novices, including village health volunteers or healthcare professionals.[9,10] Previously, the older adults considered the severity of hyperkyphosis using the number of blocks, including mild (1 block), moderate (2 blocks), and severe (at least 3 blocks) hyperkyphosis.[11] While Chokphukiao *et al.* reported that the number of blocks of at least two might indicate hyperkyphosis in older adults, they did not report its sensitivity and specificity.[12] Although the block method reported acceptable reliability and validity compared to Cobb's method (ICC_s=0.82-1.00, p<0.001 and r_s =0.63-0.64, p<0.05, respectively),[9, 10] it is at a constant block height. Such discrete data might result in errors in judgment and low sensitivity in detecting changes in the hyperkyphosis angle.

Therefore, the OBD (application of patent number 2203001689) was developed based on the block method protocol. The OBD can measure the distance from occiput to bed using continuous data. In our preliminary study, the OBD showed excellent intra- and inter-rater reliability ($ICC_{3,3}$=0.993-0.998, p<0.001) after proper training. Moreover, when applied in the older populations, it had an acceptable correlation to the occiput-wall distance (OWD) (r=0.779, p=0.008).[13] Although this method might thus be practically applied for hyperkyphosis screening in older individuals, the interpretability of the tool must be considered to confirm its stability.[14] Moreover, the practical cut-off scores to determine the risk of hyperkyphosis are crucial to promoting the effectiveness of community health care services. Therefore, the purpose of the present study was to investigate the test-retest reliability and minimal detectable change (MDC) with standard error of measurement (SEM). Additionally, the practical cut-off scores to determine the hyperkyphosis risk were explored and compared to a standard Cobb's method.

Materials and methods

Study design and participants

The study was cross-sectionally conducted in the older adults living in rural and semi-rural communities in Northeast Thailand. Eligible participants were aged at least 60 years with hyperkyphosis determined by at least three blocks and able to understand simple commands to complete the study protocol.[12] However, participants were excluded if they had scoliosis and abnormal supine symptoms, such as orthopnea. Other signs and symptoms that might influence the study were also excluded, such as dizziness, acute illness or injury, unstable heart disease (e.g., angina), and uncontrolled hypertension.

The sample size was separately calculated to assess the study's interpretability and cut-off scores of OBD. According to the number of participants in the study's interpretability, the level of significance (α) and estimate precision (e) were determined as 0.05 and 0.1, respectively. Therefore, at least 30 participants were needed for the interpretability study.[15] Furthermore, the number of participants for the cut-off scores was calculated from a pilot study (59 participants; r=0.61, p<0.001) and provided a sensitivity of 88%. Thus, the sample size required at least 96 participants to study cut-off scores. The protocols of this study were approved by the Khon Kaen University Ethics Committee for Human Research (HE644003). All participants completed a written informed consent form before participating in the study.

Research protocols

All eligible participants were assessed hyperkyphosis in two phases; each phase was separated within seven days. In the first phase, they were interviewed regarding their demographic characteristics, type, time, intensity, and frequency of physical activity and exercise. Additionally, a verbal numerical rating scale was used for pain. Then, the participants were assessed for their hyperkyphosis using the OBD. In the second phase, the participants were evaluated for their hyperkyphosis using OBD again; in addition, lateral spinal radiographic examination (Cobb's method) was performed at a hospital to determine the interpretability (test-retest reliability) and appropriate cut-off scores of OBD for the hyperkyphosis. Before both assessments, the participants were re-interviewed regarding their physical activity, exercise, and pain scale to confirm that there were no changes in these parameters affecting their hyperkyphosis condition since the baseline measurements were taken.

Hyperkyphosis measurement

OBD

The participants were asked to lie supine with their hands and arms beside the body on a firm bed with eyes gazing forward at the ceiling; the head was placed on the OBD plate in normal alignment in the horizontal view. If participants were lying flat without neck extension, they were deemed regular participants without hyperkyphosis. On the other hand, participants lying flat with neck extension were included as exhibiting hyperkyphosis.[13] Then, the height of the OBD plate was adjusted until the forehead and ear lines were parallel to the table.

The measurement was repeated in three trials, with a period of sufficient rest in a sitting position between the trials (at least 1 minute). The average distance of three trials was converted from a millimeter to a centimeter scale.

Cobb's method

A Cobb's method is commonly used as a gold standard for hyperkyphosis measurement. The participants were asked to stand in a neutral position for lateral thoracic spine radiography. During the radiography process, the participants were required to flex the shoulder at 90 degrees.[16] The Cobb angle was analyzed by drawing a straight line along the upper border of the fourth thoracic spine (T4) and another at the lower border of the twelfth thoracic spine (T12). Then, the Cobb angle was automatically computed based on the intersection between the 2 lines using the Surgimap Spine program, version 1.2 software.[17]

Participants with Cobb angles more than 40 degrees were arranged into the hyperkyphosis group.

Statistical analysis

The SPSS for Windows was used to analyze the collected data (SPSS Statistics version 18.0, IBM Corporation). Descriptive statistics were utilized to explain the demographic characteristics of participants and the study findings. A two-way random model ($ICC_{3,3}$) 's intraclass correlation coefficients were used to derive test-retest reliability. The standard error of measurement (SEM) was calculated using the following formula: SEM = (SD) × $\sqrt{1-r}$, where SD is the pooled SD of two trials and *r* is the test-retest reliability. The minimal detectable change at the 95% confidence interval (MDC_{95}) was calculated using the following formula: MDC_{95} = 1.96 × $\sqrt{2}$ × SEM. In addition, the Receiver-Operated Characteristics (ROC) curve was used to identify appropriate cut-off scores with particular sensitivity, specificity, and area under the curve (AUC) of OBD to indicate the risk of hyperkyphosis. In general, an AUC was interpreted as no discrimination (less than 0.5), poor discrimination (0.5-0.6), acceptability (0.7-0.8), excellence (more than 0.8-0.9), and outstanding (more than 0.9).[18] The significance level was set at less than 0.05.

Results

Characteristics of participants

Four hundred and eight participants were interested in participating in the study. However, 307 participants were excluded based on missing the inclusion criteria (Figure 1). After considering the exclusion criteria, 101 eligible participants met the inclusion criteria; however, five could not go to the hospital for radiography. Therefore, 96 participants (41 females and 55 males) were involved in the study to determine the cut-off score of OBD. Their average age was 72.05±5.95 years (range 60-86 years), and average Cobb angles of 37.76±9.95 degrees. In addition, the first 30 participants (18 females and 12 males) were appointed on day 7 to indicate the test-retest reliability for the interpretability study; their average age was 71.57±6.45 years old, and body mass index was 25.44±3.74 kg/m^2 (as shown in Table 1).

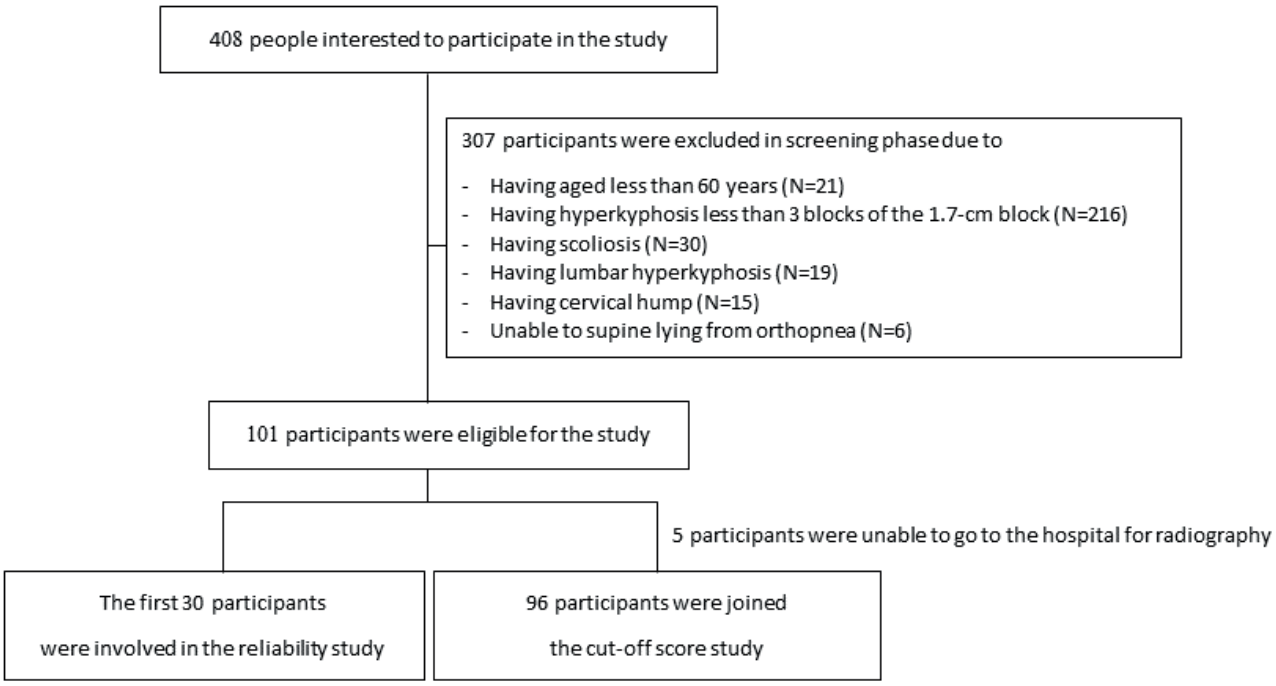


Figure 1. *Participant flowchart.*

Table 1 Demographic characteristics and hyperkyphosis data of participants.

Variable	Interpretability (N=30)	Group of participants[a] (N=96)		*p* value
		Normal (N=57)	Hyperkyphosis (N=39)	
Sex (female)[b]	18 (60.00)	29 (50.88)	12 (30.77)	0.050
Age (year)[c]	71.57±6.45 (69.16-73.98)	72.07±6.25 (70.41-73.73)	72.03±5.57 (70.22-73.83)	0.972
Weight (kg)[c]	59.86±10.61 (55.82-63.90)	61.69 ±10.01 (59.04-64.35)	60.75±8.88 (57.87-63.63)	0.637
Height (cm)[c]	153.27±7.68 (150.35-156.20)	156.01±7.98 (153.90-158.13)	157.26±7.79 (154.73-159.78)	0.451
BMI (kg/m^2)[c]	25.44±3.74 (24.02-26.86)	25.33±3.47 (24.41-26.26)	24.63±3.52 (23.48-25.77)	0.333
Number of 1.7-cm block (block)[b]				0.013*
3	17 (56.67)	29 (50.88)	7 (17.95)	
4	11 (36.67)	22 (38.60)	21 (53.85)	
5	2 (6.66)	5 (8.77)	9 (23.08)	
6	0 (0)	1 (1.75)	1 (2.56)	
7	0 (0)	0 (0)	1 (2.56)	
OBD (cm)[c]	6.41±2.20 (5.59-7.23)	6.50±1.72 (6.04-6.96)	8.36±2.33 (7.60-9.11)	<0.001*
Cobb angle (degree)[c]	-	31.26±6.44 (29.55-32.97)	47.26±5.52 (45.47-49.05)	<0.001*

*Note: [a] The groups are categorized using radiologic data, [b] Data are demonstrated using the number of participants (percentage), [c] The data are presented using the mean±SD (95% confidence interval), * Statistically significant. BMI: body mass index, OBD: occiput-bed distance.*

Test-retest reliability

There was no statistically significant difference in the OBD measurement between baseline and within seven days. The finding indicated that the OBD had excellent test-retest reliability (as shown in Table 2). In addition, the SEM and MDC_{95} values were 0.75 and 2.08 cm, respectively.

Table 2 Test-retest reliability of the OBD (N=30).

Measurement	Mean±SD (95%CI)		*p* value	Test-retest reliability (95%CI)
	Day-1 (cm)	Day-7 (cm)		
OBD (cm)	6.41±2.20 (5.59-7.23)	6.53±2.26 (5.68-7.37)	0.198	0.887* (0.762-0.946)

*Note: *Statistically significant at 0.001. OBD: occiput-bed distance, SD: standard deviation, 95%CI: 95% confidence interval.*

Cut-off scores to indicate risk of hyperkyphosis

The finding reported that 39 participants (40.63%) revealed more than 40 degrees of Cobb angle. These participants had statistically significant longer distances using OBD and larger degrees using Cobb's method than those without hyperkyphosis (p<0.001, as shown in Table 1). In addition, the cut-off score of OBD was reported to be at least 7.40 cm (sensitivity 71.80%, specificity 73.70%, and AUC=0.734) to indicate the risk of hyperkyphosis.

Discussion

This study aimed to quantify test-retest reliability and MDC_{95}; additionally, the cut-off scores were indicated by the individuals with or without risk of hyperkyphosis. The study reported excellent test-retest reliability ($ICC_{3,3}$=0.887, p<0.001) with small SEM (0.75 cm) and MDC_{95} showed 2.08 cm. The OBD was reported to have an appropriate cut-off score to identify the risk of hyperkyphosis with a cut-off score of at least 7.40 cm (sensitivity 71.80%, specificity 73.70%, and AUC=0.734).

The present study revealed excellent test-retest reliability ($ICC_{3,3}$=0.887, p<0.001) with a small amount of SEM (0.75 cm), which could confirm the practicality and stability method of OBD when applied over some time. Evidence supports that the SEM indicates a possible margin of measurement errors in an original unit, and the MDC

is a minimal change that falls outside the measurement errors in the test results. Thus, these values suggested the absolute reliability of the tools.[19] Since the OBD was developed from the protocol of the 1.7-cm block to measure hyperkyphosis in a lying position, its principles were similar to measuring distance from the wall in a standing position, known as the OWD.[20,21] Wiyanad *et al*. found smaller SEM and MDC values in people aged at least 10 years compared to the present study (0.44 cm and 1.21 cm, respectively).[21] It might occur due to the characteristics of the spine in older adults being more rigid compared to younger people, which contributes to more severe hyperkyphosis in these individuals.[2,9] However, Suwannarat and colleagues showed a similar level of test-retest reliability ($ICC_{3,3}$ =0.85, p<0.001) with smaller SEM (0.20 cm) and MDC (0.55 cm) than the present study, when they measured the distance from a wall using the seventh cervical vertebra wall distance (C7WD) in older adults.[22] The small value, even in similar participants' characteristics, could be due to the equation utilized for calculating the MDC, which depends on ICC, SD, and SEM values. The previous study showed lower SEM and SD than the present study, which might be related to lower MDC. Although prior studies have reported smaller SEM and MDC in both younger and older age groups,[21,22] evidence suggests that the number of participants for the study's reliability should be at least 30 with a level of significance (α) and estimate precision (e) of 0.05 and 0.1, respectively.[15] Previous research included 15 participants, while the present study had 30 participants. Nevertheless, more prominent participants may have a longer SD, thus it is necessary to determine these values.

Furthermore, a different position and landmark measuring in a standing position might influence back muscle activation for controlling posture against gravity and affect spinal loading.[23] As a result, the distance of OBD in the present study (6.41±2.20 cm) was lower than that of C7WD in the prior report (7.66±1.93 cm).[22] In addition, the MDC_{95} of the OBD in the present study was reported as 2.08 cm, around 1-2 blocks of 1.7-cm block, similar to Huang *et al*. They reported that the cut-off score of the block method was at least one block, which could determine mild hyperkyphosis.[11] Although it is only mild hyperkyphosis, it can affect physical ability, and future falls by 1.5 times compared to those without hyperkyphosis.[11,24] This information confirmed that a small amount of SEM and MDC_{95} were appropriate for monitoring and providing a suitable intervention in the early phase. The acceptable level of test-retest reliability of the present study might occur due to the simple protocol to adjust the spinal curve, and it does not require expertise to identify spinal bony landmarks in those with hyperkyphosis.

The findings revealed the appropriate cut-off score of OBD at least 7.40 cm (sensitivity 71.80%, specificity 73.70%, and AUC=0.734), which indicated the risk of hyperkyphosis in older people. Of all participants who completed the study, 39 participants (40.63%) had thoracic hyperkyphosis angle of more than 40 degrees. However, Wiyanad *et al.* reported the OWD cut-off score of at least 6.50 cm (sensitivity 71.4%, specificity 76.6%, and AUC=0.846) to indicate hyperkyphosis.[21] The different results between OBD and OWD might occur due to the previous study in younger people with more flexible spine (aged at least 10 years).[9,21] On the other hand, Suwannarat *et al*. have reported the cut-off score of C7WD at least 7.50 cm (sensitivity 71.80%, specificity 82.40%, and AUC=0.85) to indicate hyperkyphosis.[22] When considering a body structure in a sagittal plane, the bony prominence of C7 is closer to the body midline than the occiput.[25] As a result, when standing erect against a wall and utilizing the C7 as a landmark, the distance from the wall was more significant than the occiput landmark. Moreover, evidence supported that an acceptable level of AUC should be at least 0.7.[18] This information confirmed the utility of a cut-off score of OBD at least 7.40 cm (average Cobb angle of 47.26±5.52 degrees) to screen the risk of hyperkyphosis in older adults.

Nonetheless, this study has some limitations. The study included only older adults who reported high-incidence hyperkyphosis conditions. This may limit the generalizability of the findings for other age groups with a more flexible spinal column, such as children or younger adults. A further study should include participants with various characteristics and analyze for other variables that influence the diagnostic properties to improve the benefit of the OBD in clinical and research settings. Moreover, this study did not confirm the application in older adults with vertebral fractures, which is one of the impact consequences of hyperkyphosis. Thus, further study should confirm the ability of the OBD to indicate vertebral fracture in older populations.

Conclusion

The OBD method was verified as having excellent test-retest reliability with a small amount of SEM and an acceptable cut-off score to indicate the risk of hyperkyphosis in older adults. Thus, this method represented an alternative means of hyperkyphosis screening that could be practical to apply in older individuals, particularly those who cannot stand or stand for long; additionally, people who could not access radiology to screen their hyperkyphosis condition. The OBD would benefit health officers' consideration by applying an easy tool to identify their patients' problems.

Conflict of interest

The authors declare no conflict of interest.

Funding

This study was supported by the Research Fund for Supporting Lecturer to Admit High Potential Student to Study and Research on His Expert Program Year 2019 of Khon Kaen University [622JH209], and Research and Graduate Studies, Khon Kaen University, Thailand [RP66-4-002].

Acknowledgements

The authors thank all participants for participating in the study and the community leaders in the Northeast of Thailand for their help during data collection in their areas.

Reference

[1] Katzman WB, Harrison SL, Fink HA, Marshall LM, Orwoll E, Barrett-Connor E, et al. Physical function in older men with hyperkyphosis. J Gerontol A Biol Sci Med Sci. 2015; 70(5): 635-40. doi: 10.1093/gerona/glu213.

[2] Kado DM, Prenovost K, Crandall C. Narrative review: hyperkyphosis in older persons. Ann Intern Med. 2007; 147(5): 330-8. doi: 10.7326/0003-4819-147-5-200709040-00008.

[3] Ailon T, Shaffrey CI, Lenke LG, Harrop JS, Smith JS. Progressive spinal kyphosis in the aging population. Neurosurgery. 2015; 77(Suppl 4): S164-72. doi: 10.1227/NEU.0000000000000944.

[4] Katzman WB, Wanek L, Shepherd JA, Sellmeyer DE. Age-related hyperkyphosis: its causes, consequences, and management. J Orthop Sports Phys Ther. 2010; 40(6): 352-60. doi: 10.2519/jospt.2010.3099.

[5] Hinman MR. Comparison of thoracic kyphosis and postural stiffness in younger and older women. J Spine. 2004; 4(4): 413-7. doi: 10.1016/j.spinee.2004.01.002.

[6] Azadinia F, Kamyab M, Behtash H, Ganjavian M, Javaheri R. The validity and reliability of noninvasive methods for measuring kyphosis. J Spinal Disord Tech. 2014; 27: 212-8. doi: 10.1097/BSD.0b013 e31829a3574

[7] Amatachaya P, Wongsa S, Sooknuan T, Thaweewannakij T, Laophosri M, Manimanakorn N, et al. Validity and reliability of a thoracic kyphotic assessment tool measuring distance of the seventh cervical vertebra from the wall. Hong Kong Physiother J. 2016; 35: 30-6. doi: 10.1016/j.hkpj.2016.05.001.

[8] Wongsa S, Amatachaya S. Kyphosis assessments. Arch AHS. 2014; 26(2): 105-16.

[9] Chokphukiao P, Wiyanad A, Suwannarat P, Amatachaya S, Mato L, Wattanapan P, et al. Validity and reliability of the block method for measuring thoracic hyperkyphosis. Physiother Theory Pract. 2022; 38(13): 3037-44. doi: 10.1080/09593985.2021.

[10] Tran TH, Wing D, Davis A, Bergstrom J, Schousboe JT, Nichols JF, et al. Correlations among four measures of thoracic kyphosis in older adults. Osteoporos Int. 2016; 27(3): 1255-9. doi: 10.1007/s00198-015-3368-7.

[11] Huang MH, Barrett-Connor E, Greendale GA, Kado DM. Hyperkyphotic posture and risk of future osteoporotic fractures: the Rancho Bernardo study. J Bone Miner Res. 2006; 21(3): 419-23. doi: 10.1359/JBMR.051201.

[12] Chokphukiao P, Wiyanad A, Suwannarat P, Amatachaya S, Wattanapan P, Mato L, et al. Discriminative ability of 1.7-cm block for kyphosis measurement. Proceedings of the national and international graduate research conference 2017; 2017 March 10; Pote Sarasin Building, Khon Kaen University, Thailand, 2017. Available from: https://gsbooks.gs.kku.ac.th/60/nigrc2017/pdf/MMP21.pdf.

[13] Intaruk R, Ruangsri J, Noiphum L, Amatachaya P, Amatachaya S, Thaweewannakij T. Concurrent validity of the Occiput-Bed Distance for measuring hyperkyphosis. Thai J Phys Ther. 2023; 45(2): 135-45.

[14] Portney LG, Watkins MP. Foundations of clinical research application to practice. 3rd Ed. Upper Saddle River: Pearson Prentice Hall; 2009.

[15] Morrow JR, Jackson AW. How "significant" is your reliability? Res Q Exerc Sport. 1993; 64(3): 352-5. doi: 10.1080/02701367.1993.10608821.

[16] de Oliveira TS, Candotti CT, La Torre M, Pelinson PPT, Furlanetto TS, Kutchak FM, et al. Validity and reproducibility of the measurements obtained using the flexicurve instrument to evaluate the angles of thoracic and lumbar curvatures of the spine in the sagittal plane. Rehabil Res Pract. 2012; 2012: 186156. doi: 10.1155/2012/186156.

[17] Suwannarat P, Wattanapan P, Wiyanad A, Chokphukiao P, Wilaichit S, Amatachaya S. Reliability of novice physiotherapists for measuring Cobb angle using a digital method. Hong Kong Physiother J. 2017; 37: 34-8. doi: 10.1016/j.hkpj.2017.01.003.

[18] Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol. 2010; 5(9): 1315-6. doi: 10.1097/JTO.0b013e3181ec173d.

[19] de Vet HC TC, Mokkink LB, Knol DL. Measurement in Medicine. 1st Ed. The United States of America by Cambridge University Press: New York; 2011.

[20] Siminoski K, Warshawski RS, Jen H, Lee KC. The accuracy of clinical kyphosis examination for detection of thoracic vertebral fractures: comparison of direct and indirect kyphosis measures. J Musculoskelet Neuronal Interact. 2011; 11(3): 249-56.

[21] Wiyanad A, Chokphukiao P, Suwannarat P, Thaweewannakij T, Wattanapan P, Gaogasigam C, *et al*. Is the occiput-wall distance valid and reliable to determine the presence of thoracic hyperkyphosis? Musculoskelet Sci Pract. 2018; 38: 63-8. doi: 10.1016/j.msksp.2018.09.010.

[22] Suwannarat P, Amatachaya P, Sooknuan T, Tochaeng P, Kramkrathok K, Thaweewannakij T, et al. Hyperkyphotic measures using distance from the wall: validity, reliability, and distance from the wall to indicate the risk for thoracic hyperkyphosis and vertebral fracture. Arch Osteoporos. 2018; 13(1): 25. doi: 10.1007/s11657-018-0433-9.

[23] Lewis JS, Valentine RE. Clinical measurement of the thoracic kyphosis. A study of the intra-rater reliability in subjects with and without shoulder pain. BMC Musculoskelet Disord. 2010; 11: 39. doi: 10.1186/1471-2474-11-39.

[24] Thaweewannakij T, Wongsa S, Kamruecha, Khaengkhan J, Wongkuanklom J, Konkamtan, *et al*. Validity and discriminative ability on physical impairment relating to kyphosis using 1.7-cm block. Asia Pac J Sci Technol. 2012; 17: 660-70.

[25] Schafer R. Body Alignment, posture, and gait. 2nd Ed. Baltimore: Williams & Wilkins; 1983.

Journal of Associated Medical Sciences 2024; 57(3): 36-42

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

Executive function performance in persons with non-syndromic cleft lip and palate

Supaporn Chinchai Thanasak Kalaysak* **Piyawan Jareontonyakorn Natwipa Wanicharoen Kalyanee Makarabhirom**

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 7 May 2024
Accepted as revised 14 June 2024
Available online 17 June 2024

Keywords:
Executive function, cleft palates, speech and language pathologists.

ABSTRACT

Background: Executive functions (EFs) are crucial cognitive functions that mature from birth to adolescence. They are vital for daily task execution and overall success and also influence language and communication development. Children with EFs deficits often experience delays in language and speech abilities. These impairments are particularly prevalent among individuals with cleft lip and palate. Consequently, speech and language pathologists must address these impairments through assessments and interventions. Despite this urgent need for action, there is a scarcity of research on executive function performance in this population in Thailand, prompting an investigation to address this issue. This study explores executive function performance in this population to enhance the quality of life for individuals with cleft lip and palate.

Materials and methods: Using a survey-based approach, executive function performance was assessed in 5- to 15-year-old volunteer with non-syndromic cleft lip and palate attending the speech therapy camp provided by the Princess Sirindhorn IT Foundation Craniofacial Center at Chiang Mai University in April 2024. Parents completed the Behavior Rating Inventory of Executive Function (Parent form), with scores ≥65 indicating executive function difficulties.

Results: The study involved 29 participants, 14 males (48.28%) and 15 females (51.72%), with a mean age of 8 years and 9 months. Average scores for executive function abilities in BRI, MI, and GEC were 52.21, 56.48, and 58.90, respectively. There are several participants with abnormal executive function in each age group, along with their average T-scores across different domains. Children aged 5, 6-8, and 9-11 have T-scores for executive function performances falling into problematic levels for 1, 2, and 5 individuals, respectively.

Conclusion: Most of the sample group demonstrated executive function skills within the normal range. However, a certain number of individuals experienced issues with executive function. These findings offer guidance for speech and language pathologists and emphasize the importance of executive function in individuals with cleft palates.

Introduction

Executive Functions (EFs) are crucial cognitive functions encompassing critical higher-level brain processes essential for cognitive functioning, including focused attention, decision-making, impulse control, and task perseverance.[1] EFs are composed of two primary components: the Behavioral Regulation Index (BRI), which involves 1) inhibition, 2) shifting/flexibility, and 3) emotional control, and the Metacognition Index (MI), comprising 1) initiation, 2) working memory, 3) planning/organizing, 4) organization of materials, and 5) monitor/self-monitoring.[2] The development of EFs commences from early in life to

* **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.2024.045
E-ISSN: 2539-6056

adolescence, with significant advancements occurring between the ages of 3 and 6. Working memory skills undergo rapid development, gradually emerging from late infancy. Emotional control and inhibition typically exhibit substantial progress between ages 4 and 9, persisting steadily into adolescence, with full maturation typically achieved by age 25 upon entering adulthood. Subsequently, a gradual decline ensues with advancing age into the elderly years.[3]

Executive functions are indispensable in the seamless execution and success of various daily activities and lay the foundation for basic skills across three key areas. [1] Firstly, working memory involves retaining and retrieving information, beginning in infancy and relying heavily on attention. Secondly, inhibitory control entails the restraint of impulsive behaviors and the ability to think before acting, typically developing between ages 3 and 3.5, often correlating with the maturation of working memory. Thirdly, cognitive flexibility, emerging between ages 4 and 4.5, facilitates adaptability and problem-solving, relying on working memory and inhibitory control. Cognitive flexibility hinges on the ability to inhibit responses and recall past information for adjustment and planning, thereby building upon the foundations of working memory and inhibitory control. The harmonious alignment of these EF skills fosters creativity and the cultivation of higher-level abilities such as emotional regulation, planning, reasoning, and problem-solving.[1] EFs significantly contribute to success across various learning domains, from academic pursuits to social interactions, underscoring its pivotal role in navigating daily life.[4]

EFs also impact development, and research has shed light on the intricate relationship between EF deficits and language components, particularly evident in individuals with autism spectrum disorder (ASD). Friedman and Sterling[5] revealed that deficits in EF skills, including behavioral flexibility and inhibitory control, can adversely affect semantic language components, impacting vocabulary comprehension and usage. Similarly, when evaluated by teachers in school settings, limited vocabulary comprehension was associated with deficits in emotional control, cognitive flexibility, emotional regulation, initiation, and self-monitoring.[6] Akbar *et al*.[7] further demonstrated a correlation between lower language proficiency and deficits in working memory skills among children with ASD. Additionally, deficits in executive and developmental functions hindered language-related abilities such as working memory, syntactic components, and lexical semantics, exacerbating communication challenges.[7] In terms of pragmatics, they found that children with ASD often struggle with communication skills and social interaction. They lack flexibility and exhibit repetitive behaviors. Notably, deficits were noted in critical skills such as joint attention, initiating conversations, and working memory.[8,9] Studies on executive function skills have also found that fundamental executive function skills, such as attentional focusing, are crucially linked to language-related skills and speech. Specifically, skills in semantic processing directly impact one's ability to comprehend meaning and communicate effectively, particularly in environments where one must selectively attend to multiple stimuli, such as environments with distracting noises or when there is a need to simultaneously process information from various sources. Furthermore, it has been observed that the root cause of these issues stems from deficits in executive function skills related to the ability to inhibit behavior. The lack of flexibility in cognitive processing to selectively attend to only the desired stimuli has been identified as a contributing factor.[10]

Moreover, studies have investigated EF deficits in children with cleft lip and palate conditions. Bodoni *et al*.[11] identified deficits in intelligence and cognitive understanding in these individuals, and there were impairments in attentional focusing, working memory, and planning. Structural brain imaging revealed cortical thinning in specific brain regions, correlating with cognitive deficits. [11, 12] Similarly, Conrad *et al*. [13] observed lower verbal memory abilities in children with cleft lip and palate than controls, highlighting EF deficits' impact on language-related skills. In summary, executive function skills significantly affect language and speech abilities. A strong foundation in executive function skills helps develop and enhance learning abilities.

Researchers in the capacity of speech-language pathologists are tasked with the assessment and treatment of individuals with developmental language and speech impairments, including those with unclear speech and voice disorders, particularly prevalent among individuals with cleft lip and palate conditions. We recognize the potential contribution of speech-language pathologists in preventing cognitive and communicative deficits by enhancing various foundational skills relevant to cognitive management, memory, comprehension, and language used for communication and problem-solving. Consequently, our interest lies in investigating executive function abilities among individuals with cleft lip and palate conditions, which has been relatively unexplored in the Thai context. This study explores the executive function performance of individuals with cleft lip and palate conditions. The ultimate goal is to enhance the quality of life for individuals affected by cleft lip and palate.

Methodology

The researchers selected a specific and targeted sample group of parents of individuals aged 5-15 years with non-syndromic cleft lip and palate conditions. The inclusion criteria included: 1) participation in the Speech Camp for Children with Cleft Lip and/or Palate conducted by the Princess Sirindhorn IT Foundation Craniofacial Center Chiang Mai University, Chiang Mai Province. 2) absence of other accompanying disabilities, and 3) obtained consent and permission from the parents to participate in the research, documented through signed consent forms. For children aged 7 years and older, their consent was also required. The exclusion criteria were that the parents could not complete the Behavior Rating Inventory of Executive Function (Parent Form) assessment (BRIEF-Parent form). The researchers conducted outreach

to the parents of individuals aged 5-15 years with cleft lip and palate conditions who were receiving services at the speech camp in April 2024. Providing detailed information about the research project alongside the promotion of the speech camp, allowing parents time to decide whether to participate in the research project. The researchers conducted another promotion round on the day of the speech camp, ensuring parents had independent decision-making rights. Subsequently, interested individuals applied to join the research project, and consent forms were sent to parents and children aged 7 years and older to sign. The researchers distributed the BRIEF-Parent form in Thai to parents and guided them through its completion to ensure comprehension. Parents then selected responses that accurately reflected the executive function abilities of their children. In cases where a parent could not read Thai, the researcher read the form aloud and recorded the results reported by the parent. The results were promptly reported back to the parents for verification and accuracy. The BRIEF-Parent form, translated into Thai, is a standardized assessment comprising 86 questions from two major components: the Behavior Regulation Index (BRI) and the Metacognition Index (MI).[14] The BRI includes items related to 1) inhibit/inhibitory control, 2) shifting/flexibility, and 3.) emotional control. At the same time, the MI comprises items on 1) initiation, 2) working memory, 3) planning/organizing, 4) organization of materials, and 5) monitoring/self-monitoring. This assessment is applicable for children aged 5-18 years, including both typically developing children and those with neurological impairments such as learning disabilities and attention deficit hyperactivity disorder (ADHD), among others. It demonstrates high internal consistency ranging from 0.80 to 0.98 and possesses strong validity and reliability. Internationally recognized and extensively employed, this tool is utilized by parents and educators across diverse environments, such as homes and schools, where children participate in their daily routines. It offers valuable perspectives on children's executive function capacities as applied in practical, real-world scenarios.[15] Completion of the entire assessment typically takes approximately 10-15 minutes.

The data obtained was translated into scores for each domain according to the assessment criteria and calculated for the Global executive composite (GEC), the overarching summary score incorporating all of the BRIEF clinical scales calculated from the BRI combined with MI. Subsequently, statistical analysis was conducted to interpret the results, followed by a summary of the research findings. Interpreting executive function abilities involves comparing the obtained scores for each domain with age and gender-specific norms to derive T-scores. Scores equal to or exceeding 65 indicate significant problems in executive function within those specific domains.[2]

Results

In this study, 29 parents participated and provided data. The details are presented in Table 1.

The results of executive function performance and T-scores for the sample group of children are presented in Table 2.

Table 1. General Characteristics of the Individuals with non-syndromic cleft lip and palate (N=29).

Characteristics	Number	Percentage
Sex		
Male	14	48.28
Female	15	51.72
Age		
5 years 0 months - 5 years 11 months	4	13.79
6 years 0 months - 6 years 11 months	2	6.90
7 years 0 months - 7 years 11 months	5	17.24
8 years 0 months - 8 years 11 months	4	13.79
9 years 0 months - 9 years 11 months	4	13.79
10 years 0 months - 10 years 11 months	6	20.69
11 years 0 months - 11 years 11 months	1	3.45
12 years 0 months - 12 years 11 months	2	6.90
13 years 0 months - 13 years 11 months	1	3.45
Lowest Age: 5 years 8 months		
Highest Age: 13 years 3 months		
Average Age: 8 years 9 months		

Table 2. Average T-Scores of executive function performances by gender (N=29).

Scale/index	Female (N=15)	Male (N=14)		SD
Inhibitory control	50.13	45.57	47.93	12.37
Shifting/flexibility	54.00	50.50	52.31	15.76
Emotional control	53.20	45.29	49.38	14.19
Behavior regulation index (BRI)	55.26	48.93	52.21	11.32
Initiation	49.73	48.43	49.10	13.27
Working memory	52.73	53.57	53.14	14.17
Planning/organizing	52.66	53.79	53.21	13.56
Organization of materials	46.93	50.14	48.48	13.21
Self-monitoring	53.00	51.14	52.10	15.02
Metacognition index (MI)	56.13	56.86	56.48	11.92
Global executive composite (GEC)	60.33	57.36	58.90	21.40

The scores obtained for each domain are compared with age- and gender-specific norms to derive T-scores. However, the average score showed that the children with non-syndromic cleft lip and palate executive function skills are within the normal range. However, upon scrutinizing the scores segregated within each domain of individual participants, it is observed that some participants exhibit T- scores exceeding 65, as depicted in Table 3.

Table 3. Number of participants exhibiting abnormal and normal executive function performance by each domain (N=29).

Scale/Index	Abnormal (N)	Percentage	Normal (N)	Percentage
Inhibitory control	2	6.90	27	93.10
Shifting/flexibility	6	20.69	23	79.31
Emotional control	4	13.79	25	86.21
Behavior regulation index (BRI)	5	17.24	24	82.76
Initiatation	3	10.34	26	89.66
Working memory	6	20.69	23	79.31
Planning/organizing	7	24.14	22	75.86
Organization of materials	3	10.34	26	89.66
Self-monitoring	4	13.79	25	86.21
Metacognition index (MI)	5	17.24	24	82.76
Global executive composite (GEC)	8	27.59	21	72.41

Data concerning the executive function performances of individuals across different age groups have been analyzed, and the findings are illuminated in detail (Table 4). Table 4 details the number of participants with abnormal and normal executive function in each age group, along with the average T-scores ($\bar{x}$) categorized by domain. Upon examining the details within the individual sample groups, we find that children aged 5, 6-8, and 9-11 have T-scores for executive function performances falling into problematic levels for 1, 2, and 5 individuals, respectively.

Table 4. The number of participants with abnormal and normal executive function in each age group, average T-Scores ($\bar{x}$) categorized by the domain.

Scale/Index	5 years (N=4)			6-8 years (N=11)			9-11 years (N=12)			12-15 years (N=2)		
	abnormal	$\bar{x}$	normal	abnormal	$\bar{x}$	normal	abnormal	$\bar{x}$	normal	abnormal	$\bar{x}$	normal	$\bar{x}$
Inhibitory control	0	-	4	0	-	11	2	72	10	0	-	2	42.5
Shifting/flexibility	1	76.0	3	1	71	10	4	70.25	8	0	-	2	39.5
Emotional control	0	-	4	1	71	10	3	72	9	0	-	2	37
Behavior regulation index (BRI)	0	-	4	1	68	10	4	71.25	8	0	-	2	38
Initiation	0	-	4	1	67	10	2	65.5	10	0	-	2	41
Working memory	2	67.5	2	1	69	10	3	71.33	9	0	-	2	40
Planning/organizing	2	67	2	3	68.33	8	2	71	10	0	-	2	44.5
Organization of materials	0	-	4	1	66	10	2	67	10	0	-	2	44.5
Self-monitoring	1	69	3	2	69.5	9	1	67	11	0	-	2	48
Metacognition index (MI)	1	68	3	2	76.5	9	2	71.5	10	0	-	2	43
Global executive composite (GEC)	1	66	3	2	104.5	9	5	76.6	7	0	-	2	40

Discussion

This study examined the executive function performance of individuals aged 5 to 15 years with non-syndromic cleft lip and palate, with data provided by their parents. 29 participants, comprising 14 males and 15 females, were assessed using the BRIEF-Parent form in Thai. The results revealed that the average BRI, MI, and GEC scores were 52.21, 56.48, and 58.90, respectively (see Table 2). When interpreting these scores based on the assessment's criteria, which indicate scores equal to or greater than 65 as indicative of executive function difficulties, it was found that the sample group of children with non-syndromic cleft lip and palate exhibited a standard range of executive function skills.

Research studies such as those by Sarakai support the rationale behind this study.[16-18] The children in this study were those currently attending age-appropriate classrooms. As children progress into adolescence, school attendance becomes increasingly vital for developing executive function skills. Starting from kindergarten and progressing through higher grades, consistent school attendance is crucial for developing these skills. The school environment serves as a stimulus for efficient brain function, providing opportunities for children to learn and play beyond the confines of their homes. This, in turn, promotes the accumulation of life experiences through learning, teaching, socializing, and engaging in various activities, as discussed in articles by the Center on the Developing Child at Harvard University.[15] Moreover, social interaction in the school context, which is a daily routine for children attending school, aids in the development of self-regulation skills. It helps children control their reactions to distractions, manage their emotional expressions and behaviors towards peers, adhere to rules, regulations, and teacher instructions, and maintain discipline throughout school attendance. Consequently, this enhances children's executive function skills in all aspects. Furthermore, the sample group in this study consisted of individuals with non-syndromic cleft lip and palate. This condition arises not from genetic mutations but from failures in connecting various structural components. Thus, there is a tendency for these children not to have cognitive impairments, as elucidated by Phrathanee.[18] This synthesis of research findings underscores the importance of schooling in fostering executive function skills among children, particularly those with cleft lip and palate conditions. It highlights the school environment's role as a holistic cognitive development facilitator. It emphasizes the potential impact of social interaction and consistent school attendance on executive function abilities. Hence, this may result in most of the study sample group having an executive function within the normal range. Furthermore, it is essential to consider that all participants in this study received intensive speech therapy through speech camp activities. Li *et al.*[12] found that in a group that underwent speech correction, changes occurred in cortical plasticity in brain areas related to language, auditory processing, articulation, planning, and executive functions. These changes are believed to result from the neuroplasticity

mechanism, where nerve cells in the brain are stimulated through rehabilitation activities.

However, upon considering T-score values on an individual subject within each age group, we found that children with non-syndromic cleft lip and palate in the sample group aged 5 to 11 still exhibit some executive function difficulties. Specifically, the problematic areas across all age ranges include shifting/flexibility, working memory, planning/organizing, and self-monitoring. These challenges impact the MI index and GEC composite, with T-scores falling into problematic levels (Tables 3 and 4). partials of the results of this study are consistent with Bodoni *et al*.,[11] and Conrad *et al*.[13] may stem from a combination of factors. These can include the physical challenges associated with the cleft condition, such as hearing loss or speech difficulties, which can affect communication and social interaction. Additionally, the psychological impact of having a visible difference can lead to social stigma and emotional stress, further complicating executive function development. Research shows that children with non-syndromic cleft lip and palate may have lower scores on tasks measuring executive functions compared to their peers without the condition.[13,20] This includes shifting, working memory, planning/organizing, and self-monitoring challenges.

Hence, speech rehabilitation contributes to developing language, auditory processing, articulation, and executive function skills in individuals with cleft lip and palate defects. Therefore, individuals with cleft lip and palate defects should receive intensive speech therapy from speech-language pathologists to positively impact executive function skills development.

Conclusion

This study on executive function skills of individuals with cleft lip and palate defects, aged 5-15 years, with parents as the informants, found that most of the sample group had executive function skills within the normal range. However, some of the participants still had problems in executive function. Even though the study is a pioneering investigation in Thailand on the executive function abilities of individuals with non-syndromic cleft lip and palate aged 5-15 years, the findings may serve as a guideline for speech therapists to recognize the importance of executive function skills in individuals with cleft lip and palate defects in the future.

Limitations

This pilot study was conducted on children with non-syndromic cleft lip and palate. Future studies should include children with cleft lip and palate defects with other associated conditions, such as syndromic cleft lip and palate, Tracher-Collins syndrome, and Van De Woude syndrome. Additionally, future research should incorporate a larger sample size that encompasses all regions of Thailand to offer a more comprehensive understanding of the study results. Furthermore, conducting studies where data is collected from teachers who oversee children in school settings would provide extensive insight into children's executive function skills in various contexts.

Ethic approval

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional human experimentation committees and the Helsinki Declaration of 1975 (revised in 2013). Before being included in the study, all subjects gave written informed consent. This study was approved by the Human Research Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University (AMSEC-66FB-005).

Conflict of interest statement

The author(s) declared no potential conflicts of interest regarding this article's research, authorship, and/or publication.

Funding Sources

This research was funded by the Faculty of Associated Medical Sciences, Chiang Mai University.

References

[1] Diamond A. Executive functions. Annu Rev Psychol. 2013; 64: 135-68. doi: 10.1146/annurev-psych-113011-143750.

[2] Gioia, G.A. *et al*. Professional Manual Behavior Rating Inventory of Executive Function. Florida: Psychological Assessment Resources; 2000.

[3] Chutabhakdikul N. Executive function. Srinakharinwirot Univ J Educ Stud. 2014; 48: 62-9.

[4] Chutabhakdikul N. Executive functions and readiness for learning in preschool children. In: EDUCA 2015: The 8th Education Symposium for Teacher Professional Development; 2015; Nonthaburi. p. 35-8.

[5] Friedman L, Sterling A. A Review of Language, Executive Function, and Intervention in Autism Spectrum Disorder. Semin Speech Lang. 2019; 40(4): 291-304. doi: 10.1055/s-0039-1692964.

[6] Cascia J, Barr JJ. Associations Among Vocabulary, Executive Function Skills and Empathy in Individuals with Autism Spectrum Disorder. J Appl Res Intellect Disabil. 2017; 30(4): 627-37. doi: 10.1111/jar.12257.

[7] Akbar M, Loomis R, Paul R. The interplay of language on executive functions in children with ASD. Res Autism Spectr Disord. 2013; 7(3): 494-501. doi: 10.1016/j.rasd.2012.09.001.

[8] Gilotty L, Kenworthy L, Sirian L, Black DO, Wagner AE. Adaptive skills and executive function in autism spectrum disorders. Child Neuropsychol. 2002; 8(4): 241-8. doi: 10.1076/chin.8.4.241.13504.

[9] McEvoy RE, Rogers SJ, Pennington BF. Executive function and social communication deficits in young autistic children. J Child Psychol Psychiatry. 1993; 34(4): 563-78. doi: 10.1111/j.1469-7610.1993.tb01036.x.

[10] Perrone-Bertolotti M, Tassin M, Meunier F. Speech-in-speech perception and executive function involvement. PLoS One. 2017; 12(7). doi.org/10.1371/journal.pone.0180084.

[11] Bodoni PSB, Leoni RF, do Vale AB, da Silva PHR,

Meira Junior SG, Richieri Costa A, Tabaquim MLM. Neuropsychological functioning and its relationship with brain anatomical measures of children and adolescents with non-syndromic cleft lip and palate. Child Neuropsychol. 2021; 27(1): 2-16. doi.org/10.1080/09297049.2020.1776240.

[12] Li Z, Zhang W, Li C, Wang M, Wang S, Chen R, Zhang X. Articulation rehabilitation induces cortical plasticity in adults with non-syndromic cleft lip and palate. Aging (Albany NY). 2020; 12(13): 13147-59. doi.org/10.18632/aging.103402.

[13] Conrad AL, Richman L, Nopoulos P, Dailey S. Neuropsychological functioning in children with non-syndromic cleft of the lip and/or palate. Child Neuropsychol. 2009; 15(5): 471-84. doi.org/10.1080/09297040802691120.

[14] Thai Health Promotion Foundation. Development and standardization of assessment tools for executive function in preschool children [Internet]. 2017 [cited 2023 Sep 8]. Available from: http://kb.hsri.or.th/dspace/handle/11228/4650?locale-attribute=en.

[15] Center on the Developing Child at Harvard University. Building the brain's "air traffic control" system: How early experiences shape the development of executive function: Working paper no. 11 [Internet]. 2011 [cited 2023 Sep 10]. Available from: https://developingchild.harvard.edu/wp-content/uploads/2011/05/How-Early-Experiences-Shape-the-Development-of-Executive-Function.pdf.

[16] Sarakai S. The study of executive function in children aged 7-12 years old using Behavior Rating Inventory of Executive Function (BRIEF) [Term paper]. Chiang Mai University; 2017. Available from: http://202.28.25.187/termpaper/termrequest/fulltext/2560/OT1291.pdf.

[17] O'Meagher S, Norris K, Kemp N, Anderson P. Parent and teacher reporting of executive function and behavioral difficulties in preterm and term children at kindergarten. Appl Neuropsychol Child. 2020; 9(2): 153-64. doi: 10.1080/21622965.2018.1550404.

[18] Germano GD, Brito LB, Capellini SA. The opinion of parents and teachers of students with learning disorders regarding executive function skills. Rev CEFAC. 2012; 19(5): 674-82. doi: 10.1590/1982-0216201719510817.

[19] Prathanee B. Problems of language and speech disorders in individuals with cleft lip and palate. In: Prathanee B, editor. Language speech and auditory in individuals with cleft lip and palate: Screening. Khon Kaen: Department of Otolaryngology, Faculty of Medicine, Khon Kaen University; 2008. p. 11-24.

[20] Sándor-Bajusz KA, Dergez T, Molnár E, Hadzsiev K, Till Á, Zsigmond A, Vástyán A, Csábi G. Cognitive functioning and clinical characteristics of children with non-syndromic orofacial clefts: A case-control study. Front Psychol. 2023; 14: 1115304. doi: 10.3389/fpsyg.2023.1115304.

Journal of Associated Medical Sciences 2024; 57(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

Development of tongue strength and endurance measurement device: A pilot study in healthy adults

Palita Yaemsuan[1] Piyawat Trevittaya[1*] Nipon Theera-Umpon[2]

[1]Communication Sciences and Disorders Division, Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Biomedical Engineering and Innovation Research Center, Biomedical Engineering Institute, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 8 March 2024
Accepted as revised 14 June 2024
Available online 19 June 2024

Keywords:
Tongue, tongue strength, tongue endurance, swallowing, device.

ABSTRACT

Background: Safe and efficient swallowing relies on adequate tongue strength and endurance. Thus, tongue strength and endurance assessments are essential for swallowing rehabilitation for individuals with swallowing difficulties due to tongue structure and function abnormalities. These tongue functions can be objectively measured using a standard device. However, in Thailand, assessing tongue strength and endurance using standard devices is not widespread in clinical practice due to the high cost of importing these devices.

Objective: This study aimed to develop a precise, accurate, and reliable tongue function measurement device for the clinical assessment of tongue strength and endurance.

Materials and methods: This study was divided into three phases: 1) Development of a tongue strength and endurance measurement device in a laboratory setting and administration of a satisfaction questionnaire, 2) Trial of the prototype device with six participants, and 3) Assessment of the test-retest reliability of the developed device and investigation of tongue strength and endurance values with twenty participants.

Results: As a result of this development, a novel device for measuring tongue strength and endurance was obtained, which provides measurements in units of newtons (N) and kilopascals (kPa) for strength and seconds (s) for endurance. The device's development cost was significantly lower compared to imported commercially available devices while maintaining the performance standards for medical measurement devices. This was demonstrated by its accuracy ranging from 96.40% to 100%, high precision with a Coefficient of Variation (% CV) of 0.90% to 4.21%, and moderate to excellent reliability with an Intraclass Correlation Coefficient (ICC) of 0.56 to 0.93. Furthermore, statistically significant differences (p<0.01) were observed between genders, especially in anterior tongue strength.

Conclusion: The tongue strength and endurance measurement device developed through this study can be utilized for clinical tongue function assessment, giving patients more access to objective evaluations of tongue strength and endurance at a lower examination cost.

∗ **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: piya.trevit@cmu.ac.th
doi: 10.12982/JAMS.2024.046
E-ISSN: 2539-6056

Introduction

The tongue is an organ located in the oral cavity that is crucial to the oral and pharyngeal phases of swallowing. After food has been placed in the mouth, the tongue carries the food to the post-canine region for chewing to reduce its size, then softens the chewed food with saliva to achieve an optimal consistency and form a bolus.

The tongue propels the bolus from the oral cavity to the pharyngeal cavity. Then, it retracts against the pharyngeal walls to aid in bolus passage through the pharynx and into the upper esophageal sphincter.[1-3] Safe and efficient swallowing relies on adequate tongue strength and endurance.[4] Subsequently, reduced tongue strength and endurance can lead to dysphagia symptoms, including poor bolus formation, oral residue, and premature spillage of the bolus into the pharynx.[1,2,5] Dysphagia resulting from any of these problems causes longer meal times, reduces food consumption, and affects the quality of life.[6,7]

Tongue strength and endurance assessments are essential for swallowing rehabilitation for individuals with swallowing difficulties due to tongue structure and function abnormalities. This includes elderly adults with sarcopenia,[8] individuals suffering from neurological disorders such as stroke, Traumatic Brain Injury (TBI), or Parkinson's disease,[9,10] and individuals with head and neck cancer who have undergone surgery, radiation therapy, or chemotherapy.[11,12] The tongue function assessment is divided into two types: subjective measurement, which involves the observation of tongue movement capabilities, and objective measurement, which uses standard devices to quantify tongue strength and endurance numerically, allowing for repeatable measurement. Assessing tongue function using standard devices is common in clinical practice and research globally, with different regions preferring various instruments. For example, in North America, studies on tongue function predominantly use devices such as the Iowa Oral Performance Instrument (IOPI), the KayPENTAX Digital Swallowing Workstation, and the Madison Oral Strengthening Therapeutic (MOST), which was the early model before changing to SwallowSTRONG later.[13,14] In Europe and Asia, the IOPI is the preferred device, but in Japan, the JMS tongue measurement device and the SwallowSCAN device have been developed and are widely utilized.[15] In Thailand, objective tongue measurement is not widespread in clinical practice due to the high cost of importing these devices. Moreover, the literature review and the researcher's clinical service experiences have identified certain limitations in various aspects of device usage. Examples include the lack of stability of the tongue bulb position and imprecise indications of the force exerted during tongue measurements. Recognizing this issue, the researcher saw an opportunity to develop a new tongue strength and endurance measurement device for clinical tongue function assessment, allowing patients to have increased access to objective evaluation of tongue strength and endurance at a lower examination cost.

Materials and methods

This study employed a developmental research design divided into three phases as follows:

Phase 1: Development of a tongue strength and endurance measurement device in a laboratory setting and a satisfaction questionnaire

First, the operational principles of the current standard devices used to assess tongue strength and endurance were studied before developing the conceptual framework of the prototype device. This included devices in both the air-filled bulb type, such as the Iowa Oral Performance Instrument (IOPI® PRO model 3.1, IOPI Medical LLC, USA),[16] the JMS tongue measurement device (Orarize® TPM-02, JMS Co., Ltd. Japan),[17] and the KayPENTAX Digital Swallowing Workstation (KSW model 7200, PENTAX Medical, USA),[18] and the electric sensor type such as the SwallowSTRONG (SwallowSTRONG®, Swallow Solutions LLC, USA),[19] the SwallowSCAN (Swallow SCAN, Nitta Co., Japan),[20] and the OroPress, developed by McCormack *et al*.[14] Additionally, the design process involved a comprehensive review of the relevant literature concerning tongue measurement using standard devices and the researcher's clinical service experiences. Subsequently, these conceptualizations were discussed and refined in collaboration with the Biomedical Engineering Institute (BMEI), Chiang Mai University, to develop the prototype device.

Second, the researcher evaluated the prototype device's accuracy and precision in the laboratory by testing it with standard masses weighing 100 and 200 grams. Each weight was placed on the tongue force sensor 10 times.[21,22] The measured values were then used to calculate the percentage of accuracy (% accuracy) and percentage Coefficient of Variation (%CV).

Then, a satisfaction questionnaire for assessing tongue strength and endurance was developed using the prototype device with six participants in phase 2. The researcher adapted questions from relevant studies[23,24] and examined the content validity with five experts with at least 10 years of knowledge and expertise in dysphagia assessment and rehabilitation or experience in tongue function assessment using standard devices. Questions with an Item-Objective Congruence (IOC) value between 0.50 and 1 were deemed acceptable, whereas those with an IOC value below 0.50 were considered for revision or removal.[25] The researcher refined the questionnaire based on the feedback and submitted it to the experts for further review before implementation.

Phase 2: Trial of the prototype device with six participants

Participants

In phase 2, the participants consisted of 6 healthy adults, comprising 3 females and 3 males, who were divided into three age groups based on 20-year intervals: younger adults aged 20-39 years, middle-aged adults aged 40-59 years, and older adults aged 60-79 years, respectively. Each group included 1 female and 1 male participant and were selected using quota sampling. All participants were recruited voluntarily from the Faculty of Associated Medical Sciences, Chiang Mai University, and informed consent was obtained from each participant before the study.

All participants reported being in good general health with no current or previous swallowing difficulties, as confirmed through interviews conducted by the researcher. Furthermore, each participant underwent an oral motor examination and the 3-oz water swallowing

test to ensure their oral mechanism and swallowing-related reflexes were within normal limits.[26]

Participants meeting any of the following criteria were excluded from the study: current or historical medical conditions affecting oral structure and function, potentially impacting swallowing and speech, including neurological diseases, head and neck cancer, obstructive sleep apnea diagnosed by a physician, or a history of surgical procedures within the oral cavity, excluding wisdom tooth extraction; individuals who wore full dentures or experienced issues with ill-fitting dentures; and participants with impairments in posture maintenance, including the inability to sit upright continuously and maintain head positioning for at least 15 minutes.

Procedure

Data collection in this study adhered to the COVID-19 prevention guidelines.[27,28] Before tongue measurement, participants were instructed to rinse their mouths with 0.12% chlorhexidine (C-20 mouthwash) for 30 seconds. The researcher ensured the proper setup, conducted a calibration check, and initiated the measurement process. In phase 2, tongue strength and endurance were measured in a single round, including three measurements each for anterior tongue strength, posterior tongue strength, anterior tongue endurance, and posterior tongue endurance.

1) Tongue strength measurement

The researcher positioned the tongue force sensor on the palate, ensuring correct alignment, and held the handle throughout. Anterior tongue strength was measured first, with participants pressing their tongue tip against the sensor at the midline upper alveolar ridge for 3 seconds, followed by a 30-second rest, and repeating this process three times. After a 5-minute break, posterior tongue strength was measured at the junction of the hard and soft palates. Once both measurements were done, there was a 10-minute break before measuring tongue endurance. Results were recorded as the highest of the three measurements and presented in N (the unit commonly used to measure muscle strength) and kPa (the unit widely used by standard devices)

2) Tongue endurance measurement

Participants sustained tongue pressure at 50% of their maximum tongue force, starting at the anterior region. Circular visual feedback indicating their current pressure level appeared on the application screen. After each round, there was a 2-minute break, and this process was repeated three times, followed by a 10-minute rest before measuring posterior tongue endurance. Endurance was measured until pressure dropped below 50% of the maximum. Results were averaged and presented in seconds. Afterward, the participants completed a satisfaction questionnaire for prototype device improvement.

Phase 3: Assessment of the test-retest reliability of the developed device and investigation of tongue strength and endurance values with twenty participants.

The researcher assessed the test-retest reliability of the modified prototype device on 20 participants aged 20-39 years, equally divided into 10 males and 10 females, selected using quota sampling and adhering to the same criteria used in phase 2. Due to safety concerns about the significant pressure required for measuring tongue strength, the Human Research Ethics Committee recommended starting with young adults. If no risks related to sudden changes in blood pressure are observed, middle-aged and elderly volunteers will be included in future studies.

The methods and sequence for measuring tongue strength and endurance in these participants followed the same protocol in Phase 2. Tongue measurements in phase 3 were repeated twice, with a 15-minute interval. The researcher analyzed the test-retest reliability and studied the tongue strength and endurance values using data from both measurement rounds.

Data analysis

All data underwent statistical analysis using SPSS (IBM Corp. Released 2022. IBM SPSS Statistics for Mac, Version 29.0. Armonk, NY: IBM Corp.).[29] This analysis included:

1. Descriptive statistics presented the mean and standard deviation of the participant's demographic data and tongue measurements by gender and age.
2. Accuracy assessment by calculating the percent error (%error) and deriving the percent accuracy (%accuracy).
3. Precision assessment by calculating the Coefficient of Variation (%CV) using the mean and standard deviation.
4. Test-retest reliability assessment using the Intraclass Correlation Coefficient (ICC) with a two-way mixed effects model for multiple raters/measurements.

Results

Phase 1: Development of a tongue strength and endurance measurement device in a laboratory setting and creation of a satisfaction questionnaire

1.1 Prototype device development

The device is divided into three main components: the tongue force sensor, processing unit, and display/recording unit, as shown in Figure 1. The details of each element are as follows:

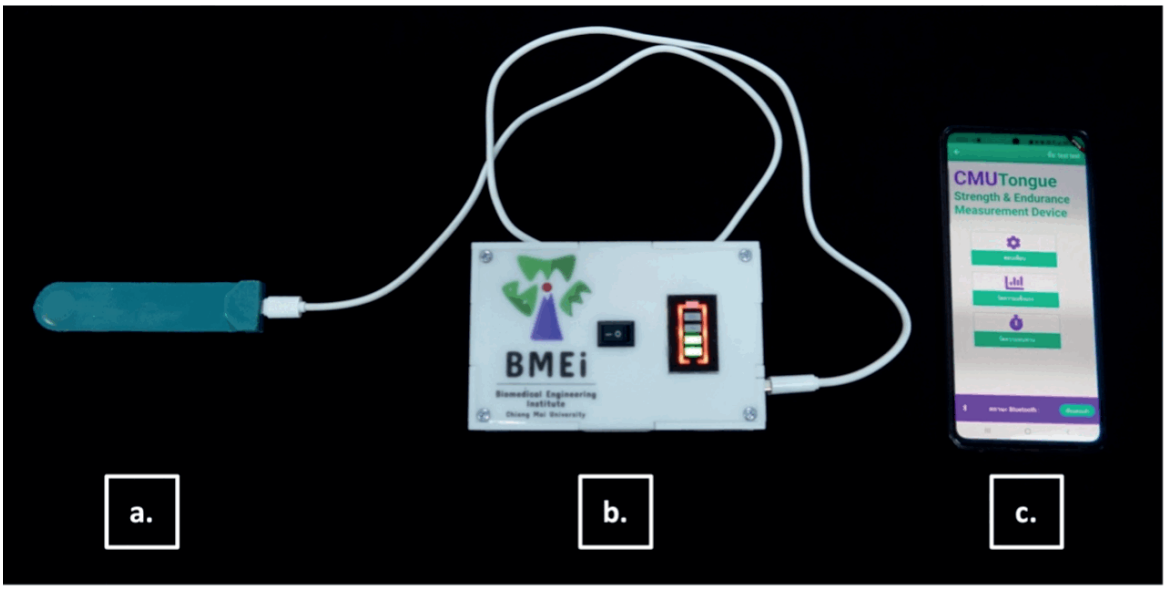


Figure 1. *Tongue strength and endurance measurement device.*
a: tongue force sensor, b: processing unit, c: display/recording unit.

1) Tongue force sensor

The researcher used the circular Force Sensing Resistor® (FSR) sensor 402 model,[30] encased in food-grade silicone. This sensor has an overall diameter of 23 mm. The tongue force sensor comprises two subparts: the force-receiving area and the handle. The force-receiving area, which is in contact with the tongue, has a smooth surface, while the area contacting the palate is slightly raised, approximately 1 mm higher, creating a 6 mm-wide ridge. This surface difference helped identify where force should be exerted by the tongue.

The handle is 5 mm thick and comfortably fits into the mouth for the posterior tongue region, reducing the risk of position changes during measurements. Additionally, the researcher increased the thickness of the handle's end to 9 mm, reducing the risk of cable damage when plugging in or unplugging the device each time, as shown in Figure 2.

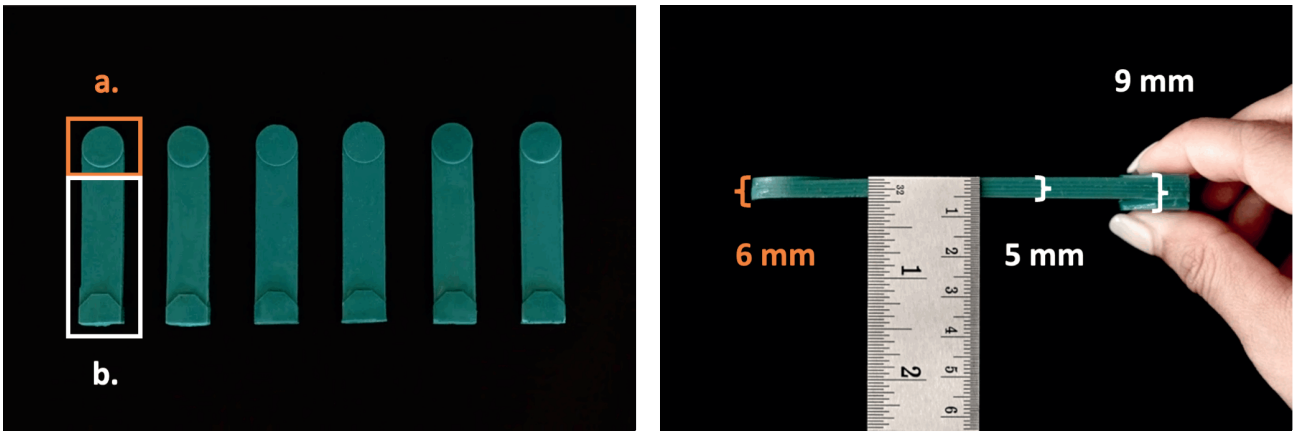


Figure 2. *The tongue force sensor consists of two subparts.*
a: the tongue force-receiving area, b: the handle.

2) Processing unit

The processing unit components are stored inside a waterproof and dust-resistant acrylic plastic box. It consists of two subparts:

2.1) Data processing equipment:

The researcher chose the DOIT ESP32 DevKit v1 Development Board[31] microcontroller for device circuitry. This compact microcontroller with Bluetooth 4.2 effectively transmits signals to display the application results.

2.2) Power supply:

The researcher chose a rechargeable lithium battery as the power source, using an adapter and a USB Type-A to Type-B Micro USB cable for charging. On top of the box, a lithium battery voltage indicator LED display was installed to indicate the battery status. The first version of our application was developed.

3) Display/recording unit

This device displays the results on a mobile application for the Android operating system, named 'CMU Tongue Strength & Endurance Measurement Device,' or simply 'CMU Tongue' (version 1.0). The application's functionality comprises three sections, as shown in Figure 3. The details of each section are as follows:

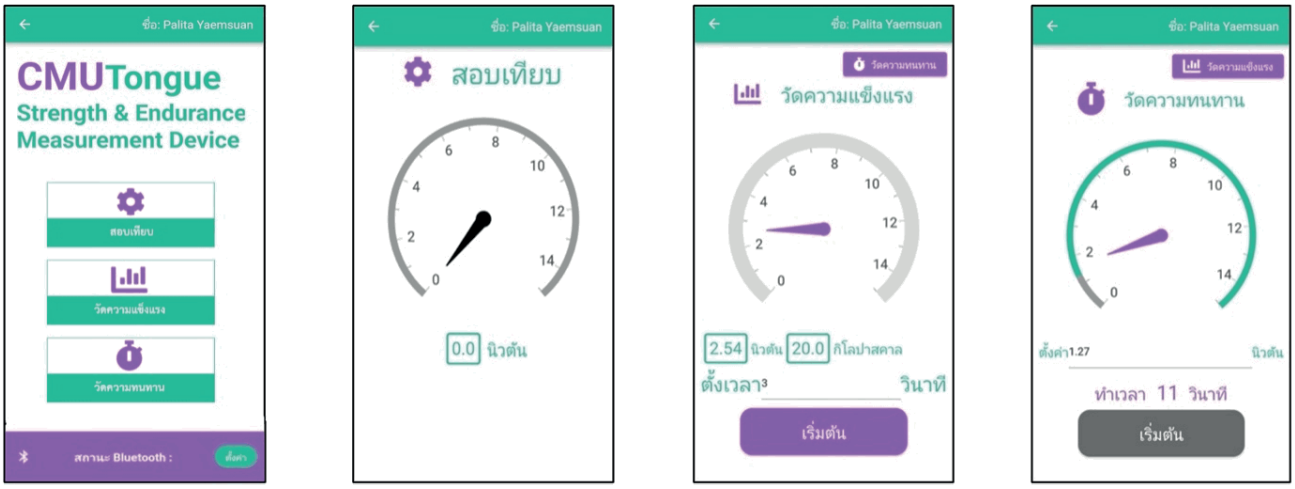


Figure 3. *An example of a 'CMU Tongue' (Version 1.1) application is displayed in each section.*

3.1) Calibration section:

Before each measurement, the researcher navigated to the 'Calibration' menu. A light finger press was applied to the force-receiving area, and the changing values were observed.

3.2) Display section:

The 'Measure Strength' function on the main menu was selected to measure tongue strength. A countdown was initiated, and participants were instructed to exert maximum force on the force-receiving area for 3 seconds. The resulting tongue strength was displayed in both N and kPa units. Subsequently, the endurance measurement was conducted. The researcher returned to the main menu and entered the 'Measure Endurance' tab. The researcher set the initial force level at 50% of the maximum tongue strength achieved in the previous strength measurement. If participants failed to sustain the tongue force within this range, the needle dropped, triggering the end of the test, and the endurance value was displayed in seconds.

3.3) Results recording section:

The researcher initiated the measurement process by registering user accounts, and each measurement was recorded under the respective user account. The data was stored internally on the device and backed up on the Firebase platform, enabling Google to store data online. Data can be downloaded as CSV or XLSX files for offline analysis using Microsoft Excel (version 2007 or higher). This study utilized Microsoft Excel for Mac version 16.81.

1.2 Accuracy testing

The device demonstrated accuracy within the 96.40% to 100% range, indicating its ability to measure tongue force close to the standard weight measures used in testing. The obtained percentages are within an acceptable threshold.[32]

1.3 Precision testing

Precision ranged from 0.90% to 4.21%, which is well within the acceptable limit of 5%, highlighting the high precision of the tongue pressure measurement component and its suitability for application as a reliable instrument in physiological measurements.[33]

1.4 Satisfaction questionnaire creation

All questionnaire items achieved a content validity index of 0.80 or higher.[34] After adjustments to improve clarity and incorporation of expert suggestions, the revised questionnaire was deemed suitable for further use.

Phase 2: Trials of the prototype device with six participants

The satisfaction scores from the participants regarding the intraoral component (tongue force sensor) were rated very satisfactory, ranging from 4.17 to 4.67. Similarly, satisfaction with the extraoral component (application) scores ranged from 3.83 to 4.67.[34] After considering participant feedback, the researcher and the BMEI team adjusted the tongue force sensor and updated the 'CMU Tongue' application from Version 1.0 to Version 1.1. These adjustments and improvements included:

- The silicone rim surrounding the tongue force sensor was modified. This adjustment involved modifying the silicone mixture to create a softer edge while maintaining the shape and preventing excessive tongue force absorption.
- The initial scale ranged from 0 to 30 N. Participants found it challenging to see their tongue force clearly, feeling they were applying less force. Therefore, the scale was adjusted to 0 to 15 N, and the display segments were made larger to improve

the visibility of the exerted force.

- Connectivity issues between the application and the processing unit were identified and resolved. This involved extending the Bluetooth connection time to ensure seamless data transmission between the app and the device. Additionally, a new protocol was implemented to automatically pair the app with the device whenever the researcher switched between measuring tongue strength and endurance.

Phase 3: Assessment test-retest reliability of the developed device and the investigation of tongue strength and endurance values with twenty participants

3.1 Maximum tongue strength and endurance variables in 20 participants

Table 1 shows the results of tongue strength measurements: females have an average anterior strength of 1.98±0.36 N (15.52±2.83 kPa) and a posterior strength of 2.29±0.80 N (17.94±6.27 kPa). Conversely, males exhibit greater tongue strength, with an anterior strength of 2.56±0.56 N (20.05±4.45 kPa) and a posterior strength of 2.51±0.96 N (19.67±7.50 kPa).

Regarding endurance, females have an average anterior endurance of 18.10±8.10 seconds and posterior endurance of 8.80±7.44 seconds. In contrast, males show markedly greater endurance, with anterior endurance of 28.90±17.32 seconds and posterior endurance of 13.30±5.95 seconds.

Table 1. Maximum tongue strength and endurance in healthy young adults, categorized by gender. (N=20)

Gender	Age	Tongue Strength				Tongue Endurance	
		Anterior (N)	Posterior (N)	Anterior (kPa)	Posterior (kPa)	Anterior (second)	Posterior (second)
Female (N=10)	23.90±3.41	1.98±0.36	2.29±0.80	15.52±2.83	17.94±6.27	18.10±8.10	8.80±7.44
Male (N=10)	26.60±4.20	2.56±0.56	2.51±0.96	20.05±4.45	19.67±7.50	28.90±17.32	13.30±5.95
Total (N=20)	25.25±3.97	2.27±0.55	2.40±0.87	17.79±4.31	18.81±6.79	23.50±14.28	11.05±6.95

Note: values are presented as mean±SD.

Table 2 shows statistical analysis using an independent t-test for the variance of tongue strength and endurance, using Levene's test with a 95% confidence interval, revealed that the variance of males and females was equal. The comparison of mean values of tongue strength and endurance between males and females revealed statistically significant differences (p<0.01), specifically in anterior tongue strength.

Table 2. Difference in tongue strength and endurance between genders. (n=20)

Measurement		Female (N=10)	Male (N=10)	t	Sig. (2-tailed)	Mean difference	95% CI	
							Upper	Lower
Tongue strength (N)	Anterior	1.98±0.36	2.56±0.56	2.74	0.01*	0.58	0.13	1.02
	Posterior	2.29±0.80	2.51±0.96	0.57	0.58	0.22	-0.61	1.05
Tongue endurance (second)	Anterior	18.10±8.10	28.90±17.32	1.79	0.09	10.80	-1.91	23.51
	Posterior	8.80±7.44	13.30±5.95	1.50	0.15	4.50	-1.83	10.83

**The mean difference is significant at the 0.01 level.*

3.2 Test-retest reliability

Table 3 shows the test-retest reliability analysis, which revealed that the intraclass correlation coefficients (ICCs) for anterior tongue strength, posterior tongue strength, anterior tongue endurance, and posterior tongue endurance were 0.61, 0.90, 0.93, and 0.56, respectively. These values indicate that the measurements obtained from the developed device were consistent across repeated administrations ranging from moderate to excellent.[35]

Table 3. Intraclass Correlation Coefficients (ICCs) for tongue strength and endurance.

	Intraclass Correlation[b]	95% CI		F Test with True Value 0			
		Lower	Upper	Value	df1	df2	Sig
Anterior tongue strength	0.61[a]	0.25	0.83	4.18	19	19	0.00
Posterior tongue strength	0.90[a]	0.76	0.96	18.42	19	19	0.00
Anterior tongue endurance	0.93[a]	0.83	0.97	26.56	19	19	0.00
Posterior tongue endurance	0.56[a]	0.17	0.80	3.55	19	19	0.00

Note: Two-way mixed effects model where people effects are random, and measures effects are fixed. a: the estimator is the same, whether the interaction effect is present, b: type C intraclass correlation coefficients using a consistency definition. The between-measure variance is excluded from the denominator variance.

Discussion

In designing and developing the device, the researchers considered various factors, including the device production cost, practical utilization, maintenance requirements, and measurement quality.

First, to minimize production costs, the researchers used commercially available components, including force sensing resistors, a microcontroller, and USB cables, which were adjusted to meet the operational requirements. Some elements, namely the sensor silicone casing and processing box, were newly developed to meet specific requirements. The production cost of the prototype device was approximately 15,000 baht, significantly lower than the imported device set priced at 137,075 baht (as confirmed by the quotation from the distributor in Thailand to the Department of Occupational Therapy, Faculty of Associated Medicine Sciences, Chiang Mai University, on November 29, 2021).

Second, in terms of usability, limitations were identified in the current standard device's tongue bulb and handle, which might lead to imprecision in measurements or difficulties in measuring the strength and endurance of the posterior tongue region. To address this, the researchers designed a flat circular tongue force-receiving area with a diameter not exceeding the average width between the maxillary intercanine teeth.[36] This design allows for tongue strength and endurance measurement across all palate sizes without the need for individualized sensor molding. Additionally, the handle was designed to be approximately 5 mm thick, consistent with the findings of Arakawa *et al*.,[37] who compared the IOPI and JMS instruments and highlighted the advantages of the JMS handle, providing a more secure grip compared to the plastic tubing handle of the IOPI device. Upon comparison with this developed device, it was found that the tongue force sensor handle, approximately 5 mm thick, was comparable to the JMS handle, contributing to a secure grip and minimizing position changes during measurement. Furthermore, the researchers designed the thickness difference in the silicone covering of the tongue force-receiving area and handle to help identify the appropriate position during measurement, allowing the participants to exert tongue force more accurately. Finally, the overall design of the entire device comprises a few components, with a simple method of connecting the device to a smartphone, enabling other therapists and patients to readily use the device immediately after receiving instructions and demonstrations.

Next, the maintenance factor was considered by designing each component to be detachable for easy cleaning. The tongue force sensor is classified as a medium-risk device for infection. After detachment, it can be easily cleaned by rinsing with clean water and then disinfected by immersion in Opal® solution containing ortho-phthalaldehyde (OPA) at a concentration of 0.57%, for 6 minutes at 20 °C.[38] This disinfectant solution is commonly used in hospitals and clinics. As for the processing box and smartphone used for displaying and recording the results through the application, these components can be wiped clean with 70% ethanol solution after detachment.

Finally, regarding the device measurement quality testing, it was found that all six tongue force sensors exhibited acceptable accuracy,[32] with the Coefficient of Variation (%CV) suggesting a high level of consistency[33] and the test-retest reliability indicating consistent measurements over repeated tests.[35] These findings confirm that the developed device is reliable and suitable for clinical assessment and rehabilitation planning for tongue function.

When considering the tongue function values obtained from 20 healthy young adults, the preliminary findings revealed that males exhibited higher tongue strength and endurance in both the anterior and posterior regions than females. However, the statistical significance was only observed in anterior tongue strength. Previous studies on gender-related factors influencing tongue strength have presented varying results. Moreover, this study was consistent with Arakawa's findings,[15] suggesting significant differences in tongue strength between genders in individuals under 60, possibly due to males having greater overall muscle mass, including tongue muscles. Additionally, there were no statistically significant gender differences in tongue endurance values, which was consistent with prior studies.[39,40] Thus, these tongue strength and endurance values provide preliminary support for this developed device and serve as a basis for further research using larger sample groups in the future.

Limitation

The present study did not investigate the correlation between tongue strength and endurance measurements obtained from the developed device and those obtained

from imported standard devices. This limitation was due to the constraints of the study timeline, and the study collected data during the severe outbreak situation of the COVID-19 pandemic.

Conclusion

The developed tongue strength and endurance measurement device through this study has a lower cost of development compared to imported commercial devices while maintaining performance standards and demonstrating acceptable accuracy, precision, and reliability. This device has high potential for use in clinical settings, allowing patients more access to objective assessments of tongue strength and endurance at a lower examination cost. This device's objective evaluation may help develop efficient rehabilitation programs to improve tongue function in the elderly and individuals with tongue-related disorders.

Conflict of interest

The authors have declared that no competing interests existed at the time of publication.

Ethical approval

The study was approved by the Research Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University (Approval ID: AMSEC-64FB-004). All participants received all necessary information related to the research, and informed written consent was obtained before enrolling.

Acknowledgements

This work was supported by the Faculty of Associated Medical Sciences and Biomedical Engineering Institute, Chiang Mai University, Thailand. Sincere thanks to all participants for their willingness to participate in this study.

References

[1] Matsuo K, Palmer JB. Anatomy and physiology of feeding and swallowing: normal and abnormal. Phys Med Rehabil Clin N Am. 2008; 19(4): 691-707. doi 10.1016/j.pmr.2008.06.001.

[2] Youmans SR, Stierwalt JA. Measures of tongue function related to normal swallowing. Dysphagia. 2006; 21(2): 102-11. doi 10.1007/s00455-006-9013-z.

[3] Groher ME. Normal swallowing in adults. In: Groher ME, Crary MA, editors. Dysphagia: clinical management in adults and children. 2nd Ed. St. Louis, Missouri: Elsevier; 2016. p. 19-40.

[4] Robbins J, Levine R, Wood J, Roecker EB, Luschei E. Age effects on lingual pressure generation as a risk factor for dysphagia. J Gerontol A Biol Sci Med Sci. 1995; 50(5): 257-62. doi 10.1093/gerona/50a.5.m257.

[5] Youmans SR, Youmans GL, Stierwalt JAG. Differences in tongue strength across age and gender: Is there a diminished strength reserve? Dysphagia. 2009; 24(1): 57-65. doi 10.1007/s00455-008-9171-2.

[6] Kays SA, Hind JA, Gangnon RE, Robbins J. Effects of dining on tongue endurance and swallowing-related outcomes. J Speech Lang Hear Res. 2010; 53(4): 898-907. doi 10.1044/1092-4388(2009/09-0048).

[7] Namasivayam AM, Steele CM, Keller H. The effect of tongue strength on meal consumption in long term care. Clin Nutr. 2016; 35(5): 1078-83. doi 10.1016/j.clnu.2015.08.001.

[8] Maeda K, Akagi J. Decreased tongue pressure is associated with sarcopenia and sarcopenic dysphagia in the elderly. Dysphagia. 2015; 30(1): 80-7. doi 10.1007/s00455-014-9577-y.

[9] Stierwalt JAG, Youmans SR. Tongue measures in individuals with normal and impaired swallowing. Am J Speech Lang Pathol. 2007; 16(2): 148-56. doi 10.1044/1058-0360(2007/019).

[10] Minagi Y, Ono T, Hori K, Fujiwara S, Tokuda Y, Murakami K, *et al*. Relationships between dysphagia and tongue pressure during swallowing in Parkinson's disease patients. J Oral Rehabil. 2018; 45(6): 459-66. doi 10.1111/joor.12626.

[11] Lazarus CL, Logemann JA, Pauloski BR, Rademaker AW, Larson CR, Mittal BB, et al. Swallowing and tongue function following treatment for oral and oropharyngeal cancer. J Speech Lang Hear Res. 2000; 43(4): 1011-23. doi 10.1044/jslhr.4304.1011.

[12] Hasegawa Y, Sugahara K, Fukuoka T, Saito S, Sakuramoto A, Horii N, et al. Change in tongue pressure in patients with head and neck cancer after surgical resection. Odontology. 2017; 105(4): 494-503. doi 10.1007/s10266-016-0291-0.

[13] Adams V, Mathisen B, Baines S, Lazarus C, Callister R. A systematic review and meta-analysis of measurements of tongue and hand strength and endurance using the Iowa Oral Performance Instrument (IOPI). Dysphagia. 2013; 28(3): 350-69. doi 10.1007/s00455-013-9451-3.

[14] McCormack J, Casey V, Conway R, Saunders J, Perry A. OroPress a new wireless tool for measuring oro-lingual pressures: a pilot study in healthy adults. J Neuro Engineering Rehabil. 2015; 12(1):32.

[15] Arakawa I, Igarashi K, Imamura Y, Müller F, Abou-Ayash S, Schimmel M. Variability in tongue pressure among elderly and young healthy cohorts: A systematic review and meta-analysis. J Oral Rehabil. 2021;48(4): 430-448. doi 10.1111/joor.13076.

[16] IOPI Medical. Product information. 2020 [cited 2020 Dec 26]. Available from: https://iopimedical.com/medical-professionals/.

[17] JMS Co. Ltd. Product information. 2011 [cited 2020 Dec 26]. Available from: http://orarize.com/zetsuatsu/product.html.

[18] AIM Technologies LLC. Digital Swallowing Workstation™ (DSW), model 7200. 2020 [cited 2020 Dec 24]. Available from: http://aimtech.ru/en/catalog/101.

[19] Swallow Solutions LLC. SwallowSTRONG device. 2017 [cited 2020 Dec 24]. Available from: http://www.swallowsolutions.com/product-information/swallow-strong-device.

[20] Hori K, Ono T, Tamine K, Kondo J, Hamanaka S, Maeda Y, *et al*. Newly developed sensor sheet for measuring

tongue pressure during swallowing. J Prosthodont Res. 2009; 53(1): 28-32. doi 10.1016/j.jpor.2008.08.008.

[21] Sonnanta S. Measurement instruments and electrical measurements. Bangkok: SE-Education; 2000. (in Thai).

[22] Kanyawattana N, Nakham N. Guidelines for the verification of applicability of chemical testing methods. Bangkok: Department of Primary Industries and Mines; 2012. (in Thai).

[23] Pitsanlayanon M, Sitthithamwilai W, Pitsittaksak C, Waliprakon P, Itthasakul P, Pitithammaphon A. Effectiveness of Rama disposable bite block in patients undergoing electroconvulsive therapy: A pilot study. J Psychiatr Assoc Thailand [internet]. 2021 Jan 9 [cited 2021 Feb 15]; 65(3): 279-88. Available from: https://he01.tci-thaijo.org/index.php/JPAT/article/view/242152 (in Thai)

[24] Hongsanun W, Insuk S. Quality assessment criteria for mobile health apps: A systematic review. WJST. 2020; 17(8): 745-59. doi 10.48048/wjst.2020.6482

[25] Turner RC, Carlson L. Indexes of item-objective congruence for multidimensional items. Int J Test. 2003; 3(2): 163-71. doi 10.1207/S15327574IJT0302_5.

[26] Edmiaston J, Connor LT, Loehr L, Nassief A. Validation of a dysphagia screening tool in acute stroke patients. Am J Crit Care. 2010; 19(4): 357-64. doi 10.4037/ajcc2009961.

[27] The Dental Association of Thailand. Guidelines for the control and prevention of COVID-19 infections in dentistry version 1. 2020 [cited 2021 Apr 5]. Available from: https://www.thaidental.or.th/main/download/upload/upload-204211344051371.pdf.

[28] World Health Organization (WHO). Cleaning and disinfection of environmental surfaces in the context of COVID-19. 2020 [cited 2021 Apr 5]. Available from: https://www.who.int/publications/i/item/cleaning-and-disinfection-of-environmental-surfaces-in-the-context-of-covid-19.

[29] IBM. How to cite IBM SPSS Statistics or earlier versions of SPSS 2014 [cited 2024 May 24]. Available from: https://www.ibm.com/support/pages/how-cite-ibm-spss-statistics-or-earlier-versions-spss.

[30] Interlink Electronics Inc. FSR 400 series data sheet. 2019 [cited 2022 May 25]. Available from: https://files.seeedstudio.com/wiki/Grove-Round_Force_Sensor_FSR402/res/FSR402.pdf.

[31] Zerynth documentation. DOIT Esp32 DevKit v1. 2021 [cited 2021 Jan 15]. Available from: https://docs.zerynth.com/latest/reference/boards/doit_esp32/docs/.

[32] Webster JG, Eren H, Webster JG, Eren H. Measurement, instrumentation, and sensors handbook: spatial, mechanical, thermal, and radiation measurement. 2nd Ed. Boca Raton, FL: CRC Press, Taylor & Francis Group; 2014.

[33] Saengnil K, Posunnont P."Reliability, validity, accuracy and precision on exercise physiology research" Method agreement and measurement error in the physiology of exercise. Veridian E-journal, Silpakorn University, Science and Technology Branch [internet]. 2018 Dec 12 [cited 2021 Jan 24]; 5(6): 1-19. Available from: https://ph01.tci-thaijo.org/index.php/VESTSU/article/view/159850 (in Thai).

[34] Srisatitnarakul B. Development and verification of research instruments: psychometric measurement properties. Bangkok: Chulalongkorn University Printing House; 2012. (in Thai).

[35] 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.

[36] Sedtasuppana A, Nunthayanon K, Prasitsak T, Tanaslarak R, Satrawaha S, Piyapattamin T. Comparison of arch widths measurements made on digital and plaster models. NUJST [internet]. 2017 Sep 15 [cited 2021 Jan 23]; 25(4): 9-16. Available from https://ph03.tci-thaijo.org/index.php/ahstr/article/view/1653

[37] Arakawa I, Abou-Ayash S, Genton L, Tsuga K, Leles CR, Schimmel M. Reliability and comparability of methods for assessing oral function: Chewing, tongue pressure and lip force. J Oral Rehabil. 2020; 47(7): 862-71. doi 10.1111/joor.12976.

[38] Center for Disease Control and Prevention (CDC). Guideline for disinfection and sterilization in healthcare facilities. 2008 [cited 6 Apr 2021]. Available from: https://www.cdc.gov/infection-control/media/pdfs/Guideline-Disinfection-H.pdf

[39] Hao G, Chih Y-C, Ni A, Harada K, Thompson J, Chen S-C, *et al*. Maximum isometric tongue strength and tongue endurance in healthy adults. Oral Sci Int. 2023;20(2):115-24. doi 10.1002/osi2.1158

[40] Jeong DM, Shin YJ, Lee NR, Lim HK, Choung HW, Pang KM, *et al*. Maximal strength and endurance scores of the tongue, lip, and cheek in healthy, normal Koreans. J Korean Assoc Oral Maxillofac Surg. 2017; 43(4): 221-8. doi 10.5125/jkaoms.2017.43.4.221

Journal of Associated Medical Sciences 2024; 57(3): 52-64

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

A genetically engineered mouse/human chimeric antibody targeting CD99 enhances antibody-dependent cellular phagocytosis against human mantle cell lymphoma Z138 cells

Kamonporn Kotemul[1] Supansa Pata[1,2] Witida Laopajon[1,2] Watchara Kasinrerk[1,2] Nuchjira Takheaw[1,2*]

[1]Division of Clinical Immunology, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Biomedical Technology Research Center, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency at the Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 2 April 2024
Accepted as revised 19 June 2024
Available online 21 June 2024

Keywords:
Cancer immunotherapy, antibody therapy, mantle cell lymphoma, anti-CD99 mAb, chimeric antibody.

ABSTRACT

Background: Mantle cell lymphoma (MCL) is an aggressive form of B-cell non-Hodgkin lymphoma. The elimination of MCL cells via phagocytosis is essential for cancer eradication. Therefore, discovering novel targeted antibodies that can induce phagocytosis is needed. We have demonstrated that our in-house-produced mouse anti-CD99 mAb clone MT99/3 could induce potent anticancer activities against MCL cell lines in both *in vitro* and *in vivo* mouse xenograft models. Nevertheless, for use in humans, the mouse mAb needs to be transformed into a mouse/human chimeric mAb that contains a human Fc region to activate human immune effector functions, especially macrophage-mediated phagocytosis. Antibody-dependent cellular phagocytosis (ADCP) mediated by mouse/human chimeric mAb MT99/3 against MCL has not been previously reported.

Objective: This study aimed to genetically engineer a mouse/human chimeric antibody against human CD99 derived from mouse mAb MT99/3 and to evaluate its effect in mediating the ADCP mechanism for eradicating MCL cells *in vitro* using monocyte-derived macrophages.

Materials and methods: The expression plasmid to produce chimeric anti-CD99 antibody, ChAbMT99/3, was constructed by fusing the variable domains of mouse mAb MT99/3 with the constant domains of human IgG1 and the constant domains of kappa light chain. ChAbMT99/3 was expressed in the stable human expression system based on HEK293T cells. ChAbMT99/3 was purified from the culture supernatant of ChAbMT99/3-expressing HEK293T cells using Protein G chromatography. The purity and structure of ChAbMT99/3 were verified by SDS-PAGE and western blotting. The binding specificity and activity were determined by staining with cells expressing recombinant and native human CD99. The anticancer activity of ChAbMT99/3 in mediating the ADCP mechanism against MCL cell line Z138 using human monocyte-derived macrophages was evaluated.

Results: We successfully constructed the plasmid to produce ChAbMT99/3. Human HEK293T cells stably expressing ChAbMT99/3 were established. The ChAbMT99/3-expressing HEK293T cells could secrete ChAbMT99/3 into the culture supernatant. The high purity and complete IgG structure of ChAbMT99/3 were obtained from the purification process. Crucially, this chimeric antibody retained its binding reactivity to recombinant and native human CD99. In addition, the produced ChAbMT99/3, upon binding to MCL cells, significantly enhanced ADCP against MCL cell line Z138 in a dose-dependent manner.

Conclusion: The production of a mouse/human chimeric antibody against human CD99 derived from mouse mAb MT99/3 was successful. The engineered antibody could mediate ADCP activity against MCL cells. The produced ChAbMT99/3 might be a promising therapeutic candidate for MCL treatment.

∗ **Corresponding contributor.**
Author's Address: *Division of Clinical Immunology, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: nuchjira.t@cmu.ac.th
doi: 10.12982/JAMS.2024.047
E-ISSN: 2539-6056

Introduction

Mantle cell lymphoma (MCL) is a rare and aggressive form of B-cell non-Hodgkin lymphomas (NHLs). MCL is genetically characterized by the *translocation of t(11;14) (q13;q32)*, leading to the overexpression of cyclin D1 and subsequent dysregulation of the cell cycle.[1] MCL accounts for 3-10% of all NHL cases, with a higher prevalence in middle-aged to older patients.[2] The treatment for MCL depends on age, overall performance status, and underlying co-morbidities.[3] Standard regimens include intensified chemotherapy in combination with rituximab, a chimeric antibody against CD20.[4] Most MCL patients respond to rituximab-based regimens, and remarkable clinical efficacy has been observed. However, patients often relapse due to drug resistance.[5] Therefore, novel targeted antibodies need to be established as therapeutic options for mantle cell lymphoma.

In the current landscape of cancer treatment, monoclonal antibodies (mAbs) have emerged as a promising therapeutic option due to their specificity and diverse mechanisms of action.[6,7] The Fc region of an antibody plays a crucial role in activating immune effector functions.[6] One of the essential Fc-related mechanisms for cancer eradication is antibody-dependent cellular phagocytosis (ADCP).[8] This mechanism is involved in activating Fc-gamma receptors (FcγRs) on macrophage surfaces by Fc region of antibody that recognizes cancer cells, resulting in the internalization and degradation of the cancer cells.[9] Several therapeutic antibodies demonstrated a significant increase in macrophage-mediated phagocytosis.[10] These antibodies are recognized as effective drugs in various cancers. However, it has been reported that MCL cells upregulate the "don't eat me" CD47 molecule, leading to resistance to macrophage-mediated phagocytosis.[11] Hence, therapeutic antibodies that mediate phagocytosis against MCL are still required.

As CD99 expression on MCL was reported, we have investigated the anticancer activities of mouse anti-CD99 mAb clone MT99/3 against MCL.[12] We demonstrated that the mAb MT99/3 could activate host immune effectors via antibody-dependent cellular cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC) mechanisms *in vitro*, leading to the killing of MCL cell lines. Moreover, it could inhibit the growth of MCL in a mouse xenograft model *in vivo*.[12] These results indicated the potential of mAb MT99/3 in treating mantle cell lymphoma. However, the anticancer activity of mAb MT99/3 in the eradication of MCL cells by the ADCP mechanism has not been investigated.

For human use, mouse mAb is not appropriate. The mouse mAb, upon treatment, will be a foreign substance and induce adverse effects.[13,14] Therefore, the transformation of mouse mAb into a mouse/human chimeric mAb is required. The chimeric mAb contains a human Fc region, while the variable domains are still a mouse mAb to retain its binding reactivity. The transformed antibody could also be used to investigate the effect of an antibody in mediating phagocytosis by macrophage.

In this study, we generated a genetically engineered mouse/human chimeric antibody against human CD99 named ChAbMT99/3 by fusing the variable domain of mouse mAb MT99/3 with the constant domains of human IgG1 and kappa light chain. We evaluated the ADCP mediated by the generated chimeric antibody ChAbMT99/3 against the MCL cell line Z138. Our results demonstrated the potential of ChAbMT99/3 as an applicable therapeutic option for eradicating mantle cell lymphoma by phagocytosis.

Materials and methods

Antibodies

The anti-human CD99 mAb clone MT99/3 (mouse IgG2a) was generated in our laboratory.[15] PE/cyanine7-conjugated anti-CD3 mAb, PE-conjugated anti-CD19 mAb, PE-conjugated anti-CD56 mAb, PerCP-conjugated anti-CD14 mAb were purchased from BioLegend (San Diego, CA, USA). PE-conjugated anti-CD4 mAb, and APC-conjugated anti-CD8 mAb were purchased from BD Bioscience (San Jose, CA, USA)

Cell lines

Mantle cell lymphoma Z138 cells were obtained from JCRB cell bank, Osaka, Japan (gift from Prof. Dr. Seiji Okada, Kumamoto University, Kumamoto, Japan). Jurkat E6.1 cells were purchased from the American Type Culture Collection (ATCC). Both cell lines were cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum (FBS) and antibiotics. CD99-expressing myeloma cells were generated in-house.[16] Myeloma and CD99-expressing myeloma cells were cultured in a 10% FBS-IMDM medium. HEK293T cells were cultured in a 10% FBS-DMEM medium. All cell lines were maintained at 37 °C in a 5% CO_2 incubator.

Construction of expression plasmid for chimeric antibody against human CD99

The plasmid pTRIOZ-hIgG1 (InvivoGen, San Diego, CA, USA) was utilized for construction. Nucleotide sequences of the variable heavy chain (VH) and variable light chain (VL) obtained from mouse mAb MT99/3-producing hybridoma clone were codon-optimized. These sequences were designed to be inserted into the pTRIOZ-hIgG1 plasmid using SnapGene software (GSL Biotech LLC, San Diego, CA, USA). The codon-optimized VH chain sequences were synthesized with AgeI and NheI restriction sites, while VL chain sequences had SgrAI and BsiWI sites at the 5′ and 3′ ends, respectively. The VL and VH genes were cloned into the plasmid pTRIOZ-hIgG1 in front of genes for constant kappa light chain (CL) and constant heavy chain (CH), respectively, using specific restriction enzymes. The expression plasmid for ChAbMT99/3 was named pTRIOZ-ChAbMT99/3. After construction, the VH and VL nucleotide sequence was verified using Sanger sequencing. The codon optimization, gene synthesis, construction, and verification of the pTRIOZ-ChAbMT99/3 plasmid were carried out by GenScript Biotech (Nanjing, China).

Transformation and verification of the constructed plasmid

Using the heat-shock method, the constructed plasmid pTRIOZ-ChAbMT99/3 was transformed into competent *Escherichia coli (E. coli)* DH5-α. Plasmid-harboring *E. coli* were selected by spreading onto LB agar (BD Biosciences) containing 25 µg/mL of zeocin. A single colony was selected and inoculated into LB broth (BD Biosciences) with zeocin. Plasmids were purified using the QIAprep Spin Miniprep Kit (QIAGEN, Germany) according to the manufacturer's instructions. Purified plasmids were verified by digestion with *Bam*HI and *Eco*RI, and the size was determined by agarose gel electrophoresis.

Transfection and verification of antibody expression in transfected HEK293T cell line

The plasmid pTRIOZ-ChAbMT99/3 was transfected into the HEK293T cells. In brief, HEK293T cells (2×10^5) were seeded into a 6-well plate (Corning Inc., NY, USA) and cultured for 2 days. Transfection was performed by mixing 0 ng (as control) or 500 ng of plasmids with 3 µL of Lipofectamine® 2000 (Invitrogen, Carlsbad, CA, USA), following the manufacturer's procedure. Then, cells were incubated at 37°C in a 5% CO_2 incubator for 3 days. The transfected cells were intracellular immunofluorescence stained to detect antibody expression. In brief, transfected cells (1×10^5) were fixed with 4% paraformaldehyde for 15 mins at room temperature (RT), then permeabilized using 0.1% saponin. Then, the cells were blocked the Fc receptors with 10% FBS and stained with Alexa Fluor 488-conjugated goat anti-human IgG antibody (Jackson ImmunoResearch Laboratories, West Grove, PA, USA) or Alexa Fluor 488-conjugated goat anti-rabbit IgG antibody (negative control) (Life technology, Eugene, OR, USA), for 30 mins on ice. The stained cells were assessed by flow cytometry (Accuri™ C6 plus flow cytometer, BD Biosciences).

Additionally, culture supernatant of un-transfected and transfected HEK293T cells was collected for testing binding reactivity on CD99-positive Jurkat E6.1 cells. In brief, Jurkat E6.1 cells (5×10^5) were incubated with 10% FBS, followed by staining with culture supernatant. Then, Alexa Fluor 488-conjugated goat anti-human IgG antibody (Jackson ImmunoResearch Laboratories) was added, with subsequent analysis using the Accuri™ C6 plus flow cytometer.

Establishment of HEK293T cells stably expressing chimeric anti-CD99 antibody

The transfected HEK293T cells were seeded at a density of 100 cells per well in 150 µL of 10% FBS-DMEM with 100 µg/mL of zeocin in a 96-well plate. Every 3 days, a medium with zeocin at a concentration of 300 µg/mL was added into the wells at a volume of 50 µL. Subsequently, surviving cells in the selected wells were harvested to determine the expression of the chimeric antibody using intracellular immunofluorescence staining.

Production and purification of chimeric antibody against human CD99

The selected ChAbMT99/3-expressing HEK293T clone was cultured in a 75 cm^2 T-flask and maintained in 10% FBS-DMEM with 100 µg/mL of zeocin. Upon reaching 80% confluence, the cells were washed with DMEM. The culture medium was changed to 293-SFM II media (Invitrogen) with 100 µg/mL of zeocin. After that, cells were cultured by frequent mixing for 5 days. The culture supernatant was harvested, centrifuged, and filtered through a 0.2 µm filter. Antibody purification was performed using a Hitrap protein G column with an ÄKTA start purification system, followed by buffer exchange with PBS. Concentrations of the purified antibody were determined by a Nanodrop® 2000 spectrophotometer. Twenty micrograms of protein per condition were used to assess the purity and structure by SDS-PAGE and western blotting. For western blotting, the proteins from SDS-PAGE were transferred to PVDF membranes. The antibody structure was determined using HRP-conjugated goat anti-human IgG, Fcγ fragment-specific antibodies (Jackson ImmunoResearch Laboratories), followed by detection with SuperSignal™ West Pico Plus Chemiluminescent Substrate (Thermo Fisher Scientific, Waltham, MA, USA).

Verifying the binding reactivity of the purified chimeric antibody against human CD99

Myeloma cells and CD99-expressing myeloma cells (5×10^5), were incubated with 10% FBS. Then, cells were stained with the purified ChAbMT99/3 or an isotype-matched control Ab (human IgG). After washing, Alexa Fluor 488-conjugated goat anti-human IgG antibody was added. The stained cells were then analyzed using the Accuri™ C6 plus flow cytometer.

Mantle cell lymphoma Z138 cells were incubated with 10% FBS. The cells were stained with ChAbMT99/3 and mAb MT99/3 at final concentrations of 0.2, 1 and 5 µg/mL. Consequently, cells were incubated with Alexa Fluor 488-conjugated goat anti-human IgG antibody for the detection of bound ChAbMT99/3 or Alexa Fluor 488-conjugated goat anti-mouse Igs antibody for the detection of bound mouse mAb MT99/3 (Invitrogen, Eugene, OR, USA). The stained cells were analyzed using the Accuri™ C6 plus flow cytometer.

For human peripheral blood mononuclear cells (PBMCs) staining, PBMCs were prepared using Ficoll-Hypaque gradient centrifugation. PBMCs were blocked with 20% human AB serum and then incubated with biotinylated ChAbMT99/3 or a biotinylated human IgG (isotype-matched control Ab). After washing, cells were stained with FITC-conjugated streptavidin (BioLegend) along with antibodies for the identification of $CD4^+$ T cells ($CD3^+CD4^+$), $CD8^+$ T cells ($CD3^+CD8^+$), B cells ($CD19^+$), NK cells ($CD3^-CD56^+$), NKT cells ($CD3^+CD56^+$) and monocytes ($CD14^+$). The stained cells were analyzed using the Accuri™ C6 plus flow cytometer.

Determination of binding affinity of chimeric antibody against human CD99

To determine the binding affinity of ChAbMT99/3 and compare it with its parental mAb MT99/3, streptavidin-coated biosensors (Pall Life Sciences, Port Washington, NY, USA) were employed to immobilize the biotinylated CD99 extracellular domain (ECD) peptide.[17] The loading step involved pre-wetting the biosensor tip with 0.05% Tween -PBS (PBST), followed by immersion in a solution containing 10 µg/mL of CD99 ECD peptide for 120 sec. The excess peptide was subsequently removed by immersing the tip in PBST. For the association step, the peptide-coated biosensors were incubated with a serial two-fold dilution of ChAbMT99/3 or mouse MT99/3, ranging from 10 µg/mL to 0.625 µg/mL, for 120 sec. This step was followed by dissociation in PBST for 300 sec. The entire experiment was conducted using an Octet® BLI System (Pall ForteBio), and the acquired data were analyzed using ForteBio data analysis software version 9.0 to determine the equilibrium dissociation constant (K_D) of antibodies.

Monocyte-derived macrophages

Monocyte-enriched PBMCs were prepared using Percoll gradient centrifugation. In brief, PBMCs at a concentration of 1×10^7 cells/mL were overlaid on Percoll working solution (GE Healthcare). Cells were centrifuged at 895×g, 40 min, with break-off at RT. The ring of monocyte-enriched PBMCs was collected. The cell numbers were counted using Turk's solution. The percentages of monocytes were determined by flow cytometry. The monocyte-enriched PBMCs were plated in a 24-well plate at 5×10^5 monocytes per well and incubated at 37°C in a 5% CO_2 incubator for 1 hr. Thereafter, non-adherent cells were removed by washing with PBS. The adherent cells were stimulated with human macrophage colony-stimulating factor (M-CSF) in 10% FBS-RPMI at 50 ng/mL, 400 µL for 6 days. Of note, fresh medium containing M-CSF was renewed every 2 days.

Assay for antibody-dependent cellular phagocytosis

Z138 cells used as target cells were labeled with CFSE at a concentration of 2 µM. CFSE-labeled Z138 were incubated with a final concentration of isotype-matched control Ab (human IgG) and ChAbMT99/3 at 0.2, 1, or 5 µg/mL or medium without antibody. After incubation at RT for 15 mins, the excess antibodies were removed by centrifugation at 2100×g for 4 mins. The antibody-treated Z138 cells were co-cultured with monocyte-derived macrophages (effector cells) at an effector-to-target ratio (E:T) of 1:4 and incubated at 37 °C in a 5% CO_2 incubator for 4 hs. Next, the cells were harvested and stained with PerCP-conjugated anti-CD14 mAb. The percentage of phagocytosis was assessed by flow cytometry.

Statistical analysis

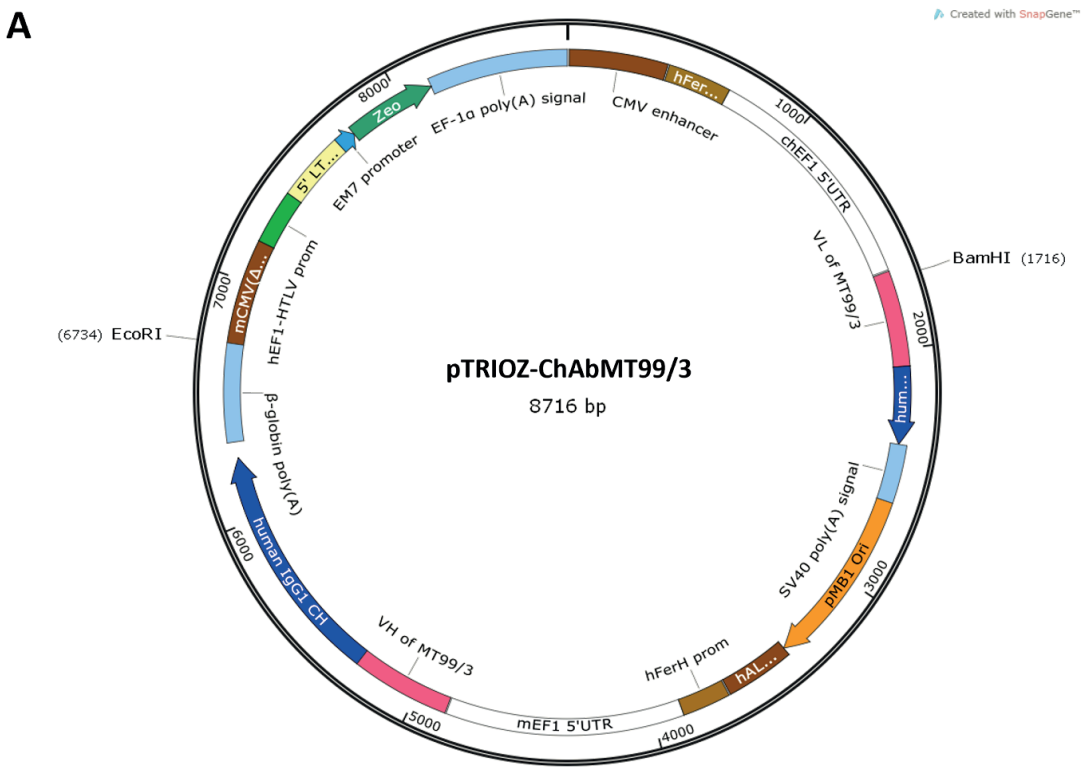
The data are expressed as mean±SD. The data were analyzed by One-way ANOVA with Tukey's multiple comparisons using GraphPad Prism Version 9.5.1 (GraphPad Software, San Diego, CA). $p<0.05$ was considered statistically significant.

Results

Construction of expression plasmid for chimeric anti-CD99 antibody production

The variable domains of mouse mAb MT99/3 were genetically engineered by fusing them with the constant domains of human IgG1 and kappa light chain. The VH and VL genes of the mouse mAb MT99/3 were codon-optimized, synthesized, and cloned into heavy/light chain cassettes of plasmid pTRIOZ-hIgG1, as illustrated in Figure 1A. DNA sequencing of VH and VL genes in the constructed plasmid pTRIOZ-ChAbMT99/3 was performed to validate the nucleotide sequences after synthesis and cloning. DNA sequencing results were translated into amino acid sequences before alignment with original amino acid sequences derived from mouse mAb MT99/3. As expected, amino acid sequences of the synthesized VH and VL genes showed a 100% identity match to the parental sequences (Figure 1B). The results confirmed that the plasmid pTRIOZ-ChAbMT99/3 was successfully constructed.

Plasmids were transformed into competent *E. coli* to amplify the constructed pTRIOZ-ChAbMT99/3, and zeocin was used for clonal selection. Purified plasmids were verified by digestion with *Bam*HI and *Eco*RI. DNA bands at approximately 3698 bp and 5018 bp were observed (Figure 1C), corresponding to the predicted sizes. The result indicated that the amplified plasmid could be further used for antibody expression.



B Amino acid sequence of variable heavy chain

Score	Expect	Method	Identities	Positives	Gaps
287 bits(734)	3e-105	Compositional matrix adjust.	136/136(100%)	136/136(100%)	0/136(0%)

```
Query  1    MKCSWVIFFLMAVVTGVNSDVQLQQSGAELVKPGASVRLSCTLSGFNIKDTYIHWVNQRP  60
            MKCSWVIFFLMAVVTGVNSDVQLQQSGAELVKPGASVRLSCTLSGFNIKDTYIHWVNQRP
Sbjct  1    MKCSWVIFFLMAVVTGVNSDVQLQQSGAELVKPGASVRLSCTLSGFNIKDTYIHWVNQRP  60

Query  61   EQGLEWIGRIDPGNGNIKYDPKFQGKATITADTSSNKVYLHLSSLTSEDTAVYYCARSGG  120
            EQGLEWIGRIDPGNGNIKYDPKFQGKATITADTSSNKVYLHLSSLTSEDTAVYYCARSGG
Sbjct  61   EQGLEWIGRIDPGNGNIKYDPKFQGKATITADTSSNKVYLHLSSLTSEDTAVYYCARSGG  120

Query  121  YDFDYWGQGTTLTVSS  136
            YDFDYWGQGTTLTVSS
Sbjct  121  YDFDYWGQGTTLTVSS  136
```

Amino acid sequence of variable light chain

Score	Expect	Method	Identities	Positives	Gaps
270 bits(691)	7e-99	Compositional matrix adjust.	131/131(100%)	131/131(100%)	0/131(0%)

```
Query  1    MKLPVRLLVLMFWIPASSSDVVMTQTPLSLPVSLGYQVSISCRSSQSLVHSNGNTYLHWY  60
            MKLPVRLLVLMFWIPASSSDVVMTQTPLSLPVSLGYQVSISCRSSQSLVHSNGNTYLHWY
Sbjct  1    MKLPVRLLVLMFWIPASSSDVVMTQTPLSLPVSLGYQVSISCRSSQSLVHSNGNTYLHWY  60

Query  61   LQKPGQSPKLLIYTVSNRFSGVPDRFSGSGSGTDFTLKISRVEAEDLGVYFCSQSTYVPY  120
            LQKPGQSPKLLIYTVSNRFSGVPDRFSGSGSGTDFTLKISRVEAEDLGVYFCSQSTYVPY
Sbjct  61   LQKPGQSPKLLIYTVSNRFSGVPDRFSGSGSGTDFTLKISRVEAEDLGVYFCSQSTYVPY  120

Query  121  TFGGGTKLEKK  131
            TFGGGTKLEKK
Sbjct  121  TFGGGTKLEKK  131
```

C

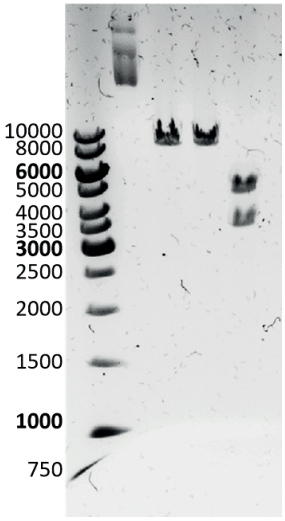


Figure 1. *Expression plasmid for chimeric anti-CD99 antibody production. A: a graphical map of plasmid pTRIOZ-ChAbMT99/3 was created by SnapGene software. Genes encoding variable heavy chain (VH) and variable light chain (VL) against human CD99 were cloned into the heavy/light chain cassettes of plasmid pTRIOZ-hIgG1, pink color, B: amino acid sequence alignment of VH and VL in the constructed plasmid pTRIOZ-ChAbMT99/3. Parental VH/VL amino acid sequences of mAb MT99/3 (Query) were aligned with those of ChAbMT99/3 (Sbjct), C: agarose gel electrophoresis of the constructed plasmid pTRIOZ-ChAbMT99/3 after amplification. The plasmids were digested with restriction enzymes BamHI (lane 2), EcoRI (lane 3), BamHI and EcoRI (lane 4), or without enzyme (lane 1). Sizes of standard markers (base pair) are shown on the left.*

Establishment of HEK293T stably expressing chimeric anti-CD99 antibody

To generate ChAbMT99/3-expressing human cells, HEK293T cells were transfected with the constructed pTRIOZ-ChAbMT99/3 plasmid. Three days post-transfection, intracellular immunofluorescence staining with Alexa Fluor 488-conjugated anti-human IgG antibody was performed and showed positive reactivity with transfected cells (Figure 2A). The results demonstrated the successful expression of ChAbMT99/3 in human HEK293T cells.

In addition, the culture supernatant of transfected HEK293T cells also exhibited positive reactivity with CD99-positive Jurkat E6.1 cells, while the culture supernatant of un-transfected cells showed no reactivity (Figure 2B). The results indicated that ChAbMT99/3 produced by HEK293T cells could be secreted into the culture supernatant. Subsequently, to establish HEK293T stably expressing ChAbMT99/3, transfected cells were cultured in a medium containing zeocin. Transfected cells survived in zeocin treatment, while un-transfected cells perished within 5 days. These transfected cells were collected from 20 wells to assess the intracellular expression of ChAbMT99/3. Cells in all selected wells showed positive reactivity with Alexa Fluor 488-conjugated anti-human IgG antibody (Figure 2C). However, expression levels varied among tested wells (Figure 2C). A transfected HEK293T clone 2F6 demonstrating the highest geometric MFI of the homogenous pattern was selected for ChAbMT99/3 production.

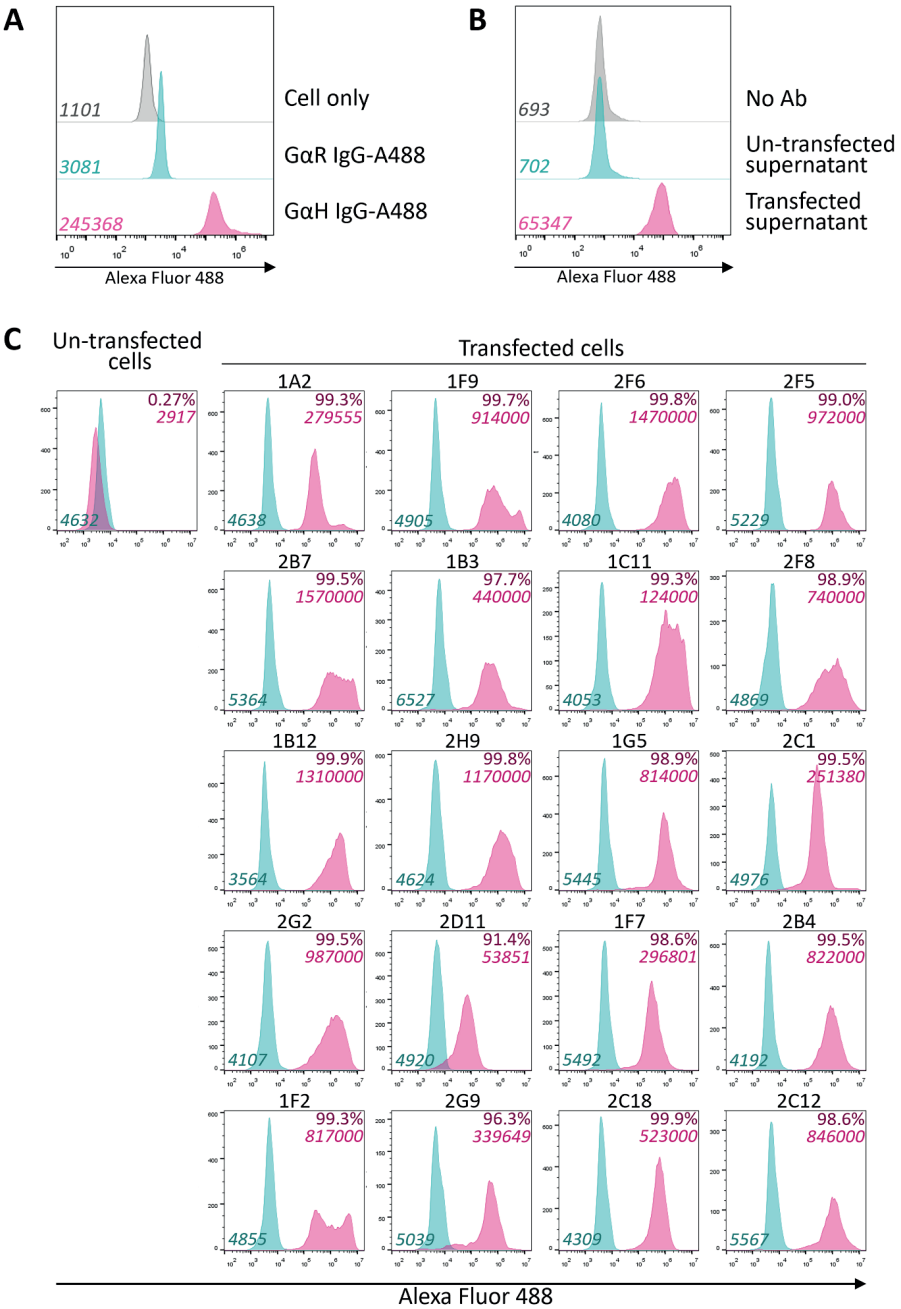


Figure 2. *Expression of chimeric anti-CD99 antibody in transfected HEK293T cells. A: human HEK293T cells were transfected with plasmid pTRIOZ-ChAbMT99/3. The transfected cells were performed intracellular staining with Alexa Fluor 488-conjugated goat anti-human IgG (GαH IgG-A488), Alexa Fluor 488-conjugated goat anti-rabbit IgG (GαR IgG-A488), or without antibodies (Cell only), B: culture supernatants obtained from the un-transfected or transfected HEK293T cells, or medium (No Ab) were subjected to stain Jurkat E6.1 cells followed by Alexa Fluor 488-conjugated goat anti-human IgG antibody, A-B: geometric mean fluorescence intensity (MFI) is indicated in the overlay histogram plots, C: un-transfected and transfected cells after zeocin selection were intracellularly stained using Alexa Fluor 488-conjugated goat anti-human IgG antibody (depicted in pink) or Alexa Fluor 488-conjugated goat anti-rabbit IgG antibody (displayed in green). The geometric MFI values and the percentage of positive cells are indicated in the overlay histogram plots.*

Large-scale production and purification of chimeric anti-CD99 antibody

To increase the amount of ChAbMT99/3, the HEK293T expressing stable ChAbMT99/3 clone 2F6 was cultured in a serum-free medium containing zeocin. The culture supernatant was purified using a protein G column. Approximately 1 mg of antibody was obtained from approximately 200 mL of the culture supernatant. The purity and structure of the purified ChAbMT99/3 were assessed using SDS-PAGE and western blotting. By SDS-PAGE, high purity of the purified ChAbMT99/3 was obtained. Protein bands were observed at approximately 55 kDa and 25 kDa under reducing conditions, corresponding to the heavy and light chains of the antibody, respectively. In non-reducing conditions, a protein band at approximately 150 kDa, representing the whole structure of the human IgG antibody, was detected (Figure 3A). For western blot analysis, positive bands were observed at 55 kDa under reducing conditions and 150 kDa under non-reducing conditions, which were the expected positions for the heavy chain and the whole structure of IgG antibody, respectively (Figure 3B). However, the major bands appeared above 180 kDa, suggesting a multimeric form where two or more antibody molecules were linked via inter-chain disulfide bonds (Figure 3A-B). The results indicated the successful production and purification of ChAbMT99/3 with high purity. The structure of ChAbMT99/3 was drawn as an intact molecule of mouse/human chimeric antibody harboring human IgG1 constant domains, as shown in Figure 3C.

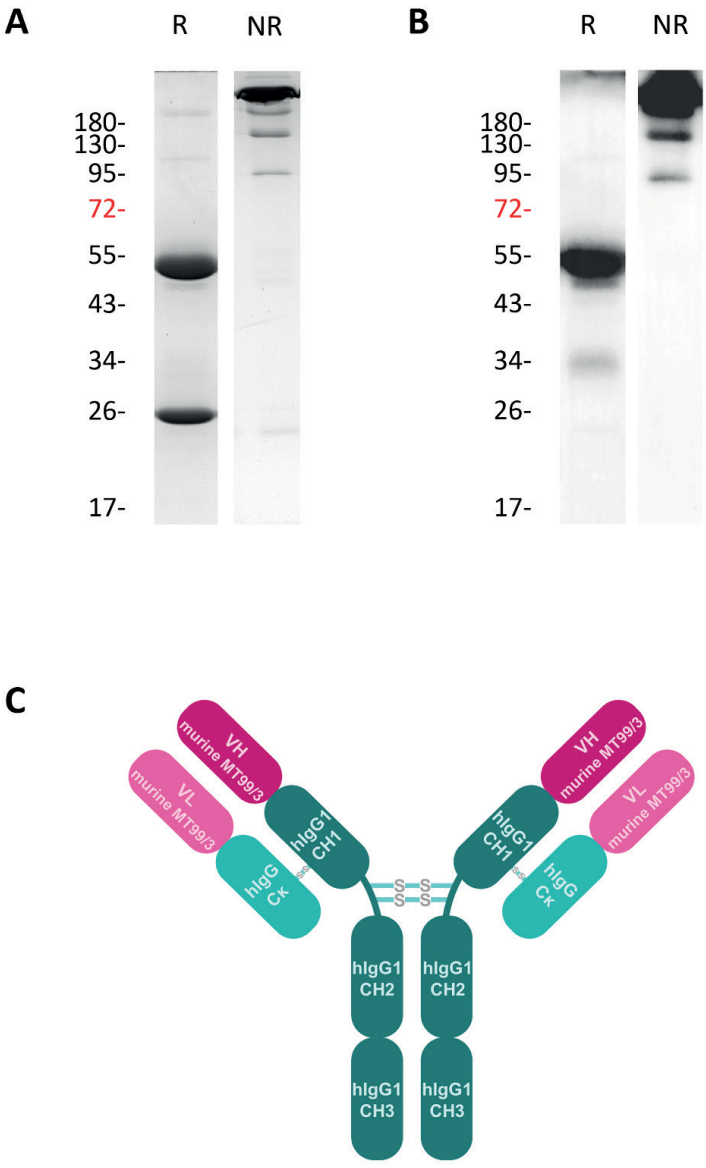


Figure 3*. The purity and structure of purified ChAbMT99/3. A: purified ChAbMT99/3 was subjected to 10% SDS-PAGE under reducing (R) and non-reducing (NR) conditions. Protein bands were stained with Coomassie blue, B: Western blotting was analyzed using HRP-conjugated goat anti-human IgG, Fcγ fragment specific Abs, C: schematic drawing of ChAbMT99/3 structure. VH and VL from murine-derived mAb MT99/3 (pink color) combined with the constant domain of human IgG1 and the constant domain of the kappa light chain (green color).*

Binding reactivity of chimeric anti-CD99 antibody

To confirm the binding specificity of ChAbMT99/3 against human CD99, ChAbMT99/3 was stained with recombinant human CD99-expressing myeloma cells. The results showed positive reactivity with CD99-expressing myeloma cells and negative reactivity with normal myeloma cells (Figure 4A). These results confirmed that the ChAbM99/3 retained its binding specificity to human CD99. Subsequently, to assess the binding activity of ChAbMT99/3 against native human CD99, surface staining of ChAbMT99/3 with MCL cell line Z138 and human PBMCs was carried out. ChAbMT99/3 could recognize Z138 cells that expressed native human CD99 on cell surfaces comparable with mouse mAb MT99/3 (Figure 4B). Moreover, the fluorescence intensity was increased in a dose-dependent manner (Figure 4B). For PBMCs, ChAbMT99/3 showed positive reactivity with all subpopulations in different patterns of CD99 expression levels, as previously reported.[18] In that order, the highest levels of CD99 expression were found in NKT cells, NK cells, monocytes, T cells, and B cells (Figure 4C). Notably, the produced ChAbMT99/3 demonstrated good binding activity to the native form of human CD99, suggesting a potential tool for studying anticancer activities. Moreover, BLI data revealed that the binding affinity of ChAbMT99/3 compared with the parental mouse mAb MT99/3 using CD99 EDC peptide exhibited very similar K_D values, $4.17x10^{-10}$ M and $4.62x10^{-10}$ M, respectively (Figure 5). This suggested that the binding capability of the produced ChAbMT99/3 and the parental mouse MT99/3 with CD99 molecule was comparable.

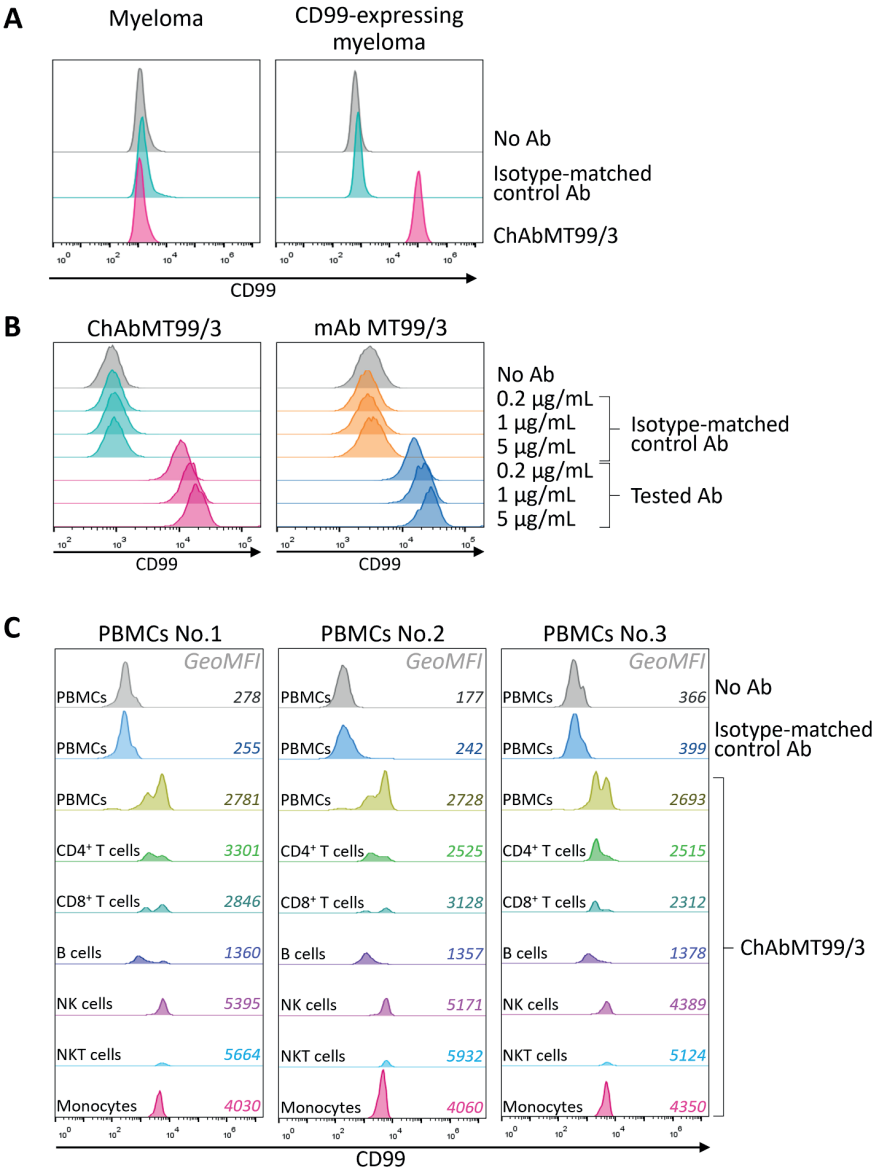
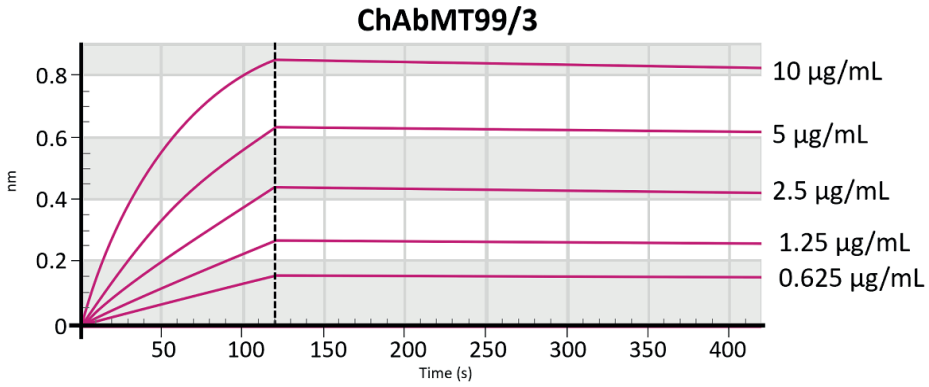
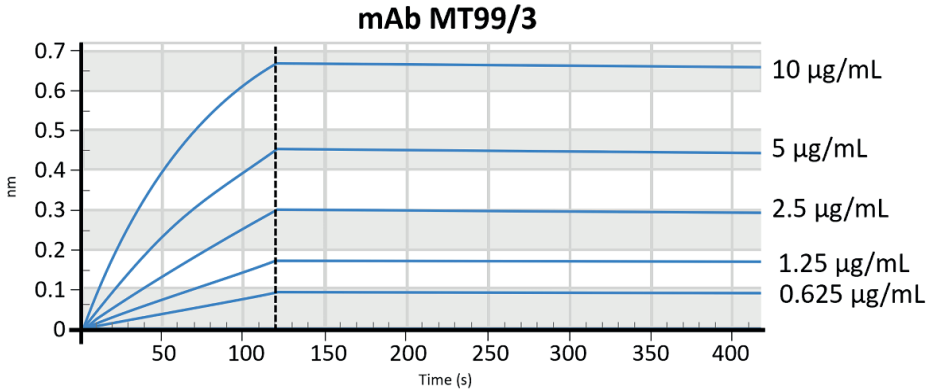


Figure 4. *Binding reactivity of chimeric anti-CD99 antibody. A: Myeloma cells and CD99-expressing myeloma cells were stained with 10 µg/mL of ChAbMT99/3, an isotype-matched control antibody (human IgG), or without antibody (No Ab), B: human mantle cell lymphoma Z138 cell line was stained with the indicated concentrations of ChAbMT99/3, mAb MT99/3, isotype-matched control Ab, or without antibody (No Ab), C: human PBMCs (3 donors) were stained with 10 µg/mL of biotinylated ChAbMT99/3, a biotinylated isotype-matched control antibody (human IgG), or without antibody (No Ab). The antibody binding on the cell surface was detected by specific conjugates. The numbers shown in the histogram indicate geometric MFI.*



K_D (M)	k_a (1/Ms)	k_d (1/s)
4.17×10^{-10}	2.5×10^{5}	1.04×10^{-4}



K_D (M)	k_a (1/Ms)	k_d (1/s)
4.62×10^{-10}	1.75×10^{5}	8.08×10^{-5}

Figure 5. *The binding affinity of ChAbMT99/3 and mouse mAb MT99/3. Biotinylated CD99 ECD peptides were loaded onto a streptavidin-coated biosensor and associated with the indicated doses of ChAbMT99/3 or mouse mAb MT99/3, followed by dissociation in PBST. The graphs are shown in the fitting view. k_a: the association rate constant, k_d: the dissociation rate constant, K_D: the equilibrium dissociation constant (k_d/k_a).*

ChAbMT99/3 enhances macrophages-mediated phagocytosis on mantle cell lymphoma

Phagocytosis is an essential mechanism for the elimination of lymphoma cells.[19,20] Thus, ADCP mediated by ChAbMT99/3 against MCL was evaluated. Z138 cells were treated with various concentrations of ChAbMT99/3, isotype-matched control antibody, or medium without antibody. After that, the treated cells were co-cultured with monocyte-derived macrophages. The results demonstrated that ChAbMT99/3 could enhance the eradication of Z138 cells by mediating the ADCP mechanism in a dose-dependent manner (Figure 6), correlated with antibody binding on the cell surface (Figure 4B). The macrophage-mediated phagocytosis in the ChAbMT99/3 treatment at 1 and 5 µg/mL concentrations was significantly increased compared with the isotype-matched control Ab treatment (Figure 6B). These results indicated the potential of ChAbMT99/3 as a therapeutic option for eradicating MCL by the ADCP mechanism.

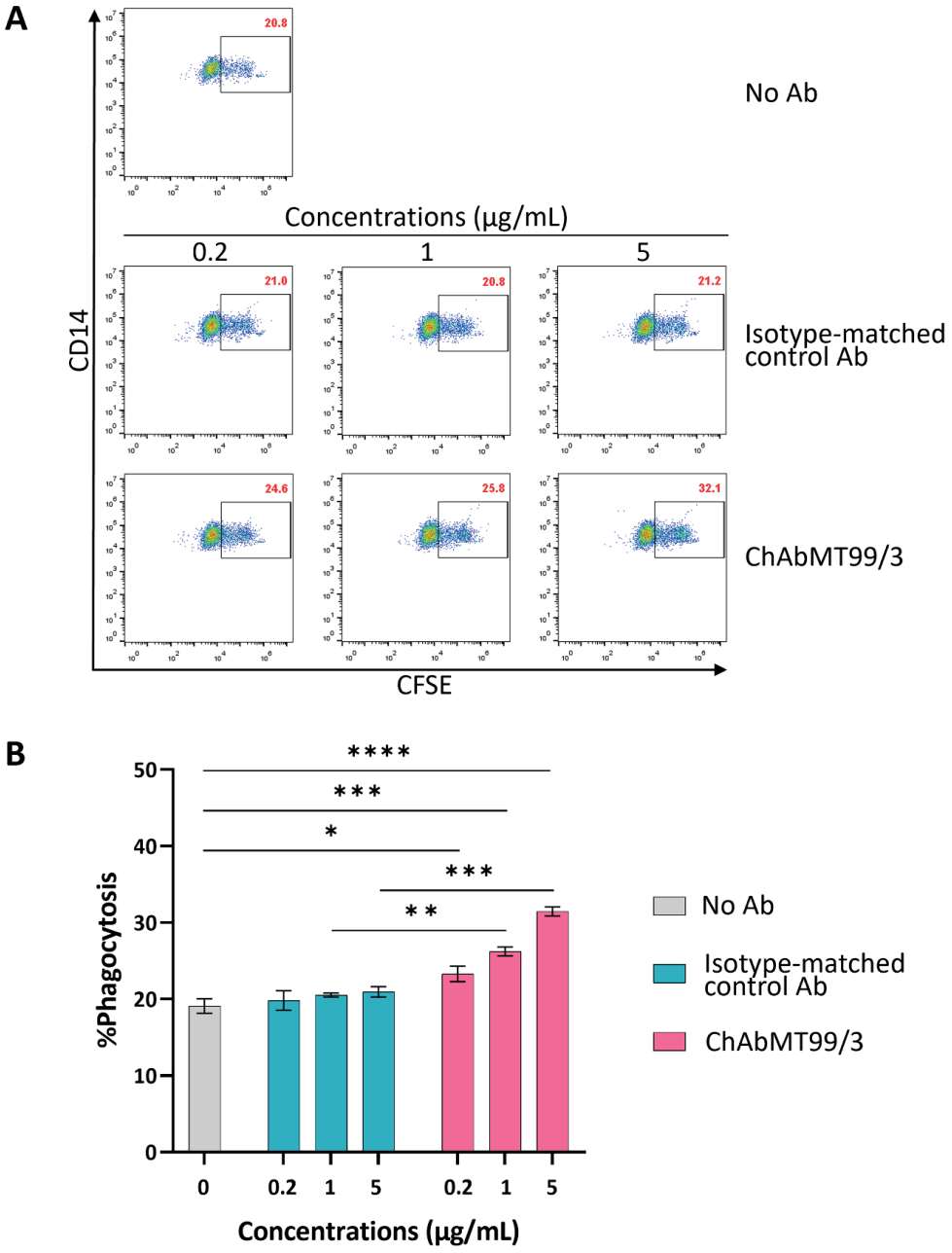


Figure 6. *ADCP mechanism mediated by ChAbMT99/3. A-B: Monocyte-derived macrophages (3 donors) were used as effector cells and incubated with CFSE-labelled Z138 in the presence of isotype-matched control Ab (human IgG) or ChAbMT99/3 at 0.2, 1 and 5 µg/mL or medium (0 µg/mL, No Ab). The percentages of phagocytosis ($CD14^+CFSE^+$) were analyzed by flow cytometry. $CD14^+$ cells were gated for further analysis of double positive of $CD14^+CFSE^+$ cells. A: Dot plots are shown as a representative of flow cytometric data. B: The bar graphs are shown as mean ± SD. One-way ANOVA followed by Tukey's multiple comparison test was used for statistical comparison. *P<0.05, **P<0.01, ***P<0.001 ****P<0.0001 represent statistically significant values.*

Discussion

Mantle cell lymphoma is an aggressive type of mature B-cell lymphoma. It can resist rituximab by inducing the internalization of CD20 molecules, consequently decreasing levels of CD20 expression on the cancer cell surface after rituximab treatment.[21] Therefore, discovering novel antibodies is crucial for improving therapeutic efficacy in mantle cell lymphoma. Our previous results indicated that the in-house produced anti-CD99 mAb clone MT99/3 is a promising antibody for treating mantle cell lymphoma. The mouse mAb MT99/3 could mediate MCL eradication by inducing ADCC and CDC mechanisms *in vitro* and anticancer activities *in vivo*.[12] However, the ADCP mechanism of mAb MT99/3 has not been investigated. The Fc region of mouse mAb was inappropriate for activating certain Fc receptors on human immune effectors.[22] Moreover, for clinical use, mouse mAb is inappropriate. After being introduced into the human body, the mouse mAb acts as a foreign antigen, effectively triggering an immune response. The reactions against mouse mAb will cause several adverse effects and destroy the treated antibody. This will reduce the efficacy of treatment.[13, 14] A chimeric antibody was requested for

clinical use to increase humanness and reduce its potential to cause an immune response.

Therefore, in this study, we engineered the mouse mAb MT99/3 into the mouse/human chimeric antibody named ChAbMT99/3, which contained the human Fc region. This engineered antibody was evaluated for the ADCP mechanism against MCL using human monocyte-derived macrophages. The pTRIOZ-hIgG1 plasmid containing the IgG1 constant domain and zeocin resistance genes was employed to construct expression plasmid pTRIOZ-ChAbMT99/3. The VH and VL genes of mouse mAb MT99/3 were cloned upstream of constant heavy chain and constant light chain genes, respectively. Therefore, this antibody consisted of the variable domain of mouse mAb MT99/3 connecting the constant domains of human IgG1 and kappa light chain. Human IgG1 exhibits the highest affinity for FcγR binding compared with other subclasses, so it is a potent activator of the ADCP mechanism.[23,24] Notably, human IgG1 antibodies commonly serve as the framework for Fc engineering strategies to boost immune effector functions.[25,26] HEK239T cells were employed as host cells in this study as they have been widely used for manufacturing research-grade proteins, with the benefit of producing fully human post-translational modifications (PTMs).[27] Moreover, the use of HEK293T avoids problems of possible immunogenicity caused by the presence of non-human PTMs.[28] Additionally, the ease of transfectability and relatively high protein productivity of HEK293T contribute to its popularity for the production of recombinant proteins.[29] In this study, HEK293T cells were transfected with pTRIOZ-ChAbMT99/3 using Lipofectamine. Flow cytometry analysis showed 100% transfectability. Stable cells with high antibody expression levels were obtained after zeocin drug selection. The purified ChAbMT99/3 was achieved from the culture supernatant using protein G affinity chromatography. High purity and full IgG structure of ChAbMT99/3 were obtained. These results demonstrated that a genetically engineered mouse/human ChAbMT99/3 was successfully produced in a stable human expression system. Furthermore, the produced ChAbMT99/3 retained its binding specificity to human CD99, showing positive reactivity with human CD99-expressing myeloma cells. Importantly, ChAbMT99/3 could recognize native human CD99 on the cell surface of MCL cell line Z138, the same as its parental mouse mAb MT99/3. However, indirect immunofluorescence assay could not be directly used to compare the binding capability of ChAbMT99/3 and mouse mAb MT99/3 with CD99 due to differences in conjugates. Therefore, we conducted BLI to measure the binding affinity of both antibodies using CD99 ECD peptide.[17] Our findings revealed that the K_D values of ChAbMT99/3 and mouse mAb MT99/3 were very close. This suggested that the binding affinities of ChAbMT99/3 and mouse mAb MT99/3 were comparable. Moreover, the ChAbMT99/3 could recognize CD99 on human PBMCs in the same pattern as other antibodies against CD99.[18]

Macrophage-mediated ADCP is one of the most essential mechanisms in cancer eradication.[8] The upregulation of CD47, the "don't eat me" molecule, in mantle cell lymphoma, hinders the elimination of cancer cells via phagocytosis.[11] Several studies demonstrated that anti-CD47 mAbs must interfere with the "don't eat me" signal and restore phagocytosis against MCL.[11,30] Therefore, antibodies that are capable of inducing the ADCP mechanism are pivotal. In our experiment, surprisingly, ChAbMT99/3 could significantly enhance the macrophage-mediated phagocytosis of an MCL cell line Z138 in a dose-dependent manner. The increase in phagocytosis was correlated with the number of bound antibodies on the surface of target cells. Furthermore, in the context of ADCP, the engagement of "eat-me" and "don't eat me" signals is pivotal.[31,32] ChAbMT99/3, which recognized CD99 on target cells, might alter these signals by downregulating "don't eat me" molecules or upregulating "eat me" molecules, resulting in higher sensitivity of MCL toward macrophage-mediated ADCP. Moreover, the internalization rate upon antibody binding with an antigen on target cells activates immune functions.[33,34] ChAbMT99/3 might induce a slow internalization rate and retain many bound antibodies on the cell surface to activate the Fc receptors on macrophages. However, the mechanism by which ChAbMT99/3 could overrule the function of CD47 expressed in mantle cell lymphoma is still unknown. Nevertheless, the combination of rituximab with anti-CD47 mAb exhibited the synergistic effect of phagocytosis.[11, 30] Therefore, the combination treatment of ChAbMT99/3 and anti-CD47 mAbs may be better than a single treatment and is attractive for further investigation.

Conclusion

The results demonstrated the successful production of a genetically engineered mouse/human chimeric antibody targeting human CD99 using a stable human expression system based on HEK293T cells. Notably, the produced chimeric antibody derived from mouse mAb MT99/3 could retain its binding reactivity to recombinant and native human CD99. Interestingly, the ChAbMT99/3 significantly enhanced antibody-dependent cellular phagocytosis against mantle cell lymphoma. Consequently, ChAbMT99/3 might be a promising therapeutic option for eradicating mantle cell lymphoma by the ADCP mechanism. However, further investigations are necessary to explore its potential in other anti-cancer activities.

Conflict of interest

The authors declare no conflict of interest.

Ethics approval

The study was approved by the ethics committees of the Faculty of Associated Medical Sciences, Chiang Mai University (AMSEC-66EX-014).

Funding

This work was financially supported by the Office of the Permanent Secretary, Ministry of Higher Education, Science, Research and Innovation (OPS MHESI), Thailand Science Research and Innovation (TSRI), and Chiang Mai University to NT (Grant numbers RGNS 64-068). KK

received Teaching Assistant and Research Assistant (TA/RA) Scholarships from Graduate School, Chiang Mai University. This work was also partially supported by the Targeted Research, Chiang Mai University.

Acknowledgments

We are grateful to Mrs. Supaporn Bourjai for technical assistance.

References

[1] Campo E, Raffeld M, Jaffe ES. Mantle-cell lymphoma. Semin Hematol. 1999; 36(2): 115-27.

[2] Dabaja B, Ha CS, Cox JD. Chapter 34 - Leukemias and Lymphomas. In: Cox JD, Ang KK, editors. Radiation Oncology (Ninth Edition). Philadelphia: Mosby; 2010. p. 875-911.

[3] Hanel W, Epperla N. Emerging therapies in mantle cell lymphoma. J Hematol Oncol. 2020; 13(1): 79. doi: 10.1186/s13045-020-00914-1

[4] Fischer L, Jiang L, Bittenbring JT, Huebel K, Schmidt C, Duell J, *et al*. The addition of rituximab to chemotherapy improves overall survival in mantle cell lymphoma-a pooled trials analysis. Ann Hematol. 2023; 102(10): 2791-801. doi: 10.1007/s00277-023-05385-1

[5] Roué G, Sola B. Management of Drug Resistance in Mantle Cell Lymphoma. Cancers (Basel). 2020; 12(6): 1565. doi: 10.3390/cancers12061565

[6] Zahavi D, Weiner L. Monoclonal Antibodies in Cancer Therapy. Antibodies (Basel). 2020; 9(3): 34. doi: 10.3390/antib9030034

[7] Grossbard ML, Press OW, Appelbaum FR, Bernstein ID, Nadler LM. Monoclonal Antibody-Based Therapies of Leukemia and Lymphoma. Blood. 1992; 80(4): 863-78. doi: 10.1182/blood.V80.4.863.863

[8] Cao X, Chen J, Li B, Dang J, Zhang W, Zhong X, *et al*. Promoting antibody-dependent cellular phagocytosis for effective macrophage-based cancer immunotherapy. Sci Adv. 2022; 8(11): eabl9171. doi: 10.1126/sciadv.abl9171

[9] Kamen L, Myneni S, Langsdorf C, Kho E, Ordonia B, Thakurta T, *et al*. A novel method for determining antibody-dependent cellular phagocytosis. J Immunol Methods. 2019; 468: 55-60. doi: 10.1016/j.jim.2019.03.001

[10] Van Wagoner CM, Rivera-Escalera F, Jaimes-Delgadillo NC, Chu CC, Zent CS, Elliott MR. Antibody-mediated phagocytosis in cancer immunotherapy. Immunological Reviews. 2023; 319(1): 128-41. doi: 10.1111/imr.13265

[11] Chao MP, Alizadeh AA, Tang C, Myklebust JH, Varghese B, Gill S, *et al*. Anti-CD47 Antibody Synergizes with Rituximab to Promote Phagocytosis and Eradicate Non-Hodgkin Lymphoma. Cell. 2010; 142(5): 699-713. doi: 10.1016/j.cell.2010.07.044

[12] Takheaw N, Sittithumcharee G, Kariya R, Kasinrerk W, Okada S. Anti-human CD99 antibody exerts potent antitumor effects in mantle cell lymphoma. Cancer Immunol Immunother. 2021; 70(6): 1557-67. doi: 10.1007/s00262-020-02789-0

[13] Carter P. Improving the efficacy of antibody-based cancer therapies. Nature Reviews Cancer. 2001; 1(2): 118-29. doi: 10.1038/35101072

[14] Khan FH. Chapter 25 - Antibodies and Their Applications. In: Verma AS, Singh A, editors. Animal Biotechnology. San Diego: Academic Press; 2014. p. 473-90.

[15] Kasinrerk W, Tokrasinwit N, Moonsom S, Stockinger H. CD99 monoclonal antibody induce homotypic adhesion of Jurkat cells through protein tyrosine kinase and protein kinase C-dependent pathway. Immunology Letters. 2000; 71(1): 33-41. doi: 10.1016/S0165-2478(99)00165-0

[16] Takheaw N, Laopajon W, Chuensirikulchai K, Kasinrerk W, Pata S. Exploitation of human CD99 expressing mouse myeloma cells as immunogen for production of mouse specific polyclonal antibodies. Protein Expr Purif. 2017; 134: 82-8. doi: 10.1016/j.pep.2017.02.015

[17] Takheaw N, Kotemul K, Chaiwut R, Pata S, Laopajon W, Rangnoi K, *et al*. Transcriptome Analysis Reveals the Induction of Apoptosis-Related Genes by a Monoclonal Antibody against a New Epitope of CD99 on T-Acute Lymphoblastic Leukemia. Antibodies. 2024; 13(2): 42. doi: 10.3390/antib13020042

[18] Shi J, Zhang Z, Cen H, Wu H, Zhang S, Liu J, *et al*. CAR T cells targeting CD99 as an approach to eradicate T-cell acute lymphoblastic leukemia without normal blood cells toxicity. Journal of Hematology & Oncology. 2021; 14(1): 162. doi: 10.1186/s13045-021-01178-z

[19] Glennie MJ, French RR, Cragg MS, Taylor RP. Mechanisms of killing by anti-CD20 monoclonal antibodies. Molecular Immunology. 2007; 44(16): 3823-37. doi: 10.1016/j.molimm.2007.06.151

[20] Overdijk MB, Verploegen S, Bögels M, van Egmond M, Lammerts van Bueren JJ, Mutis T, *et al*. Antibody-mediated phagocytosis contributes to the anti-tumor activity of the therapeutic antibody daratumumab in lymphoma and multiple myeloma. MAbs. 2015; 7(2): 311-21. doi: 10.1080/19420862.2015.1007813

[21] Beers SA, French RR, Chan HT, Lim SH, Jarrett TC, Vidal RM, et al. Antigenic modulation limits the efficacy of anti-CD20 antibodies: implications for antibody selection. Blood. 2010; 115(25): 5191-201. doi: 10.1182/blood-2010-01-263533

[22] Doevendans E, Schellekens H. Immunogenicity of Innovative and Biosimilar Monoclonal Antibodies. Antibodies (Basel). 2019; 8(1): 21. doi: 10.3390/antib8010021

[23] Yu J, Song Y, Tian W. How to select IgG subclasses in developing anti-tumor therapeutic antibodies. Journal of Hematology & Oncology. 2020; 13(1): 45. doi: 10.1186/s13045-020-00876-4

[24] Kretschmer A, Schwanbeck R, Valerius T, Rösner T. Antibody Isotypes for Tumor Immunotherapy. Transfus Med Hemother. 2017; 44(5): 320-6. doi: 10.1159/000479240

[25] Wirt T, Rosskopf S, Rösner T, Eichholz KM, Kahrs A, Lutz S, *et al*. An Fc Double-Engineered CD20 Antibody with Enhanced Ability to Trigger Complement-

Dependent Cytotoxicity and Antibody-Dependent Cell-Mediated Cytotoxicity. Transfus Med Hemother. 2017; 44(5): 292-300. doi: 10.1159/000479978

[26] Carter PJ. Potent antibody therapeutics by design. Nat Rev Immunol. 2006; 6(5): 343-57. doi: 10.1038/nri1837

[27] Dumont J, Euwart D, Mei B, Estes S, Kshirsagar R. Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. Crit Rev Biotechnol. 2016; 36(6): 1110-22. doi: 10.3109/07388551.2015.1084266

[28] Tan E, Chin CSH, Lim ZFS, Ng SK. HEK293 Cell Line as a Platform to Produce Recombinant Proteins and Viral Vectors. Front Bioeng Biotechnol. 2021; 9: 796991. doi: 10.3389/fbioe.2021.796991

[29] Abaandou L, Quan D, Shiloach J. Affecting HEK293 Cell Growth and Production Performance by Modifying the Expression of Specific Genes. Cells. 2021; 10(7): 1667. doi: 10.3390/cells10071667

[30] Aroldi A, Mauri M, Ramazzotti D, Villa M, Malighetti F, Crippa V, et al. Effects of blocking CD24 and CD47 'don't eat me' signals in combination with rituximab in mantle-cell lymphoma and chronic lymphocytic leukaemia. J Cell Mol Med. 2023; 27(20): 3053-64. doi: 10.1111/jcmm.17868

[31] Khalaji A, Yancheshmeh FB, Farham F, Khorram A, Sheshbolouki S, Zokaei M, et al. Don't eat me/eat me signals as a novel strategy in cancer immunotherapy. Heliyon. 2023; 9(10): e20507. doi: 10.1016/j.heliyon.2023.e20507

[32] Liu Ye, Wang Y, Yang Y, Weng L, Wu Q, Zhang J, *et al*. Emerging phagocytosis checkpoints in cancer immunotherapy. Signal Transduction and Targeted Therapy. 2023; 8(1): 104. doi: 10.1038/s41392-023-01365-z

[33] Liao-Chan S, Daine-Matsuoka B, Heald N, Wong T, Lin T, Cai AG, *et al*. Quantitative assessment of antibody internalization with novel monoclonal antibodies against Alexa fluorophores. PLoS One. 2015; 10(4):e 0124708. doi: 10.1371/journal.pone.0124708

[34] Jin H, D'Urso V, Neuteboom B, McKenna SD, Schweickhardt R, Gross AW, *et al*. Avelumab internalization by human circulating immune cells is mediated by both Fc gamma receptor and PD-L1 binding. OncoImmunology. 2021; 10(1): 1958590. doi: 10.1080/2162402X.2021.1958590

Journal of Associated Medical Sciences 2024; 57(3): 65-74

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Development of a simple HPLC method for the determination of urinary O-aminohippuric acid (OAH) and an establishment of OAH reference interval

Shoon Lae Maw[1] Natthawat Semakul[2] Khanittha Punturee[3*]

[1]Division of Clinical Chemistry, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai Province, Thailand.
[3]Cancer Research Unit of Associated Medical Sciences (AMS-CRU), Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 7 May 2024
Accepted as revised 19 June 2024
Available online 29 June 2024

Keywords:
HPLC, OAH, reference intervals

ABSTRACT

Background: O-aminohippuric acid (OAH) is considered a low-abundance urinary fluorescent metabolite with the potential to be an innovative lung cancer biomarker. There is a lack of simple methods for measuring urinary OAH metabolite, and the measurement needs to be normalized by urinary creatinine, which is produced at a constant rate. Thus, the newly developed method must be able to determine urinary creatinine and OAH simultaneously in the same run.

Objective: This work aimed to develop and validate a simple high-performance liquid chromatography (HPLC) method for measuring urinary OAH and to establish the reference intervals of OAH in healthy individuals.

Materials and methods: We synthesized OAH standard in a simple route and optimized simple HPLC method for simultaneous measurement of creatinine and OAH in a single run. Analysis was performed on a RP-C18 column with a gradient elution system of acetonitrile - ammonium acetate buffer (pH 6.5, 100 mM). After implementing optimal conditions, the procedure was compiled according to the International Council for Harmonisation (ICH) validation parameters. The developed method was used for the establishment of reference intervals of a total of 120 random urine samples of healthy individuals.

Results: The linear range of the calibration curve for creatinine and OAH were 1-1000 µg/mL and 0.1-100 µg/mL, respectively. The recoveries ranged for both metabolites were between 91.35 % and 109.12%. The relative standard deviations (RSDs) for the intra-day and inter-day results ranged from 0.11-0.66 % to 0.16-1.73 %, respectively. The limits of detection (LOD) and quantification (LOQ) were 0.258 µg/mL and 0.783 µg/mL for creatinine, while OAH was 0.045 µg/mL and 0.137 µg/mL, respectively. The method was successfully applied to establish reference intervals of OAH in healthy individuals and was defined as 0.420-2.287 mmol/mol creatinine.

Conclusion: According to various validated parameters, the proposed method was proven to quantify urinary OAH and creatinine in a single run. It can also be analyzed noninvasively without additional sample processing. Reported herein is the first establishment of OAH reference intervals in healthy individuals, which may benefit the utilization of OAH as a noninvasive biomarker for lung cancer detection in the future.

∗ **Corresponding contributor.**
Author's Address: *Cancer Research Unit of Associated Medical Sciences (AMS-CRU), Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: khanittha.taneyhill@cmu.ac.th
doi: 10.12982/JAMS.2024.048
E-ISSN: 2539-6056

Introduction

The composition of O-aminohippuric acid (OAH) includes a metabolite of tryptophan and glycine, and it is believed to originate from kynurenine (KYN) through anthranilic acid.[1] Tryptophan plays a vital role as an essential

amino acid in the synthesis of proteins and serves as a pivotal regulator in cancer progression.[2] Its metabolites undergo processing through the kynurenine pathways, significantly impacting immune response regulation.[3,4] The main enzymes, indolamine-pyrrole 2,3-dioxygenase (IDO)1, IDO2, and tryptophan 2,3-dioxygenase (TDO) are responsible for the KYN pathway activities.[5,6] Moreover, cancer cells predominantly rely on glycine for their growth and multiplication. Lung tumor-initiating cells demonstrate increased glycine levels derived from enhanced glycolysis and glutaminolysis, subsequently converting into deleterious metabolites.[7,8]

The extensive use of tryptophan and glycine represents cancer cells' potential metabolic reprogramming trait.[9,10] However, the exact pathway of OAH formation in the urine of lung cancer patients is still unknown. OAH could be generated through tryptophan metabolism, producing anthranilic acid, which is then conjugated with glycine. Therefore, elevated OAH levels in lung cancer patients may be attributed to increased tryptophan and glycine metabolism in cancer cells.[1]

Metabolomics offers a direct measurement of the metabolic profile of an organism, offering a more accurate approach to detecting changes in metabolism. In recent years, studying metabolites in biological specimens has been a subject of great interest.[11,12] Metabolomics experiments can be subdivided into targeted (limited and predefined subset) and untargeted analyses (broader approach).[10] Urine become the most frequently used biological fluid for metabolomics research for different cancer types since it has additional benefits over blood, such as being rich in metabolites, easy to handle, and accessible in substantial quantities, all without the need for invasive collection procedures.[12,13] Besides, it contains diagnostically important metabolite OAH, and its levels in urine samples from lung cancer patients were significantly higher than those in non-cancer.[1] Under a stable condition, the overall daily excretion of creatinine is directly proportional to the total body creatinine content and the total mass of skeletal muscle.[14] To increase the accuracy of the analysis, the concentration of urine constituents was standardized using urine creatinine levels to normalize urine volume variations.[15]

Mass spectrometry (MS) and nuclear magnetic resonance (NMR) based techniques are currently utilized in metabolomics studies.[7] The quantification of OAH in the previous survey utilized high-performance liquid chromatography coupled with mass spectrometry and nuclear magnetic resonance (HPLC-MS/NMR).[1] These methods used separation via HPLC before detection, and OAH levels were adjusted by creatinine with an enzymatic reaction in all subjects. These approaches, however, include ion suppression, interference with other ions, and low throughput.[1,15,16] Given the current state of technologies and the diverse physiological properties of metabolites, no single separation or detection technique can simultaneously identify all metabolites within a urine sample.[3,8]

The present work aimed to develop and validate high-performance liquid chromatography (HPLC) to determine urinary metabolites OAH and creatinine simultaneously. This process consists of a simple sample preparation enabling chromatographic separation. There were no reports of OAH reference intervals in non-cancer patients, and the previous study just compared the levels statistically; the developed method was successfully applied to establish reference intervals of OAH in healthy individuals, which might be helpful to differentiate healthy and diseased patients in clinical settings.[17-20]

Materials and methods

Chemicals and reagents

Creatinine, isatoic anhydride, glycine hydrochloride, and deuterated methanol were purchased from Sigma-Aldrich, St.Louis, MO, USA. HPLC grade acetonitrile (purity, minimum 99.8%) was obtained from RCI Labscan Limited, Thailand. Ammonium acetate was purchased from BDK Chemicals Limited, Poole, England. ARIOSO water purifier (Human Corporation, Korea) was the source of type I water, which was employed in preparing samples and mobile phase throughout the study.

Instrumentation

HPLC was performed with Agilent Infinity Lab LC series 1260 Infinity II, USA, equipped with a diode array detector (G7115A), fluorescent detector (G7121A), vial sampler (G7129A), and a quaternary pump (G7111B). The analytical column used was an EC-C18 reversed-phase column with a length of 250 mm, an internal diameter of 4.6 mm, and a bead size of 4 µm. ChemStation version 1.1.1 system control software was used for data. Microcentrifugation of samples was carried out in a Kubota 5200 Japan centrifuge. Thin-layer chromatography (TLC) was performed with Merck precoated silica-gel 60 F_{254} plates. NMR spectra were measured with 500-MHz Bruker spectrometer FT-IR (Fourier Transform Infrared Spectroscopy).

Preparation of Standard Solution for Creatinine and OAH

Creatinine (0.03 gm) and OAH (0.003 gm) were transferred into a 10 mL volumetric flask and filtrated by nylon syringe filter pore size 0.22 µm. The stock standards were stored at 4 °C until further use. A series of calibration standards were prepared in the concentration range of 0.1-1000 µg/mL for creatinine and 0.1-100 µg/mL for OAH. Quality controls for creatinine were 1000 µg/mL (high-level quality control, HQC), 500 µg/mL (middle-level quality control, MQC), and 250 µg/mL (low-level quality control, LQC). For OAH, quality controls were 1 µg/mL (HQC), 0.5 µg/mL (MQC), and 0.25 µg/mL (LQC). All solutions were diluted with ultra-pure water and stored at 4 °C. The standards on HPLC were further analyzed for different validation parameters.

Sample Collection and Preparation for HPLC Analysis

The healthy volunteers, who had normal results in urine analysis, were selected. The left-over urine samples

were collected from the PROMT Health Center, Chiang Mai University. Ethical approval was obtained from the Ethics Review Committee, Faculty of Medicine, University of Chiang Mai, Thailand**.** Demographic characteristics were provided by the clinic. One hundred and twenty left-over random urine samples from healthy volunteers were collected and centrifuged at 315 g for 10 minutes. The supernatant was transferred to sterile tubes and was diluted by 10-fold dilution with deionized water. After that, a 2-mL aliquot of this mixture was passed through a 0.22- µm nylon syringe filter, and the filtrate was collected for HPLC analysis.

Method Development and Validation

Firstly, a retrosynthetic analysis of O-aminohippuric acid has been performed. Then, the optimized method for the simultaneous estimation of creatinine and OAH has been validated as per the ICH guidelines, consisting of evaluating system suitability, specificity, linearity, and range, the lower limit of detection (LOD), the lower limit of quantification (LOQ), accuracy, precision (intra- and inter-day), and robustness of creatinine and OAH.[21,22]

Establishment of Reference Intervals

Using the validated HPLC method, the current process was employed to establish a reference interval of potential lung cancer urinary biomarkers, OAH levels, in healthy populations.

Statistical Analysis

Linear regression tests were performed using Microsoft Excel 2018, and the respective calibration curve equation was concluded. The data were expressed as the mean±SD and 95% confidence interval. Statistical analysis was performed using SPSS 19.0. The Kolmogorov-Smirnov test evaluated the data distribution with $p>0.05$. If the data were normally distributed, the reference intervals were expressed as mean±1.96 SDs.

Results

Chemical Synthesis of OAH Standard

The previous synthesis of O-aminohippuric acid relied on multi-step synthetic routes.[23,24] As shown in Figure 1, we identified two commercially available precursors, isatoic anhydride and glycine hydrochloride, as key starting materials.

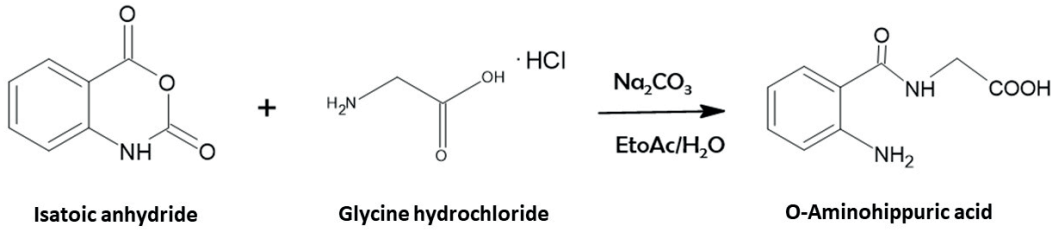


Figure 1*. Retrosynthetic analysis of O-aminohippuric acid using isatoic anhydride and glycine hydrocholride.*

After subjecting isatoic anhydride and glycine hydrochloride in the presence of sodium carbonate, TLC analysis of a crude product showed two new spots with retention factor (R_f) of 0.23 and 0.94, respectively. The spot with R_f of 0.94 corresponds to the hydrolysis product of isatoic acid, anthranilic acid. The spot with R_f of 0.23 corresponds to the targeted product, a higher polar compound. In addition, the R_f of 0.71 in the first lane corresponds to the starting. The three spots in the co-sport lane serve as a control, and all these spots were visualized under ultraviolet (UV) light 254 nm in the lamp chamber material shown in Figure 2(A). We purified the target compound by a simple wash with ethyl acetate and hexane. After washing with ethyl acetate and hexane several times, TLC analysis of a washed sample showed only one spot with R_f of 0.37 illustrated in Figure 2(B). The purity of a purified sample was rechecked with high-performance liquid chromatography. To our delight, the chromatogram of a purified sample showed one significant peak of O-aminohippuric acid at 13.38 retention time (RT) in Figure 2(C). In contrast, anthranilic acid impurity was insignificant at 10.44 retention time, with a peak area of less than 10 mAU*s, accounting for 95% purity. Despite a small amount of impurities, the chromatogram results from TLC and HPLC agree with each other after purification and can be accepted.

Finally, the structure of OAH was analyzed by ^{1}H-NMR spectroscopy using methanol-d_4 (CD_3OD) as a solvent in which the ^{1}H-NMR spectrum is shown in Figure 2 (D). The chemical shift (δ) at 7.52 (dd, *J*=7.9, 1.5 Hz, 1H), 7.21 (dd, *J*=8.4, 7.1, Hz, 1H), 6.77 (d, *J*=8.2 Hz, 1H), 6.65 (t, *J*=7.5 Hz, 1H) are ascribed to aromatic protons. The chemical shift (δ) of 4.06 (s, 2H) corresponds to methylene protons of glycine units. It should be noted that the protons of the amino group ($-NH_2$) could not be observed by ^{1}H-NMR spectroscopy. Altogether, ^{1}H-NMR spectroscopy could be used to confirm the structure of synthetic O-aminohippuric acid and the purity of our synthetic compound, which is sufficient for use in HPLC analysis.

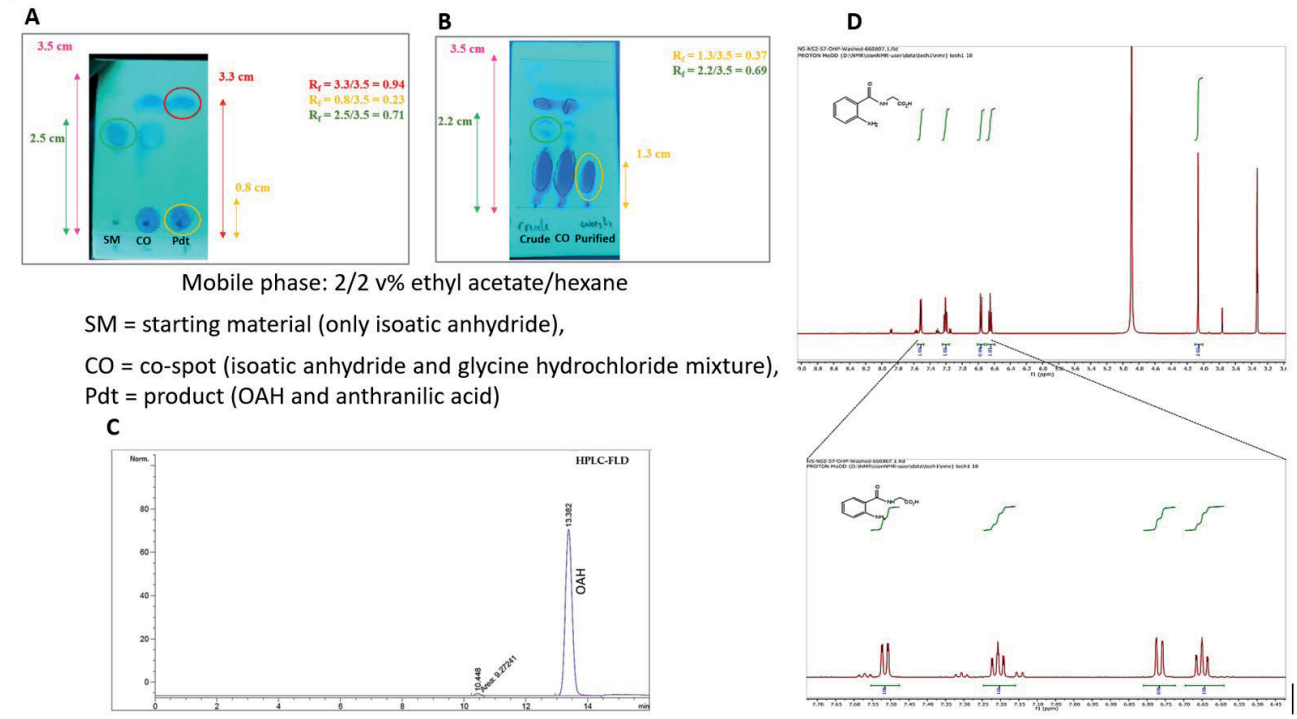


***Figure 2**. Chemical synthesis for OAH standard. A: TLC chromatogram of crude OAH, B: TLC chromatogram of washed sample OAH, C: HPLC chromatogram of washed sample OAH, D: ^{1}H-NMR spectrum (500MHz, CD_3OD) of OAH.*

HPLC method development and validation

After many trials had been tested, the gradient program, mobile phase composition, buffer pH, and column temperatures were adjusted, and the best-optimized condition was obtained. It was applied from 0 to 30% solvent B (acetonitrile) in solvent A (100 mM ammonium acetate, pH 6.5) with a 27-min gradient time (time-B (%): 0-1, 12-1, 13-30, 19-30, 20-1, 27-1). The flow rate was 1.0 mL/min at 40 °C. The detection wavelengths for both analytes were obtained the maximum sensitivity. For OAH detection using a fluorescent detector, the excitation wavelength was 340 nm, and the emission wavelength was 430 nm. For creatinine detection using diode array detector, the wavelength was 235 nm. The column temperature was 40 °C, and the sample injection volume was 10 µL.[1,18,25-27] Then, the developed method was validated according to ICH guidelines.

System Suitability

The relative standard deviation (RSD) (%) values of RT and peak area for creatinine and OAH were <2.0, indicating excellent repeatability of replicate injections on the integral HPLC system used. The tailing factor never exceeded 2 in all peaks, demonstrating reasonable peak regularity (acceptance limit is <2.0); and the number of theoretical plates was always >2000 in all chromatography runs, ensuring good columns efficacy throughout the developed separation process reported in Table 1.

Table 1. Results of system suitability parameter

Parameters	Value		% RSD	
	Creatinine	OAH	Creatinine	OAH
	Mean±SD (N=6)	Mean±SD (N=6)		
Theoretical plate	19549.2±331.7	23522.2±380.70	1.69	1.62
Tailing factor	0.76±0.013	1.18±0.01	1.71	1.18
RT (min)	3.848±0.03	14.04±0.11	0.78	0.78
Peak area	2257±18.42	440.5±5.06	0.82	1.15

Note: RSD: relative standard deviation, RT: retention time, SD: standard deviation.

Specificity

The optimized analytical method successfully detected and evaluated OAH even in the presence of minor impurities. The chromatographic specificity of the process was demonstrated by the absence of significant interfering peaks at the RT of creatinine and OAH (3.87 and 14.05 mins) upon comparison of blank chromatograms. This outcome suggests that the peaks corresponding to the analytes were pure, thereby confirming the specificity of the method.

Linearity and Range

Linearity was verified through the analytical curve using four concentration levels and evaluated in five injections.[21,28,29] The calibration curves were linear in the specified ranges of 0.1-1000 µg/mL for creatinine and 0.1-100 µg/mL for OAH. The correlation coefficient (r^2) for creatinine and OAH was 1 with linear regression equations Y =37.051 and Y =7.6284, respectively, in Figure 3.

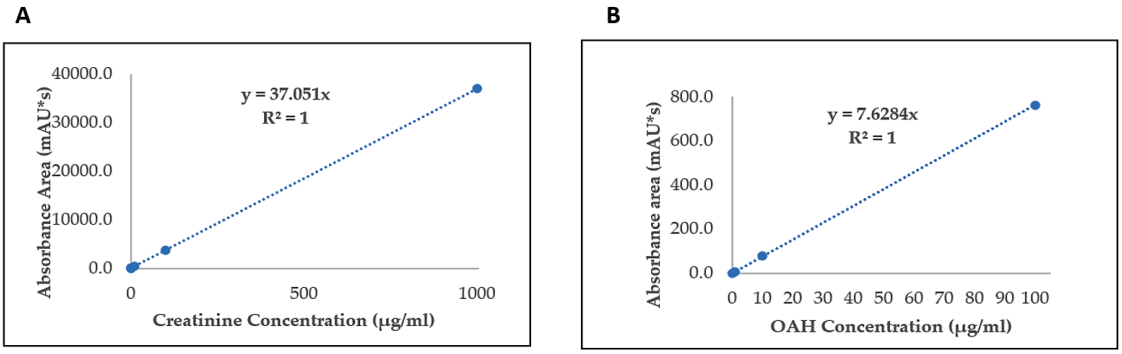


Figure 3*. Calibration curves of standard creatinine (A) and OAH (B).*

Limit of detection (LOD) and limit of quantification (LOQ)

The LOD and LOQ are used to calculate the sensitivity of the developed method.[29,30] The LOD were 0.258 and 0.045 µg/mL for creatinine and OAH, respectively. Moreover, the LOQ was 0.783 and 0.137 µg/mL for OAH and creatinine, respectively. These values are adequate for accurate and precise detection and quantification of OAH and creatinine.

Accuracy

Accuracy represents the closeness of the agreement between the mean values obtained from the series of measurements by the method and the actual value.[28,31] The corresponding HQC, MQC, and LQC values of creatinine and OAH were spiked into the normal urine sample, and percentage recovery was calculated along with standard deviation, which fell into 109.12%, 97.68%, 97.1% for creatinine and 91.35%, 94.0%, 96.39% for OAH, respectively. The relative standard deviation (RSD) was less than 2%, indicating a high accuracy of this method.

Precision

Precision represents the closeness of agreement among a series of measurements obtained from multiple sampling.[28] The peak areas were obtained by injecting HQC, MQC, and LQC for inter-day and intra-day studies. The results of both intra-day and inter-day findings ensure the high repeatability and precision of the developed method. All data were expressed in % RSD, less than the acceptable limit (<2.0%). Results for intra-and inter-day precision for OAH and creatinine are given in Table 2.

Table 2. Results of the precision parameter.

Parameters	Intra-day (Repeatability)		Inter-day (Intermediate precision)	
	Mean area±SD (N=6)	RSD (%)	Mean area±SD (N=6)	RSD (%)
Creatinine				
HQC	34344.0±52.16	0.15	34308.9±53.39	0.16
MQC	17141.1±18.94	0.11	17161.6±29.03	0.17
LQC	8533.2±10.05	0.12	8518.8±14.48	0.17
OAH				
HQC	7.1±0.05	0.66	7.2±0.09	1.73
MQC	3.6±0.02	0.65	3.6±0.05	1.67
LQC	1.8±0.01	0.60	1.8±0.01	1.29

Note: HQC: high-level quality control, MQC: middle-level quality control, LQC: low-level quality control.

Robustness

In an analytical method, robustness measures its reliability, characterized by its capacity to remain unaffected by minor, intentional changes in method parameters.[32] In all cases, creatinine and OAH peaks were symmetric, and the RSD (%) of a number of the theoretical plates (NTP) and tailing factor (TF) values were <2% with varying parameters. Results for robustness are presented in Table 3.

Table 3. Overview of robustness parameter.

Variables	Value	Conc. (µg/mL)	Peak area Mean±SD (N=3)	RSD (%)	RT (min) Mean±SD (N=3)	RSD (%)
Creatinine						
Flow rate (mL/min)	0.9	300	13747.1±113.88	0.83	4.16±0.05	1.23
	1	300	12408.57±44.6	0.36	3.80±0.04	1.15
	1.1	300	11229.53±80.65	0.72	3.68±0.01	0.27
Wavelength	233	300	12434.4+24.07	0.19	3.88±0.01	0.15
	235	300	12408.57±44.6	0.36	3.80±0.04	1.15
	237	300	12143.30±73.71	0.61	3.85±0.05	1.31
pH	6	300	11400.67±104.36	0.92	3.82±0.01	0.30
	6.5	300	12404.63±43.47	0.35	3.80±0.04	1.15
	7	300	12486.43±51.25	0.41	3.82±0.02	0.55
OAH						
Flow rate (mL/min)	0.9	100	813.87±3.35	0.41	14.51±0.06	0.42
	1	100	753.47±3.25	0.43	14.26±0.24	0.98
	1.1	100	659.60±0.87	0.13	13.59±0.03	0.19
Wavelength	338, 428	100	809.67±0.84	0.10	14.15±0.04	0.33
	340, 430	100	753.47±3.25	0.43	14.26±0.14	0.98
	342, 432	100	717.30±11.17	1.56	14.17±0.15	1.08
pH	6	100	756.93±6.38	0.84	13.57±0.15	1.13
	6.5	100	753.47±3.25	0.43	14.26±0.14	0.98
	7	100	796.07±12.12	1.52	13.61±0.08	0.57

Stability

The percentage of the original amount of both analytes remaining in the urine sample was calculated for different days at different temperatures. More than 80% of these two metabolites at various concentration levels remained in the urine sample for 20 days at -20 °C and -80 °C. However, HQC and LQC of OAH levels reduced to less than 80% at 4°C for 20 days.

Establishment of reference intervals

Table 4 summarizes the demographic characteristics of healthy population involved in this study. The optimized conditions were applied to quantify the OAH level of 120 urine sample. An example of HPLC chromatogram for the measurement of OAH and creatinine is shown in Figure 4.

Table 4. Sample characteristics of all samples presented in the study.

	Total N=120	Male N=55 (45.8%)	Female N=65 (54.2%)	p^a
Age	47.5 (35.25-57.0)	48 (37-60)	45 (34-55.5)	0.142
21-60	100 (83.33%)	43 (35.83%)	57 (47.5%)	
≥61	20 (16.67%)	12 (10%)	8 (667%)	
Smoking Status				
Yes	16 (13.33%)	16 (13.33%)	0 (0%)	
No	104 (86.67%)	39 (32.5%)	65 (54.17%)	
Exercise				
Yes	41 (34.16%)	20 (22.5%)	21 (17.5%)	
No	79 (64.16%)	35 (23.33%)	44 (36.67%)	
Alcohol				
Occasionally	43 (35.84%)	27 (22.5%)	16 (13.34%)	
Never	77 (64.16%)	28 (23.33%)	49 (40.83%)	
BMI	22.1 (20.10-23.97)	23 (20.80-24.20)	21 (19.60-23.90)	0.253
Normal weight (≤22.9)	69 (57.5%)	27 (22.5%)	42 (35%)	
Overweight (≥23)	51 (42.5%)	28 (23.33%)	23 (19.17%)	

Note: BMI: body mass index, N: number of samples. Data are presented as N (%). The median (interquartile range) is also given for age and BMI. [a]: Independent t test, p<0.05 is statistically significant different between male and female.

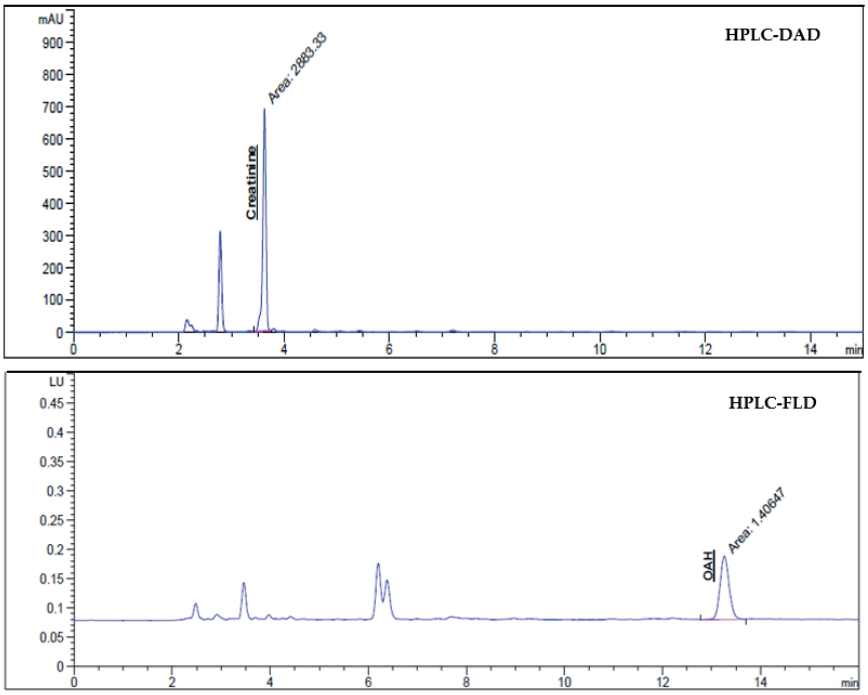


Figure 4. *HPLC chromatogram of human urine. Creatinine at retention time 3.62 mins detected by DAD detector, OAH at retention time 13.26 mins detected by FLD detector.*

The established reference intervals (RIs) for males was 0.469-2.244 mmole/mole creatinine, and females was 0.373-2.329 mmole/mole creatinine as shown in Figure 5. No difference was found when stratified by gender and age. Therefore, we expressed RIs for both, which were 0.420-2.287 mmole/mole creatinine.

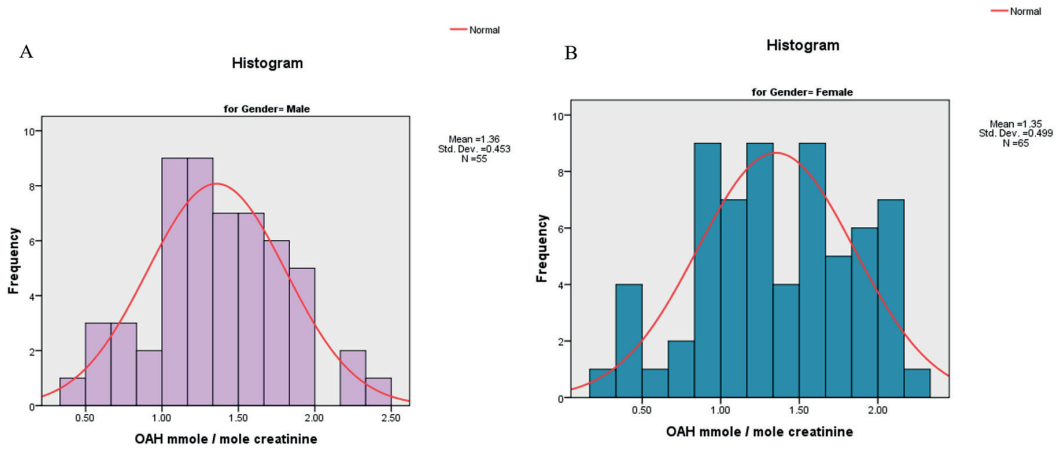


Figure 5. *Determination of reference intervals for OAH in males (A) and females (B).*

Discussion

This study aimed to identify and establish urinary biomarker OAH reference intervals with a simple method of high-performance liquid chromatography (HPLC). Although first-morning void urine samples are a better choice than random urine samples, the choices between them depend on the specific goals of the metabolomics study and practical considerations.[33] Since the current study aimed to establish reference intervals of OAH metabolite in a large-scale survey, first-morning void urine may be less convenient for participants. Random urine samples are easy to collect and often most cost-effective because they require less coordination and a close approach to routine analysis.

The previous synthesis of O-aminohippuric acid relied on multi-step synthetic routes and the use of highly toxic chemicals. In this study, we synthesized OAH with a simple route inspired by previous work. The synthesis was designed to eliminate the deprotection step of toxic carboxylic acid. In short, an efficient synthesis of OAH has been developed, and it was about 95% pure enough to use in HPLC analysis.

The recent method was never validated per guidelines, focusing on quantifying OAH with long analytical run times and low throughput. They mainly focused on the identification of OAH as a low-abundance urinary fluorescent metabolite as a potential lung cancer biomarker. Several steps are required for sample preparation, such

as solid phase extraction, precipitation, dilution, and filtration.[1] The creatinine concentration was analyzed using an enzyme-based method to normalize in each sample. We used simple HPLC method and rapid sample preparations, which allowed us to detect OAH, and creatinine concentration in each sample simultaneously, saving time and cost effectively.

Due to no significant difference being found by gender, reference intervals of OAH were expressed for both genders in the 0.420-2.287 mmole/mole creatinine range. The median concentration of OAH in healthy urine detected by HPLC was 1.335 mmole/mole creatinine. The median concentration of OAH in a previous study detected by mass spectrometry was 0.05 mmole/mole creatinine.[1] This is probably because detection could not differentiate other similar naturally fluorescent metabolites in urine from various factors such as food, drugs, and other amino acid tryptophan metabolites.[34,35] We selected participants according to our criteria and collected urine samples left over from the clinic that were reported as normal urine from urine analysis. However, samples were not collected under stringent conditions of diets, exercise, drugs, or hydration, which can be influenced by variations even though we did questionnaires or applied creatinine normalization.

A previous study identified a peak with a mass per charge ratio of mass spectrometry and confirmed using a commercially available standard.[1] Thus, a nuclear magnetic resonance (NMR) study is needed to confirm the presence of only OAH metabolites using selected peak fragments in HPLC. This is the first study to quantify and establish reference intervals of OAH levels in the urine samples of healthy people. However, the utility of these metabolites has not been evaluated in other cancers, and its potential to aid early diagnosis of lung cancer remains to be further assessed. The study's strength is simple sample preparation and immediate processing after urine collection to avoid storage bias. We have some limitations in proving the OAH level detection in lung cancer patient's urine. This would require us to obtain and analyze some time points. Thus, further studies are needed to quantify noninvasive urinary biomarkers OAH in lung cancer patients because these fluorescent metabolites may screen lung cancer and other cancers effectively.

Conclusion

In summary, we developed a simple and accurate RP-HPLC method for the simultaneous measurement of creatinine and OAH. The validation results proved that the process is accurate, precise, sensitive, and reproducible for targeted metabolomics studies. We interrogated urine for cancer biomarkers, and the method requires minimal preparation before being subjected to HPLC. The reference interval of OAH in healthy individuals was 0.420-2.287 mmole/mole creatinine. The results of this investigation suggest that the established approach may be evaluated in further steps to measure these metabolites in healthy individuals and other cancer patients in preclinical or clinical trials.

Conflicts of Interest

The authors declare no conflict of interest.

Funding

This work was funded by the Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Thailand.

Acknowledgements

The first author is grateful for the CMU Presidential Scholarship, which supported the master's degree program in Medical Technology at the Faculty of Associated Sciences, Chiang Mai University.

References

[1] Funai K, Honzawa K, Suzuki M, Momiki S, Asai K, Kasamatsu N, *et al*. Urinary fluorescent metabolite O-aminohippuric acid is a useful biomarker for lung cancer detection. Metabolomics. 2020; 16(10): 101. doi: 10.1007/s11306-020-01721-y.

[2] An Z, Chen Y, Zhang R, Song Y, Sun J, He J, *et al*. Integrated ionization approach for RRLC-MS/MS-based metabonomics: finding potential biomarkers for lung cancer. J Proteome Res. 2010; 9(8): 4071-81. doi: 10.1021/pr100265g.

[3] Wang Q, Liu D, Song P, Zou MH. Tryptophan-kynurenine pathway is dysregulated in inflammation, and immune activation. Front Biosci (Landmark Ed). 2015; 20(7): 1116-43. doi: 10.2741/4363.

[4] Chuang SC, Fanidi A, Ueland PM, Relton C, Midttun O, Vollset SE, *et al*. Circulating biomarkers of tryptophan and the kynurenine pathway and lung cancer risk. Cancer Epidemiol Biomarkers Prev. 2014; 23(3): 461-8. doi: 10.1158/1055-9965.Epi-13-0770.

[5] Walczak K, Wnorowski A, Turski WA, Plech T. Kynurenic acid and cancer: facts and controversies. Cell Mol Life Sci. 2020; 77(8): 1531-50. doi: 10.1007/s00018-019-03332-w.

[6] Mor A, Tankiewicz-Kwedlo A, Pawlak D. Kynurenines as a novel target for the treatment of malignancies. Pharmaceuticals (Basel). 2021;14(7): 606. doi: 10.3390/ph14070606.

[7] Schmidt DR, Patel R, Kirsch DG, Lewis CA, Vander Heiden MG, Locasale JW. Metabolomics in cancer research and emerging applications in clinical oncology. CA Cancer J Clin. 2021; 71(4): 333-58. doi: 10.3322/caac.21670.

[8] Armitage EG, Southam AD. Monitoring cancer prognosis, diagnosis and treatment efficacy using metabolomics and lipidomics. Metabolomics. 2016; 12(9): 146. doi: 10.1007/s11306-016-1093-7.

[9] Kannampuzha S, Mukherjee AG, Wanjari UR, Gopalakrishnan AV, Murali R, Namachivayam A, *et al*. A systematic role of metabolomics, metabolic pathways, and chemical metabolism in lung cancer. Vaccines (Basel). 2023;11(2):381. doi: 10.3390/vaccines11020381.

[10] Liesenfeld DB, Habermann N, Owen RW, Scalbert A, Ulrich CM. Review of mass spectrometry-based metabolomics in cancer research. Cancer Epidemiol

Biomarkers Prev. 2013; 22(12): 2182-201. doi: 10.1158/1055-9965.Epi-13-0584.

[11] Lee KB, Ang L, Yau WP, Seow WJ. Association between metabolites and the risk of lung cancer: a systematic literature review and meta-analysis of observational studies. Metabolites. 2020; 10(9): 362. doi: 10.3390/metabo10090362.

[12] Gasparri R, Sedda G, Caminiti V, Maisonneuve P, Prisciandaro E, Spaggiari L. Urinary biomarkers for early diagnosis of lung cancer. J Clin Med. 2021; 10(8): 1723. doi: 10.3390/jcm10081723.

[13] Jordaens S, Zwaenepoel K, Tjalma W, Deben C, Beyers K, Vankerckhoven V, *et al*. Urine biomarkers in cancer detection: a systematic review of preanalytical parameters and applied methods. Int J Cancer. 2023; 152(10): 2186-205. doi: 10.1002/ijc.34434.

[14] Kalantari K, Bolton WK. A good reason to measure 24-hour urine creatinine excretion, but not to assess kidney function. Clin J Am Soc Nephrol. 2013; 8(11): 1847-9. doi: 10.2215/cjn.09770913.

[15] Khamis MM, Adamko DJ, El-Aneed A. Mass spectrometric based approaches in urine metabolomics and biomarker discovery. Mass Spectrom Rev. 2017; 36(2): 115-34. doi: 10.1002/mas.21455.

[16] Yang Q, Shi X, Wang Y, Wang W, He H, Lu X, *et al*. Urinary metabonomic study of lung cancer by a fully automatic hyphenated hydrophilic interaction/RPLC-MS system. J Sep Sci. 2010; 33(10): 1495-503. doi: 10.1002/jssc.200900798.

[17] Pearson MA, Lu C, Schmotzer BJ, Waller LA, Riederer AM. Evaluation of physiological measures for correcting variation in urinary output: Implications for assessing environmental chemical exposure in children. J Expo Sci Environ Epidemiol. 2009; 19(3): 336-42. doi: 10.1038/jes.2008.48.

[18] Wijemanne N, Soysa P, Wijesundara S, Perera H. Development and validation of a simple high performance liquid chromatography/UV method for simultaneous determination of urinary uric acid, hypoxanthine, and creatinine in human urine. Int J Anal Chem. 2018; 2018: 1647923. doi: 10.1155/2018/1647923.

[19] Garde AH, Hansen AM, Kristiansen J, Knudsen LE. Comparison of uncertainties related to standardization of urine samples with volume and creatinine concentration. Ann Occup Hyg. 2004; 48(2): 171-9. doi: 10.1093/annhyg/meh019.

[20] Boeniger MF, Lowry LK, Rosenberg J. Interpretation of urine results used to assess chemical exposure with emphasis on creatinine adjustments: a review. Am Ind Hyg Assoc J. 1993; 54(10): 615-27. doi: 10.1080/15298669391355134.

[21] Sutariya V, Wehrung D, Geldenhuys WJ. Development and validation of a novel RP-HPLC method for the analysis of reduced glutathione. J Chromatogr Sci. 2012; 50(3): 271-6. doi: 10.1093/chromsci/bmr055.

[22] Bijukumar G, Maloyesh B, Sampat S, Bhirud SB, Rajendra A. Efficient synthesis of sivelestat sodium hydrate. Synthetic Communications. 2008; 38: 1718-24. doi: 10.1080/00397910801982373

[23] Locke D, Bevans CG, Wang LX, Zhang Y, Harris AL, Lee YC. Neutral, acidic, and basic derivatives of anthranilamide that confer different formal charge to reducing oligosaccharides. Carbohydr Res. 2004; 339(2): 221-31. doi: 10.1016/j.carres.2003.10.020.

[24] Morimoto K, Nishimura K, Miyasaka S, Maeta H, Taniguchi I. The effect of sivelestat sodium hydrate on severe respiratory failure after thoracic aortic surgery with deep hypothermia. Ann Thorac Cardiovasc Surg. 2011; 17(4): 369-75. doi: 10.5761/atcs.oa.10.01555.

[25] Resines JA, Arín MJ, Díez MT. Determination of creatinine and purine derivatives in ruminants' urine by reversed-phase high-performance liquid chromatography. J Chromatogr. 1992; 607(2): 199-202. doi: 10.1016/0021-9673(92)87075-j.

[26] Yokoyama Y, Tsuchiya M, Sato H, Kakinuma H. Determination of creatinine and ultraviolet-absorbing amino acids and organic acids in urine by reversed-phase high-performance liquid chromatography. J Chromatogr. 1992; 583(1): 1-10. doi: 10.1016/0378-4347(92)80338-q.

[27] Zuo Y, Wang C, Zhou J, Sachdeva A, Ruelos VC. Simultaneous determination of creatinine and uric acid in human urine by high-performance liquid chromatography. Anal Sci. 2008; 24(12): 1589-92. doi: 10.2116/analsci.24.1589.

[28] Muthusamy S, Palanisamy S, Ramalingam S. Exposure of bisphenol A in breast cancer patients-quantitatively assessed by sensitivity-enhanced high-performance liquid chromatography coupled with fluorescence detection: A case-control study. Biomed Chromatogr. 2021; 35(9): e5137. doi: 10.1002/bmc.5137.

[29] Pawul-Gruba M, Kiljanek T, Madejska A, Osek J. Development of a high performance liquid chromatography with diode array detector (HPLC-DAD) method for determination of biogenic amines in ripened cheeses. Molecules. 2022; 27(23): 8194. doi: 10.3390/molecules27238194.

[30] Martínez-Navarro EM, Cebrián-Tarancón C, Moratalla-López N, Lorenzo C, Alonso GL, Salinas RM. Development and validation of an HPLC-DAD method for determination of oleuropein and other bioactive compounds in olive leaf by-products. J Sci Food Agric. 2021; 101(4): 1447-53. doi: 10.1002/jsfa.10758.

[31] Solano-Cueva N, Figueroa JG, Loja C, Armijos C, Vidari G, Ramírez J. A Validated HPLC-UV-ESI-IT-MS method for the quantification of carnosol in lepechinia mutica, a medicinal plant endemic to ecuador. Molecules. 2023; 28(18): 6701. doi: 10.3390/molecules28186701.

[32] Jîtcă G, Fogarasi E, Ősz BE, Vari CE, Tero-Vescan A, Miklos A, et al. A simple HPLC/DAD method validation for the quantification of malondialdehyde in rodent's brain. Molecules. 2021; 26(16): 5066. doi: 10.3390/molecules26165066.

[33] Witte EC, Lambers Heerspink HJ, de Zeeuw D, Bakker SJ, de Jong PE, Gansevoort R. First morning voids are more reliable than spot urine samples to assess microalbuminuria. J Am Soc Nephrol. 2009; 20(2):

436-43. doi.org/10.1681/asn.2008030292.

[34] Birková A, Valko-Rokytovská M, Hubková B, Zábavníková M, Mareková M. Strong dependence between tryptophan-related fluorescence of urine and malignant melanoma. Int J Mol Sci. 2021; 22(4): 1884. doi: 10.3390/ijms22041884.

[35] Bouatra S, Aziat F, Mandal R, Guo AC, Wilson MR, Knox C, *et al*. The human urine metabolome. PLoS One. 2013; 8(9): e73076. doi: 10.1371/journal.pone.0073076.

Journal of Associated Medical Sciences 2024; 57(3): 75-82

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 development and psychometric properties of the Chiang Mai Aphasia Screening Test for stroke

Atchaphan Ruangsuk 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 20 May 2024
Accepted as revised 1 July 2024
Available online 7 July 2024

Keywords:
Aphasia, speech, language, screening test, stroke.

ABSTRACT

Background: The aphasia screening test detects language and speech impairments, clarifying individuals' language and speech abilities before administering a standardized aphasia diagnosis test.

Objective: This study aimed to develop and validate an aphasia screening test for suspected cerebrovascular accident (CVA) patients with communication difficulties.

Materials and methods: The study underwent two phases: developing and assessing psychometric properties. Five experts established content validity across receptive language, expressive language, reading, and writing. The Chiang Mai Aphasia Screening Test (CMAST) was evaluated on 14 CVA patients with and 14 without aphasia.

Results: The content validity showed item-objective congruence ranging from 0.80 to 1.00. Sensitivity and specificity were 96.30% and 69%, respectively, with a maximum Youden's Index at 65.30% and a cut-off point of 43 points. Concurrent validity was high (phi coefficient =0.67), and significant score differences (p<0.001) in construct validity confirmed the tool's ability to distinguish aphasic from non-aphasic patients. Inter-rater reliability (ICC=0.99) and internal consistency (Cronbach's alpha =0.97, 95% CI 0.95-0.98) were observed.

Conclusion: The CMAST, comprising 45 items, exhibits sufficient validity and reliability for screening individuals suspected of aphasia due to CVA.

Introduction

Aphasia, characterized by language abnormalities affecting receptive, expressive, and overall linguistic abilities, often arises from neurological dysfunction.[1] These difficulties can profoundly impact various facets of life, including self-care, education, work, social engagement, and leisure activities.[2] The prevalence of aphasia is notable among individuals with traumatic brain injury, meningitis, and cerebrovascular accidents (CVAs), with an incidence ranging from 4-20% among CVA patients.[1, 3]

According to statistics from the Division of Non-Communicable Disease, Ministry of Public Health, Thailand, the prevalence of CVA patients in Thailand has consistently increased from 2016 to 2018, with an annual rate of 3.39 percent. In 2018, the incidence rates of CVA patients in Thailand and Chiang Mai province were as high as 506.20 and 490.59 per 100,000 population, respectively.[4]

Assessment and management of aphasia are integral aspects of rehabilitation, and Speech-Language

* **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: supaporn.c@cmu.ac.th
doi: 10.12982/JAMS.2024.049
E-ISSN: 2539-6056

Pathologists (SLPs) play a crucial role in these processes. Evaluation typically involves bedside screening and standardized tests, which may take 1 to 6 hours.[5-12] However, the exhaustive nature of these evaluations can pose challenges for patients experiencing weakness or fatigue during recovery. As such, briefer screening tests, lasting between 3 to 15 minutes, are often conducted when symptoms stabilize or within 2-3 days post-stroke onset.[13] These screenings aimed to differentiate between CVA patients with aphasia and those without, informing treatment planning, rehabilitation approaches, and patient and caregiver advice.

Various screening tests are currently available for assessing aphasia in different languages, such as the Frenchay Aphasia Screening Test (FAST), Mississippi Aphasia Screening Test (MAST), ScreeLing, Ullevaal Aphasia Screening Test (UAS), Bedside of Western Aphasia Battery-Revised, and Quick Aphasia Battery (QAB).[5-12] However, Speech-Language Pathologists in Thailand face limited options, with the Aphasia Screening Test Saraburi SD-SLP-01 being the only available tool.[14] Consequently, there is a pressing need to develop a new test or translate an existing international test into Thai to expand the range of assessment options for language and speech disorders within the Thai context.

In response to this need, the Chiang Mai Aphasia Screening Test (CMAST) has been developed as an assessment tool to address the identified gaps in screening options. The primary objective of this study is to create a Thai-language version of an aphasia screening test that comprehensively covers all four language aspects: receptive language (including auditory comprehension, word recognition, and sequential commands), expressive language (spontaneous speech, verbal fluency, repetition, and naming), as well as reading and writing abilities. By rigorously evaluating the psychometric properties of CMAST, including its validity and reliability, this initiative aims to significantly enhance the repertoire of assessment tools available to Speech-Language Pathologists in Thailand. This expansion of options will enable SLPs to conduct more comprehensive evaluations of language and speech abilities among individuals with neurological conditions, thereby facilitating more accurate identification and management of communication difficulties.

Materials and methods

The study involves research and development of an aphasia screening test, assessing its psychometric properties in two phases.

Phase 1 Developing a screening test and evaluating content validity.

The researchers established objectives, identified target groups, and defined established criteria. We then explored information, including concepts, theories, and existing literature, to shape the content of the screening test and clinical expertise and experiences. Finally, we designed and created the first version of the aphasia screening test.[15] The screening questionnaire is formulated using concise and clear language. Each question is designed to have a singular and unambiguous interpretation, utilizing the Thai language appropriately. The questions that cover all four language aspects,[15] namely (1) receptive language comprises auditory comprehension 37 items, word recognition 17 items, and sequential commands 11 items (2) expressive language comprises spontaneous speech 5 items, verbal fluency 12 items, repetition 13 items, and naming 17 items (3) reading 12 items and (4) writing 10 items. Total 134 items. Scoring criteria: 1 point for correct, 0 for incorrect. Content validity is evaluated by five experts, each with over ten years of clinical experience. The panel comprises two SLPs specialized in aphasia screening, one communication disorders professor, one Occupational Therapist focusing on CVA care, and one rehabilitation physician. Subsequently, questions with an IOC value of 0.8 or higher will be retained, and those below 0.8 will be excluded. Questions in each category were chosen through the quota sampling method. The revised communication disorder screening questionnaire will be administered to individuals with 3 CVA without aphasia to explore image size and clarity issues. Based on these results, a comprehensive version of the aphasia screening test called the Chiang Mai Aphasia Screening Test (CMAST), will be created in Appendix 1.

Phase 2 The validation phase assesses psychometric properties.

Thai individuals aged 20 and above in the province of Chiang Mai, Thailand, who have been diagnosed with CVA by a rehabilitation physician and are currently receiving speech therapy. These facilities include 1) the Speech Therapy Clinic, Faculty of Associated Medical Sciences, Chiang Mai University; 2) PROMPT Health Center, Chiang Mai University; 3) the Community Health Clinic, Nong Pa Khrang; and 4) the Disability Service Center; Nong Kwai. The computed sample size is 12 participants. To enhance the study's reliability and address potential dropouts, we added 15% more participants. Consequently, the final sample size comprises 14 individuals with aphasia and 14 without aphasia.

CVA without aphasia. Inclusion criteria: (1) fluency in Thai language, (2) literacy in Thai before CVA, (3) normal or corrected vision and hearing, and (4) determination of non-aphasia through the Saraburi Aphasia Screening Test: SD-SLP-01, with SLPs requiring a score of 27 or higher.[14] Exclusion criteria: (1) neurological abnormalities unrelated to cerebral vascular disease, (2) incomplete assessment during research, (3) desire to withdraw from ongoing research, (4) inability to write due to hand muscle issues, and (5) blood pressure exceeds 180/110 mmHg.[13]

CVA with aphasia. Inclusion criteria: (1) fluency in Thai language, (2) literacy in Thai before CVA, (3) normal or corrected vision and hearing, (4) aphasia is determined through screening with the Saraburi Aphasia Screening Test: SD-SLP-01,[16] and (5) diagnosed with language and/or speech abnormalities resulting from CVA for at least 6 months.[17, 18] Exclusion criteria: same as described above.

Procedures

Before commencing data collection, it is essential to obtain consent and permission from the subjects to participate in the research, which should be documented through signed consent forms. The research project has received ethical approval from the ethical committees of the Faculty of Associated Medical Sciences, Chiang Mai University, with reference number AMSEC-64FB-005. A comprehensive assessment involved 14 individuals with aphasia, evaluated using CMAST and the Thai version of the Western Aphasia Battery (Thai WAB),[19] and 14 non-aphasia individuals who underwent CMAST evaluation. Two SLPs assessed the video recordings of all 28 participants. An expert in aphasia screening scored recordings using a 1-point criterion, while Thai WAB [19] determined aphasia presence. Statistical analysis assessed reliability and validity.

Data analysis

SPSS Version 25 is used for analysis, employing descriptive statistics for demographic characteristics. Validation involves sensitivity, specificity, and construct validity using the Known-groups technique analyzed with the Mann-Whitney U test and concurrent validity. Reliability is assessed using the Intraclass Correlation Coefficient (ICC) with the Two-way random effect model and internal consistency using Cronbach's alpha coefficient.

Results

The demographic characteristics of each group are presented in Table 1. Participants were divided into two groups: one with aphasia and one without aphasia. The aphasia group included 8 males and 6 females, with mean ages of 60 (SD=10.64) for males and 60 (SD=10.36) for females. All had been diagnosed with cerebrovascular accident (CVA) for over 6 months. The non-aphasia group also consisted of 8 males and 6 females, with mean ages of 59.75 years (SD=8.41) for males and 46.33 years (SD=17.65) for females. None of these individuals had aphasia.

Table 1 Demographic characteristics.

Participant group	N	Type of aphasia	Gender	Number of participants	Mean age (year)	Medical diagnosis	Onset (month)
With aphasia	01	Broca's	Male	8	60.70 (SD=10.36)	Cerebrovascular Accident (CVA)	>6
	02	Broca's					
	03	Anomic					
	04	Wernicke's					
	05	Broca's					
	06	Anomic					
	07	Global					
	08	Transcortical Motor	Female	6			
	09	Anomic					
	10	Global					
	11	Anomic					
	12	Anomic					
	13	Anomic					
	14	Global					
Without aphasia			Male	8	59.75 (SD=8.41)		
			Female	6	46.33 (SD=17.65)		
Mean age of the 28 participants =57.04 years (SD=12.56) with age range 20-76 years.							

Phase 1: Develop a screening test and evaluate content validity and trials.

The questions that cover all four language aspects, namely (1) receptive language comprises auditory comprehension 37 items, word recognition 17 items, and sequential commands 11 items (2) expressive language comprises spontaneous speech 5 items, verbal fluency 12 items, repetition 13 items, and naming 17 items (3) reading 12 items and (4) writing 10 items. Total 134 items. The study results indicate that this screening tool demonstrated content validity in receptive language, expressive language, reading, and writing, with an IOC of 0.80-1.00. Evaluations were conducted in receptive language, expressive language, reading, and writing to assess language and speech abilities in individuals with aphasia. Additionally, the CMAST was tested on those with CVA without aphasia. The findings indicated a clear understanding of questions, legible letter reading, and appropriately sized images, leading to using the CMAST.

Phase 2 The validation phase assesses psychometric properties.

The CMAST was compared to the diagnosis results obtained from the gold standard using the Thai WAB.[19] This comparison aimed to calculate sensitivity and specificity. The sensitivity and specificity results indicate the accuracy of diagnosing individuals with and without the condition, as demonstrated in Table 2. The highest Youden's index value[20] of 65.30% and a cut-off point of 43 points indicate CMAST's 96.30% sensitivity in detecting aphasia and 69% specificity in distinguishing individuals with and without aphasia.

The Phi correlation statistic yielded 0.67, indicating CMAST's high concurrent validity with gold standard examination results at a significant level.[21] Mann-Whitney U Test results (Table 3) revealed a $p<0.001$, signifying significant differences in Mean Rank between aphasia and non-aphasia groups.

Table 2. Cut-off, sensitivity, specificity, Youden, and phi value of CMAST.

Cut-off	Sensitivity (%)	Specificity (%)	Youden (Se+Sp-1) (%)	Phi value
24	7.4	100	7.4	0.2
25	14.8	100	14.8	0.29
26	22.2	100	22.2	0.36
28	25.9	100	25.9	0.39
33	29.6	100	29.6	0.42
34	44.4	100	44.4	0.54
36	48.1	100	48.1	0.57
38	55.6	89.7	45.3	0.48
39	70.4	89.7	60.1	0.61
40	77.8	86.2	64.0	0.64
41	81.5	79.3	60.8	0.61
42	85.2	75.9	61.1	0.61
43	96.3	69.0	65.3	0.67
44	100	55.2	55.2	0.61

Table 3. Result of construct validity by using the Mann-Whitney U test.

Participant group	N	Mean rank	Z	*p* value
With aphasia	14	8.75	3.74	<.001
Without aphasia	14	20.25		
Total	28			

Thus, CMAST demonstrates construct validity, effectively discerning between CVA with and without aphasia.[22] The results showed inter-rater reliability for each question, with ICC values ranging from 0.76 to 1.00, signifying good to excellent agreement. The overall CMAST score's ICC value was 0.99, indicating highly consistent inter-rater reliability,[23] detailed in Table 4. Internal consistency assessment yielded a Cronbach's alpha coefficient of 0.97 for CMAST (Table 5), indicating excellent consistency among item scores,[24] indicating consistent content across the items.

Table 4. Result of inter-rater reliability.

Item number	Intraclass correlation[b]	95% Confidence interval	
		Lower bound	Upper bound
Spontaneous speech			
1	0.862	0.700	0.936
2	0.884	0.749	0.946
Verbal fluency			
3	0.881	0.745	0.945
4	1.000		
5	1.000		
Auditory comprehension			
6	0.760	0.479	0.889
7	1.000		
8	1.000		
9	0.794	0.559	0.904
10	1.000		
11	0.794	0.559	0.904
12	0.834	0.639	0.924
13	0.867	0.715	0.938
14	1.000		
15	1.000		
16	1.000		
17	1.000		
18	1.000		
19	1.000		
20	1.000		
Sequential commands			
21	1.000		
22	0.794	0.559	0.904
23	0.802	0.572	0.908
24	0.794	0.559	0.904
Naming			
25	0.932	0.854	0.968
26	0.942	0.876	0.973
27	1.000		
28	0.832	0.637	0.922
29	0.932	0.854	0.968
30	0.765	0.499	0.891
Repetition			
31	1.000		
32	1.000		
33	1.000		
34	1.000		
35	0.838	0.654	0.925
Word recognition			
36	1.000		
37	1.000		
38	1.000		
39	0.794	0.565	0.903
40	1.000		
41	0.794	0.565	0.903
Reading			
42	1.000		
43	0.962	0.918	0.982
Writing			
44	1.000		
45	0.964	0.923	0.983
Total	**0.993**	**0.985**	**0.997**

Note: a: two-way random effects model where both people and measures effects are random,
b: type A intraclass correlation coefficients using an absolute agreement definition.

Table 5. Result of internal consistency.

Subtypes of CMAST	Number of items	Cronbach's alpha coefficient	internal consistency	95% Confidence interval	
				Lower bound	Upper bound
Expressive language	16	0.95	Excellent	0.92	0.97
Receptive language	25	0.97	Excellent	0.94	0.98
Reading and Writing	4	0.88	Good	0.79	0.94
Total	**45**	**0.97**	**Excellent**	**0.95**	**0.98**

Discussion

The CMAST has an IOC value of 0.80-1.00, with a sensitivity of 96.30% in detecting aphasia and a specificity of 69%. Furthermore, it can be compared to the results of the Saraburi Aphasia Screening Test: SD-SLP-01,[14] which has sensitivity and specificity values of 92.50% and 84.62%, respectively[18] It's important to note that sensitivity and specificity depend on the population tested. Testing in a patient group compared to a control group of normal individuals would likely yield higher true positives in the patient group and true negatives in the control group, resulting in higher sensitivity and specificity than in reality.[27] Concurrent validity is high, as indicated by a substantial phi coefficient of 0.67 from the correlation analysis. This can be compared to the results of the Saraburi Aphasia Screening Test: SD-SLP-01,[14] which has a concurrent validity of 0.97.[18]

In summary, CMAST demonstrates high concurrent validity, showing a strong relationship with the diagnosis results obtained from the gold standard examination at a significant level.[23] It can be compared to the results of the Saraburi Aphasia Screening Test: SD-SLP-01.[14] The $p<0.001$ indicates significant differences between the aphasia and non-aphasia groups. Therefore, CMAST demonstrates construct validity by distinguishing between individuals with CVA and with or without aphasia.[24] The inter-rater reliability for the overall CMAST score with an ICC value of 0.99 indicates highly consistent inter-rater reliability.[25] This suggests that CMAST is reliable, allowing for consistent results even when assessed by different evaluators. The internal consistency using Cronbach's Alpha yielded a value of 0.97, indicating excellent reliability.[26] Additionally, it highlights the comparison with the Saraburi Aphasia Screening Test: SD-SLP-01,[14] where both tools demonstrate consistent alignment within their content, as indicated by similar Cronbach's Alpha values.[18] The psychometric properties of CMAST, which demonstrate content validity, sensitivity, specificity, construct validity, concurrent validity, and reliability, are excellent. This includes excellent inter-rater reliability (r=0.99) and high internal consistency (Cronbach's Alpha =0.97). Therefore, CMAST is deemed suitable for screening individuals with CVA and aphasia.

Conclusion

The newly developed CMAST is a valuable tool for screening language skills, including all four language aspects, namely (1) receptive language comprises auditory comprehension, word recognition, and sequential commands (2) expressive language comprises spontaneous speech, verbal fluency, repetition, and naming (3) reading and (4) writing. The criteria are well-defined, making result interpretation easy, and the screening takes only 10-15 minutes. CMAST has undergone thorough testing, ensuring it is both valid and reliable. This makes it suitable for clinical use in screening CVA with aphasia. These findings are crucial for shaping treatment plans and rehabilitation approaches and guiding patients and their caregivers.

Limitation

Collecting data from a small sample group of individuals with CVA in Chiang Mai, using the Thai language for screening, may result in assessment differences compared to evaluating patients in different contexts. Factors such as geographical region and hospital characteristics (government or private) can contribute to variations in the assessment outcomes. Moreover, it was observed that when screening individuals with CVA with aphasia, despite framing questions with predefined responses or speech targets, patients' responses varied. Therefore, when used for screening in individual cases, the answers obtained may differ from the anticipated responses. Scoring in that specific subtest may depend on the evaluator's discretion. The researcher suggests conducting further studies with a larger sample size of CVA patients with aphasia to differentiate between types of aphasia and assess severity. This comparison could be made using the Thai WAB.[13]

Statement of ethics

This study was approved by the Research Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University (Approval ID: AMSEC-64FB-005)

Conflict of interest

The authors have declared that no competing interests existed at the time of publication.

References

[1] Owens REJ. Introduction to communication disorders: a lifespan evidence-based perspective. 4th Ed. Boston: Pearson Education; 2011.

[2] Cherney LR, Patterson JP, Raymer AM. Intensity of aphasia therapy: evidence and efficacy. Curr Neurol Neurosci Rep. 2011; 11(6): 560-9. doi: 10.1007/s11910-011-0227-6.

[3] Rhea P, Norbury C, Gosse C. Language disorders from infancy through adolescence: listening, speaking, reading, writing, and communicating. 5th Ed. Maryland Heights, MO: Elsevier/Mosby; 2018.

[4] Ministry of Public Health. Number and rate of stroke patients in the years 2016-2018. In: Non-communicable diseases division, Ed. 2019.

[5] Enderby P, Crow E. Frenchay Aphasia Screening Test: validity and comparability. Disabil Rehabil. 1996; 18(5): 238-40. doi: 10.3109/09638289609166307.

[6] Enderby PM, Wood VA, Wade DT, Hewer RL. The Frenchay Aphasia Screening Test: a short, simple test for aphasia appropriate for non-specialists. Int Rehabil Med. 1987; 8(4): 166-70. doi: 10.3109/03790798709166209.

[7] Nakase-Thompson R, Manning E, Sherer M, Yablon S, Gontkovsky S, Vickery C. Brief assessment of severe language impairments: initial validation of the Mississippi aphasia screening test. Brain Inj. 2005; 19(9): 685-91. doi: 10.1080/02699050400025331.

[8] Doesborgh SJ, van de Sandt-Koenderman WM, Dippel DW, van Harskamp F, Koudstaal PJ, Visch-Brink EG. Linguistic deficits in the acute phase of stroke. J Neurol. 2003; 250(8): 977-82. doi: 10.1007/s00415-003-1134-9.

[9] El Hachioui H, Van de Sandt-Koenderman M, Dippel D, Koudstaal P, Visch-Brink E. The ScreeLing: occurrence of linguistic deficits in acute aphasia post-stroke. J Rehabil Med. 2012; 44: 429-35. doi: 10.2340/16501977-0955.

[10] Thommessen B, Thoresen GE, Bautz-Holter E, Laake K. Screening by nurses for aphasia in stroke-the Ullevaal Aphasia Screening (UAS) test. Disabil Rehabil. 1999; 21(3): 110-5. doi: 10.1080/096382899297846.

[11] El Hachioui H, Visch-Brink EG, de Lau LM, van de Sandt-Koenderman MW, Nouwens F, Koudstaal PJ, *et al*. Screening tests for aphasia in patients with stroke: a systematic review. J Neurol. 2017; 264(2): 211-20. doi: 10.1007/s00415-016-8170-8.

[12] Wilson SM, Eriksson DK, Schneck SM, Lucanie JM. A quick aphasia battery for efficient, reliable, and multidimensional assessment of language function. PLoS One. 2018; 13(2): 1-29. doi: 10.1371/journal.pone.0192773.

[13] Institute of Neurology. Clinical practice guideline for stroke rehabilitation. 3rd Ed. Bangkok: Institute of Neurology; 2016.

[14] Supawatrajarugon R, Lorthanawatthapong P, Eua-siriratnaphaisan A. Content validity and reliability of the Saraburi Aphasia Screening Test: SD-SLP-01. In: Proceedings of the annual academic conference 2013, Faculty of Medicine, Srinakharinwirot University; February 3-4, 2014; Nakhon Nayok: Faculty of Medicine, Srinakharinwirot University; 2014.

[15] Chapey R. Language intervention strategies in aphasia and related neurogenic communication disorders. 5th Ed. Philadelphia: Lippincott Williams & Wilkins; 2008.

[16] Supawatrajarugon R, Eua-siriratnaphaisan A. Standardization of the Saraburi Aphasia Screening Test: SD-SLP-01 in stroke patients and control group. In: Proceedings of the annual academic conference 2016, Thai Association of Speech and Hearing Sciences; 2016; Bangkok: Ramada Plaza Menam Riverside Hotel.

[17] Kiran S. What is the nature of poststroke language recovery and reorganization? ISRN Neurol. 2012; 2012: 786872. doi: 10.5402/2012/786872.

[18] Breitenstein C, Martus P, Willmes K, Ziegler W, Baumgaertner A. Intensive speech and language therapy after stroke - Authors' reply. Lancet. 2017; 390(10091): 228-9. doi: 10.1016/S0140-6736(17)31801-9.

[19] 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.

[20] Schisterman E, Perkins N, Liu A, *et al*. Optimal cut-point and its corresponding Youden index to discriminate individuals using pooled blood samples. Epidemiology (Cambridge, Mass) 2005; 16: 73-81. doi: 10.1097/01.ede.0000147512.81966.ba.

[21] Yule GU. On the methods of measuring association between two attributes. J R Stat Soc. 1912; 75(6): 579-652. doi:10.1111/J.2397-2335.1912.

[22] Cohen RJ, Swerdlik ME. Psychological testing and assessment: An introduction to tests and measurement. 6th Ed. New York: McGraw-Hill; 2005.

[23] 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.

[24] Gliem J, Gliem R. Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. In proceedings of Midwest research to practice conference in adults, continuing and community education; 2003 Oct 8-10; Columbus: The Ohio State University; 2003. p.82-8.

Appendix 1

Example of Chiang Mai Aphasia Screening Test (CMAST)

ภาคผนวก ข

แบบคัดกรองภาวะเสียการสื่อความฉบับสมบูรณ์

ข้อคำถาม	เกณฑ์การให้คะแนน		หมายเหตุ
	ถูก (1)	ผิด (0)	
การพูดโต้ตอบ			
1. บอกที่อยู่อาศัย			
2. อธิบายรูปภาพ (ภาพกิจกรรมในครัว)			
การพูดคล่อง			
3. นับเลข 1-10			
4. ไก่งามเพราะขน คนงามเพราะ..(แต่ง)..			
5. เกลือมีรส...(เค็ม)...			
การฟังเข้าใจภาษา			
6. คุณชื่อจันทร์ใช่หรือไม่			
7. คุณชื่อ.....(ชื่อผู้ป่วย)..... ใช่หรือไม่			
8. คุณอาศัยอยู่ที่...(ไม่ใช่จังหวัดที่ผู้ป่วยอาศัย)....ใช่หรือไม่			
9. ที่นี่คือโรงพยาบาลใช่หรือไม่			
10. คุณเป็นหมอใช่หรือไม่			
11. ไก่ตัวใหญ่กว่าวัวใช่หรือไม่			
12. พระอาทิตย์ขึ้นทางทิศตะวันตกใช่หรือไม่			

Journal of Associated Medical Sciences 2024; 57(3): 83-89

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Factors associated with musculoskeletal pain in ambulatory individuals with spinal cord injury

Narongsak Khamnon[1,2] Sugalya Amatachaya[2,3] Thiwabhorn Thaweewannakij[2,3] Lugkana Mato[2,3*]

[1]Department of Physical Therapy, School of Integrative Medicine, Mae Fah Luang University, Chiang Rai Province, Thailand.
[2]Improvement of Physical Performance and Quality of Life (IPQ) Research Group, Khon Kaen University, Khon Kaen Province, Thailand.
[3]School of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen Province, Thailand.

ARTICLE INFO

Article history:
Received 14 March 2024
Accepted as revised 3 July 2024
Available online 7 July 2024

Keywords:
Musculoskeletal disorders, factors, pain, neurological deficits.

ABSTRACT

Background: Such myriad consequences of musculoskeletal pain were common and distorted the level of functioning after spinal cord injury (SCI). No known data identifies pain-related factors in ambulatory individuals with SCI.

Objective: To examine the risk factors for musculoskeletal pain in ambulatory patients with SCI.

Materials and methods: A total of 138 ambulatory participants with SCI were interviewed and evaluated for their demographics, SCI traits, levels of locomotor disability (Functional Independence Measure-Locomotor), and data on musculoskeletal pain, including area, cause, and severity of pain using a body chart diagram and Visual Analogue Scale. Then, logistic regression was used to analyze risk factors associated with musculoskeletal pain.

Results: Precisely, 55.07% of the 138 ambulatory individuals with SCI reported musculoskeletal pain. The common top three areas of pain include the lower back (27%), hip (23%) and shoulder (15%). Factors associated with musculoskeletal pain include being a man and FIM-L 6 in which males had OR=3.56 (95% CI, 1.62-7.84; p=0.002) and FIM-L 6 had OR=2.57 (95% CI, 1.08-6.10; p=0.032).

Conclusion: The present findings revealed that musculoskeletal pain is highly prevalent in ambulatory individuals with SCI. Ambulatory individuals with SCI who are males and able to walk at least 50 meters while using gait devices are most concerned about musculoskeletal pain problems.

Introduction

Spinal cord injury (SCI) directly results in permanent sensorimotor and autonomic deterioration.[1] In addition, an increasing number of coronaviruses in 2019 (COVID-19) affected adaptation to the new normal medical services model, which limited the number of people admitted to rehabilitation care.[2] Therefore, ambulation with disabilities without health care follow-up after SCI may increase complications.[3]

Musculoskeletal pain is a chronic condition commonly reported in individuals with SCI.[4] It is characterized by injury related to the overuse, strain, and wear and tear of the musculoskeletal structure.[5] It usually gets worse with movement and better with rest.[5] A previous study showed that musculoskeletal pain frequently occurred in individuals with SCI in the first 5 years after injury[4]. Another study reported that 68 of 100 individuals with SCI experienced musculoskeletal pain.[5] The problem with pain significantly

* **Corresponding contributor.**
Author's Address: *School of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen Province, Thailand.*
E-mail address: yui@kku.ac.th
doi: 10.12982/JAMS.2024.050
E-ISSN: 2539-6056

interfered with mental and physical health, which led to a significant increase in the rate of hospitalization and cost of care, as well as a decline in quality of life.[5-8] Despite evidence that musculoskeletal pain is a common problem for individuals with SCI, the various factors associated with the development of musculoskeletal pain in SCI among ambulatory individuals are unknown.

Regarding concerns about advancing age, which declines in the musculoskeletal system, daring behaviors in males, having an excessive BMI, experiencing high levels of locomotor disability, as well as the cause, level, severity, and stage of the injury, are most likely to be associated with pain. However, existing evidence has not yet supported these hypotheses in ambulatory individuals with SCI.[5-14] The data regarding factors associated with musculoskeletal pain could provide for ambulatory individuals with SCI who have a high risk of developing this issue. Therefore, the study aimed to investigate risk factors associated with musculoskeletal pain in ambulatory individuals with SCI.

Materials and methods

Participants

Ambulatory individuals with SCI who visited community hospitals and tertiary rehabilitation centers were invited to participate in this study. The inclusion criteria were ambulatory individuals with SCI, both traumatic and non-traumatic causes, aged at least 18 years, and capable of walking independently, with or without a walking device, for at least 15 meters.[15,16] Participants were excluded if they had joint or extremity deformities, had present medical conditions that might limit walking ability, could not understand and follow commands, and had any underlying diseases that caused pain.[14] The ethical committee (HE 591462) authorized the trial, and eligible subjects provided written informed permission before participating in the study.

Protocol

Participants were interviewed and assessed by one physiotherapist for their demographic information (including gender, age, body weight, and height) and SCI features, such as post-injury duration, cause, severity, and levels of SCI diagnosis using the International Standards for Neurological Classification of SCI (ISNCSCI).[17] Subsequently, the physiotherapist assessed participants present level of musculoskeletal pain using dichotomous variables (yes /no pain), areas of pain using a body chart diagram, and intensity of pain using the Visual Analogue Scale (VAS), ranging from 0 to 10.[18] The participants were first interviewed about their musculoskeletal pain. The area of pain was localized using a body chart diagram. The participants evaluated pain intensity using the VAS scale by placing a mark at a point on a 10 cm line to rate the pain intensity; the far left indicated no pain, and the far right indicated the most intense pain imaginable.[18,19] The definition of musculoskeletal pain in this study was bone, joint, or muscle pain caused by repetitive strain or overuse.[4] In addition, participants who suffer from neuropathic pain often describe the sensation as being like an electric shock or a burning sensation.[20] active, passive, and accessory movement tests were used to differentiate musculoskeletal pain from neurological pain. Furthermore, we distinguish neuropathic pain from musculoskeletal pain using this essential feature of neuropathic pain.[4,20] Individuals with SCI with a VAS scale ≥ 1 were allocated to the musculoskeletal pain group, and those with a VAS scale <1 were in the no-pain group.[21]

The Functional Independence Measure-Locomotor (FIM-L) criteria were used to evaluate the participants' levels of locomotor disability.[16] The FIM-L5 is an exception (household locomotion), in which participants walk only short distances (a minimum of 15 meters independently with or without a device). The activity takes more than a reasonable amount of time, or there is safety consideration. With FIM-L6, patients have modified independence. They can walk at least 50 meters but use a brace (orthosis) or prosthesis on the leg, special adaptive shoes, cane, crutches, or walker; they take more than a reasonable amount of time to complete the activity; or there are safety considerations. With FIM-L7, patients have complete independence. They can walk a minimum of 50 meters without an assistive device. The patients perform the activity safely.[16]

Statistical analysis

The study data was analyzed using STATA ReleSTATA10. The normal distribution was verified through a Kolmogorov-Smirnov test. Descriptive statistics were applied to evaluate demographic factors, SCI features, and the prevalence of musculoskeletal pain. The severity of SCI was categorized into 4 subgroups: 1) C1-C4 AIS A, B and C; 2) C5-C8 AIS A, B and C; 3) T1-S3 AIS A, B and C; 4) AIS D at any injury level according to international SCI core data set with the reference. Mann-Whitney U tests for continuous data and the Chi-square test for categorical data were used to compare group results. Logistic regression analysis (odds ratio [OR]) was used to analyze risk factors of musculoskeletal pain. The factors associated with musculoskeletal pain in ambulatory individuals with SCI included age (<65, ≥65 years), gender (male and female), BMI (normal and overweight), levels of locomotor disability (FIM-L5, 6, and 7), causes (traumatic and non-traumatic), levels of SCI diagnosis (tetraplegia and paraplegia) and stages of injury (sub-acute and chronic). The significant difference was set as a p-value of < 0.05.

Results

At the beginning, 174 individuals with SCI were interested in participating in this research. Nevertheless, 36 participants were excluded because they were unable to ambulate (N=15), had osteoarthritis of the knee (N=12), had brain involvement (n=5), and had pressure sores (N=4). Finally, a sample of 138 completed this research. The demographic and SCI characteristics data were non-normal distribution and presented in median values, as shown in Table 1. Most participants were males, and the median age was 54.5 years (40.0-65.0) at the chronic stage, during which the median time since injury was 42

months (13.00-91.25). Injury levels among the individuals ranged from C4 to the cauda equina with AIS C and D.

Seventy-six of the 138 participants complained of moderate musculoskeletal pain, with an average VAS score of 5 (the median value of the intensity of pain score was between 4 and 6). The most common areas of pain were the lower back (27%), hip (23%), and shoulder (15%), as shown in Figure 1. Logistic regression analysis found that factors associated with musculoskeletal pain were male gender (OR, 3.56; 95% CI, 1.62-7.84; *p*=0.002) and FIM-L 6 (OR, 2.57; 95% CI, 1.08-6.10; *p*=0.032), as presented in Table 2.

Table 1. Demographics and characteristics of participants.

Variable	Total (N=138)	Pain (N=76)	No pain (N=62)	*p* value
Age at injury (years) [a]	54.50 (40.00-65.00)	53.50 (40.25-66.75)	55.00 (37.50-64.25)	0.550
Time since injury (months)[a]	42.00 (13.00-91.25)	42.00 (18.25-83.75)	42.50 (8.75-42.50)	0.583
Body mass index (kg/m^2)[a]	21.95 (18.97-25.20)	21.95 (18.82-25.07)	21.90 (19.10-25.62)	0.966
Gender: male; N (%)[b]	90 (65)	57 (75)	33 (53)	0.008*
Etiology: traumatic; N (%)[b]	59 (43)	31 (41)	28 (45)	0.606
Levels and severity of SCI; N (%)[b]				
C1-4 AIS A, B and C	1(1)	0 (0)	1 (2)	0.068
C5-8 AIS A, B and C	6 (4)	6 (8)	0 (0)	
T1-S3 AIS A, B, and C	44 (32)	26 (34)	18 (29)	
AIS D at any NLI	87 (63)	44 (58)	43 (69)	
Walking devices used; N (%)[b]				
Independence	48 (34)	22 (46)	26 (54)	0.417
Cane	12 (37)	4 (61)	8 (33)	
Crutches	12 (9)	8 (67)	4 (33)	
Walker	50 (20)	29 (58)	21 (42)	
Level of locomotor disability; N (%)[b]				
FIM-L7	46 (33)	20 (43)	26 (57)	0.141
FIM-L6	76 (55)	47 (62)	29 (38)	
FIM-L5	16 (12)	9 (56)	7 (44)	

Note: NLI: neurological level of injury, FIM-L: functional Independent Measure-Locomotion score FIM-L5: participants walk only short distances (a minimum of 17 m) independently with or without a device, FIM-L6: participants walk a minimum of 50 m (150 ft) but use a brace (orthosis) or prosthesis on the leg, special adaptive shoes, cane, crutches, or walker, FIM-L7: participants walk a minimum of 50 m (150 ft) without assistive device, [a]data were presented in terms of the median (interquartile range: Q1-Q3). The findings between the groups were compared using the Mann-Whitney U test, [b]these variables were categorized data, and a chi-square test was used to compare differences between groups according to the following criteria, gender: male/ female, etiology: non-traumatic/ traumatic,[*]indicates a significant difference between groups (p<0.05).

Table 2. Factors associated with musculoskeletal pains among ambulatory individuals with spinal cord injury (SCI).

Variable	Total (N)	Pain, N (%)	No pain, N (%)	aOR (95% CI)	*p* value
Gender					
Female	48	19 (25)	29 (47)	1	
Male	90	57 (75)	33 (53)	3.56 (1.62-7.84)	0.002*
Age					
<65 years	87	48 (63)	39 (63)	1	
≥65 years	51	28 (37)	23 (37)	0.81 (0.36-1.79)	0.607
Cause					
Traumatic	59	31 (41)	28 (45)	1	
Non-traumatic	79	45 (59)	34 (55)	1.45 (0.67-3.13)	0.340
Body mass index					
Normal (18-23.0 kg/m^2)	79	46 (61)	33 (53)	1	
Overweight (≥23.0 kg/m^2)	59	30 (39)	29 (47)	0.96 (0.87-1.06)	0.130
Stage of injury					
Sub-acute stage (<12 months)	32	15 (20)	17 (27)	0.90 (0.29-1.44)	
Chronic stage (≥12 months)	106	61 (80)	45 (73)	0.56 (0.23-1.31)	0.181
Level of SCI diagnosis					
Tetraplegia	23	11 (14)	12 (19)	1	
Paraplegia	115	65 (86)	50 (81)	0.49 (0.18-1.31)	0.157
Levels of locomotor disability					
FIM-L7	46	20 (26)	26 (42)	1	
FIM-L6	76	47 (62)	29 (47)	2.57 (1.08-6.10)	0.032*
FIM-L5	16	9 (12)	7 (11)	2.30 (0.62-8.60)	0.214

Note: *SE: standard error, aOR: adjusted odds ratio, 95%: CI95% confidence interval, AIS: American Spinal Injury Association (ASIA) Impairment Scale, [a]aOR: significant difference from the reference group for which the value was set at 1.0 (p<0.05), *significant difference between groups (p<0.05).*

Discussion

Few is known about musculoskeletal pain in ambulatory individuals with SCI; therefore, this study aimed to explore factors contributing to musculoskeletal pain in people with SCI. The findings demonstrated that 55.07% of ambulatory individuals with SCI reported moderate pain, and the top three areas of suffering were lower back, hip, and shoulder suffering (Figure 1). In addition, ambulatory individuals with SCI who were male gender and had a high level of locomotor disability (FIML-6) were associated with an increased risk of musculoskeletal pain (Table 2).

Approximately 55% of individuals with SCI experience musculoskeletal pain; the pain data showed the association between chronic SCI with ambulation and various levels of locomotor disability, including FIM-L6 (62%), FIM-L7 (43%), and FIM-L5 (56%) (Table 1). The lower back, hip, and shoulder were most affected (Figure 1). Different impairments between the lower limbs and the use of a walking device to compensate for these impairments are important causes of asymmetrical walking in individuals with neurological disorders, which may encourage patients to learn to limit the asymmetrical use of their limbs, increasing their risk of musculoskeletal pain and joint damage in the lower limb of greater use.[22] Moreover, the extreme use of the upper limbs to compensate for loss of strength and/or mobility in lower limbs might induce an injury to the upper extremity structures.[22,23] Therefore, areas of pain were shown in the lower back, hip, and shoulder (Figure 1). However, previous studies reported that the area most affected by pain was the shoulder.[5,22] The reason might be that most of the participants were wheelchair users.[24] The present findings focused only on ambulatory individuals with SCI who could walk and who had a high risk of developing musculoskeletal pain. Therefore, medical rehabilitation could find an effective strategy to prevent negative effects for these individuals with SCI.

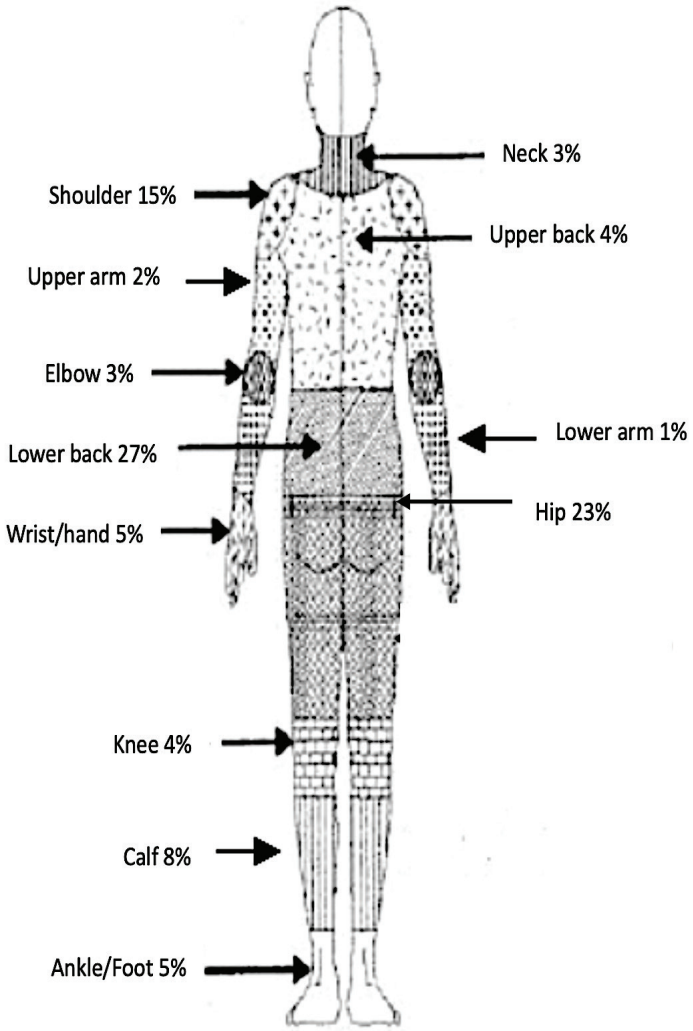


Figure 1. *Percent reported area of pain (The percent of pain in each area was derived from the total 191 pain areas and each area was included from both sides)*

This study proves the hypothesis that male individuals with SCI are at a greater risk of musculoskeletal pain than females with SCI. This finding aligns with the study of Hassan&Kamrujjaman, who reported an increased prevalence of pain in males with SCI than females with SCI.[25] However, other studies have indicated that females with SCI are more likely than males to experience musculoskeletal pain.[12,13,25] Barbetta *et al*. found that women are twice as likely to suffer from upper extremity pain due to physiological, and structural variances in their shoulders and because more women use their forearms to propel their wheelchairs than men.[12] In the present study, we hypothesize that male challenges and daring behaviors improve self-confidence to execute everyday tasks more than females, which promotes the risk of musculoskeletal injury. According to van Drongelen *et al.,* participants with higher functioning levels measured by FIM who perform too many activities of daily living (ADL) are more likely experience complaints from overuse. In contrast, participants with lower functioning levels who perform fewer tasks are less likely to experience overuse pain complaints.[24] Furthermore, Phonthee *et al*. reported that improved functional ability may lead to greater integration of walking into everyday activities and reduced reliance on a wheelchair. High ability levels might also be accompanied by increased confidence and decreased movement awareness.[14] Thus, challenge and courage in male ambulatory with SCI may explain a higher risk of pain than those in females.

Furthermore, the current findings demonstrate that participants with FIM-L6 (able to walk at least 50 meters with AAD) have a higher risk of musculoskeletal pain than those with FIM-L7 (able to walk at least 50 meters independently) and FIM-L5 (able to walk maximum 17 meters with or without AAD). Participants with FIM-L6 are a group of individuals who can walk a long distance or can walk to participate in their community, but they also need AAD. A previous study found that persons with arthritis who frequently use canes or walkers to decrease weight on their legs are especially susceptible to joint inflammation caused by repetitive forces.[22] Therefore, long-distance walking with AAD may promote the risk of musculoskeletal pain more than long-distance without AAD and short distances with or without AAD. This finding indicates that even though patients with SCI can walk independently, AAD use also induces the risk of pain.

Therefore, intervention or follow-up to wean off of AAD after SCI is crucial.

Finally, this study did not find an association between musculoskeletal pain and age, the SCI diagnosis level, and the injury stage (Table 2). The researcher supposes that most of the study participants were at a chronic stage and may have almost full functional recovery. In addition, most of them had a lower level of SCI diagnosis. Therefore, they could preserve musculoskeletal structure during walking and contributed to showing non-significant in age, level of SCI diagnosis, and stage of injury in this study. However, this study has some limitations regarding its findings' clinical implications. The design of this study was cross-sectional; therefore, it was unable to confirm the causality of pain. Thus, further exploration of long-term follow-up musculoskeletal pain is needed.

Conclusion

The finding found that male individuals with SCI and FIM-L6 had a high risk of developing musculoskeletal pain. Daring behaviors in men and AAD-related long-distance walking may increase the incidence of musculoskeletal pain in ambulatory patients with spinal cord injury.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Ethical approval

The participants gave their informed consent before enrolling in the study, which was approved by the Institute Ethics Committee for Human Research (no. 591462).

Acknowledgements

We sincerely thank the Improvement of Physical Performance and Quality of Life (IPQ) research group for financial support.

References

[1] Groah SL, Charlifue S, Tate D, Jensen MP, Molton IR, Forchheimer M, et al. Spinal cord injury and aging: Challenges and recommendations for future research. Am J Phys Med Rehabil. 2012; 91: 80-93. doi: 10.1097/PHM.0b013e31821f70bc.

[2] Zullo S, Ingravallo F, Crespi V, Cascioli M, D'Alessandro R, Gasperini M, et al. The impact of the COVID-19 pandemic on people with neurological disorders: An urgent need to enhance the health care system's preparedness. Neurol Sci. 2021; 42(3): 799-804. doi: 10.1007/s10072-020-04984-4.

[3] Haisma JA, van der Woude LH, Stam HJ, Bergen MP, Sluis TA, Post MW, et al. Complications following spinal cord injury: Occurrence and risk factors in a longitudinal study during and after inpatient rehabilitation. J Rehabil Med. 2007; 39: 393-8. doi: 10.2340/16501977-0067.

[4] Siddall PJ, McClelland JM, Rutkowski SB, Cousins MJ. A longitudinal study of the prevalence and characteristics of pain in the first 5 years following spinal cord injury. Pain. 2003; 103:249-57. doi: 10.1016/S0304-3959(02)00452-9.

[5] Das A, Equebal A, Kumar S. Incidence of musculoskeletal pain and its impact on daily and functional activities in Indian spinal cord injury patients. Int J Physioth Res. 2013; 3: 99-106.

[6] Hassanijirdehi M, Khak M, Afshari-Mirak S, Holakouie-Naieni K, Saadat S, Taheri T, et al. Evaluation of pain and its effect on quality of life and functioning in men with spinal cord Injury. Korean J Pain. 2015; 28: 129-36. doi: 10.3344/kjp.2015.28.2.129.

[7] Dryden DM, Saunders LD, Rowe BH, May LA, Yiannakoulias N, Svenson LW, et al. Utilization of health services following spinal cord injury: A 6-year follow-up study. Spinal Cord. 2004; 42: 513-25. doi: 10.1038/sj.sc.3101629.

[8] Ragnarsson K. Management of pain in persons with spinal cord injury. J Spinal Cord Med. 1997; 20: 186-99. doi:10.1080/10790268.1997.11719468.

[9] Rintala DH, Loubser PG, Castro J, Hart KA, Fuhrer MJ. Chronic pain in a community-based sample of men with spinal cord injury: prevalence, severity, and relationship with impairment, disability, handicap, and subjective well-being. Arch Phys Med Rehabil. 1998; 79: 604-01. doi: 10.1016/s0003-9993(98)90032-6.

[10] Kennedy P, Frankel H, Gardner B, Nuseibeh I. Factors associated with acute and chronic pain following traumatic spinal cord injuries. Spinal Cord. 1997; 35: 814-17. doi: 10.1038/sj.sc.3100569.

[11] Tsuritani I, Honda R, Noborisaka Y, Ishida M, Ishizaki M, Yamada Y. Impact of obesity on musculoskeletal pain and difficulty of daily movements in Japanese middle-aged women. Maturitas. 2002; 42: 23-30. doi: 10.1016/s0378-5122(02)00025-7.

[12] Barbetta DC, Lopes AC, Chagas FN, Soares PT, Casaro FM, Poletto MF, et al. Predictors of musculoskeletal pain in the upper extremities. Spinal Cord. 2016; 54: 145-9. doi: 10.1038/sc.2015.126.

[13] Müller R, Brinkhof MW, Arnet U, Hinrichs T, Landmann G, Jordan X, et al. Prevalence and associated factors of pain in the Swiss spinal cord injury population. Spinal Cord. 2017; 55: 346-54. doi: 10.1038/sc.2016.157.

[14] Phonthee S, Saengsuwan J, Siritaratiwat W, Amatachaya S. Incidence and factors associated with falls in independent ambulatory individuals with spinal cord injury: A 6-month prospective study. Phys Ther. 2012; 93: 1061-72. doi: 10.2522/ptj.20120467.

[15] Saensook W, Phonthee S, Srisim K, Mato L, Wattanapan P, Amatachaya S. Ambulatory assistive devices and walking performance in patients with incomplete spinal cord injury. Spinal Cord. 2014; 52: 216-9. doi: 10.1038/sc.2013.120.

[16] Behrman AL, Harkema SJ. Locomotor training after human spinal cord injury: A series of case studies. Phys Ther. 2000; 80(7): 688-700. doi:10.1093/PTJ/80.7.688.

[17] Kirshblum SC, Waring W, Biering-Sorensen F, Burns SP, Johansen M, Schmidt-Read M, et al. Reference

for the 2011 revision of the International Standards for Neurological Classification of Spinal Cord Injury. J Spinal Cord Med. 2001; 34: 547-54. doi: 10.1179/107902611X13186000420242.

[18] Reed MD, van Nostran W. Assessing pain intensity with the visual analog scale: A plea for uniformity. J Clin Pharmacol. 2014; 54(3): 241-4. doi: 10.1002/jcph.250.

[19] Weigl K, Forstner T. Design of paper-based visual analogue scale Items. Educ Psychol Meas. 2021 ;81(3):595-611. doi: 10.1177/0013164420952118.

[20] Crawford B, Bouhassira D, Wong A, Dukes E. Conceptual adequacy of the neuropathic pain symptom inventory in six countries. Health Qual Life Outcomes. 2008; 6:62. doi: 10.1186/1477-7525-6-62.

[21] Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res. 2011; 63: S2040-52. doi: 10.1002/acr.20543.

[22] Kumprou M, Amatachaya P, Sooknuan T, Thaweewannakij T, Mato L, Amatachaya S. Do ambulatory patients with spinal cord injury walk symmetrically? Spinal Cord. 2017; 55: 204-7. doi: 10.1038/sc.2016.149.

[23] Bateni H, Maki BE. Assistive devices for balance and mobility: benefits, demands, and adverse consequence. Arch Phys Med Rehabil. 2005; 86: 134-45. doi: 10.1016/j.apmr.2004.04.023.

[24] van Drongelen S, de Groot S, Veeger HE, Angenot EL, Dallmeijer AJ, Post MW, et al. Upper extremity musculoskeletal pain during and after rehabilitation in wheelchair-using persons with a spinal cord injury. Spinal Cord. 2006; 44: 152-9. doi: 10.1038/sj.sc.3101826.

[25] Hassan R, Kamrujjaman M. A cross-sectional study of spinal cord injury-induced musculoskeletal Pain. Curr Med Res Opin. 2019; 2(12): 367-71. doi: 10.15520/jcmro.v2i12.234.

Journal of Associated Medical Sciences 2024; 57(3): 90-97

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Association between sensory processing patterns and stress among community-dwelling people with metabolic syndrome

Ilada Pomngen[1] Tiam Srikhamjak[1] Warunee Kumsaiyai[2] Anuchart Kaunnil[1] Pornpen Sirisatayawong[1*]

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

ARTICLE INFO

Article history:
Received 16 May 2024
Accepted as revised 3 July 2024
Available online 7 July 2024

Keywords:
Sensory processing patterns, stress, metabolic syndrome, sensory preferences, sensory arousals.

ABSTRACT

Background: Stress is a risk factor for metabolic syndrome (MetS). High and low sensory stimulation can trigger high stress, but no research exists on the relationship between sensory processing patterns and stress among people with MetS.

Objective: This study examined the association between sensory processing patterns and stress.

Materials and methods: A total of 117 people with MetS in the Nam Phrea subdistrict, Hang-Dong District, Chiang-Mai Province, Thailand, aged 35-85, completed the Thai Sensory Patterns Assessment-Adult Version (TSPA) and the Thai Stress Test (TST-24). Demographic and metabolic variable information was gathered. Descriptive statistics were used to summarize the demographic characteristics and stress. Spearman's correlation and regression analysis examined the associations between the sensory processing patterns in each sensory modality and stress scores.

Results: A total stress score was significantly correlated with preferences in visual and auditory senses, as well as with arousal levels in visual and smell-taste senses (r =-0.397, -0.199, -0.358, 0.268, and p<0.05). Regression analysis revealed that stress can be predicted by preferences and arousal levels in visual and smell-taste senses (R^2 =0.156, 0.039, 0.174, and 0.050, respectively; p<0.05).

Conclusion: The results suggest that sensory processing patterns might associated with stress, which is a predictor of MetS. Health professionals can utilize acquired knowledge to implement sensory-based interventions for individuals with MetS to address their stress issues.

Introduction

Metabolic syndrome (MetS) is defined as a cluster of metabolic abnormalities, such as insulin resistance, visceral obesity, hypertension, and dyslipidemia.[1] Over the last decades, MetS has been considered a worldwide epidemic because of its increased prevalence in the general population, estimated at 25% of the world population[2, 3] and 20–27% of Thai adults.[4] Individuals with MetS have a greater risk of developing cardiovascular disease (CVD) in the next 5-10 years than those without MetS.[5] Thus, identifying and understanding the risk factors for MetS are important for both early screening and providing early intervention to prevent the increasing severity of MetS.

Recently, evidence has revealed psychological and behavioral risks for MetS, including unhealthy eating,[6] a sedentary lifestyle,[7,8] and stress.[9,10] Recent treatments have also focused directly on addressing these risks, but most treatments are not successful enough to minimize the

∗ **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.2024.051
E-ISSN: 2539-6056

severe progression of MetS. Thus, a deeper understanding of these behavioral and psychological risks is necessary for understanding the mechanisms underlying unhealthy behaviors and emotions linked to the development of MetS.

Sensory Processing Patterns refer to how the brain receives, organizes, and responds to sensory information obtained from the environment in daily life.[11] These patterns are a crucial aspect of how individuals interact with the world around us, affect how they perceive and react to various stimuli, and shape human emotions and behaviors in healthy and unhealthy ways.[11,12] Importantly, sensory processing patterns influence an individual's ability to handle external and internal stimuli, which can impact their stress levels.[13] For instance, individuals with high arousal or sensitivity to sensory stimuli may experience higher stress levels due to their heightened sensitivity to stimuli, making it difficult for them to tolerate certain environments.[13] Conversely, those with low arousal or sensitivity to sensory stimuli may become stressed from the need to exert more effort to seek out sufficient stimuli to feel engaged.[11] Therefore, understanding an individual's sensory processing pattern is essential for developing effective strategies to manage stress.[13]

Moreover, evidence highlights the role of sensory processing patterns in many risks of MetS, especially stress.[14-18] Previous research has shown that people with sensory sensitivity is related to high stress.[18-20] and increased cortisol levels.[21] This is because individuals in this group are quicker to detect and react to specific sensory stimuli than those with typical sensory processing, making them more susceptible to feeling overwhelmed and stressed. The physiological response to stress, sustained over an extended period, can trigger metabolic abnormalities through the sympathetic-adrenal-medullary (SAM) and hypothalamic-pituitary-adrenal (HPA) axes, ultimately leading to the development of MetS.[10] Moreover, behavioral responses to stress are also associated with MetS, such as increased food consumption and reduced physical activity.[22] Therefore, investigating sensory processing patterns as determining stress factors among individuals with MetS is imperative for targeted interventions that consider sensory aspects in managing stress and, subsequently, the severity of MetS.

However, there is no research regarding the relationship between sensory processing patterns in specific modalities and stress among people with MetS. Specifying what sensory processing patterns predict stress in a particular sense is important for tailoring sensory-based interventions to support both physical and mental well-being to minimize the progression of MetS. Therefore, the aim of this study was to examine the relationships between sensory processing patterns in specific modalities and stress among community-dwelling people with MetS.

Materials and methods

This study used a correlational research design and received ethics approval from the Research Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University (No. AMSEC-65EX-071).

Participants

This study recruited 117 people with MetS (42 males and 75 females) aged 35 to 85 years (mean = 54.31±10.77) from the Nam Phrae subdistrict health promotion hospital, Hang-Dong District, Chiang-Mai province, Thailand. The inclusion criteria encompassed the presence of three metabolic risk factors, as delineated by the Harmonized Criteria for MetS,[1] which included elevated fasting blood glucose (FPG) levels ≥100 mg/dL, elevated blood pressure (BP) ≥130/85 mmHg, and waist circumference (WC) exceeding >90 cm (for males) and >80 cm (for females). The exclusion criteria were the presence of serious mental illness, sensory impairment, chronic diseases (e.g., thyroid, respiratory, liver, kidney, and cardiovascular diseases), or cognitive impairments screened by the Mental State Examination Thai 10 (MSET-10).[23]

Data collection materials

Demographic questionnaire

This questionnaire was used to gather data, including age, gender, education level, career status, and marital status.

Measures of metabolic risk factors

Capillary fasting blood glucose (CFBG) was measured by a nurse in the early morning (between 7:00 and 8:00 am) following a 12-hour fasting period. Blood pressure (BP) was assessed using an automatic sphygmomanometer with the participant seated following a 10-minute resting period. The mean of the two measurements was used to calculate the estimated blood pressure. Waist circumference (WC) was measured with a nonelastic circumference measuring tape. At the same time, the individual stood, rounded to the nearest 0.1 cm, and taken midway between the lowest rib and the iliac crest.

The Thai Sensory Patterns Assessment-Adult Version (TSPA)

The self-reported TSPA-adult version comprises 60 items divided into two parts: sensory preferences (35 items) and sensory arousals (25 items); each is divided into six categories based on sensory modalities. As confirmed by factor analysis, these include visual, auditory, smell and taste, tactile, proprioceptive, and vestibular senses. This questionnaire uses a 5-point Likert scale, in which participants rate how frequently they respond to sensory events in daily life.

In part of sensory preferences, there are 35 items for assessing the type of sensory stimulus that tends to make individuals feel affirmed, comfortable, and pleasurable. The items are divided into 6 categories, including visual (4 items), auditory (5 items), smell and taste (9 items), tactile (4 items), proprioceptive (8 items), and vestibular (5 items) senses. An example of sensory preference is: "I like eating sweets or smelling food." Participants responded to each item using a 5-point Likert scale: 1 = never, 2 = seldom, 3 = occasionally, 4 = frequently, 5 = always.

In part of sensory arousal, there are 25 items for assessing the alertness of the nervous system, which is affected by sensory input in daily life. The items are

also divided into 6 categories, including visual (3 items), auditory (6 items), smell and taste (3 items), tactile (3 items), proprioceptive (3 items), and vestibular (7 items) senses. An example item for sensory arousal is: "I can eat food that has a strong or pungent smell." Participants responded to each item using a 5-point Likert scale. The scale for items indicating high arousal was: 1 = never, 2 = seldom, 3 = occasionally, 4 = frequently, 5 = always. The scale was reversed for items indicating low arousal: 5 = never, 4 = seldom, 3 = occasionally, 2 = frequently, 1 = always.

The scores for each sensory modality in both two parts are aggregated and reported as a percentage, indicating low (below 25%), moderate (25-75%), or high (above 75%) sensory preferences or levels of sensory arousal. The TSPA has acceptable validity (IOC = 0.60 to 1.00 in parts 1 and 2) and reliability (ICC = 0.93, 0.77, Cronbach's alpha = 0.89, and 0.62 in parts 1 and 2).[24]

Thai Stress Test (TST-24)

The self-reported TST-24 consists of 24 items to assess both positive and negative psychological reactions to daily life events, and each item is rated on a 3-point scale from 1 to 3 (never = 0, sometimes = 1, and often = 3). The items were divided into negative (items 1-12) and positive (items 13-24) feelings, each of which were separately summed, with both having a possible range of 0 to 36 (higher scores indicating greater stress). The scores were then compared to a matrix table to obtain an index score, classifying the individual's stress into four levels (excellent mental health, normal mental health, mild anxiety, and stress). The Cronbach's alpha for the total test was 0.84, and the sprit half reliability coefficient for the total test was 0.88.[25]

Procedure

After obtaining approval from the Research Ethics Committee, the participants were recruited from the Nam Phrae subdistrict health promotion hospital from January 1st, 2023, to April 30th, 2023. Purposive sampling was employed for recruited participants. Initially, a physician screens individuals with MetS during the annual health checkup and screening for metabolic diseases, and then village health volunteers establish initial contact with eligible participants. A total of 145 individuals who met the diagnostic criteria for MetS were invited to participate in the study through advertisements via posted flyers and word-of-mouth done by the village health volunteers. Those who expressed interest in participating in this study directly contacted the researcher or the village health volunteers via phone. An assessor used the MSET-10 and asked them to complete a demographic questionnaire to ensure that they met the inclusion criteria. Finally, the 117 participants who met the inclusion criteria and agreed to participate in the study were asked to sign an informed consent form and then complete the questionnaire for gathering data on TSPA and TST-24.

Data analysis

The data was analyzed utilizing SPSS version 20 (SPAA Inc., USA) and STATA. Descriptive statistics were used to analyze demographic information, metabolic risk variables (CFBG, SBP, DBP, and WC), and stress scores. The Kolmogorov- Smirnov test was used to determine the normality of the distribution of variables. The test revealed that the sensory processing patterns and stress data did not follow a normal distribution. Therefore, the Spearman correlation test was employed to determine the relationships between sensory processing patterns in each specific modality and stress. Nonparametric regression or Kernel regression methods were used to determine whether stress can be predicted by sensory patterns in specific sensory modalities. The significance was set at the 0.05 level or below.

Results

Demographic characteristics

The general sample characteristics and metabolic risks are presented in Table 1. The participants were between 35 and 84 years old. Most participants were female (64.1%). Most participants have completed primary school (51.3%) and were married (65.8%). The means of the metabolic risk variables were as follows: capillary fasting blood glucose (CFPG) was 115.61 mg/dL, systolic blood pressure (SBP) was 135.06 mmHg, diastolic blood pressure (DBP) was 86.88 mmHg, and wrist circumference (WC) was 91.36 cm.

Table 1. Demographical characteristics and metabolic risk variables in people with MetS (N=117).

Variables	N (%)
Gender	
Male	42 (35.9)
Female	75 (64.1)
Education Levels	
Under primary school	17 (14.5)
Primary school	60 (51.3)
High school	32 (27.4)
Bachelor's degree	8 (6.8)
Postgraduate	0 (0)
Marriage status	
Single	18 (15.4)
Married	77 (65.8)
Divorced	9 (7.7)
Widowed	13 (11.1)
Age	
35-44 years	19 (16.2)
45-64 years	72 (61.6)
65 year and over	26 (22.2)
mean ± Std. Deviation	54.31±10.77
Metabolic risk variables	**Mean±SD**
Metabolic risks	
CFPG (mg/dL)	115.61±18.17
SBP (mmHg)	135.06±6.56
DBP (mmHg)	86.88±3.30
WC (cm)	91.36±9.37

Note: CFPG: capillary fasting blood glucose, SBP: systolic blood pressure, DBP: diastolic blood pressure, WC: wrist circumferences.

Stress

The stress scores are presented in Table 2. The mean negative stress score exceeded the mean positive (mean = 8.77 and 6.31, respectively). Additionally, most participants had low-stress levels (43.60%), as shown in Table 2.

Table 2. Stress scores and stress levels of people with MetS (N=117).

Variables	Minimum	Maximum	Mean	SD
Positive stress scores	0	24	6.31	4.93
Negative stress scores	0	36	8.77	7.34
Total stress scores	0	43	15.09	9.24
Stress levels n (%)				
Level 1: Good mental health	8 (6.80)			
Level 2: Normal	25 (21.40)			
Level 3: Low stress	51 (43.60)			
Level 4: High stress	33 (28.20)			

Correlation between sensory patterns and stress

As presented in Table 3, negative stress scores were negatively correlated with preferences in visual (r=-0.297, *p*=0.001) and auditory senses (r=-0.213, *p*=0.021), as well as with arousal levels in visual (r=-0.373, *p*=0.000), auditory (r=-0.244, *p*=0.008), and proprioceptive senses (r=-0.229, *p*=0.013). This indicates that individuals who do not prefer visual or auditory sensory inputs or had low arousal levels in visual, auditory, or proprioceptive senses were likelier to report experiencing negative stress.

Table 3. Correlations between sensory patterns and stress (N=117).

Sensory patterns	Negative stress score		Positive stress score		Total stress score	
	r	*p* value	r	*p* value	r	*p* value
P-visual	-0.297	**0.001***	-0.319	**0.000***	-0.397	**0.000***
P-auditory	-0.213	**0.021***	-0.140	0.133	-0.199	**0.031***
P-smell and taste	0.017	0.856	-0.233	**0.011***	-0.168	0.071
P-tactile	0.143	0.125	-0.103	0.271	0.009	0.921
P-vestibular	0.110	0.237	0.118	0.204	0.173	0.062
P-proprioceptive	0.117	0.208	-0.078	0.406	-0.016	0.868
A-visual	-0.373	**0.000***	-0.257	**0.005***	-0.358	**0.000***
A-auditory	-0.244	**0.008***	0.035	0.705	-0.085	0.360
A-smell and taste	0.164	0.078	0.260	**0.005***	0.268	**0.003***
A-tactile	0.145	0.119	-0.113	0.152	-0.036	0.701
A-vestibular	0.168	0.069	-0.195	**0.035***	-0.076	0.418
A-proprioceptive	-0.229	**0.013***	-0.004	0.969	-0.109	0.243

Note: P: preferences, A: arousal, p values were calculated using the Spearman correlation test, p≤0.05.*

A positive stress score was negatively correlated with preferences in visual (r=-0.319, *p*=0.000) and smell-taste senses (r=-0.233, *p*=0.011) and with arousal levels in visual (r=-0.257, *p*=0.005) and vestibular senses (r=-0.195, *p*=0.035), except for arousal levels in smell-taste senses, which showed a positive correlation (r=0.260, *p*=0.005). This indicates that individuals who do not prefer visual or smell-taste sensory inputs or had low arousal levels in visual or vestibular senses were likelier to report experiencing positive stress. However, individuals with high arousal levels in the smell-taste sense were more likely to report experiencing positive stress.

The total stress scores were negatively correlated with preferences for the visual (r=-0.397, *p*=0.000) and auditory senses (r=-0.199, *p*=0.031) and with arousal levels in the visual sense (r=-0.358, *p*=0.000). In contrast, they were positively correlated with arousal levels in the smell-taste senses (r=0.268, *p*=0.003). This indicates that individuals who do not prefer visual or auditory sensory inputs or who have low arousal levels in a visual sense were more likely to report experiencing stress. However, individuals with high arousal levels in the smell-taste sense were also more likely to report experiencing stress.

Predicting stress by sensory patterns

As presented in Table 4, the Kernel regression analysis revealed that both preferences and arousal levels in visual and smell-taste senses were significant predictors of stress (R^2=0.156, 0.039, 0.174, 0.050; *p*=<0.001, 0.040, <0.001, 0.026, respectively), as well as preference in vestibular sense (R^2=0.044; *p*=0.016). However, the R^2 values or the coefficient of determination, which indicate the proportion of the variance in stress that can be explained by sensory pattern predictors, were relatively low (ranging from 0.039-0.174). This means that the sensory processing patterns could predict 3.9%-17.4% of the variability in the stress (dependent variable). This suggests that these predictors could explain or predict a smaller portion of the variation in stress levels.

Table 4. Sensory patterns predictors of total stress score (N=117).

Dependent variable	Predictor variables	Coefficient	Bootstrap estimate of SE	R^2	*p* value
Total stress scores	P-visual	-0.221	0.043	0.156	**<0.001***
	P-auditory	-0.101	0.065	0.097	0.118
	P-smell and taste	-0.152	0.074	0.039	**0.040***
	P-tactile	-0.001	0.115	0.071	0.993
	P-vestibular	0.176	0.073	0.044	**0.016***
	P-proprioceptive	0.012	0.067	0.002	0.859
	A-visual	-0.177	0.040	0.174	**<0.001***
	A-auditory	-0.094	0.088	0.030	0.289
	A-smell and taste	0.112	0.050	0.050	**0.026***
	A-tactile	-0.016	0.059	0.041	0.784
	A-vestibular	-0.023	0.059	0.015	0.693
	A-proprioceptive	0.060	0.123	0.083	0.627

Note: P: preferences, A: arousal, p values were calculated the kernel regression analysis, p ≤ 0.05.*

Discussion

The main findings of the present study are that sensory processing patterns, both in terms of sensory preferences and levels of sensory arousal, are associated with stress among people with MetS.

Relationship between levels of sensory arousal and stress sensory processing patterns

The findings showed that levels of sensory arousal in smell-taste senses were positively correlated with stress. This means that individuals who more quickly register smell-taste stimuli were likelier to report experiencing higher stress levels. This finding aligns with previous studies on the relationship between sensory sensitivity (high levels of sensory arousal) and high stress,[18-20, 26,27] but prior studies were often performed only in healthy adults. This study has shown that higher sensory arousals are also associated with higher stress in adults with MetS. Furthermore, while most previous studies did not clearly establish whether specific sensory modalities were related to stress, this finding demonstrated that higher stress was associated with higher arousal levels, especially in smell and taste. The link between the senses of smell and taste and the emotional system is well established. The olfactory and gustatory system is closely linked to the limbic system in the brain, which plays a crucial role in processing emotions and memories.[28,29] Therefore, individuals with high arousal to smell and taste stimuli are more prone to feeling overwhelmed, especially in environments with strong or unpleasant stimuli.

Moreover, the findings also showed that stress was negatively correlated with arousal levels in the visual sense. This means that lower arousal levels in response to visual stimuli were associated with higher stress levels, or, in other words, individuals who experience low arousal from visual stimuli tend to report high stress. This finding is consistent with earlier studies, which have shown that a problem in noticing or detecting sensory input, especially a low registration, is also related to stress.[18] However, this study specified specific senses associated with greater stress levels, including lower sensory arousals in visual, auditory, vestibular, and proprioceptive senses. According to Cox and Mackay,[30] stress is a human perception resulting from comparing a person's ability to cope with environmental demands. People with low arousal levels in the abovementioned senses might put extra effort or pressure on themselves to avoid threats or dangers in everyday situations, causing them stress. Furthermore, because the abovementioned sensory modalities are involved in human survival, a lack of sensory awareness might contribute to feelings of insecurity, ultimately leading to increased stress responses.

Therefore, the finding suggested that both under and over-sensory stimulation from certain stimuli can trigger stress.

Associations between low sensory preference and high stress

The new aspects highlighted in this study, which also expand upon previous studies by including an assessment of sensory preferences in specific modalities, have shown that stress was negatively correlated with preferences for visual, auditory, and smell-taste senses. This means that individuals with lower preferences for or do not prefer visual, auditory, and smell-taste stimuli in their daily lives are more likely to report experiencing higher stress levels. Generally, each person has different sensory experiences and develops individualized memories across their life span.[11] The memories make every person unique in the sense they prefer or do not prefer. Exposure to unpleasant sensory stimuli makes people more likely to become overwhelmed and have heightened stress responses.[31,32] Previous studies showed that unpleasant odor stimuli can cause stress,[33,34] while pleasant sensory stimuli promote healing and relaxation in mammals.[33] Therefore, when providing intervention to deal with stress, an individual's sensory preferences should be considered, especially visual, auditory, and smell-taste senses.

Although the correlations between sensory processing patterns and stress were found at low to medium levels (R-values ranged from -0.397 to 0.268), they still highlight the importance of understanding individuals' sensory processing patterns in managing stress. According to the findings, under-stimulation, over-stimulation, and unpleasant stimulation from certain sensory modalities, particularly visual, auditory, and smell-taste stimuli, could contribute to stress levels. Therefore, health professionals should incorporate sensory processing assessments into their evaluation and treatment processes to understand how individuals respond to sensory stimuli. By identifying whether an individual is more sensitive (high arousal) or less sensitive (low arousal) to these stimuli, health professionals can tailor interventions to either minimize sensory overload or provide necessary sensory stimulation. This personalized approach can help individuals manage stress levels and improve overall well-being.

Predicting stress by sensory processing patterns

The findings highlighted the possibility of using sensory processing patterns as predictors of stress, specifically low preferences for visual and smell-taste senses, low arousal levels in the visual sense, and high arousal levels in the smell-taste sense. Based on these findings, people with those sensory patterns are at high risk of suffering from chronic stress triggered by sensory stimuli in daily life. Neuroendocrine hormones released during stress may cause increased activation of the SAM and HPA axes, altering metabolic abnormalities and increasing the severity of MetS.[10] People in those groups should receive sensory-based interventions that consider their sensory needs to modulate high-stress levels triggered by those sensory stimuli in daily life in order to prevent serious mental health problems related to the development of chronic diseases.

However, the findings showed that the coefficient of determination (R^2), which indicates the proportion of the variance in the dependent variable (stress) that is predictable from the predictors (sensory processing

patterns) in a regression model, was relatively low (R^2 values ranged from 0.039 to 0.174). This means that the sensory processing pattern variables could predict 3.9% to 17.4% of the variability in the stress variable. This finding suggests that while sensory preferences and arousal levels in visual, auditory, smell-taste, and vestibular senses were significant predictors of stress, other factors not examined in this study and not included in the current regression model might play substantial roles in influencing the stress experiences of participants. The factors could consist of individual factors or environmental influences.[35] Furthermore, the stress responses may also be influenced by the dynamic nature of sensory experiences and their interaction with individual perceptions, coping mechanisms, and socio-economical and situational contexts.[13,29,35] Therefore, future research could explore these factors in greater depth to provide a more comprehensive understanding of how sensory processing patterns impact stress responses across different populations and contexts. Additionally, investigations into multifaceted factors contributing to stress experiences beyond sensory modalities alone are needed.

The finding showed an association between sensory processing patterns and stress. However, it is important to note that implementing sensory-based interventions should be personalized due to individual differences in sensory processing patterns.[11] Therefore, health professionals should collaborate with patients to specify whether sensory stimuli can trigger a stress response. This information can help therapists suggest sensory environments, coping strategies, activities, or routines that can support sensory needs and address stress in daily life while being appropriate with what the individual truly wants or needs to do to support physical and mental health.

Limitations

A small sample size could limit the generalizability of the findings, which may not represent the larger population. Furthermore, relying on self-reported participant data introduces subjectivity and potential reporting bias. Future research should consider using a larger and more diverse sample to increase the robustness and applicability of the findings. In addition, incorporating multiple data collection methods, such as objective assessments based on biochemical testing for stress hormones (e.g., salivary cortisol testing or blood cortisol testing), could also provide more comprehensive and validated results.

Conclusion

The results of this study indicated that sensory processing patterns, especially preferences and arousal levels in the visual and smell-taste senses, were associated with and potentially predictors of stress. The findings may assist health professionals in tailoring sensory-based environments, strategies, and interventions to support stress reduction and minimize MetS progression.

Conflict of interest

The authors declare no conflict of interest.

Funding

This research was supported by the Faculty of Associated Medical Sciences, Chiang Mai University, under Grant AMS-2023.

Acknowledgements

This research was supplied by the Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University.

References

[1] Alberti KGMM, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, *et al*. Harmonizing the Metabolic Syndrome. Circulation. 2009; 120(16): 1640-5. doi: 10.1161/CIRCULATIONAHA. 109.192644.

[2] O'Neill S, O'Driscoll L. Metabolic syndrome: A closer look at the growing epidemic and its associated pathologies. Obes Rev. 2015; 16(1): 1-12. doi: 10.1111/obr.12229.

[3] Guembe MJ, Fernandez-Lazaro CI, Sayon-Orea C, Toledo E, Moreno-Iribas C. Risk for cardiovascular disease associated with metabolic syndrome and its components: A 13-year prospective study in the RIVANA cohort. Cardiovasc Diabetol. 2020; 19(1): 195. doi: 10.11 86/s12933-020-01166-6.

[4] Ngoc HN, Kriengsinyos W, Rojroongwasinkul N, Aekplakorn W. Prevalence of Metabolic Syndrome and Its Prediction by Simple Adiposity Indices in Thai Adults. J Health Sci Med Res. 2021; 39(4): 13. doi: 10.31584/jhsmr.2021791.

[5] Li X, Zhai Y, Zhao J, He H, Li Y, Liu Y, *et al*. Impact of Metabolic Syndrome and It's Components on Prognosis in Patients With Cardiovascular Diseases: A Meta-Analysis. Front Cardiovasc Med. 2021; 8: 1-13. doi: 10.3389/fcvm.2021.704145.

[6] Li Y, Sun Y, Wu H, Yang P, Huang X, Zhang L, *et al*. Metabolic syndromes increase significantly with the accumulation of bad dietary habits. J Nutr Health Aging. 2024; 28(2): doi: 100017. 10.1016/j.jnha.2023.100017.

[7] Lemes IR, Sui X, Fernandes RA, Blair SN, Turi-Lynch BC, Codogno JS, *et al*. Association of sedentary behavior and metabolic syndrome. Public Health. 2019; 167: 96-102. doi: 10.1016/j.puhe.2018.11.007.

[8] Yeo Y, Cho IY, Sim MS, Song HG, Song Y-M. Relationship Between Daily Sedentary Behaviors and Metabolic Syndrome in Middle-Aged Adults: Results from a Health Survey in Taean-Gun, Republic of Korea. Metab Syndr Relat Disord. 2020; 19(1): 48-55. doi: 10.1089/met.2020.0021.

[9] Tamashiro KL, Sakai RR, Shively CA, Karatsoreos IN, Reagan LP. Chronic stress, metabolism, and metabolic syndrome. Stress. 2011; 14(5): 468-74. doi: 10.3109/10253890.2011.606341.

[10] Kuo WC, Bratzke LC, Oakley LD, Kuo F, Wang H, Brown RL. The association between psychological stress

and metabolic syndrome: A systematic review and meta-analysis. Obes Rev. 2019; 20(11): 1651-64. doi: 10.1111/obr.12915.

[11] Dunn W. Living sensationally: Understanding your senses. London: Jessica Kingsley; 2008.

[12] Dunn W. The sensations of everyday life: empirical, theoretical, and pragmatic considerations. Am J Occup Ther. 2001; 55(6): 608-20.

[13] Harrold A, Keating K, Larkin F, Setti A. The association between sensory processing and stress in the adult population: A systematic review. Applied Psychology: Health and Well-Being. 2024: 1-31. doi: 10.1111/aphw.12554

[14] Benham G. The highly sensitive person: Stress and physical symptom reports. Pers Individ Dif. 2006; 40(7): 1433-40. doi: 10.1016/j.paid.2005.11.021.

[15] Corbett BA, Schupp CW, Levine S, Mendoza S. Comparing cortisol, stress, and sensory sensitivity in children with autism. Autism Res. 2009; 2(1): 39-49. doi: 10.1002/aur.64.

[16] Kongngern T. Relationship between sensory patterns and stresses of relapsed alcohol dependence clients in Suansaranroom Psychiatric Hospital, Suratthani Province [Thesis]. Faculty of Associated Medical Sciences: Chiang Mai University; 2012. [in Thai]

[17] Chang M, Turner T, Kwee S, Shaul J, Stanbach T, Stringari J. The Effects of Sensory Processing Patterns on Perceived Stress and Sleep Quality Among College Students. Am J Occup Ther. 2019; 73(4_Supplement_1): 7311505182p1-p1. doi: 10.5014/ajot.2019.73S1-PO7022.

[18] van den Boogert F, Spaan P, Sizoo B, Bouman YHA, Hoogendijk WJG, Roza SJ. Sensory Processing, Perceived Stress and Burnout Symptoms in a Working Population during the COVID-19 Crisis. Int J Environ Res Public Health. 2022; 19(4): 2043. doi: 10.3390/ijerph19042043.

[19] Engel-Yeger B, Dunn W. The Relationship between Sensory Processing Difficulties and Anxiety Level of Healthy Adults. Br J Occup Ther. 2011; 74(5): 210-6. doi: 10.4276/030802211X130 46730116407.

[20] Wu X, Zhang R, Li X, Feng T, Yan N. The moderating role of sensory processing sensitivity in the link between stress and depression: A VBM study. Neuropsychologia. 2021; 150: 107704. doi: 10.1016/j.neuropsychologia.2020.107704.

[21] Corbett BA, Muscatello RA, Blain SD. Impact of Sensory Sensitivity on Physiological Stress Response and Novel Peer Interaction in Children with and without Autism Spectrum Disorder. Front Neurosci. 2016; 10: 278. doi: 10.3389/fnins.2016.0027.

[22] Das UN. Metabolic syndrome pathophysiology: The role of essential fatty acids. UK: Wiley-Blackwell; 2010.

[23] Boongerd P. Interesting topic about dementia. Dementia Association of Thailand, Newsletter. 2018; 10: 1-4.

[24] Pomngen I, Srikhamjak T, Putthinoi S. Development of the Sensory Patterns Assessment [Thesis]. Faculty of Associated Medical Sciences, Chiang Mai University; 2020. [in Thai]

[25] Phattharayuttawat S, Ngamthipwattana T, Sukhatungkha K. The Development of the Thai Stress Test. J Psychiatr Assoc Thailand. 2000; 45(3): 237-50 doi: 10.1.1.522.2455&rep=rep1&type =pdf.

[26] Engel-Yeger B, Dunn W. The Relationship between Sensory Processing Difficulties and Anxiety Level of Healthy Adults. Br J Occup Ther. 2011; 74(5): 210-6. doi: 10.4276/030802211X130467 30116407

[27] Carr M, Matthews E, Williams J, Blagrove M. Testing the theory of Differential Susceptibility to nightmares: The interaction of Sensory Processing Sensitivity with the relationship of low mental wellbeing to nightmare frequency and nightmare distress. J Sleep Res. 2021; 30(3): e13200. doi: 10.1111/jsr.13200.

[28] Lundy-Ekman L. Neuroscience: Fundamentals for rehabilitation. 4th ed. St Louis: Elsevier; 2013.

[29] Masuo Y, Satou T, Takemoto H, Koike K. Smell and Stress Response in the Brain: Review of the Connection between Chemistry and Neuropharmacology. Molecules. 2021; 26(9). 2571. doi: 10.3390/molecules26092571.

[30] Cox T, Mackay C. A psychological model of occupational stress: A paper presented to the medical research council. London: Mental Health in Industry; 1976.

[31] Payne RA, Donaghy M. Payne's handbook of relaxation techniques: A practical guide for the health care professional. New York: Elsevier; 2010.

[32] Pagliano. The multisensory handbook. New York, NY: Taylor & Francis Group; 2012.

[33] 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.

[34] Hirasawa Y, Shirasu M, Okamoto M, Touhara K. Subjective unpleasantness of malodors induces a stress response. Psychoneuroendocrinology. 2019; 106: 206-15. doi: 10.1016/j.psyneuen. 2019.03.018.

[35] Cheon Y, Park J, Jeong BY, Park EY, Oh J-K, Yun EH, *et al*. Factors associated with psychological stress and distress among Korean adults: the results from Korea National Health and Nutrition Examination Survey. Scientific Reports. 2020; 10(1): 15134. doi: 10.1038/s41598-020-71789-y

Journal of Associated Medical Sciences 2024; 57(3): 98-106

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Effectiveness of telespeech therapy on language abilities in people with aphasia: A pilot study

Tanyasiri Prasertsrisak Phuanjai Rattakorn*

Communication Disorders,Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand

ARTICLE INFO

Article history:
Received 7 March 2024
Accepted as revised 3 July 2024
Available online 10 July 2024

Keywords:
Telepractice, aphasia, intervention, speech therapy.

ABSTRACT

Background: Aphasia is a condition that happens when certain areas of the brain responsible for language are damaged, causing difficulties in communicating. Treatment involving speech and language interventions is essential for rehabilitating communication abilities. However, due to accessibility and distancing challenges, access to these medical services has been restricted, especially during COVID-19. Telepractice was introduced as an alternative approach to speech and language therapy. However, there currently needs to be more research on its application, specifically within the aphasia population in Thailand.

Objective: This study aimed to investigate the effectiveness of telepractice in aphasia intervention to enhance the general practice standard and to explore satisfaction with using telepractice in speech therapy among aphasia patients and their caregivers.

Materials and methods: Sixteen participants were recruited for this study, comprising eight aphasia patients and eight caregivers. A standardized aphasia test, the Thai Adaptation of the Western Aphasia Battery (TWAB), was used for pre- and post-assessments. Additionally, satisfaction surveys were employed to gauge significant satisfaction levels among participants. A one-month telepractice intervention (12 sessions in total) was conducted between the pre- and post-assessments to determine the effectiveness of telepractice based on its impact on TWAB test scores.

Results: Overall, this study revealed a significant improvement in the Aphasia Quotient (AQ) as measured by the TWAB test (p=0.011). However, only the repetition and naming sub-tests showed significant improvement between pre- and post-assessment (p=0.019 and p=0.011). The satisfaction levels were reported as high to very high.

Conclusion: This study demonstrated the effectiveness of telepractice in aphasia intervention, particularly in improving naming and repetition skills and eliciting a high to very high level of satisfaction among patients and their caregivers.

Introduction

Aphasia, a condition impacting communication, results from left hemisphere brain pathology,[1] particularly affecting the primary language cortex. Its severity varies, affecting speaking, word retrieval, comprehension, sentence construction, reading, and writing.[2] The prevalence of aphasia after stroke indicates that it occurs more frequently in older individuals compared to younger ones. About 15% of people under 65 years old who experience a stroke will develop aphasia, whereas the likelihood increases to 43% in individuals over 85 years old.[3] Speech and language pathologists (SLPs) are pivotal in evaluating

* **Corresponding contributor.**
Author's Address: *Communication Disorders, Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand*
E-mail address: phuanjai.rattakorn@cmu.ac.th
doi: 10.12982/JAMS.2024.052
E-ISSN: 2539-6056

and rehabilitating individuals with aphasia, aiming for continuous, effective, standardized therapy overseen by certified professionals. However, obstacles like transportation costs, rural residency, and lack of specialized care, including access to speech-language pathologists, hinder access to services.[4]Additionally, the COVID-19 pandemic poses challenges, particularly for elderly aphasia patients, hindering in-person therapy attendance due to close contact risks. Adapting therapy becomes imperative, with telepractice emerging as a crucial solution, offering increased access while minimizing risks.[5]

Telepractice employs conferencing technology, fostering multidisciplinary collaboration to enhance therapy continuity, frequency, and duration. Synchronous and asynchronous modes, primarily through videoconferencing, are common.[6] For speech therapy, optimal effectiveness requires at least three sessions per week, with a minimum of nine sessions necessary to achieve significant progress. Studies show that telepractice is as effective as in-person sessions and reduces costs.[4] Collaboration between speech-language pathologists (SLPs) and caregivers plays an essential role in telepractice in overcoming the lack of tactile cues. Although there is limited research in Thailand, this study aims to investigate the effectiveness of telepractice in improving speech and language therapy for individuals with aphasia. This will help to increase accessibility to treatment for people with communication difficulties.

Materials and methods

Materials

This study utilized assessment tools and survey forms to comprehensively evaluate individuals with aphasia and their caregivers. The SD-SLP-01 Screening test for aphasia was used initially to identify individuals with aphasia, followed by the Thai Adaptation of Western Aphasia Battery (TWAB) to provide a detailed assessment of their linguistic function and severity levels of aphasia. Satisfaction surveys were designed and utilized to measure the satisfaction levels of participants with aphasia and their caregivers concerning the telepractice interventions. Importantly, no fees were collected for telepractice at either location, in line with the hospitals' regulations. As a result, there was no bias in satisfaction levels due to cost, ensuring a fair comparison of participant satisfaction with telepractice.

Screening test for aphasia (SD-SLP-01)[7]

This study uses the SD-SLP-01 Screening test for aphasia to screen individuals for receptive and expressive language skills. The test consists of a 30-item list, and individuals are assessed based on their ability to complete each test item. Those who score less than 27 are considered to have aphasia. This tool is included in the study's inclusion criteria to determine whether individuals have aphasia.

Thai Adaptation of Western Aphasia Battery (TWAB)[8]

Thai Adaptation of Western Aphasia Battery (TWAB) comprises four subtests, each serving as a distinct measure of linguistic function. These subtests include Spontaneous Speech, which assesses the individual's ability to produce language; Auditory Verbal Comprehension, which evaluates the comprehension of spoken language; Repetition, focusing on the ability to repeat spoken words or phrases; and Naming, which assesses the ability to name objects or respond to naming cues. These subtests collectively provide a comprehensive evaluation of various language skills, categorizing individuals into specific types of aphasia based on their performance.

The scores obtained in each subtest can help categorize individuals into specific types of aphasia. Additionally, the TWAB test generates an Aphasia Quotient (AQ) score, which indicates the individual's auditory-verbal communication ability and the extent of aphasia severity. AQ scores range from 0 to 100, with lower scores indicating more significant language deficits. An AQ score of 94.7, the lowest observed among individuals without aphasia, can be used as a threshold to differentiate between normal individuals and those with aphasia.[8]

The study utilized the Aphasia Quotient (AQ) score from the Thai adaptation of the Thai Adaptation of Western Aphasia Battery (TWAB) as a baseline to compare with the post-test scores following the intervention.

Satisfaction survey forms

The study employed satisfaction survey forms designed by the researchers to evaluate the satisfaction levels of participants with aphasia and their caregivers regarding the telepractice interventions. The survey includes two forms, one tailored for participants with aphasia and another for their caregivers. The survey form for aphasic participants utilizes a three-point satisfaction scale (dissatisfied, neutral, satisfied) accompanied by symbolic pictures to alleviate linguistic challenges. The form for caregivers employs a five-point satisfaction scale (very dissatisfied, dissatisfied, neutral, satisfied, very satisfied). For aphasic participants, scores ranging from 1.00 to 1.66 were categorized as dissatisfied, 1.67 to 2.33 as neutral, and 2.34 to 3.00 as satisfied. Regarding caregivers, scores from 1.00 to 1.50 were considered very dissatisfied, from 1.51 to 2.50 as dissatisfied, 2.51 to 3.50 as neutral, 3.51 to 4.50 as satisfied, and 4.51 to 5.00 as very satisfied.[9]

Participants recruitment

Participants were Thais with aphasia and their caregivers. The individuals with aphasia were recruited from two hospitals in Thailand, Phaholpolpayuhasena Hospital in Kanchanaburi and the AMS Clinical Service Center (Speech Clinic) at Chiang Mai University. Inclusion criteria for individuals with aphasia were as follows: 1) Individuals assessed with SD-SLP-01, a screening test for aphasia, with a score lower than 27, indicating the presence of aphasia; 2) Attaining an auditory comprehension score of more than 42 out of 60 in the yes-no question subtest from Thai Adaptation of Western Aphasia Battery (TWAB). 3) Having chronic aphasia (at least six months post-onset); 4) Managing conditions like high blood pressure and/or

seizures with medication; 5) Being able to sit up properly during telepractice sessions; 6) Possessing sufficient cognitive ability, understanding, and attention to engage in telepractice; 7) Having Thai as their native language;(8) No visual or auditory impairments; and 9) Being proficient in reading and writing in Thai before the onset of aphasia. The exclusion criteria included complications that could impact the effectiveness of telepractice, such as experiencing recurrent strokes. Additionally, participants were not receiving on-site speech therapy.

The caregivers who participated in this study had to meet the following requirements: 1) Be at least 18 years old; 2) Have Thai as their native language; 3) Be proficient in reading and writing in Thai; 4) Have no visual or auditory impairments; 5) Possess the necessary equipment and skills to use telepractice tools, such as Zoom cloud meetings, tablets or computer, earphones, and microphones; and 6) Be able to facilitate telepractice, assist the person with aphasia, and implement advice from the speech and language pathologist.

The study included eight Thais with aphasia and their caregivers. Table 1 presents information about the participants with aphasia, while Table 2 details their caregivers. Table 3 shows TWAB scores and types of aphasia for each individual with aphasia.

Table 1. Information on participants with aphasia.

General information	Numbers (%)
Gender	
Male	6 (75.0)
Female	2 (25.0)
Age	
20-39	1 (12.5)
40-59	5 (62.5)
>60	2 (25.0)
Average age and standard deviation	54.62±10.87
Educational level	
Diploma	1 (12.5)
Bachelor's degree	3 (37.5)
Higher than bachelor's degree	4 (50.0)

Table 2. Information of caregivers of participants with aphasia

General information	Numbers (%)
Gender	
Male	3 (37.5)
Female	5 (62.5)
Age	
20-39	2 (25.0)
40-59	2 (25.0)
>60	4 (50.0)
Average age and standard deviation	54.62±14.75
Relationship to individuals with aphasia	
Relatives (sister, father)	2 (25.0)
Spouses	5 (62.5)
Children	1 (12.5)

Table 3. TWAB scores of participants with aphasia.

	P1	P2	P3	P4	P5	P6	P7	P8
Spontaneous speech								
Functional content (10)	7	8	8	5	9	10	10	7
Fluency (10)	5	4	6	4	5	8	9	4
Total (20)	**12**	**12**	**14**	**9**	**14**	**18**	**19**	**11**
Auditory verbal comprehension								
Yes/no questions (60)	42	48	51	54	60	60	51	54
Auditory word recognition (60)	58	57	57	46	59	58	55	57
Sequential command (80)	36	68	65	58	80	76	50	54
Total (10)	**6.8**	**8.65**	**8.65**	**7.4**	**9.95**	**9.7**	**7.8**	**8.25**
Repetition								
Repetition	*60*	*90*	*61*	*94*	*80*	*94*	*92*	*91*
Total (10)	*6.0*	*9.0*	*6.1*	*9.4*	*8.0*	*9.4*	*9.2*	*9.1*
Naming								
Object naming (60)	45	44	53	38	60	60	57	56
Word fluency (20)	6	3	8	5	6	13	17	8
Sentence completion (10)	6	10	8	6	10	8	8	6
Responsive speech (10)	8	6	10	6	10	10	6	10
Total (10)	**6.5**	**6.3**	**7.9**	**5.5**	**8.6**	**9.1**	**8.8**	**8.0**
Aphasia quotient (100)	62.2	71.9	73.3	62.6	81.1	92.4	89.6	72.7
Type of aphasia	Wernicke	Transcortical motor	Conduction	Transcortical motor	Anomic	Anomic	Anomic	Transcortical motor

Study design

This study is an experimental research study using a group pretest-posttest design to determine the effectiveness of telepractice on people with aphasia.[10] Participants with aphasia (with assistance from their caregivers) received 1-hour session of speech therapy three times a week for four weeks (12 sessions in total). The TWAB assessment, before and after the intervention, was conducted on-site by research assistants with at least five years of experience working as speech and language pathologists. The therapy was provided by the primary researcher, who has nine years of experience as a speech and language pathologist. After the telepractice, the study investigated the effectiveness of telepractice and the satisfaction of both participants with aphasia and their caregivers.

Data analysis

In this study, statistical analysis was conducted using Stata software version 17. Descriptive statistical analysis was employed for demographic data, calculating and representing percentages, means, and standard deviations. The primary objective was to evaluate the effectiveness of telepractice for speech therapy, reflecting an individual's language ability and the severity of aphasia. The study compared pretest and posttest scores of the Thai Adaptation of the Western Aphasia Battery (TWAB) to determine changes in the Aphasia Quotient (AQ) score. Nonparametric statistics, specifically the Wilcoxon Matched-Pairs Signed-Ranks Test, were utilized for comparison, with a significance level set at $p<0.05$. Additionally, satisfaction survey forms were used to assess the satisfaction levels of participants with aphasia and their caregivers. The average total scores of each survey form were calculated to indicate the degree of satisfaction, represented as a percentage.

Results

TWAB scores

The study found that following speech therapy telepractice, the Aphasia Quotient (AQ) scores of all the participants with aphasia significantly increased. Additionally, the overall scores for each subtest also showed an increase. The raw scores before and after treatment are presented in Table 4.

Table 4. TWAB scores of participants with aphasia before and after receiving telepractice.

	P1 Pre/Post	P2 Pre/Post	P3 Pre/Post	P4 Pre/Post	P5 Pre/Post	P6 Pre/Post	P7 Pre/Post	P8 Pre/Post
Spontaneous speech								
Functional content (10)	7/7	8/8	8/9	5/5	9/10	10/10	10/9	7/8
Fluency (10)	5/5	4/5	6/6	4/4	5/9	8/9	9/9	4/4
Total (20)	**12/12**	**12/13**	**14/15**	**9/9**	**14/19**	**18/19**	**19/18**	**11/12**
Auditory verbal comprehension								
Yes/no questions (60)	42/48	48/54	51/60	54/42	60/60	60/57	51/51	54/54
Auditory word recognition (60)	58/58	57/58	57/55	46/50	59/60	58/58	55/56	57/59
Sequential command (80)	36/30	68/76	65/74	58/60	80/80	76/80	50/66	54/42
Total (10)	**6.8/6.8**	**8.65/9.4**	**8.65/9.45**	**7.4/7.6**	**9.95/10**	**9.7/9.75**	**7.8/8.65**	**8.25/7.75**
Repetition								
Repetition	60/60	90/92	61/87	94/100	80/98	94/94	92/100	91/94
Total (10)	**6.0/6.0**	**9.0/9.2**	**6.1/8.7**	**9.4/10**	**8.0/9.8**	**9.4/9.4**	**9.2/10**	**9.1/9.4**
Naming								
Object naming (60)	45/48	44/44	53/58	38/40	60/60	60/60	57/58	56/58
Word fluency (20)	6/6	3/4	8/16	5/6	6/12	13/14	17/17	8/11
Sentence completion (10)	6/6	10/10	8/8	6/6	10/10	8/8	8/4	6/6
Responsive speech (10)	8/8	6/8	10/10	6/4	10/10	10/10	6/10	10/10
Total (10)	**6.5/6.8**	**6.3/6.6**	**7.9/9.2**	**5.5/5.6**	**8.6/9.2**	**9.1/9.2**	**8.8/8.9**	**8.0/8.5**
Aphasia quotient (100)	62.20/63.2	71.9/76.4	73.3/84.7	62.6/64.4	81.1/96	92.4/94.7	89.6/91.1	72.7/80.55

When employing the Wilcoxon test to compare pre- and post-scores, the study identified a significant improvement in the Aphasia Quotient (AQ) score: the pre-median was 73 (62.6-92.4), and the post-median was 80.55 (63.2-96), with a *p* value of 0.011.

The study results showed significant differences in the repetition subtest, with a pre-median of 90.5 (range: 60-94) and a post-median of 94 (range: 60-100), resulting in a p-value of 0.019. Significant differences were also observed in the naming subtest, with a pre-median of 7.95 (range: 5.5-9.1) and a post-median of 8.7 (range: 5.6-9.2), yielding a *p* value of 0.011. After the intervention, the median scores for Digit Span Forward increased from 6 to 7 (*p*=0.042) and Word Fluency from 7 to 11 (*p*=0.019).

No significant differences were found in the scores of the spontaneous speech subtest and auditory comprehension subtest following the telepractice intervention. Table 5 presents the differences between pre- and post-TWAB test scores for each subtest and Aphasia Quotient (AQ) scores.

Table 5. Data analysis results of TWAB scores of participants with aphasia before and after receiving telepractice.

Topics	N	pre			post			Mean difference	Median difference	Z	*p* value
		Mean	SD	Median (min-max)	Mean	SD	Median (min-max)				
Spontaneous speech											
Functional content	8	8	1.69	8 (5-10)	8.25	1.67	8.5 (5-10)	-0.25	0	-1.000	0.317
Fluency	8	5.62	1.92	5 (4-9)	6.37	2.26	5.5 (4-9)	-0.75	0	-1.723	0.085
Total	**8**	**13.62**	**3.42**	**13 (9-19)**	**14.62**	**3.74**	**14 (9-19)**	**-1**	**-1**	**-1.673**	**0.094**
Auditory verbal comprehension											
Yes/no questions	8	52.5	6	52.5 (42-60)	53.25	6.16	54 (42-60)	-0.75	0	-0.436	0.663
Auditory word recognition	8	55.87	4.15	57 (46-59)	56.75	3.15	58 (50-60)	-0.87	-1	-1.427	0.154
Sequential command	8	60.87	14.41	61.5 (36-80)	63.5	18.57	70 (30-80)	-2.62	-3	-0.912	0.361
Total	**8**	**8.4**	**1.08**	**8.45 (6.8-9.95)**	**8.67**	**1.17**	**9.02 (6.8-10)**	**-0.27**	**-0.13**	**-1.757**	**0.079**
Repetition	8	82.75	14.43	90.5 (60- 94)	90.62	13.12	94 (60-100)	-7.87	-4.5	-2.339	0.019*
Total	**8**	**8.27**	**1.44**	**9.05 (6-9.4)**	**9.06**	**1.31**	**9.4 (6-10)**	**-0.78**	**-0.45**	**-2.339**	**0.019***
Naming											
Object naming	8	51.62	8.26	54.5 (38-60)	53.25	7.99	58 (40-60)	-1.62	-1.5	-2.179	0.029*
Word fluency	8	8.25	4.59	7 (3-17)	10.75	4.92	11.5 (4-17)	-2.5	-1	-2.351	0.019*
Sentence completion	8	7.75	1.67	8 (6-10)	7.25	2.12	7 (4-10)	0.5	0	1.000	0.317
Responsive speech	8	8.25	1.98	9 (6-10)	8.75	2.12	10 (4-10)	-0.5	0	-0.656	0.511
Total	**8**	**7.59**	**1.32**	**7.95 (5.5-9.1)**	**8**	**1.44**	**8.7 (5.6-9.2)**	**-0.41**	**-0.2**	**-2.880**	**0.011***
Aphasia quotient	8	75.77	11.17	73 (62.6-92.4)	80.72	12.95	80.55 (63.2-96)	-4.95	-2.45	-2.521	0.011*

Satisfaction survey forms

After completing all 12 telepractice sessions, Only the research assistant administered a satisfaction assessment to individuals with aphasia and their caregivers. Table 6 depicts the satisfaction levels of caregivers of participants with aphasia regarding telepractice. Based on the data presented in Table 6, caregivers' satisfaction level towards telepractice is very high across all five aspects. These aspects are as follows: 1) user guide for telepractice-with an average score ranging from 4.625 to 4.875; 2) utilization of Zoom Cloud Meeting for telepractice-having an average score of 4.75; 3) telepractice sessions - with an average score ranging from 4.5 to 4.875; 4) benefits of telepractice with an average score ranging from 4.75 to 4.875; and 5) overall satisfaction-having an average score of 4.75. The caregivers find telepractice very beneficial and are highly satisfied with it.

Table 6. Satisfaction levels of caregivers of participants with aphasia regarding telepractice.

Questions	Very satisfied	Satisfied	Neutral	Disatisfied	Very dissatisfied	Average
User's guide telepractice						
1. The font size is visible and clear, making it easy to read	6 (75.0)	2 (25.0)	0	0	0	4.75
2. The size and color of the pictures are clearly visible	5 (62.5)	3 (37.5)	0	0	0	4.625
3. The user's guide is easy to read and follow	6 (75.0)	2 (25.0)	0	0	0	4.75
4. It contains enough necessary information	7 (87.5)	1 (12.5)	0	0	0	4.875
Utilization of Zoom cloud meeting for telepractice						
1. The convenience of using Zoom makes it easy to use	6 (75.0)	2 (25.0)	0	0	0	4.75
2. Utilizing Zoom for telepractice is practical	6 (75.0)	2 (25.0)	0	0	0	4.75
3. Using Zoom makes it easy to communicate smoothly with speech and language pathologists	6 (75.0)	2 (25.0)	0	0	0	4.75
Telepractice sessions						
1. Duration of the session	4 (50)	4 (50)	0	0	0	4.5
2. The activities and utensils, such as pictures	7 (87.5)	1 (12.5)	0	0	0	4.875
3. Providing advice during the session	6 (75.0)	2 (25.0)	0	0	0	4.75
Benefits of telepractice						
1. Telepractice through Zoom saves time and travel expenses	7 (87.5)	1 (12.5)	0	0	0	4.875
2. Telepractice sessions at home are more comfortable, and the environment is familiar	7 (87.5)	1 (12.5)	0	0	0	4.875
3. Reduction in concern about disease transmission during in-person sessions	7 (87.5)	1 (12.5)	0	0	0	4.875
4. Telepractice via Zoom reduces wait times for training appointments and increases the frequency of speech therapy sessions	6 (75.0)	2 (25.0)	0	0	0	4.75
Overall satisfaction						
Overall satisfaction	6 (75.0)	2 (25.0)	0	0	0	4.75

Similarly, for participants with aphasia, the satisfaction assessment results of individuals with aphasia towards telepractice are consistently high across all six aspects, namely (see Table 7): 1) convenience in training through Zoom, with an average score of 2.875; 2) telepractice through Zoom saves time and travel expenses, with an average score of 3; 3) telepractice session duration and frequency have an average score of 2.625; 4) telepractice through Zoom reduces waiting time for training appointments, with an average score of 3; 5) reduction in concern about disease transmission during in-person sessions has an average score of 2.875; and 6) overall satisfaction with telepractice has an average score of 3. Individuals with communication disorders have provided high satisfaction ratings for telepractice across these criteria.

Table 7. Satisfaction levels of participants with aphasia regarding telepractice.

Questions	Dissatisfied	Neutral	Satisfied	Average
1. Convenience in training through Zoom	0	1 (12.5)	7 (87.5)	2.875
2. Telepractice through Zoom saves time and travel expenses	0	0	8 (100)	3
3. Telepractice session duration and frequency	0	3 (37.5)	5 (62.5)	2.625
4. Telepractice through Zoom reduces waiting time for training appointments	0	0	8 (100)	3
5. Reduction in concern about disease transmission during in-person sessions	0	1 (12.5)	7 (87.5)	2.875
6. Overall satisfaction with telepractice	0	0	8 (100)	3

Discussion

The study examined the effectiveness of delivering speech and language therapy via telepractice for individuals with aphasia, assessing their language abilities before and after receiving the intervention through the Thai Adaptation of the Western Aphasia Battery (TWAB) assessment. The analysis employed nonparametric statistics, specifically the Wilcoxon Matched-Pairs Signed-Ranks Test, with a significance level set at 0.05.[11]The statistical analysis revealed significant differences in Aphasia Quotient (AQ) scores before and after participating in the study, indicating a notable improvement in the overall language abilities of each participant. This highlights the effectiveness of telepractice in enhancing the linguistic skills of individuals with aphasia.

In comparing pretest and posttest AQ scores following telepractice, the median AQ score increased from 73 at pretest to 80.55 at posttest, with a p-value of 0.011, indicating a statistically significant improvement. This finding aligns with previous research on iAphasia and the Korean version of the Western Aphasia Battery (KWAB) by Choi *et al*.,[12] which reported an approximate 24.80% increase in AQ, as well as the results of the study by Jacob *et al*.[13] Furthermore, comparing the raw scores of participants 5 and 6, whose pretest scores were 81.10 and 92.40, respectively, with posttest scores of 96 and 94.70, respectively, demonstrates significant improvement. According to the criteria established by Woranwan *et al*.,[8] an AQ score of 94.7, the lowest observed among individuals without aphasia, serves as a threshold for distinguishing between normal individuals and those with aphasia. Both participants meet the threshold based on this criterion, indicating notable progress.

Examining the scores within each subtest, the study identifies significant improvement in the repetition and naming subtests. At the same time, no notable changes were observed in the spontaneous speech and auditory comprehension subtests. One possible hypothesis for the lack of substantial improvement in the spontaneous speech and auditory comprehension subtests is that participants initially had relatively strong skills in these areas. Individuals with aphasia participating in the study might have retained specific proficiency in spontaneous speech and auditory comprehension even before the telepractice intervention. Only one participant (P2) had Wernicke's aphasia, characterized by comprehension difficulties, while the remaining participants had relatively strong skills in this area. Consequently, the lack of significant changes in these subtests may indicate that the intervention did not substantially impact participants who already demonstrated proficiency in spontaneous speech and auditory comprehension skills.

The satisfaction assessments' results for caregivers and individuals with aphasia highlight a consistently high level of contentment with the telepractice intervention. As depicted in Table 6, caregivers expressed robust satisfaction across various dimensions, including the user guide, utilization of Zoom cloud meeting, telepractice sessions, perceived benefits, and overall satisfaction. The average scores within the very high range underscore the positive impact and effectiveness of telepractice, emphasizing its user-friendly nature and perceived advantages. Similarly, as indicated in Table 7, individuals with aphasia conveyed notable satisfaction in multiple aspects, such as convenience in training through Zoom, time and cost savings, and reduced waiting times.[14,15] This study considered the bias from the service cost on satisfaction levels. The billing for telepractice in all sessions was done using the same process as the usual service. Therefore, we thought that the differences in satisfaction scores were not due to this bias. The overall satisfaction score reflects a positive response from individuals with aphasia, affirming the acceptability and success of telepractice in addressing their specific needs and concerns.[16]

The consistently high satisfaction levels from both groups underscore the feasibility and acceptability of telepractice as an effective mode of delivering speech therapy for individuals with aphasia. These findings are consistent with existing research demonstrating the effectiveness of telepractice and emphasize its potential integration into everyday clinical practice to enhance accessibility and improve outcomes in speech and language therapy. They provide valuable insights into how telepractice positively affects language skills, highlighting its potential as an accessible and effective form of communication therapy.[17]

Limitations

During the data collection phase, limitations were encountered, primarily stemming from the instability of internet connections, leading to issues such as sound loss or video lag during video conferencing sessions. These disruptions, often linked to external factors such

as heavy rainfall or nearby construction noise, prompted various troubleshooting efforts by the researcher. Adjustments included rescheduling sessions and advising participants on alternative connectivity options. Patients and speech and language pathologists needed help scheduling appointments, requiring weekend sessions and flexible time slots. A weekly confirmation system was implemented to enhance scheduling efficiency. Fatigue also posed a significant concern among aphasia patients during telepractice sessions due to prolonged screen time and the absence of tactile cues. The speech and language pathologists took proactive measures to address fatigue, such as monitoring patient responses and providing breaks or private discussions with caregivers.

For future studies, several recommendations should be considered. Firstly, During the session, we should be mindful of fatigue resulting from prolonged screen time, which can lead to increased tiredness for the patient compared to face-to-face sessions. Secondly, comparative experiments between in-person and remote training should be expanded to evaluate the effectiveness of remote methods. Caution should also be exercised to ensure that the duration between pretest and posttest TWAB assessments is at least 3 months apart, minimizing the risk of bias and preventing participants from recalling the questions. Thirdly, follow-up evaluations should be conducted one-month post-training to assess the maintenance of improvements in patients' abilities. Lastly, a remote training system should be developed to enhance access in areas without speech and language pathologists or adequate equipment, potentially through network collaborations with community hospitals, to streamline patient appointments and support, thereby minimizing missed opportunities and waiting times.

Conclusion

This study evaluated aphasia patients who underwent telepractice for speech therapy by comparing their TWAB scores before and after training, revealing improved language abilities post-training. However, it is crucial to acknowledge the study's limitation with only eight participants, suggesting that the findings may not generalize to the broader population and highlighting its preliminary nature as a pilot study. Nonetheless, these results support the hypothesis that telepractice effectively enhances the language abilities of aphasia patients.

Regarding satisfaction with telepractice, both individuals with aphasia and caregivers reported high to very high satisfaction levels. This consistently positive feedback underscores telepractice's feasibility and acceptability as a valuable tool in speech therapy. These insights are pivotal for integrating telepractice into routine clinical practice to improve accessibility and treatment outcomes for individuals with aphasia.

In conclusion, this study provides valuable insights into the positive impact of telepractice on linguistic abilities and advocates for its potential as an accessible and effective communication therapy intervention.

Furthermore, telepractice could serve as a promising alternative for individuals with various communication impairments, thereby advancing inclusivity and enhancing the efficiency of speech therapy interventions.

Conflict of interest

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

Ethical approval

This study constitutes part of the thesis and fulfills the requirements for a master's degree in communication disorders. It has obtained ethical approval from the Faculty of Associated Medical Sciences Research Ethics Committee, Chiang Mai University (Approval ID: AMSEC-64EX-123) and Phaholpolpayuhasena Hospital (Approval ID: 2023-01). Each participant was provided with comprehensive information regarding the research, and their informed written consent was obtained before participation in the study commenced.

Acknowledgements

The researcher deeply appreciates the staff at Phaholpolpayuhasena Hospital in Kanchanaburi and the AMS Clinical Service Center (Speech Clinic) at Chiang Mai University for their cooperation in recruiting participants for this study.

References

[1] Berthier ML. Poststroke aphasia. Drugs Aging. 2005; 22(2): 163-82. doi.org/10.2165/00002512-200522020-00006

[2] Parr S, Byng S, Gilpin S. Talking about aphasia: Living with loss of language afterstroke. McGraw-Hill Education (UK); 1997.

[3] American Speech-Language-Hearing Association. Aphasia. 2016[cited 2024 Jul 3]. Available from: https://www.asha.org/Practice-Portal/Clinical-Topics/Aphasia/.

[4] Knepley KD, Mao JZ, Wieczorek P, Okoye FO, Jain AP, Harel NY. Impact of telerehabilitation for stroke-related deficits. Telemed e-Health. 2021; 27(3): 239-46. doi.org/10.1089/tmj.2020.0036

[5] Fong R, Tsai CF, Yiu OY. The implementation of telepractice in speech language pathology in Hong Kong during the COVID-19 pandemic. Telemed e-Health. 2021; 27(1): 30-8. doi.org/10.1089/tmj.2020.0155

[6] Øra HP, Kirmess M, Brady MC, Winsnes IE, Hansen SM, Becker F. Telerehabilitation for aphasia–protocol of a pragmatic, exploratory, pilot randomized controlled trial. Trials. 2018; 19(1): 1-10. doi.org/10.1186/s13063-018-2702-7

[7] Suphawatjariyakul R, Lorwatanapongsa P, Euasirirattanapaisan A. Handbook for helping people with aphasia. Saraburi: Thaisiri Printing; 2012.

[8] 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.

[9] Joshi A, Kale S, Chandel S, Pal DK. Likert scale: Explored and explained. Br J Appl Sci Technol. 2015; 7(4): 396-403. doi.org/10.9734/BJAST/2015/14975

[10] Fitz-Gibbon CT, Morris LL. How to design a program evaluation. Sage; 1987.

[11] Vanichbuncha K. Statistics for research. 6th ed. Bangkok: Chulalongkorn Business School; 2013.

[12] Choi YH, Park HK, Paik NJ. A telerehabilitation approach for chronic aphasia following stroke. Telemed e-Health. 2016; 22(5): 434-40. doi.org/10.1089/tmj.2015.0100

[13] Jacobs M, Briley PM, Fang X, Ellis C. Telepractice treatment for aphasia: Association between clinical outcomes and client satisfaction. Telemed Rep. 2021; 2(1): 118-24. doi.org/10.1089/tmr.2020.0025

[14] Nguyen M, Waller M, Pandya A, Portnoy J. A review of patient and provider satisfaction with telemedicine. Curr Allergy Asthma Rep. 2020; 20(11): 1-7. doi.org/10.1007/s11882-020-00969-4

[15] Palasik S, Irani F. Evaluating telepractice: Measuring success, cost-benefit, & outcomes. In: Telepractice in Speech-Language Pathology. 2013:187.

[16] Weidner K, Lowman J. Telepractice for adult speech-language pathology services: A systematic review. Perspect ASHA Spec Interest Groups. 2020; 5(1): 326-38. doi.org/10.1044/2019_PERSP-19-00146

[17] Aueworakhunanan T, Dejket P, Phakkachok S, Punkla W. Exploring the feasibility of implementing telepractice innovation for speech-language pathologists in Thailand. Ramathibodi Med J. 2024; 47(1): 32-46. doi.org/10.33165/rmj.2024.47.1.266036

Journal of Associated Medical Sciences 2024; 57(3): 107-118

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Comparison of acute cardiovascular and perceptual responses between moderate-intensity interval and continuous exercise in inactive obese young adults

Jatuporn Phoemsapthawee* Bhuwanat Sriton Piyaporn Tumnark

Department of Sports Science, Faculty of Sports Science and Health, Kasetsart University, Nakhon Pathom Province, Thailand.

ARTICLE INFO

Article history:
Received 19 April 2024
Accepted as revised 9 July 2024
Available online 11 July 2024

Keywords:
Affective, cardiac output, cyclic aerobic activity, heart rate variability, obesity.

ABSTRACT

Background: Obesity dramatically elevates the risk of mortality and morbidity associated with cardiovascular disease (CVD). Obese people exhibit left ventricular structure abnormalities at an early age, in addition to having diminished resting systolic and diastolic function. Furthermore, there is evidence to suggest that autonomic dysfunction may play a role in the increased prevalence of CVD among obese individuals. Engaging in vigorous exercises may place excessive strain on the cardiovascular system, rendering it inappropriate for inactive, obese individuals.

Objective: The crossover design with randomization study aimed to compare the immediate effects of moderate-intensity interval (MIIE) and continuous exercises (MICE) on hemodynamics, heart rate variability (HRV), affective, and enjoyment responses in young obese men.

Materials and methods: A total of eighteen male participants, consisting of nine individuals with normal weight and nine individuals classified as obese, were involved in this study. The participants completed two 30-minute cycling protocols consisting of MIIE (3×5-minute at 90% ventilatory threshold; VT) interspersed with 3×5-minute active recovery cycling at 50% VT) or MICE at 70% VT with a 3-minute warm-up and cool-down at free load. Hemodynamic, HRV, enjoyment, and affective responses were measured at rest and at 5-, 10-, 15-, 20-, 25-, and 30-minute intervals during the exercise sessions of each protocol.

Results: There were no significant differences in stroke volume index, cardiac index, ejection fraction, or systolic blood pressure between the MICE and MIIE protocols after each stage of the exercise protocols, both in normal-weight and obese participants. The MIIE protocol led to significantly lower ratings of perceived exertion (RPE) (p<0.05) and higher enjoyment (p<0.05) and affective responses (p<0.01) relative to MICE after each stage of exercise in obese participants. In addition, the MICE protocol led to a significantly lower standard deviation of normal-to-normal intervals (p<0.01) and a very low frequency (p<0.01) when compared to the MIIE in obese participants.

Conclusion: Our study's findings indicate that the MIIE and MICE protocols elicit similar cardiac responses in normal-weight and obese participants. Furthermore, the MIIE protocol results in significantly less RPE and exercise-induced fatigue and greater levels of enjoyment and affective response. The potential for heightened enjoyment and affective responses within the MIIE protocol may have significant implications for fostering exercise adherence among individuals with obesity.

* **Corresponding contributor.**
Author's Address: *Department of Sports Science, Faculty of Sports Science and Health, Kasetsart University, Nakhon Pathom Province, Thailand.*
E-mail address: jatuporn.w@ku.th
doi: 10.12982/JAMS.2024.053
E-ISSN: 2539-6056

Introduction

The global obesity epidemic is a major public health concern. Obesity raises the risk of cardiovascular disease (CVD) mortality and morbidity.[1] Previous research has demonstrated that young, healthy, obese people have early abnormalities in left ventricular structure as

well as reduced resting systolic and diastolic function.[2] Furthermore, it has been demonstrated that young, healthy, obese individuals have a higher cardiac response but a lower arteriovenous oxygen difference during exercise. These conditions may precede obesity-related cardiac remodeling and dysfunction.[3] The severity of cardiac impairment is related to the degree of obesity.[4,5] Furthermore, a reduction in autonomic function has been proposed as a possible explanation for the higher prevalence of CVD in obese people.[6]

Regular exercise is one of the most well-known and effective strategies for managing obesity. The ideal exercise program should have a significant negative energy balance, long-term adherence, and psychological and physical benefits.[7] Obese people frequently experience high attrition and low adherence rates to moderate-intensity continuous training (MICT), which is commonly recommended as the initial step in a weight-loss exercise program.[8] High-intensity interval training (HIIT) has gained popularity due to its reputation as a time-efficiency and effective exercise modality. Systematic reviews and meta-analyses of supervised and unsupervised exercise training studies have repeatedly shown that HIIT improves cardiorespiratory fitness (CRF), often to a greater extent than conventional MICT.[8,9]

Nevertheless, recent research indicates that HIIT prescriptions for overweight and obese individuals may not be optimal in practice. Low-active, overweight, or obese participants assigned to HIIT demonstrated a slight decline in net physical activity from baseline to twelve months. In contrast, those in the MICT group became marginally more physically active.[10] Furthermore, a recent report demonstrated a lack of long-term adherence to vigorous-intensity activity among HIIT participants.[11] This suggests that these individuals may have encountered difficulties in their ability or willingness to engage in recommended intensity exercise levels. When unsupervised, more of those assigned to HIIT did not adhere to their prescription than those assigned to MICT. Due to nonadherence, initial physiological adaptations tended to diminish over the course of the follow-up.[11]

Furthermore, it is hypothesized that the anticipated affective responses to future exercise can influence exercise adherence.[12] A significant association has been observed between the level of intensity and affective responses during continuous exercise.[13] The anaerobic threshold (AT) has been identified as an extremely reliable physiological indicator of affective responses.[14] According to the Dual-Mode Model,[13] negative affective responses are more common at intensities above AT, and positive affective responses at intensities below AT. Given the low adherence and health concerns associated with vigorous exercise in sedentary and obese individuals, the efficacy of a moderate-intensity interval exercise (MIIE) protocol should be investigated. If reducing exercise intensity and incorporating varying intensity levels during sessions can still elicit physiological and perceptual responses, it is necessary to determine whether or not the protocol is still effective. Coquart *et al.* 2008[15] demonstrated previously that obese individuals perceived MIIE as less strenuous than moderate-intensity continuous exercise (MICE) protocol despite similar training relative to workload and exercise duration. This led the authors to suggest that MIIE rather than MICE be prescribed to maintain adherence to exercise programs in inactive obese populations. Nonetheless, this remains a contentious issue.[16]

It is critical to determine whether MIIE can produce similar favorable cardiovascular outcomes as MICE while preserving effective and enjoyable responses. In addition, to effectively prescribe exercise to prevent or treat cardiac dysfunction associated with obesity, a greater understanding of how cardiac function and autonomic changes during exercise occur in obese participants is required.

Thus, this study aimed to investigate the acute effects of MIIE and MICE on cardiac function, cardiac autonomic modulation, and affective and enjoyment responses in young obese men.

Materials and methods

Participants

Eighteen young men were recruited to participate in the study. Participants were sought out via word-of-mouth and the widespread distribution of flyers around the campus. All participants had sedentary lifestyles, defined as less than 1 hour of regular weekly exercise for the previous 6 months. According to body mass index (BMI), participants were grouped as normal-weight (18.5-22.9 kg/m^2) or obese (BMI≥25 kg/m^2). Participants were excluded if they had: 1) a history of hypertension, dyslipidemia, diabetes mellitus, cardiovascular disease, or endocrine disease; 2) were taking any medication that could potentially influence cardiometabolism; 3) smoked or used tobacco within the previous 6 months; 4) were unable to engage in physical activity due to musculoskeletal limitations; 5) had a psychiatric disorder; or 6) were on a weight loss diet within the previous 6 months. Based on a previous study, we estimated that nine men would have 90% power to detect a 33% increase in enjoyment after HIIT.[17] The study protocol was approved by the Kasetsart University Research Ethics Committee (COA no. COA61/076). Before signing a written informed consent form, participants were informed of potential risks and the study procedure.

Protocol

The study employed a crossover design with randomization. Participants were instructed to visit the laboratory on three occasions, including an initial visit and two experimental protocols. At least 72 hours separated the experimental protocols, which were conducted randomly. Participants were instructed to avoid alcohol and caffeine for 12 hours before each trial, to stay hydrated, and to avoid exercise for 24 hours before each trial. Measurements of anthropometry, body composition, blood pressure (BP), and peak oxygen consumption ($\dot{V}O_2$peak) were taken during the initial visit. At each exercise session, participants engaged in MIIE or MICE

protocols. Figure 1 illustrates the experimental protocol that each experiment employed. All participants were trained on the cycle ergometer and exercise protocols before the testing. Each experimental protocol consisted of a resting condition lasting 10 minutes, followed by a 36-minute MIIE or MICE protocol. Upon arrival at the laboratory for each experimental protocol, participants were given a standard meal (approximately 70 kcal, 75% carbohydrates, 12% protein, 13% fat, and 50 ml of water) to control variations in energy intake and nutrient composition that could affect exercise performance and physiological responses. The experiments began an hour after the meal had been consumed.

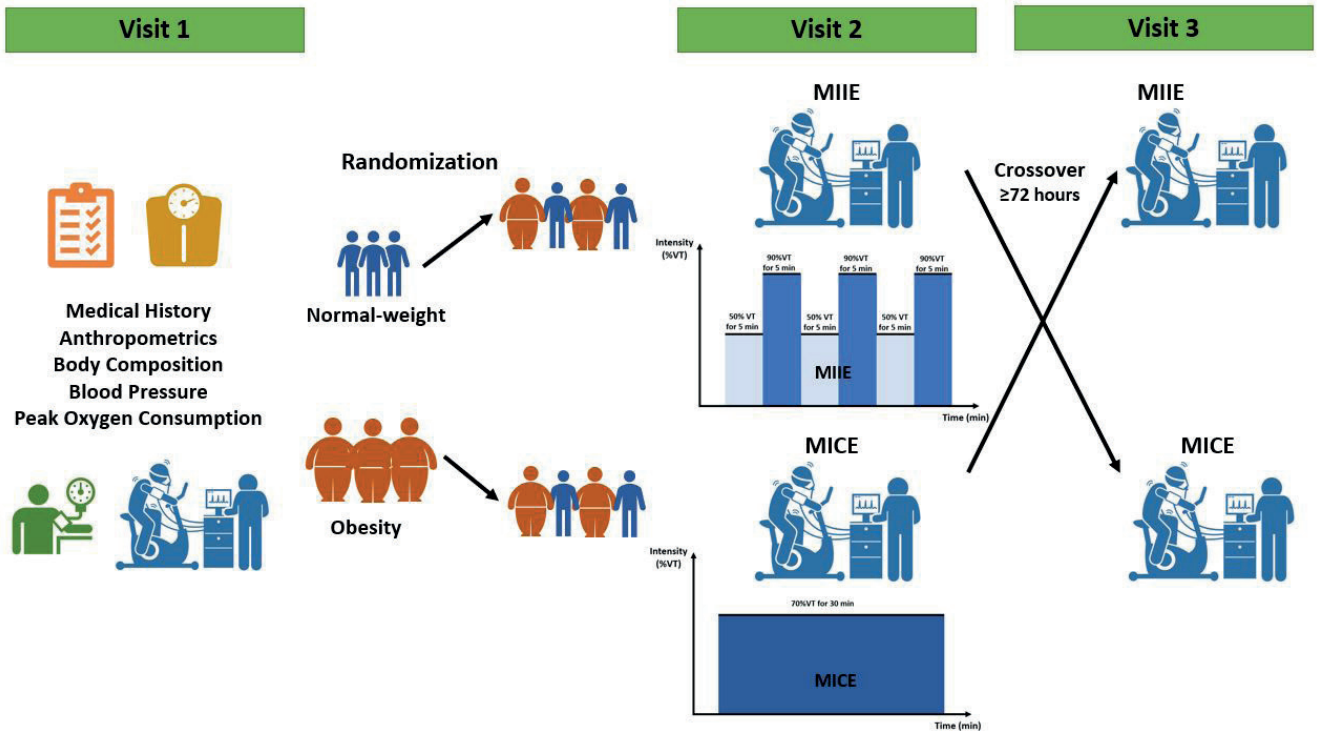


Figure 1. *Experimental design.*

MICE and MIIE protocols

The MICE protocol consisted of a 3-minute warm-up with free-load cycling and 30 minutes continuous moderate-intensity cycling at 70% ventilatory threshold (VT) with an intensity workload range of 60-115 watts. The MIIE consisted of a 3-minute warm-up at free-load cycling followed by 3 repetitions of 5 minutes of moderate-intensity interval cycling at 90% VT (3×5 min at 90% VT) with an intensity workload range of 75-146 watts, interspersed with 5 minutes of active recovery cycling at 50% VT (3×5 min at 50% VT) with an intensity workload range of 42-81 watts. The MIIE protocol thus consisted of 15 minutes of moderate-intensity interval cycling followed by 15 minutes of active recovery for a total interval exercise time of 30 minutes. Following the exercise protocols, participants performed a three-minute cool-down of free-load cycling. The exercise protocols were matched in terms of duration (30 minutes) and intensity (70% VT or ~63% $\dot{V}O_2$peak). The total energy cost of $\dot{V}O_2$ was calculated for each protocol and matched in both exercise protocols.

Measurement

Anthropometry and body composition

Body mass, fat mass (FM), percentage of body fat (%BF), and fat-free mass (FFM) were measured after an overnight fast using a bioimpedance analysis device (Inbody 720, Biospace Inc., Seoul, Korea) in light clothing and without footwear. Height was measured on a flat surface with a spring-coil measuring tape and a wall-mounted stadiometer (Bilancia Pesapersone Professionale MPE, Kern & Sohn GmbH, Balingen, Germany). BMI is calculated by dividing body mass (kg) by height squared (m^2).

Blood pressure

BP was measured using an automated brachial sphygmomanometer (Tango M2, SunTech Medical Inc., NC, USA). After a 10-minute period of seated rest, resting BP was taken three times at 2-minute intervals. During exercise protocols, BP was taken in the last 30 seconds of each 5-minute interval. The rate pressure product (RPP) is a hemodynamic parameter that assesses myocardial oxygen demand and workload. RPP was calculated using the formula RPP=heart rate (HR)×systolic blood pressure (SBP).

$\dot{V}O_2$peak

Participants performed an incremental exercise test on an electromagnetically braked cycle ergometer (VIAsprint 150P, Ergoline GmbH, Bitz, Germany). The test started with a 5-minute seated rest and 3 minutes unloaded cycling as a warm-up. The incremental protocol

began with a work rate of 50 W and increased it by 25 W every two minutes until voluntary exhaustion, after which they completed a 3-minute cool-down with a free load. Expired gas samples were collected breath-by-breath during the test using a portable metabolic device (JAEGER Oxycon Mobile, CareFusion, Hochberg, Germany). RPE was determined using the Borg scale (6-20) at the end of each.[18] $\dot{V}O_2$peak is the highest value of $\dot{V}O_2$ observed over 30 seconds during the test's final stages.

Cardiac function

Thoracic bioimpedance (PhysioFlow®, Manatec Biomedical, Poissy, France) was used to assess cardiac function. Cardiac output (CO), cardiac index (CI), stroke volume (SV), SV index, HR, and ejection fraction (EF) were continuously measured and averaged over 1 minute at rest and at 5-, 10-, 15-, 20-, 25-, and 30-minute intervals during the exercise sessions of each protocol.

Heart rate variability

Inter-beat (RR) Intervals were measured using an electrocardiogram sensor (eMotion Faros device, Mega Electronics, Kuopio, Finland) during rest, exercise, and recovery periods. The data was imported into an HRV-Scanner application, which corrected artifacts and premature beats. Time domain parameters were calculated using RR intervals, namely the standard deviation of normal-to-normal intervals (SDNN) and the root mean square of sequential deviations (RMSSD). The frequency domain parameters were also determined, including the power of integrals over the high-frequency band (HF; 0.15-0.40 Hz), the low-frequency band (LF; 0.04-0.15 Hz), and the very low-frequency power (VLF; 0.001-0.04 Hz), and the LF/HF ratio was calculated.[19]

Affective responses and enjoyment

During the exercise phases, RPE, affective response, and enjoyment were measured at 5-, 10-, 15-, 20-, 25-, and 30-minute intervals. The Borg scale (6-20) was used to assess RPE.[18] All participants reported an RPE 6, indicating no exertion before beginning the protocols. The affective response was assessed using a bipolar 11-point scale, with categories ranging from -5 to 5. The numerical categories ranged from "very bad" (-5), representing the highest level of displeasure or unpleasantness, to "neutral" (0), representing the lowest level of pleasure, and "very good" (+5), representing the highest level of pleasure.[20] The level of enjoyment was graded on a 7-point scale, with 1 indicating no enjoyment and 7 indicating a high level of enjoyment. Participants were asked to select the appropriate point on the scale to indicate their level of enjoyment in response to the question, "Use the following scale to indicate how much you are enjoying this exercise session".[21]

Statistical Analysis

Standard descriptive statistics were calculated using SPSS version 26 for Windows (SPSS Inc., Chicago, IL, USA) to assess the mean characteristics of the study's participants. To verify the assumption of normality, the Shapiro-Wilk test was administered. Log transformations were applied to variables that were not normally distributed. An independent t-test was utilized to determine the differences between groups at baseline. The dependent *t*-test or repeated-measures one-way analysis of variance was utilized to conduct within-group analyses. A two-way repeated measures analysis of variance was used to determine whether there was an interaction (condition×group) or any significant differences between conditions for cardiac variables (SV index, CI, EF, HR, SBP, and rate-pressure product; RPP), HRV, and perceptual variables (RPE, enjoyment, and affective response). In post-hoc analyses, the Bonferroni method was used. A $p<0.05$ was accepted as the minimum level of significance.

Results

Table 1 presents the descriptive characteristics of the participants. Body mass ($p<0.01$), BMI ($p<0.01$), %BF ($p<0.01$), and FM ($p<0.01$) were significantly higher in the obese group compared to the normal-weight group, whereas relative $\dot{V}O_2$peak ($p<0.01$), peak CI ($p<0.05$), and VT ($p<0.05$) were significantly lower in the obese group. Resting BP, resting and peak HR, HR at VT, RPE at peak exercise and at VT, and peak SV were similar between groups.

Table 1. Descriptive characteristics of the participants.

	Normal weight (N=9)	**Obese** (N=9)	***p* value**
Age (yrs)	20.7±1.3	20.9±1.7	0.765
Body mass (kg)	68.1±7.7	92.0±16.9	0.001**
BMI (kg/m^2)	22.2±1.4	32.5±5.9	0.000**
BF (%)	25.6±6.4	36.5±6.9	0.002**
FM (kg)	17.3±4.4	34.3±11.6	0.000**
FFM (kg)	50.8±8.1	57.7±8.1	0.072
Resting HR (bpm)	80.9±13.9	79.6±6.9	0.784
SBP (mmHg)	123.5±7.0	125.6±15.2	0.697
DBP (mmHg)	74.8±9.3	76.5±8.6	0.677
MAP (mmHg)	95.7±5.6	98.0±9.8	0.526
$\dot{V}O_2$peak (mL/kg/min)	33.5±5.3	26.1±6.1	0.009**
HR peak (bpm)	180.0±11.3	181.8±13.6	0.752
RPE at peak exercise	18.8±1.2	18.9±0.9	0.836
Peak CO (L/min)	19.2±6.6	16.0±4.0	0.205
Peak CI (L/min/m^2)	10.7±4.0	7.7±2.0	0.049*
Peak SV (mL)	107.5±35.5	100.1±25.6	0.596
Peak SV index (mL/m^2)	59.9±21.6	47.6±10.5	0.120
VT (mL/kg/min)	24.6±6.6	18.3±5.8	0.034*
VT (%$\dot{V}O_2$peak)	73.5±15.9	70.1±17.3	0.675
HR at VT (bpm)	157.6±17.9	152.2±24.5	0.581
RPE at VT	14.8±2.1	15.0±2.2	0.838
Load at VT (watt)	127.5±29.9	122.5±29.9	0.713

*Note: Data are expressed as mean±SD, BMI: body mass index, %BF: body fat percentage, FM: fat mass, FFM: fat-free mass, HR: heart rate, DBP: diastolic blood pressure, SBP: systolic blood pressure, MAP: mean arterial blood pressure, $\dot{V}O_2$peak: peak oxygen consumption, RPE: rating of perceived exertion, CO: cardiac output, CI: cardiac index, SV: stroke volume, VT: ventilatory threshold, asterisk denote a significant difference between normal-weight and obese participants (*p<0.05, **p<0.01).*

As for hemodynamic variables, there was no significant difference between the MICE and MIIE protocols in average HR, average %HR peak, average $\dot{V}O_2$, %$\dot{V}O_2$peak, average CO, %CO peak, average SV, %SV peak, average SBP, %SBP peak, average DBP, %DBP peak, average MAP, %MAP peak, or average EE (Table 2). In both exercise protocols, however, the average $\dot{V}O_2$ relative to body mass was significantly lower in obese participants than in normal-weight participants.

Table 2. Hemodynamic responses to the MICE and MIIE protocols in normal-weight and obese participants.

	MICE			MIIE			Main effect of conditions η^2 (p-value)	Condition×Group interaction η^2 (p value)
	Normal weight (N=9)	Obese (N=9)	p value	Normal weight (N=9)	Obese (N=9)	p value		
Average HR (bpm)	139.0±14.4	134.4±20.1	0.581	133.4±17.5	131.9±16.7	0.862	0.055 (0.348)	0.009 (0.708)
%HR peak	77.8±7.6	74.2±9.1	0.376	77.0±4.6	74.1±9.8	0.298	0.018 (0.593)	0.000 (0.934)
Average $\dot{V}O_2$ (mL/kg/min)	21.7±4.5	16.5±5.3	0.040*	20.9±3.0	15.8±4.1	0.008**	0.023 (0.550)	0.000 (0.982)
%$\dot{V}O_2$peak	63.9±12.8	61.8±14.3	0.750	63.0±9.5	63.0±12.9	0.998	0.000 (0.950)	0.007 (0.733)
$\dot{V}O_2$/HR	10.3±1.6	10.0±4.0	0.850	10.3±1.3	10.7±1.4	0.462	0.023 (0.546)	0.027 (0.515)
Average CO (L/min)	14.1±4.4	15.7±3.0	0.391	13.5±3.7	14.2±1.6	0.653	0.055 (0.351)	0.011 (0.680)
%CO peak	75.9±28.7	101.1±28.9	0.082	72.7±28.2	93.9±23.8	0.103	0.022 (0.560)	0.003 (0.825)
Average SV (mL)	99.8±23.7	117.6±20.1	0.105	99.8±21.0	108.1±19.2	0.397	0.024 (0.543)	0.024 (0.542)
%SV peak	97.8±38.0	121.2±32.8	0.180	96.1±35.0	111.2±20.5	0.284	0.024 (0.540)	0.013 (0.655)
Average SBP (mmHg)	151.1±9.9	151.4±19.1	0.975	149.3±13.6	150.2±15.8	0.901	0.008 (0.725)	0.000 (0.940)
%SBP peak	76.8±8.0	83.6±13.0	0.202	77.0±7.1	81.5±7.9	0.220	0.005 (0.782)	0.008 (0.731)
Average DBP (mmHg)	70.4±6.2	74.5±10.3	0.323	68.9±7.5	74.1±9.3	0.210	0.024 (0.543)	0.007 (0.733)
%DBP peak	78.6±12.2	83.9±16.0	0.441	77.9±15.0	85.6±18.9	0.358	0.001 (0.905)	0.005 (0.789)
Average MAP (mmHg)	102.9±5.5	105.9±12.9	0.523	100.9±4.5	105.0±11.2	0.314	0.039 (0.433)	0.006 (0.760)
%MAP peak	77.2±7.1	82.8±9.4	0.172	76.9±6.0	82.7±10.6	0.172	0.000 (0.944)	0.000 (0.976)
Average EE (kcal/min)	7.3±1.7	7.3±1.7	0.964	7.2±1.5	7.0±0.8	0.778	0.017 (0.608)	0.005 (0.785)

*Note: Data are expressed as mean±SD. HR: heart rate, $\dot{V}O_2$peak: peak oxygen consumption, CO: cardiac output, SV: stroke volume, SBP: systolic blood pressure, DBP: diastolic blood pressure, MAP: mean arterial blood pressure, EE: energy expenditure, η^2: eta-squared effect size analysis. Asterisks denote a significant difference between normal-weight and obese participants (*p<0.05, **p<0.01).*

There were no significant differences in SV index, CI, EF, or SBP between the MICE and MIIE protocols after each stage of the exercise protocols, both in normal-weight and obese participants (Figure 2). However, there were statistically significant differences between the MICE and MIIE protocols in HR at 5-minute intervals ($F_{(1,16)}$=25.064, η^2=0.610, p<0.01), HR at 10-minute intervals ($F_{(1,16)}$=8.292, η^2=0.341, p<0.05), and RPP at 5-minute intervals ($F_{(1,16)}$=8.461, η^2=0.346, p=0.01) (Figure 2). Post-hoc analyses revealed that the MIIE protocol led to significantly lower HR at 5-minute intervals in obese (p<0.01) and normal-weight participants (p<0.01), and at 10-minute intervals in only normal-weight participants (p<0.05) relative to MICE. In contrast, there was no difference between the MICE and MIIE protocols in HR at 15-, 20-, 25-, and 30-minute intervals in obese and normal-weight participants (Figure 2C). It was also discovered that the MIIE protocol led to significantly lower RPP at 5-minute intervals in obese participants (p<0.01) compared to the MIIE protocol but not in normal-weight participants. In contrast, there was no difference between the MICE and MIIE protocols in RPP at 10-, 15-, 20-, 25-, and 30-minute intervals in obese and normal-weight participants (Figure 2E).

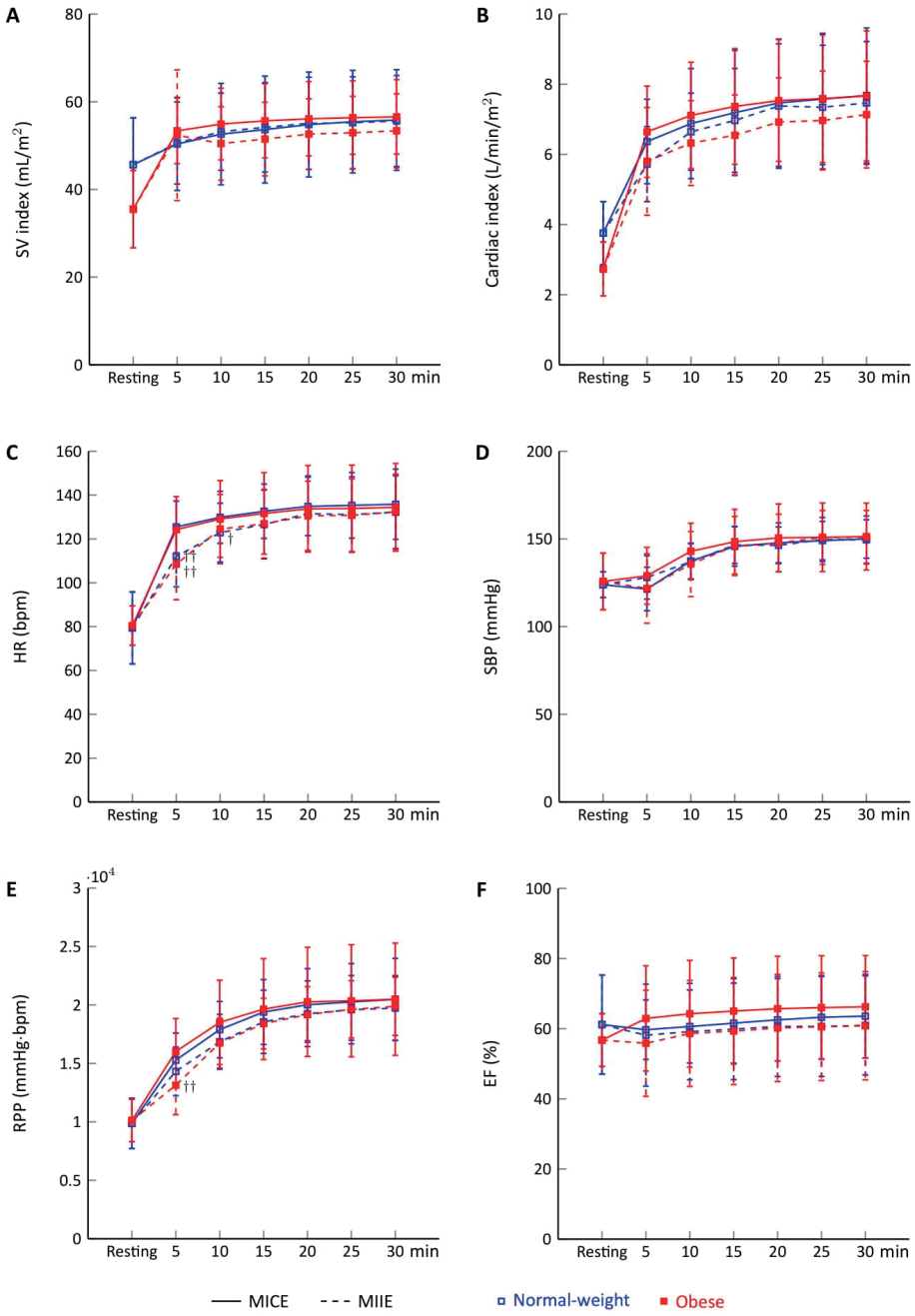


Figure 2. *Stroke volume index (A), Cardiac index (B), heart rate (C), systolic blood pressure (D), rate-pressure product (E), and ejection fraction (F) after each stage of the MICE (––) and MIIE (┈┈) in normal-weight (□) (N=9) and obese participants (■) (N=9). Data are expressed as mean±SD. SV: stroke volume, HR: heart rate, SBP: systolic blood pressure, RPP: rate-pressure product, EF: ejection fraction. Daggers denote a significant difference between MICE and MIIE within the group (†p<0.05, ††p<0.05).*

As for HRV variables, there were no significant differences in RMSSD, LnHF, LnLF, or the LF/HF ratio between the MICE and MIIE protocols during exercise, both in normal-weight and obese participants. During the exercise phase, there were statistically significant differences between the MICE and MIIE protocols in SDNN ($F_{(1,16)}$=7.194, η^2=0.310, p<0.05) and LnVLF ($F_{(1,16)}$=18.568, η^2=0.537, p<0.01) (Figure 3). Post-hoc analyses revealed that the MICE protocol led to a significantly lower SDNN (p<0.01) in obese participants relative to MIIE. In contrast, there was no difference between the MICE and MIIE protocols in normal-weight participants (Figure 3A).

It was also discovered that the MICE protocol led to significantly lower LnVLF in both normal-weight (p<0.05) and obese (p<0.01) participants when compared to the MIIE protocol. In addition, it was discovered that LnVLF (p<0.05) decreased in obese versus normal-weight participants during only the MICE protocol but not the MIIE protocol (Figure 3E). During the recovery periods, there were no significant differences in HRV variables between the MICE and MIIE protocols in both normal-weight and obese participants. However, LnHF (p<0.05) was found to be lower in obese versus normal-weight participants only in the MICE protocol but not in the MIIE protocol (Figure 3C).

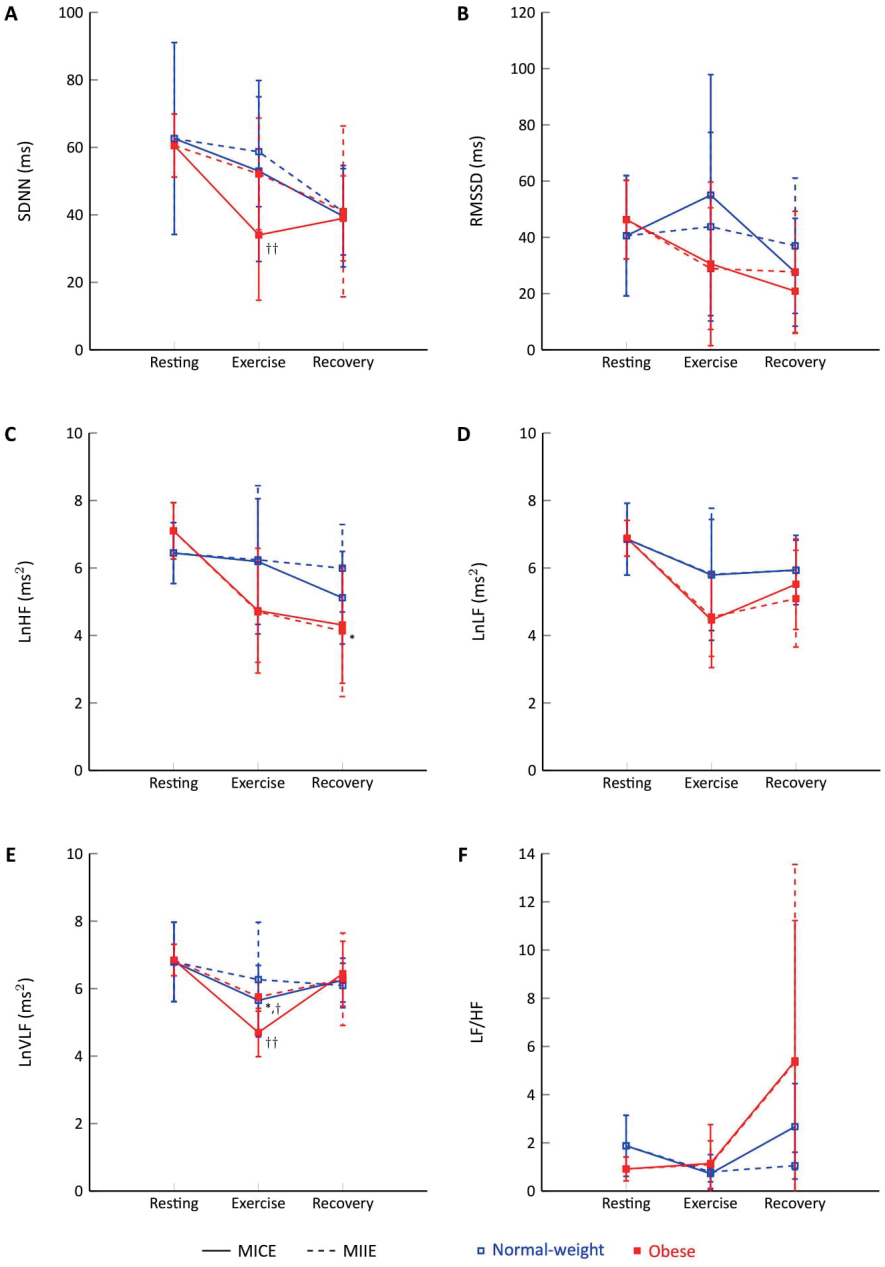


Figure 3. *SDNN (A), RMSSD (B), LnHF (C), LnLF (D), LnVLF (E), and LF/HF (F) at rest, during the exercise [MICE (––) and MIIE (·····)], and recovery phases in normal-weight (□) (N=9) and obese participants (■) (N=9). Data are expressed as mean±SD. Ln: natural logarithm, SDNN: standard deviation of all RR intervals, RMSSD: root mean square of differences of successive RR intervals, HF: high frequency, LF: low frequency, VLF: very low frequency, LF/HF ratio: ratio of absolute LF power to HF power. Asterisks denote a significant difference between normal-weight and obese participants (*p<0.05). Daggers denote a significant difference between MICE and MIIE within the group (†p<0.05, ††p<0.01).*

As for affective responses and enjoyment, there were statistically significant differences between the MICE and MIIE protocols after each stage of the exercises in RPE at 5-minute ($F_{(1,16)}$=16.840, η^2=0.513, p<0.01), 15-minute ($F_{(1,16)}$=10.318, η^2=0.392, p<0.01), and 25-minute intervals ($F_{(1,16)}$=9.590, η^2=0.375, p<001). Post-hoc analyses revealed that the MIIE protocol led to significantly lower RPE at 5-minute (p<0.05) and 25-minute intervals (p<0.05) in obese participants and at 5-minute (p<0.01), 15-minute (p<0.05), and 25-minute intervals (p<0.05) in normal-weight participants relative to MICE (Figure 4A).

After each stage of the exercise protocols, enjoyment scales showed significant differences between the MICE and MIIE protocols at 20-minute ($F_{(1,16)}$=4.585, η^2=0.223, p<0.05), 25-minute ($F_{(1,16)}$=13.444, η^2=0.457, p<0.01), and 30-minute intervals ($F_{(1,16)}$=5.046, η^2=0.240, p<0.05). Enjoyments were significantly higher in MIIE at 20-minute (p<0.05), 25-minute (p<0.01), and 30-minute intervals (p<0.05) in obese participants and at a 25-minute interval (p<0.05) in normal-weight participants relative to MICE (Figure 4B).

It was also discovered that there were statistically significant differences between the MICE and MIIE protocols after each stage of the exercises in affective responses at 15-minute ($F_{(1,16)}$=12.470, η^2=0.438, p<0.01), 20-minute ($F_{(1,16)}$=10.827, η^2=0.404, p<0.01), 25-minute ($F_{(1,16)}$=26.672, η^2=0.625, p<0.01), and 30-minute intervals ($F_{(1,16)}$=17.495, η^2=0.522, p<0.01). Post-hoc analyses revealed that the MIIE protocol led to significantly higher affective responses at 15-minute (p<0.05), 20-minute (p<0.01), 25-minute (p<0.01), and 30-minute intervals (p<0.01) in obese participants and at 15-minute (p<0.05), 25-minute (p<0.05), and 30-minute intervals (p<0.05) in normal-weight participants relative to MICE (Figure 4C).

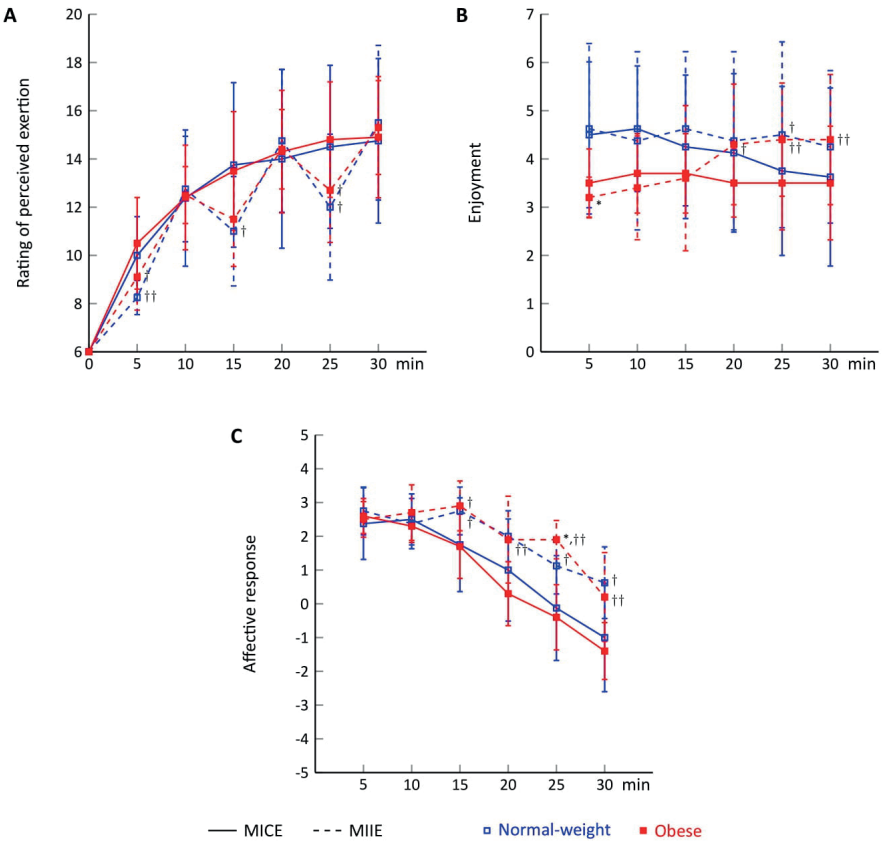


Figure 4. *Rating of perceived exertion (A), enjoyment (B), and affective response (C) after each stage of the MICE (––) and MIIE (┈┈) in normal-weight (□) (N=9) and obese participants (■) (N=9). Data are expressed as mean±SD. Asterisks denote a significant difference between normal-weight and obese participants (*p<0.05). Daggers denote a significant difference between MICE and MIIE within the group (†p<0.05, ††p<0.01).*

Discussion

This study aimed to examine the immediate physiological effects of MIIE and MICE on cardiac function and HRV, as well as the subjective perceptions of enjoyment and affective responses elicited by a single bout of MIIE and MICE in young, healthy, normal-weight, and obese men. Our findings show that males, regardless of weight status (normal or obese), prefer MIIE in terms of enjoyment and experience a smaller decrease in affective response compared to MICE. During interval exercise sessions, participants report lower RPE and exercise-induced fatigue, as indicated by SDNN and LnVLF. It is worth noting that the MIIE and MICE protocols produce comparable cardiac responses. Furthermore, this is the first study to compare MIIE and MICE's duration, average intensity, average $\dot{V}O_2$, and energy expenditure.

A total of 78% of the obese participants remained positive, and 22% of the obese participants evoked

a negative feeling at the end of their work in MICE. In contrast, 100% of the obese participants' affective responses in MIIE remained positive. Additionally, it was observed that MIIE resulted in a higher level of immediate enjoyment following a 20-minute exercise session than MICE. Also, 20 minutes after performing the MICE protocol, the participant's level of enjoyment decreased, whereas there was no such decline in the MIIE protocol. Enjoyment may be a factor associated with psychological responses to exercise. Those who enjoy physical activity may exhibit more positive affective responses than those who enjoy it less.[22] Our study's results support this hypothesis, as they reveal a statistically significant positive correlation (r=0.601, p<0.01) between the level of enjoyment and affective responses observed following the MIIE. Several studies[17,23,24] investigated the relationship between enjoyment and exercise adherence, concluding that enjoyment appears to be a relevant predictor of the intention to continue exercising, exercise habits, and adherence.[25] Thus, implementing MIIE in obese people may lead to higher adherence rates.

The MIIE protocol in this study consisted of three sets of five-minute exercise sessions performed at 90% of the subject's VT, then three sets of five-minute active recovery cycling at 50% of the subject's VT. This approach aligns with the Dual-Mode Model,[13] which states that exercise intensities exceeding the AT tend to elicit negative affective responses. In contrast, intensities below the AT are associated with positive affective responses. A previous study suggested that, despite identical physiological and perceived exertion responses, obesity is associated with diminished affective responses to exercise, even at self-paced intensities.[26] Affect-regulated prescriptions, shorter bouts of self-paced exercise, and distraction away from internal cues are required to directly increase pleasure and/or decrease the perception of effort in this population. The potential for increased enjoyment and a lesser decrease in affective response in MIIE may have implications for promoting exercise adherence among obese individuals. In addition, it has been observed that RPE serves as an indicator of affective responses. A negative correlation between RPE and affective responses has been demonstrated.[27]

This study reveals a lesser increase in RPE, and exercise-induced fatigue is observed following each stage of MIIE in comparison to MICE. The manifestation of exertional responses, such as fatigue, can be observed by the elevation of physiological indicators, such as HR. As a result, in the context of the MICE protocol, it is observed that physiological cues assume a prominent role, particularly within the initial 10-minute period of exercise. This is supported by HRV measurements, particularly SDNN and LnVLF, which indicate a smaller decrease in these markers during interval exercise sessions. According to the study, the findings indicate that the SDNN and the LnVLF exhibited significantly fewer reductions during the MIIE protocol exercise phase than the MICE protocol. This observation suggests an elevation in sympathetic activity within the autonomic nervous system, resulting in an increased HR during exercise-induced fatigue in the MICE protocol.

Both sympathetic and parasympathetic activities influence the SDNN, with a tendency toward sympathetic dominance reducing the SDNN.[28] Moreover, the intrinsic nervous system and sympathetic activity can alter the amplitude and frequency of the heart's oscillations, potentially influencing VLF power.[29] This study revealed that the MICE protocol stimulated more sympathetic activity than the MIIE protocol, particularly in obese subjects. These results are comparable to HRV findings in people who exercise strenuously.[30] Furthermore, the observed reduction in LnHF during the recovery period among obese individuals compared to those with normal weight indicated a decline in vagal activity. As a result, obese participants may require more than rest to restore cardiac autonomic balance. If they continue to exercise, they may have to increase their rest periods.

The study found that obese men (32.5±5.9 kg/m^2) had similar SV index, CI, EF, HR, SBP, and RPP after each stage of MIIE and MICE as normal-weight men (22.2±1.4 kg/m^2) of the same age and gender. Overall, the data indicates that obese participants have a similar cardiac response to MIIE and MICE as normal-weight participants. These findings suggest that the cardiovascular system can adjust to the level of exercise intensity. It is critical to understand the risks and limitations of high-intensity exercise. It may increase the risk of injury and impose more cardiovascular strain on the obese population.[31] Therefore, the MIIE program is likely safe and effective for improving cardiovascular fitness.

Limitation

Despite our findings' significant contribution, this study has some limitations. It is advisable to expand the sample size. It is worth noting, however, that all participants successfully completed all sessions, reducing the possibility of interference with the results. This finding suggests that the MIIE protocols may effectively achieve positive psychological outcomes. However, more research is needed to determine the effectiveness of MIIE in encouraging long-term adherence to exercise routines. In addition, the study included young male adult volunteers who are less likely to develop age-related CVD. The findings' generalizability is limited by a lack of representation from all age groups, including those who are more vulnerable to these conditions. Furthermore, because the study only included male participants, potential gender differences in cardiovascular health were not considered. Therefore, the findings may not apply to women.

Our findings suggest that the MIIE protocol significantly reduces RPE and exercise-induced fatigue while increasing enjoyment and affective response. Furthermore, it is worth noting that the MIIE and MICE protocols produce similar cardiac responses. The MIIE protocol's potential for greater enjoyment and affective response may have significant implications for promoting exercise adherence among inactive, obese people.

Conflict of interest

The authors have no conflict of interest to declare.

Funding

This study received no specific funding from any public, commercial, or non-profit organization.

Acknowledgements

We would like to thank the participants for their enthusiastic participation in this study.

References

[1] Zalesin KC, Franklin BA, Miller WM, Peterson ED, McCullough PA. Impact of obesity on cardiovascular disease. Endocrinol Metab Clin North Am. 2008; 37(3): 663-84. doi: 10.1016/j.ecl.2008.06.004.

[2] Peterson LR, Waggoner AD, Schechtman KB, Meyer T, Gropler RJ, Barzilai B, *et al*. Alterations in left ventricular structure and function in young healthy obese women: assessment by echocardiography and tissue Doppler imaging. J Am Coll Cardiol. 2004; 43(8): 1399-404.

[3] Vella CA, Ontiveros D, Zubia RY. Cardiac function and arteriovenous oxygen difference during exercise in obese adults. Eur J Appl Physiol. 2011; 111: 915-23. doi: 10.1007/s00421-010-1554-z.

[4] Alpert MA, Karthikeyan K, Abdullah O, Ghadban R. Obesity and cardiac remodeling in adults: mechanisms and clinical implications. Prog Cardiovasc Dis. 2018; 61(2): 114-23. doi: 10.1016/j.pcad.2018.07.012.

[5] Ferraro S, Perrone-Filardi P, Desiderio A, Betocchi S, D'Alto M, Liguori L, *et al*. Left ventricular systolic and diastolic function in severe obesity: a radionuclide study. Cardiology. 1996; 87(4): 347-53. doi: 10.1159/000177118.

[6] Kaufman CL, Kaiser DR, Steinberger J, Kelly AS, Dengel DR. Relationships of cardiac autonomic function with metabolic abnormalities in childhood obesity. Obesity. 2007; 15(5): 1164-71. doi: 10.1038/oby.2007.619.

[7] Petridou A, Siopi A, Mougios V. Exercise in the management of obesity. Metabolism. 2019; 92: 163-69. doi: 10.1016/j.metabol.2018.10.009.

[8] Guo Z, Li M, Cai J, Gong W, Liu Y, Liu Z. Effect of High-Intensity Interval Training vs. Moderate-Intensity Continuous Training on Fat Loss and Cardiorespiratory Fitness in the Young and Middle-Aged a Systematic Review and Meta-Analysis. Int J Environ Res Public Health. 2023; 20(6): 4741. doi: 10.3390/ijerph20064741.

[9] Martin-Smith R, Cox A, Buchan DS, Baker JS, Grace F, Sculthorpe N. High intensity interval training (HIIT) improves cardiorespiratory fitness (CRF) in healthy, overweight and obese adolescents: a systematic review and meta-analysis of controlled studies. Int J Environ Res Public Health. 2020; 17(8): 2955. doi: 10.3390/ijerph17082955.

[10] Jung M, Locke S, Bourne J, Beauchamp M, Lee T, Singer J, *et al*. Cardiorespiratory fitness and accelerometer-determined physical activity following one year of free-living high-intensity interval training and moderate-intensity continuous training: a randomized trial. Int J Behav Nutr Phys Act. 2020; 17(1): 1-10. doi: 10.1186/s12966-020-00933-8.

[11] Ekkekakis P, Biddle SH. Extraordinary claims in the literature on high-intensity interval training (HIIT): IV. Is HIIT associated with higher long-term exercise adherence? Psychol Sport Exerc. 2022; 64: 102295. doi: 10.1016/j.psychsport.2022.102295.

[12] Williams DM. Exercise, affect, and adherence: an integrated model and a case for self-paced exercise. J Sport Exerc Psychol. 2008; 30(5): 471-96. doi: 10.1123/jsep.30.5.471.

[13] Ekkekakis P, Parfitt G, Petruzzello SJ. The pleasure and displeasure people feel when they exercise at different intensities: decennial update and progress towards a tripartite rationale for exercise intensity prescription. Sports Med. 2011; 41: 641-71. doi: 10.2165/11590680-000000000-00000.

[14] Hall EE, Ekkekakis P, Petruzzello SJ. The affective beneficence of vigorous exercise revisited. Br J Health Psychol. 2002; 7(1): 47-66. doi: 10.1348/135910702169358.

[15] Coquart J, Lemaire C, Dubart A-E, Luttembacher D-P, Douillard C, Garcin M. Intermittent versus continuous exercise: effects of perceptually lower exercise in obese women. Med Sci Sports Exerc. 2008; 40(8): 1546-53. doi: 10.1249/mss.0b013e31816fc30c.

[16] Biddle SJ, Batterham AM. High-intensity interval exercise training for public health: a big HIT or shall we HIT it on the head? Int J Behav Nutr Phys Act. 2015; 12: 95. doi: 10.1186/s12966-015-0254-9.

[17] Bartlett JD, Close GL, MacLaren DP, Gregson W, Drust B, Morton JP. High-intensity interval running is perceived to be more enjoyable than moderate-intensity continuous exercise: implications for exercise adherence. J Sports Sci. 2011; 29(6): 547-53. doi: 10.1080/02640414.2010.545427.

[18] Borg G. Physical training. 3. Perceived exertion in physical work. Lakartidningen. 1970; 67(40): 4548-57.

[19] Task Force of the European Society of Cardiology the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation, and clinical use. Circulation. 1996; 93(5): 1043-65. doi: 10.1161/01.CIR.93.5.1043.

[20] Hardy CJ, Rejeski WJ. Not what, but how one feels: the measurement of affect during exercise. J Sport Exerc Psychol. 1989; 11(3): 304-17. doi: 10.1123/jsep.11.3.304.

[21] Stanley DM, Cumming J. Are we having fun yet? Testing the effects of imagery use on the affective and enjoyment responses to acute moderate exercise. Psychol Sport Exerc. 2010; 11(6): 582-90. doi:10.1016/j.psychsport.2010.06.010.

[22] Raedeke TD. The relationship between enjoyment and affective responses to exercise. J Appl Sport Psychol. 2007; 19(1): 105-15. doi: 10.1080/10413200601113638.

[23] Martinez N, Kilpatrick MW, Salomon K, Jung ME, Little JP. Affective and enjoyment responses to high-intensity interval training in overweight-to-obese and insufficiently active adults. J Sport Exerc Psychol. 2015; 37(2): 138-49. doi: 10.1123/jsep.2014-0212.

[24] Oliveira BR, Slama FA, Deslandes AC, Furtado ES, Santos TM. Continuous and high-intensity interval training: which promotes higher pleasure? PloS one. 2013; 8(11): e79965. doi: 10.1371/journal.pone.0079965.

[25] Jekauc D. Enjoyment during exercise mediates the effects of an intervention on exercise adherence. Psychology. 2015; 6(1): 48-54. doi: 10.4236/psych.2015.61005.

[26] Elsangedy HM, Nascimento PH, Machado DG, Krinski K, Hardcastle SJ, DaSilva SG. Poorer positive affect in response to self-paced exercise among the obese. Physiol Behav. 2018; 189: 32-39. doi: 10.1016/j.physbeh.2018.02.031.

[27] Ramalho Oliveira BR, Viana BF, Pires FO, Júnior Oliveira M, Santos TM. Prediction of affective responses in aerobic exercise sessions. CNS Neurol Disord Drug Targets. 2015; 14(9): 1214-18. doi: 10.2174/1871527315666151111121924.

[28] Umetani K, Singer DH, McCraty R, Atkinson M. Twenty-four hour time domain heart rate variability and heart rate: relations to age and gender over nine decades. J Am Coll Cardiol. 1998; 31(3): 593-601. doi: 10.1016/S0735-1097(97)00554-8.

[29] Shaffer F, McCraty R, Zerr CL. A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. Front Psychol. 2014; 5: 1040. doi: 10.3389/fpsyg.2014.01040.

[30] Valenzuela PL, Sánchez-Martínez G, Torrontegi E, Vázquez-Carrión J, Montalvo Z, Kara O. Validity, reliability, and sensitivity to exercise-induced fatigue of a customer-friendly device for the measurement of the brain's direct current potential. J Strength Cond Res. 2022; 36(6): 1605-09. doi: 10.1519/JSC.0000000000003695.

[31] Viana RB, Naves JPA, Coswig VS, De Lira CAB, Steele J, Fisher JP, *et al*. Is interval training the magic bullet for fat loss? A systematic review and meta-analysis comparing moderate-intensity continuous training with high-intensity interval training (HIIT). Br J Sports Med. 2019; 53: 655-64. doi: 10.1136/bjsports-2018-099928.

Journal of Associated Medical Sciences 2024; 57(3): 119-130

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

Estimation of organ-absorbed doses in patients undergoing chest and abdominal X-ray examinations using Monte Carlo-based software

Veerapat Thongkum Phunnawut Sermsuk Arocha Teerakeerayut Chayanit Jumpee*

Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 30 March 2024
Accepted as revised 7 July 2024
Available online 11 July 2024

Keywords:
Organ-absorbed doses, entrance surface dose (ESD), kVp, dose area product (DAP), PCXMC, Monte Carlo.

ABSTRACT

Background: Medical imaging is essential for diagnosis and treatment. However, it raises long-term radiation exposure concerns for patient safety. The measurements of entrance surface air kerma (ESAK), entrance surface dose (ESD), and dose area product (DAP) have been widely used to estimate patient doses. However, these parameters cannot represent the patient's organ-absorbed doses.

Objective: To establish a relationship between ESD and organ-absorbed doses, calculated using Monte Carlo-based software to estimate organ-absorbed doses in patients undergoing chest and abdominal X-ray examinations.

Materials and methods: The SHIMADZU RADspeed Pro X-ray machine with VacuDAP meter, Radcal, and AGMS-D+ solid-state detector with Accu-Gold software was used to measure ESD. Tissue-equivalent slab phantoms with 9.8, 15.0, and 20.0 cm thickness were used at 60, 80, and 120 kVp and 3.2 mAs. Surface field sizes varied from 7×7 cm^2 to 28×28 cm^2. Organ-absorbed doses were calculated using PCXMC 2.0 from STUK (Finland) using the same X-ray exposure techniques as the experiment.

Results: The X-ray tube voltage (kVp) is proportional to the ESD (µGy/mAs). This relationship can be described using a power equation. The ESD increased as the phantom thickness increased for each beam field size. A linear function was used to estimate the relationship between ESD and organ-absorbed doses. Due to their anatomical positions, which are most adjacent to the X-ray source, the breast, uterus, and heart have the highest organ-absorbed doses undergoing examinations of the chest (anteroposterior; AP), abdomen (AP), and chest (posteroanterior; PA), respectively.

Conclusion: The relationships between ESD and kVp were established as power functions, but the relationships with DAP were linear functions. The relationship between ESD and organ-absorbed doses was described using linear equations that varied with beam field size and patient thickness. This study provides a helpful method for estimating organ-absorbed and effective patient doses. This is important in assessing the risk of radiation exposure from diagnostic radiography.

Introduction

Medical imaging is significant in diagnosing and treating many health conditions, as it provides vital details about the internal compositions of the human body. X-ray imaging, especially abdominal and chest X-rays, extensively employs imaging modalities that are necessary in the field of clinical practice. Nevertheless, X-rays are ionizing radiation that can interact and cause harm to biological tissues. These are essential factors to consider when assessing the risks caused by exposure to X-rays.[1] As medical technology advances, evaluating and

* **Corresponding contributor.**
Author's Address: *Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand*
E-mail address: chayanit.j@cmu.ac.th
doi: 10.12982/JAMS.2024.054
E-ISSN: 2539-6056

optimizing the radiation dose delivered to patients during procedures becomes increasingly important to ensure accurate diagnosis and patient safety.

Although the X-ray dose used in medical diagnostic radiography is generally low, undergoing multiple X-ray examinations can raise the risk of stochastic effects.[2] Therefore, the concept of diagnostic reference levels (DRLs)[3], which can be defined as an investigation level or dose quantity intended to identify abnormally high radiation doses during radiological examinations, was introduced by the International Commission on Radiological Protection (ICRP) to provide radiologic technologist with an optimization tool.[4]

The quantities used to define dose reference levels (DRLs) vary depending on the modality. These include dose area product (DAP), entrance surface air kerma (ESAK), and entrance surface dose (ESD). Several studies employ DAP quantity to evaluate DRLs.[5,6] Additionally, DAP and ESAK are commonly used to assess ESD through experimental research and mathematical models.[7,8] However, ESD is used to quantify the radiation dose received by patients' surface during X-ray examinations; it does not indicate organ-absorbed dose, which offers a more detailed and organ-specific evaluation of the radiation dose that is absorbed within the body. Several studies have been carried out to estimate ESAK as the absorbed dose in internal organs, utilizing dosimeters or radiation detectors to measure the absorbed dose in a phantom. Thermoluminescent dosimeters (TLD) and optically stimulated luminescent dosimeters (OSLD) were used to measure the absorbed dose in a Rando-Alderson anthropomorphic phantom.[8,9] Furthermore, some research adopted the Monte Carlo simulation program to determine the absorbed dose in internal organs during diagnostic radiography.[10-12] Nevertheless, few studies have established the relationship between ESD (and DAP) and the absorbed dose for patients' internal organs.

This study aimed to establish a relationship between ESD and organ-absorbed doses, which are calculated using Monte Carlo-based software to estimate organ-absorbed doses in patients undergoing chest and abdominal X-ray examinations.

Materials and methods

This study focused on the organ-absorbed doses received by patients undergoing chest and abdominal X-ray examinations. The SHIMADZU diagnostic X-ray machine model RADspeed Pro (Japan), with aluminum filters of 1.0 mm and 0.5 mm for permanent and additional filtration, respectively, was utilized to produce the X-ray beam, and a VacuDAP (Germany) was used to measure the DAP value.

Determination of Entrance surface Dose (ESD)

ESD was determined using a solid-state detector designed for diagnostic X-ray range (Radcal United States, AGMS-D+) with Accu-Gold software. A 25x25 cm^2 tissue-equivalent slab phantom was placed on the patient's couch. The thickness of the phantom varied at 9.8, 15.0, and 20.0 cm, representing the ranges of patient thickness of newborn, 5 years, and adult, respectively. The solid-state detector was placed on the surface at the center of a tissue-equivalent slab phantom to measure the ESD from X-ray tube voltages of 60, 80, and 120 kVp using a constant tube current (mA) and exposure time (s) of 3.2 mAs. On the slab phantom's surface, the X-ray field sizes were adjusted to 7×7 cm^2, 14×14 cm^2, 21×21 cm^2, and 28×28 cm^2. The focus-to-image distance (FID) is 100 cm, representing chest anteroposterior (AP) and abdomen (AP) X-ray, and 180 cm (focus-to-image at the standing bucky distance) representing chest posteroanterior (PA) X-ray.

The focus-to-DAP distance (Td) is 31 cm. The geometry of the experiment setup is illustrated in Figure 1. The slab phantom object thickness is denoted as Tp, whereas the bucky thickness is 7.5 cm and represented as Tb. Three ESD measurements were conducted to determine the mean ESD value in units of µGy, incorporating measurement uncertainties. The relationship between ESD and DAP (µGy·m^2) was measured under the same exposure conditions.

The radiation dose is affected directly by the parameter of tube current-time (mAs). To properly evaluate the relationship between the organ-absorbed dose and ESD, the ESD was calculated in units of µGy/mAs.

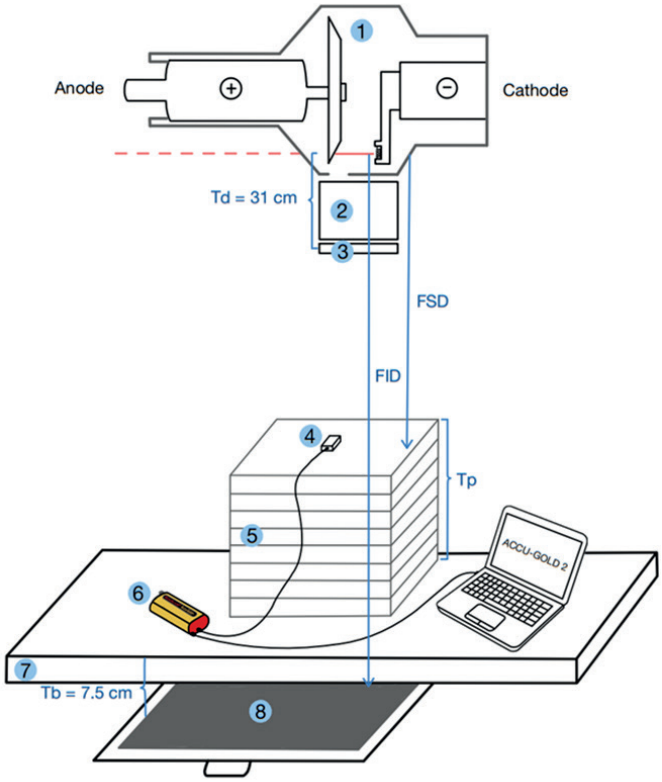


Figure 1. *Measurement geometry for determination of entrance surface dose (ESD).*
1: X-ray tube, 2: collimator, 3: DAP meter,
4: solid state detector, 5: slab phantom,
6: Accu-Gold digitizer, 7: couch, 8: image detector.

Estimation of organ-absorbed dose

PCXMC version 2.0 Monte Carlo-based software was used to simulate radiation transportation in organ-absorbed doses of patients undergoing chest (AP), chest (PA), and abdomen (AP) radiography examinations. PCXMC, a commercial program developed by the Laboratory for Medical at the Finnish Centre for Radiation and Nuclear Safety (STUK, Finland) to calculate X-ray spectra and the attenuation of X-rays in various materials, is generally applied for estimating patient doses in medical imaging procedures including radiography and computed tomography (CT).[13] This study's simulation utilized a newborn, 5-year-old, and adult phantom, which was shown in the specification details in Table 1. The examinations were simulated with tube voltages of 50, 60, 80, 120, and 150 kVp, using a tube current-time of 1 mAs. A focus-to-image distance (FID) of 100 cm was used for examinations of chest (AP) and abdomen (AP) X-ray, and a 180 cm FID was used for chest (PA) X-ray. The X-ray beam field sizes varied throughout 7×7 cm^2, 14×14 cm^2, 21×21 cm^2, and 28×28 cm^2 at the surface of the phantom. The focus-to-surface distance (FSD) was adjusted to 82.7 cm, 77.5 cm, 72.5 cm, and 152.5 cm to correspond to the phantom thicknesses of 9.8 cm, 15.0 cm, 20.0 cm, and 20.0 cm for the chest (PA), respectively. The other simulation parameters were set as follows: maximum energy of 150 keV, 1,000,000 photon histories (quantity of particles that get input into the Monte Carlo simulation program), 1.5 mm aluminum filter, and X-ray tube anode angle of 12°. Figure 2-3 shows the interface window of PCXMC version 2.0 simulation program.

The PCXMC simulation software computed both the organ-absorbed doses and the effective doses. Calculations were performed for various examinations, image field sizes, phantom thicknesses, and X-ray tube voltages to determine organ-absorbed doses (μGy/mAs) and effective dose (μSv/mAs), which utilizing the organ weighting factors from ICRP Publication 103.[14]

Table 1. The specifications of the mathematical phantom separated by ages.[15]

Age	Weight (kg)	Total height (cm)	Trunk thickness (cm)	Trunk width including arms (cm)	Leg length (cm)
Newborn	3.4	50.9	9.8	12.7	16.8
5 years	19.0	109.1	15.0	22.9	48.0
Adult	73.2	178.6	20.0	40.0	80.0

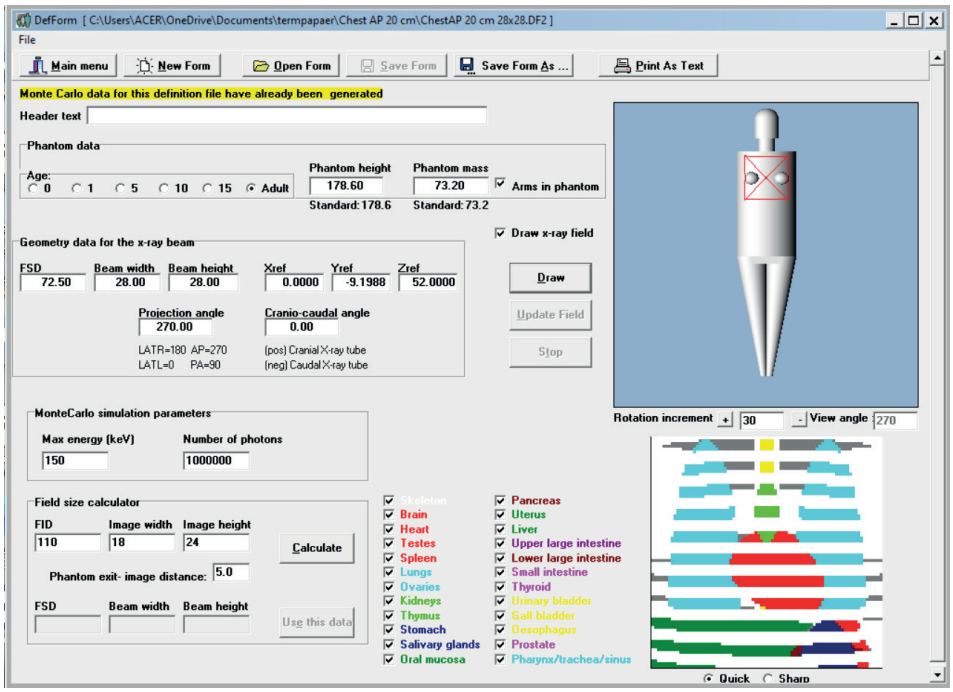


Figure 2. *Interface of program for patients undergoing chest (AP) X-ray with field size 28×28 cm^2 (phantom thickness is 20.0 cm).*

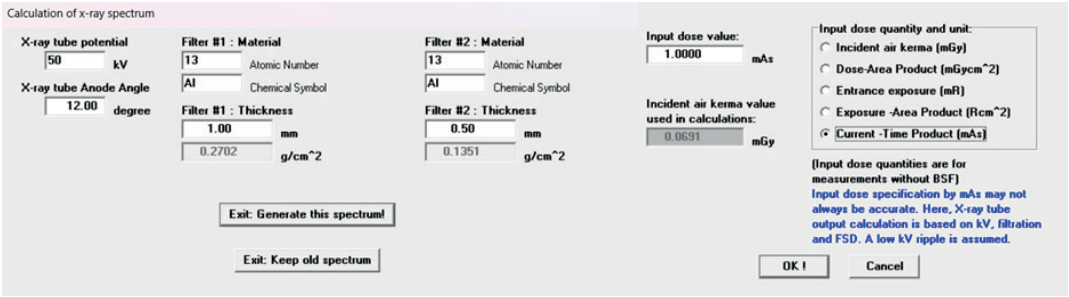


***Figure 3**. Parameters related to the simulation at 50 kVp, 1mAs.*

Results and discussion

ESD VS kVp

The ESD (µGy/mAs) is directly proportional to the X-ray tube voltages (kVp), as exhibited in Figure 4. for the X-ray beam field size of 28×28 cm^2. According to our preliminary study, a power equation related to the X-ray output (mR) proportional to kVp can mathematically describe this relationship.[2,3] The ESD increased as the phantom thickness increased for each beam field size. Due to the increasing thickness of the phantom, the focus to surface distance (FSD) decreases. Therefore, the surface of the phantom can be exposed to a higher X-ray intensity. The phantom has a thickness of 20.0 cm (PA), with an FID of 180 cm. The FID of 180 cm is the longest, which means that the surface of the phantom is exposed to the lowest intensity of X-rays. The equations that exhibit the relationships between ESD (y values) and specific kVp (x values) for different field sizes and phantom thicknesses are provided in Table 2. For an individual kVp and beam field size, ESD can be estimated using a power equation relationship evaluated for a particular thickness.

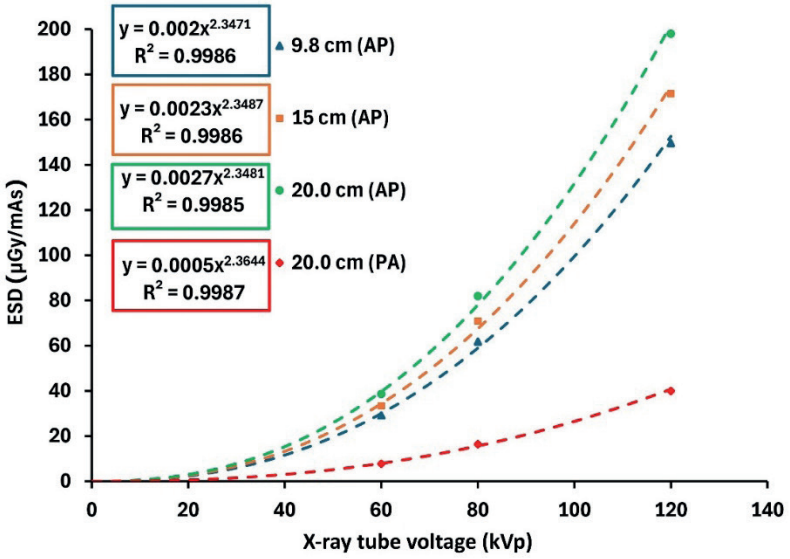


***Figure 4**. Relationship between entrance surface dose (ESD) and X-ray tube voltages (kVp) for different slab phantom thicknesses and 28x28 cm^2 beam field size.*
FID: 100 cm for thickness of 9.8 cm (AP), 15.0 cm (AP), and 20.0 cm (AP), FID: 180 cm for thickness of 20.0 cm (PA).

Table 2. Relationship equations of entrance surface dose (y) and kVp (x) for different slab phantom thicknesses and beam field sizes.

Thickness	Field sizes			
	7×7 cm²	14×14 cm²	21×21 cm²	28×28 cm²
9.8 cm (AP)	$y = 0.0018x^{2.3572}$	$y = 0.0019x^{2.3505}$	$y = 0.0020x^{2.3470}$	$y = 0.0020x^{2.3471}$
15 cm (AP)	$y = 0.0021x^{2.3566}$	$y = 0.0022x^{2.3556}$	$y = 0.0022x^{2.3517}$	$y = 0.0023x^{2.3487}$
20 cm (AP)	$y = 0.0025x^{2.3531}$	$y = 0.0018x^{2.3572}$	$y = 0.0005x^{2.3628}$	$y = 0.0027x^{2.3481}$
20 cm (PA)	$y = 0.0018x^{2.3572}$	$y = 0.0005x^{2.3628}$	$y = 0.0005x^{2.3659}$	$y = 0.0005x^{2.3644}$

***Note*:** *FID: 100 cm for thickness of 9.8 cm (AP), 15.0 cm (AP), and 20.0 cm (AP), FID: 180 cm for thickness of 20.0 cm (PA).*

ESD VS DAP

Figure 5 illustrates that the ESD is proportional to the DAP (μGy·m²). Their relationships are described by employing of a linear equation. ESD can be estimated on a DAP by utilizing the linear relationship. The results indicate a positive linear correlation between ESD and DAP. Table 3 shows the equations that represent the relationships between ESD (y values) and specific values of DAP (x values). When comparing different phantom thicknesses, it was observed that DAP increased with phantom thickness. This was because the beam field size maintained constant at the phantom surface, but as phantom thickness increased, the beam field size at DAP expanded, resulting in elevated DAP values.

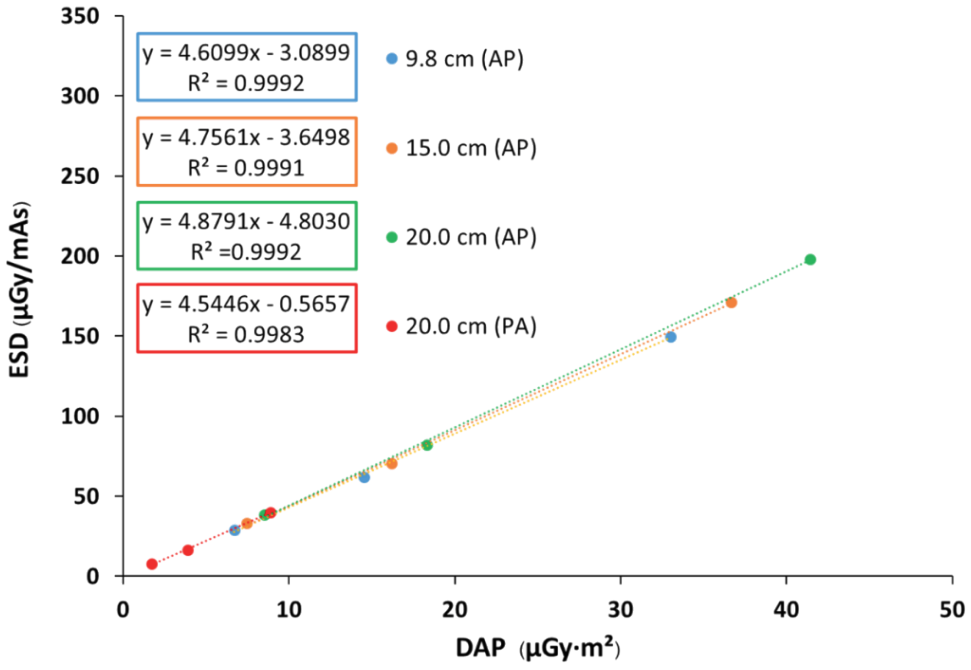


Figure 5. *Relationship between entrance surface dose (ESD) and DAP for 9.8 cm (AP), 15.0 cm (AP), 20.0 cm (AP), and 20.0 cm (PA) slab phantom thicknesses and 28×28 cm² beam field size. FID: 100 cm for a thickness of 20.0 cm (AP), FID: 180 cm for a thickness of 20.0 cm (PA).*

Table 3. Relationship equation of entrance surface dose (y) and DAP (x) for different slab phantom thicknesses and beam field sizes.

Thickness	Field sizes			
	7×7 cm²	14×14 cm²	21×21 cm²	28×28 cm²
9.8 cm (AP)	y=73.0580x+51.242	y=18.2610x-2.3775	y=8.0003x-2.9564	y=4.6099x-3.0899
15.0 cm (AP)	y=72.7710x+1.8506	y=17.7930x-2.6864	y=8.0420x-3.3217	y=4.7561x-3.6498
20.0 cm (AP)	y=75.2530x+0.2596	y=18.0740x-3.5788	y=8.0068x-3.6076	y=4.8797x-4.8030
20.0 cm (PA)	y=77.0770x+7.6973	y=18.7320x+0.4483	y=8.0830x-0.4363	y=4.5446x-0.5657

Note: *FID: 100 cm for thicknesses of 9.8 cm (AP), 15.0 cm (AP), and 20.0 cm (AP), FID: 180 cm for thickness of 20.0 cm (PA).*

Organ-absorbed doses VS ESD

The experimental parameters, phantom sizes, weights (with thicknesses of 9.8 cm, 15.0 cm, and 20.0 cm representing phantoms of newborns, 5 years, and adults, respectively), and PCXMC 2.0 simulation software were utilized to execute the simulation. The absorbed dose was reported in µGy for selected vital organs, including active bone marrow, breasts, colon, heart, ovaries, stomach, testicles, thyroid, and uterus. The effective dose was reported in µSv. The relationship between ESD and organ-absorbed doses (left-hand side axis) and effective dose (right-hand side axis) are illustrated in Figure 6. The organ-absorbed doses can be estimated by applying the relationship equation of ESD and kVp (or DAP) described in the section on ESD determination (Table 2 and 3).

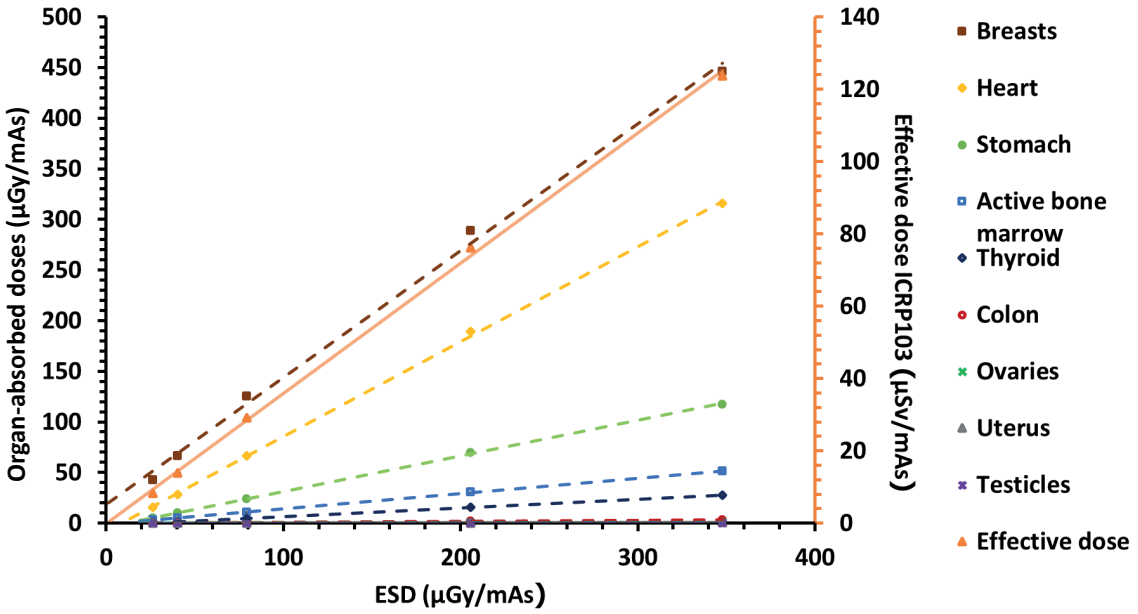


Figure 6*. Relationship between ESD and selected organ-absorbed doses (left-hand side axis) and effective dose. Right-hand side axis) for patient undergoing chest (AP) (with field size 28×28 cm², phantom thickness is 20.0 cm.*

This relationship was estimated using a linear function. In this study, the ESD has been calculated in units of µGy/mAs. Therefore, only the X-ray tube voltage (kVp) is impacted by ESD. High kVp can generate high-energy X-rays that can penetrate the patient's internal organs and transfer X-ray energy to gain the absorbed dose. The equations representing the relationships between organ-absorbed doses (or effective dose) (y values) and ESD (x values) are shown in Table 4-10. Table 4 does not provide a relationship equation for the testicles because this organ is located externally at the beam field size.

Table 4. Relationship equation of organ-absorbed doses (y) and ESD (x) for selected organs in adult phantom (20.0 cm thickness) for chest (AP) X-ray examination.

Organ	Field size			
	7×7 cm²	14×14 cm²	21×21 cm²	28×28 cm²
Active bone marrow	y=0.0127x - 0.1078	y=0.0437x - 0.2937	y=0.0921x - 0.5963	y=0.1516x - 1.0373
Breasts	y=0.0174x - 0.2877	y=0.1424x + 0.7875	y=0.7165x + 10.3610	y=1.2527x + 18.7390
Colon	y=0.0003x - 0.0109	y=0.0013x - 0.0486	y=0.0042x - 0.1472	y=0.0107x - 0.3651
Heart	y=0.2539x - 1.3770	y=0.7224x - 4.2541	y=0.8991x - 6.7143	y=0.9387x - 8.2910
Ovaries	y=0.00008x - 0.0035	y=0.0003x - 0.0135	y=0.0010x - 0.0426	y=0.0028x - 0.1183
Stomach	y=0.0048x - 0.1355	y=0.0257x - 0.6337	y=0.1288x - 1.9321	y=0.3517x - 4.0573
Testicles	NA	NA	NA	NA
Thyroid	y=0.0022x - 0.0720	y=0.0113x - 0.3347	y=0.0335x - 0.9114	y=0.0860x - 2.0396
Uterus	y=0.0001x - 0.0030	y=0.0003x - 0.0129	y=0.0010x - 0.0400	y=0.0026x - 0.0987
Effective dose ICRP103 (µSv)	y=0.0166x - 0.2091	y=0.0755x - 0.5537	y=0.2219x - 0.1841	y=0.3602x - 0.1499

Table 5. Relationship equation of organ-absorbed doses (y) and ESD (x) for selected organs in 5 years phantom (15.0 cm thickness) for chest (AP) X-ray examination.

Organ	Field size			
	7×7 cm²	14×14 cm²	21×21 cm²	28×28 cm²
Active bone marrow	y=0.0225x - 0.0840	y=0.0801x - 0.2134	y=0.1927x - 0.8973	y=0.2183x - 1.3594
Breasts	y=0.0618x - 0.2656	y=1.3731x + 41.4280	y=1.4612x + 39.2500	y=1.4395x + 41.531
Colon	y=0.0041x - 0.0948	y=0.0209x - 0.4418	y=0.0724x - 1.1834	y=0.3022x - 2.3507
Heart	y=0.7286x + 1.3595	y=1.0074x + 0.6709	y=1.1385x - 0.9243	y=1.1188x - 1.4844
Ovaries	y=0.0019x - 0.0552	y=0.0088x - 0.2071	y=0.0295x - 0.7006	y=0.0779x - 1.2997
Stomach	y=0.0356x - 0.4320	y=0.3816x - 0.7400	y=0.9222x + 0.4417	y=1.1243x + 0.1990
Testicles	NA	NA	NA	y=0.0029x - 0.1213
Thyroid	y=0.0103x - 0.1888	y=0.0499x - 0.8047	y=0.1856x - 2.1096	y=1.1882x + 16.881
Uterus	y=0.0018x - 0.0545	y=0.0098x - 0.2559	y=0.0276x - 0.7035	y=0.0670x - 1.4193
Effective dose ICRP103 (µSv)	y=0.0493x - 0.3223	y=0.3721x + 4.4672	y=0.5499x + 3.7291	y=0.6617x + 4.3785

Table 6. Relationship equation of organ-absorbed doses (y) and ESD (x) for selected organs in newborn phantom (9.8 cm thickness) for chest (AP) X-ray examination.

Organ	Field size			
	7×7 cm²	14×14 cm²	21×21 cm²	28×28 cm²
Active bone marrow	y=0.0832x + 0.1486	y=0.2699x + 0.5344	y=0.3840x + 0.5655	y=0.5703x + 1.2089
Breasts	y=1.4188x + 43.6330	y=1.5754x + 48.1180	y=1.5661x + 44.6680	y=1.4023x + 43.1290
Colon	y=0.0143x - 0.1620	y=0.0841x - 0.5643	y=0.5552x + 1.0146	y=0.9082x + 1.9839
Heart	y=1.0801x + 9.3221	y=1.2544x + 9.1571	y=1.2150x + 8.9979	y=1.2510x + 8.2577
Ovaries	y=0.0090x - 0.1430	y=0.0431x - 0.7544	y=0.1965x - 0.6335	y=0.8937x - 3.5386
Stomach	y=0.0739x - 0.1636	y=0.9502x + 6.9529	y=1.2178x + 8.8935	y=1.2939x + 9.0580
Testicles	y=0.0016x - 0.0188	y=0.0086x - 0.1816	y=0.0231x - 0.3262	y=1.2939x + 9.0580
Thyroid	y=0.1040x - 0.1276	y=1.4539x + 26.918	y=1.4648x + 26.71	y=1.4998x + 27.237
Uterus	y=0.0086x - 0.1261	y=0.0451x - 0.5223	y=0.1627x - 1.102	y=0.9435x - 0.4140
Effective dose ICRP103 (µSv)	y=0.3293x + 5.9394	y=0.6779x + 8.8983	y=0.8331x + 9.314	y=0.9742x + 9.2879

Table 7. Relationship equation of organ-absorbed doses (y) and ESD (x) for selected organs in adult phantom (20.0 cm thickness) for abdomen (AP) X-ray examination.

Organ	Field size			
	7×7 cm²	14×14 cm²	21×21 cm²	28×28 cm²
Active bone marrow	y=0.9742x + 9.2879	y=0.0307x - 0.9315	y=0.0756x - 2.0819	y=0.1150x - 3.0842
Breasts	y=0.0003x - 0.0064	y=0.0012x - 0.0288	y=0.0035x - 0.0816	y=0.0078x - 0.1807
Colon	y=0.0533x - 0.8515	y=0.2736x - 3.7100	y=0.5705x - 7.2580	y=0.6660x - 9.3648
Heart	y=0.0006x - 0.0207	y=0.0028x - 0.1012	y=0.0091x - 0.3053	y=0.0221x - 0.7020
Ovaries	y=0.0401x - 0.8932	y=0.3329x - 6.1153	y=0.5024x - 9.3793	y=0.5654x - 11.317
Stomach	y=0.0091x - 0.2261	y=0.0603x - 1.0671	y=0.2976x - 2.5364	y=0.5802x - 4.0538
Testicles	y=0.001x - 0.0332	y=0.0048x - 0.1600	y=0.0145x - 0.4352	y=0.0355x - 0.9801
Thyroid	NA	NA	NA	NA
Uterus	y=0.0631x - 1.1235	y=0.4140x - 6.1056	y=0.6405x - 9.5598	y=0.6986x - 11.671
Effective dose ICRP103 (µSv)	y=0.0141x - 0.2642	y=0.0726x - 1.1980	y=0.1756x - 2.3927	y=0.2652x - 3.3937

Table 8. Relationship equation of organ-absorbed doses (y) and ESD (x) for selected organs in 5 years phantom (15.0 cm thickness) for abdomen (AP) X-ray examination.

Organ	Field size			
	7×7 cm²	14×14 cm²	21×21 cm²	28×28 cm²
Active bone marrow	y=0.0209x - 0.4300	y=0.0837x - 1.4240	y=0.1841x - 2.3344	y=0.2382x - 2.8403
Breasts	y=0.0031x - 0.0736	y=0.0141x - 0.2461	y=0.0468x - 0.6227	y=0.1204x - 0.8938
Colon	y=0.2059x - 1.0328	y=0.6405x - 2.9750	y=0.8377x - 5.0506	y=0.9025x - 6.1625
Heart	y=0.0105x - 0.2076	y=0.0531x - 0.9012	y=0.2028x - 1.8657	y=0.7065x - 1.0829
Ovaries	y=0.0700x - 0.8931	y=0.5473x - 5.9215	y=0.7354x - 8.3526	y=0.7639x - 8.9553
Stomach	y=0.1304x - 0.4514	y=0.7974x + 1.8550	y=1.1476x + 1.0064	y=1.1609x + 0.6064
Testicles	y=0.0044x - 0.1066	y=0.0232x - 0.4684	y=0.0751x - 1.0712	y=0.2193x - 1.6206
Thyroid	y=0.0005x - 0.0160	y=0.0022x - 0.078	y=0.0095x - 0.2115	y=0.0169x - 0.4161
Uterus	y=0.0773x - 0.9108	y=0.6277x - 4.8341	y=0.8662x - 6.8445	y=0.9042x - 7.7590
Effective dose ICRP103 (µSv)	y=0.0650x - 0.4082	y=0.6277x - 4.8341	y=0.4520x - 1.9830	y=0.5449x - 2.3575

Table 9. Relationship equation of organ-absorbed doses (y) and ESD (x) for selected organs in Newborn phantom (9.8 cm thickness) for abdomen (AP) X-ray examination.

Organ	Field size			
	7×7 cm²	14×14 cm²	21×21 cm²	28×28 cm²
Active bone marrow	y=0.0623x - 0.3795	y=0.2619x - 0.1712	y=0.3641x - 0.1480	y y=0.4542x - 0.0586
Breasts	y=0.0301x - 0.1935	y=0.8078x + 21.5260	y=1.4191x + 48.3370	y=1.4278x + 42.3030
Colon	y=0.6079x + 2.1440	y=0.9864x + 2.2431	y=1.0603x + 1.8827	y=1.0954x + 1.4529
Heart	y=0.0795x - 0.2900	y=0.8587x + 6.2748	y=1.1769x + 8.7558	y=1.2361x + 8.6967
Ovaries	y=0.1872x - 0.4961	y=0.9143x - 0.7483	y=0.9544x - 3.9056	y=0.9006x - 2.2987
Stomach	y=1.0085x + 8.4574	y=1.3005x + 9.4050	y=1.2595x + 8.9823	y=1.3072x + 8.6914
Testicles	y=0.0209x - 0.2373	y=0.1102x - 0.5650	y=1.4137x + 37.8820	y=1.5260x + 34.0606
Thyroid	y=0.0074x - 0.1157	y=0.0286x - 0.3435	y=0.1024x - 0.4878	y=1.2930x + 22.4440
Uterus	y=0.1763x - 0.3997	y=1.0160x + 0.5191	y=1.0340x - 0.0747	y=1.0729x - 0.2466
Effective dose ICRP103 (µSv)	y=0.3024x + 1.3068	y=0.6980x + 4.9555	y=0.9244x + 10.114	y=1.0356x + 10.2530

Table 10. Relationship equation of organ-absorbed doses (y) and ESD (x) for selected organs in adult phantom (20.0 cm thickness) for chest (PA) X-ray examination.

Organ	Field size			
	7×7 cm²	14×14 cm²	21×21 cm²	28×28 cm²
Active bone marrow	y=0.0379x - 0.0930	y=0.1016x - 0.2131	y=0.1838x - 0.3324	y=0.3037x - 0.4838
Breasts	y=0.0063x - 0.0405	y=0.0471x - 0.2474	y=0.1895x - 0.8357	y=0.2632x - 1.2043
Colon	y=0.0002x - 0.0020	y=0.0013x - 0.0100	y=0.0037x - 0.0284	y=0.0098x - 0.0696
Heart	y=0.0719x - 0.3859	y=0.2478x - 1.1931	y=0.3481x - 1.6514	y=0.4265x - 2.0480
Ovaries	NA	NA	NA	y=0.0018x - 0.0165
Stomach	y=0.0036x - 0.0241	y=0.0198x-0.1166	y=0.0675x - 0.3606	y=0.1634x - 0.8303
Testicles	NA	NA	NA	NA
Thyroid	y=0.0016x - 0.0118	y=0.0087x - 0.0588	y=0.0262x - 0.1695	y=0.063x - 0.3931
Uterus	NA	y=0.0003x - 0.0030	y=0.0009x - 0.0071	y=0.0023x - 0.0199
Effective dose ICRP103 (µSv)	y=0.0151x - 0.0632	y=0.0659x - 0.2097	y=0.1713x - 0.4274	y=0.2726x - 0.7183

Therefore, the testicles receive a small amount of absorbed dose, making them unsuitable for determining an accurate fitting curve equation. This is also why the relationship equation for the thyroid in Table 7 and others cannot be established. The result calculated from the specific equation from Table 4-10 can be exhibited that the maximum organ-absorbed dose is in the breasts, uterus, and heart for examination of the chest (AP), abdomen (AP), and chest (PA) X-rays, respectively.

The absorbed dose of selected organs inside the beam field size increased proportionally to the increase of ESD, as shown in Figures 7-9. In Figure 7, when evaluating the organ-absorbed doses for various beam field sizes, it can be observed that the organ-absorbed doses are increasing correspondingly with the increase in beam field sizes.

During a 21×21 cm^2 field size chest (AP) examination, the maximum absorbed dose is in the heart. However, the maximum absorbed dose is in the breasts when considering the field size of 28×28 cm^2, as this field size can cover the entire breasts more effectively than a smaller field size. A selected organ-absorbed dose during an abdomen (AP) examination is shown in Figure 8. The absorbed dose to the stomach is dramatically higher when the field size increases from 7×7 cm^2 to 28×28 cm^2, as this larger field size may involve a greater stomach area. Figures 7 and 9 demonstrate that the absorbed doses in the uterus cannot be applied when using a chest X-ray examination field size of 7×7 cm^2. This is because this radiation field size is insufficient to cover the uterus in the lower abdomen.

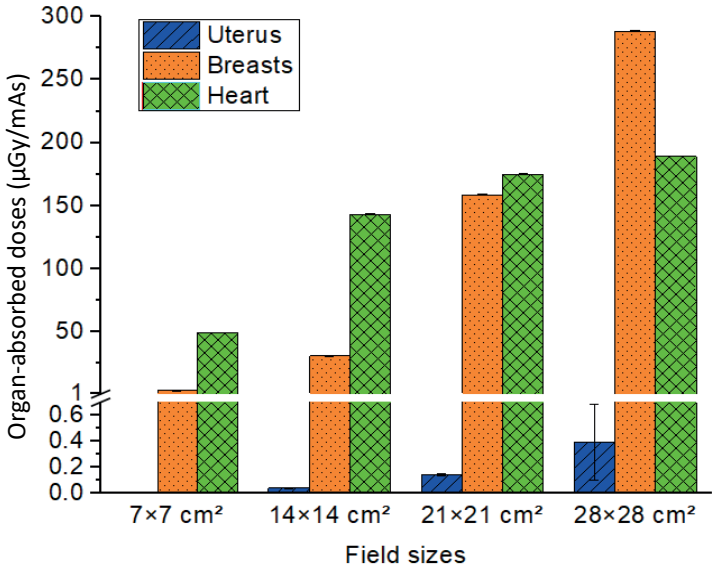


***Figure 7**. Comparison of selected organ-absorbed doses in field sizes 7×7 cm^2, 14×14 cm^2, 21×21 cm^2 and 28×28 cm^2 for patients undergoing chest (AP) (with phantom thickness is 20.0 cm).*

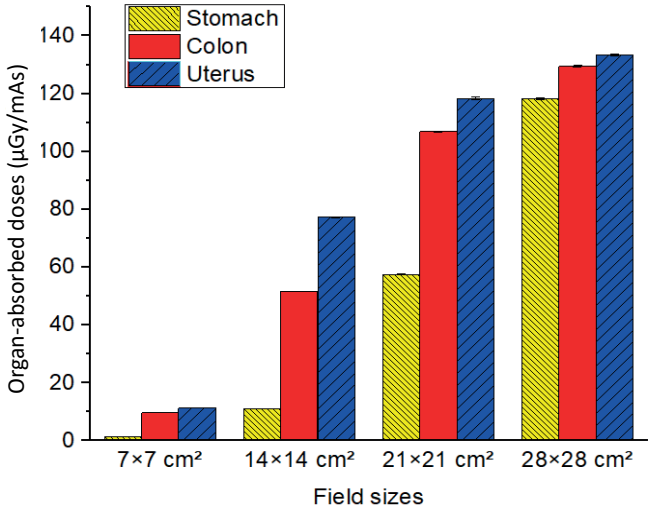


***Figure 8**. Comparison of selected organ-absorbed doses in field sizes 7×7 cm^2, 14×14 cm^2, 21×21 cm^2 and 28×28 cm^2 for patients undergoing abdomen (AP). (with phantom thickness is 20.0 cm).*

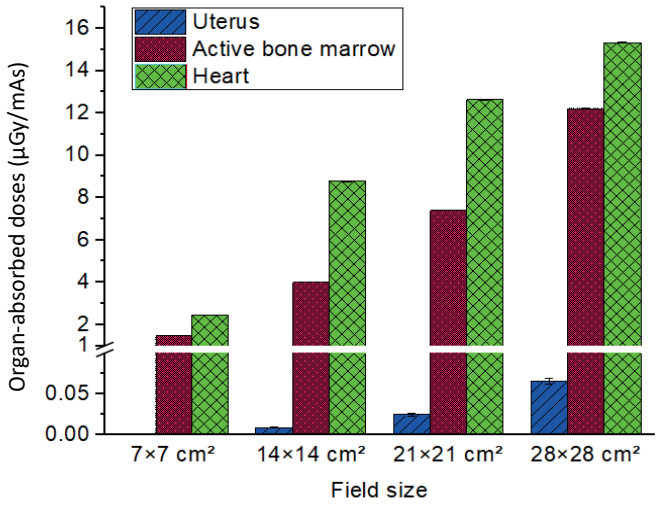


***Figure 9**. Comparison of selected organ-absorbed doses in field sizes 7×7 cm², 14×14 cm², 21×21 cm² and 28×28 cm² for patients undergoing chest (PA). (with phantom thickness is 20.0 cm).*

In Figure 10, when compared to the study carried out by Kawaura *et al*.[16] in Japan, which used photodiode dosemeters to measure organ doses (mGy) for chest (PA) examination, indicates that the present study resulted in a decrease of 18.00% and 7.00% in the active bone marrow and breasts doses, respectively. The observed variations can be related to differences in the methodology used in the two research studies. Kawaura *et al*. used an anthropomorphic phantom of a standard Japanese adult male that is 170.0 cm in height and weight 60.0 kg. As a comparison, the present study used a phantom in the PCXMC program that is 178.6 cm in height and 73.2 kg in weight. The lower organ-absorbed doses can be attributed to the variations in the thicknesses of the phantom, which is less than 20.0 cm, according to Kawamura et al. Therefore, absorbed doses in organs were determined to be higher because of the lower thickness of the phantom, resulting in a lower radiation attenuation. In addition, abdomen (AP) examination, the absorbed doses for the colon and stomach showed respective increases of 70% and 18%. Compared to Kawaura et al.'s study, the results of the present study indicate a twice increase in the uterus. The difference could result from the variation in the method of measuring doses. Kawaura *et al*. measured the absorbed doses in organs at certain points, while the PCXMC program quantified them in volume.

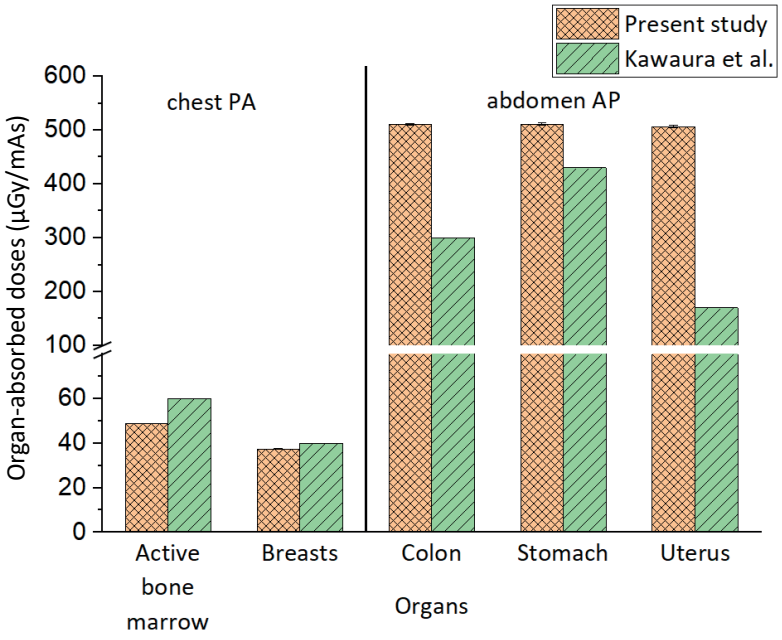


***Figure 10**. Comparison of selected organ-absorbed doses of present study with Kawaura et al.'s study for adult patients undergoing chest (PA) and abdomen (AP).*

A comparison with Kumaresan *et al*.[17] in India and Ma *et al*.[18] in Canada is shown in Figure 11. These two studies focused on determining the absorbed doses in pediatric patients undergoing chest (AP) and chest-abdomen (AP) examinations, respectively. In Kumaresan *et al*.'s study, MCNP 3.1 was used, combined with a phantom that was 15.0 kg in weight and 105 cm in height, representing 5 years Indian. The study found that the absorbed dose in the heart is lower (49.54%), whereas the dose in the thyroid is like that in the present study. In the study by Ma *et al*., they utilized TLD with a phantom of 51 cm in height and 3.5 kg in weight related to the newborn phantom. In the present study, the absorbed doses in the uterus and colon are lower, with a decrease of 31.03% and 45.61%, respectively. However, the absorbed dose in the breasts is greater, with an increase of 56.31%. The differences in absorbed doses found in these studies can be attributed to differences in the phantom type, beam field sizes, and method of dose assessments. Furthermore, various studies indicate a statistically significant increase in radiation DAP and effective dose in radiography examinations for overweight and obese patients.[19-21] It is necessary to investigate more studies on the relationship between body mass index (BMI) and internal organ-absorbed dose.

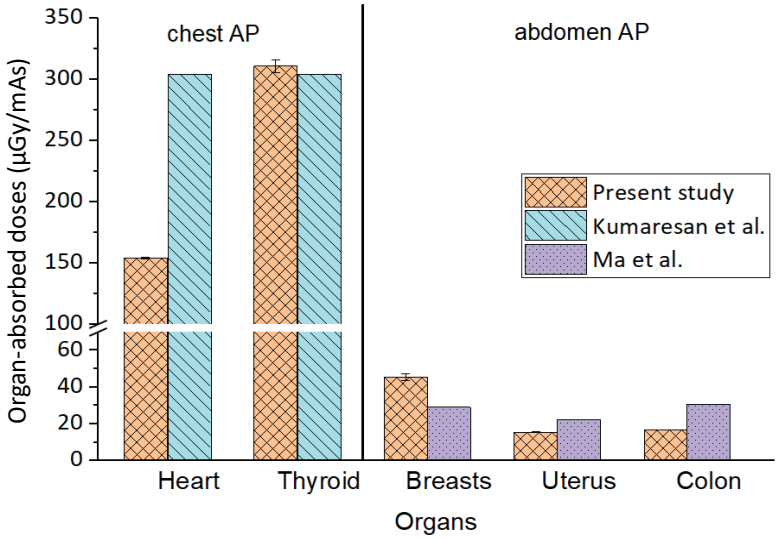


Figure 11. *Comparison of selected organ-absorbed doses of present study with Kumaresan et al.'s study and Ma et al.'s study for pediatric patients undergoing chest (AP) (5 years) and abdomen (AP) (newborn), respectively.*

Conclusion

This study investigated the absorbed and effective organ doses during chest and abdominal radiography examinations. Various parameters, such as tube voltages, beam field sizes, and patient thicknesses, were investigated using PCXMC, a Monte Carlo-based software. The relationship equations between ESD and X-ray tube voltage can be established as a power function. In contrast, the relationship equations between ESD and DAP can be established as a linear function. In addition, the absorbed doses in organs can be estimated using linear equations correlated to the ESD. The estimated organ-absorbed doses reveal variations in beam field size and patient thickness. The study would provide a practical alternative method for determining the organ-absorbed dose, including effective dose, which is important in determining the appropriate technical parameters in a diagnostic radiograph that evaluates radiation risk.

Conflict of interest

The author(s) declared no conflicts of interest regarding this article's research, authorship, and/or publication.

References

[1] Hall E, Brenner D. Cancer risks from diagnostic radiology. Br J Radiol. 2008; 81(965): 362-78. doi: 10.1259/bjr/01948454.

[2] Dainiak N. Radiation dose and stochastic risk from exposure to medical imaging. Chest. 2013; 144(5): 1431-3. doi: 10.1378/chest.13-1064.

[3] Matthews K, Brennan P. The application of diagnostic reference levels: General principles and an Irish perspective. Radiography. 2009; 15(2): 171-8. doi: 10.1016/j.radi.2008.03.001.

[4] Vaño E, Miller D, Martin C, Rehani M, Kang K, Rosenstein M, *et al*. ICRP publication 135: Diagnostic Reference Levels in Medical Imaging. Ann ICRP. 2017; 46(1): 78-80. doi: 10.1177/0146645317717209.

[5] Uniya S, Chaturvedi V, Sharma S, Raghuvanshi S. Estimation of entrance surface air kerma due to diagnostic X-ray examinations of adult patients in Uttarakhand, India and establishment of local diagnostic reference levels. Australas Phys Eng Sci Med. 2017; 40: 687-94. doi: 10.1007/s13246-017-0577-8.

[6] Jibiri N, Olowookere C. Evaluation of dose-area product of common radiographic examinations towards establishing a preliminary diagnostic reference levels

(PDRLs) in Southwestern Nigeria. J Appl Clin Med Phys. 2016; 17(6): 392-404. doi: 10.1120/jacmp.v17i6.6011.

[7] Davoodi R, Eydian M, Karampour H, Nassarpour M, Rezazadeh-Farokh R, Maraei A, *et al*. Application of Dose Area Product (DAP) to Estimate Entrance Surface Dose (ESD) in Pediatric Chest X-Rays. Mod Health Sci. 2020; 3(2): 1-6. doi: 10.30560/mhs.v3n2p1.

[8] Kaushik C, Sandhu IS, Srivastava AK, Chitkara M. Estimation of entrance surface air kerma in digital radiographic examinations. Radiat Prot Dosimetry. 2021; 193(1): 16-23. doi: 10.1093/rpd/ncab018.

[9] Jamal N, Sayed I, Syed W. Estimation of organ-absorbed dose in pediatric chest X-ray examination: A phantom study. Radiat Phys Chem. 2020; 166: 108472. doi: 10.1016/j.radphyschem.2019.108472.

[10] Ibrahim I, Hassan F, Abdulkareem N. Effective Dose Calculation for Patients Undergoing X-ray Examinations in Erbil Hospitals. Int Electron Med. 2020; 9(3): 121-3. doi: 10.34172/IEJM.2020.22.

[11] Alonso M, Barriuso T, Castañeda M, Díaz-Caneja N, Gutiérrez I, Sarmiento J, *et al*. Monte carlo estimation of absorbed dose to organs in diagnostic radiology. Health Phys. 1999; 76(4): 388-92. doi: 10.1097/00004032-199904000-00006.

[12] Massoud E, Diab H. Optimization of Dose to Patient in Diagnostic Radiology Using Monte Carlo Method. J Cell Sci Ther. 2014;5(1). doi: 10.4172/2157-7013.1000155.

[13] Tapiovaara M, Siiskonen T. PCXMC A Monte Carlo program for calculating patient doses in medical x-ray examinations 2ed. Findland2008. 49 p.

[14] ICRP. Annals of the ICRP PUBLICATION 103 "The 2007 Recommendations of the International Commission on Radiological Protection". Polestar Wheatons Ltd, Exeter, UK: International Commission on Radiological Protection, 2007. 332 p. doi: 10.1016/j.icrp.2007.10.003.

[15] Markku Tapiovaara TS. PCXMC 2.0 User's Guide. Helsinki, FINLAND: STUK • SÄTEILYTURVAKE; 2008.

[16] Kawaura C, Aoyama T, Koyama S, Achiwa M, Mori M. Organ and effective dose evaluation in diagnostic radiology based on in-phantom dose measurements with novel photodiode-dosimeters. Radiat Prot Dosimetry. 2005; 118(4): 421-30. doi: 10.1093/rpd/nci372.

[17] Kumaresan M, Kumar R, Biju K, Choubey A, Kantharia S. Measurement of entrance skin dose and estimation of organ dose during pediatric chest radigraphy. Health Phys. 2011; 100(6): 654-7. doi: 10.1093/rpd/nci372.

[18] Ma H, Elbakri I, Reed M. Estimation of organ and effective doses from newborn radiography of the chest and abdomen. Radiat Prot Dosimetry. 2013; 156(2): 160-7. doi: 10.1093/rpd/nct050.

[19] Dolenc L, Petrinjak B, Mekiš N, Škrk D. The impact of body mass index on patient radiation dose in general radiography. J Radiol Prot. 2022;42(4):041505. doi: 10.1088/1361-6498/ac9f1f.

[20] Yanch JC, Behrman RH, Hendricks MJ, McCall JH. Increased radiation dose to overweight and obese patients from radiographic examinations. Radiology. 2009; 252(1): 128-39. doi: 10.1148/radiol.2521080141.

[21] Prapan A, Chuprempri K, Fong-in P, Kheawtubtim P, Rattanarungruangchai N, Pengpan T. Radiation dose and image quality optimization in lumbar spine digital radiography for overweight and obese patients: Phantom study. J Assoc Med Sci. 2024; 57(3): 1-7. doi: 10.12982/JAMS.2024.041.

Journal of Associated Medical Sciences 2024; 57(3): 131-137

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Comparative study of erythrocyte sedimentation rate measurement between Celltac Alpha+ (MEK-1305) ESR Analyzer and standard Westergren method

Phutanes Thangvorathum[1] Kanyarak Tuentam[1] Amorntep Pichaipong[1,2] Suwit Duangmano[1,2*]

[1]Division of Clinical Microscopy, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]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 24 January 2024
Accepted as revised 10 July 2024
Available online 15 July 2024

Keywords:
ESR, Celltac Alpha+, Syllectogram, comparative study.

ABSTRACT

Background: The Westergren method, a gold standard method for erythrocyte sedimentation rate (ESR) measurement, is one of the screening tests for routine clinical hematology that can be used as an inflammation marker. Since the results can be altered by numerous factors, new methodologies have been established to manage the possible pitfalls.

Objective: This study compared erythrocyte sedimentation rate measurements obtained by the Celltac Alpha+ (MEK-1305) Automated Hematology and ESR analyzer with those obtained by the standard Westergren method.

Materials and methods: The Celltac Alpha+ (MEK-1305) Automated Hematology and ESR analyzer was used to assess the performance of ESR measurement compared to the standard Westergren method. A total of 220 random EDTA whole blood samples from patients were included and analyzed in parallel using the automated analyzer and the standard Westergren method.

Results: Spearman's rank correlation coefficient (ρ = 0.916) obtained a good correlation between the two methods. In addition, the precision was good and acceptable according to the criteria, with a %CV of less than 10% in both intra-run and inter-run precision. There was no contamination between blood samples from using a single capillary probe. Moreover, Bland-Altman mean difference plot also indicated a good agreement, with a mean bias of -1.4 (95% CI: -0.8 to 2.1), an upper LOA of 7.6 (95% CI: 6.6 to 8.7) and a lower LOA = -10.5 (95% CI: -11.6 to -9.5).

Conclusion: This comparative study indicated that the Celltac Alpha+ (MEK-1305) Automated Hematology and ESR analyzer is applicable within clinical routine practice along with proper measurement regulation. Even though the principles between these two methods are completely different, an acceptable interrelation was obtained due to the calibration process by Nihon Kohden corporation.

Introduction

Westergren method, a gold standard method for erythrocyte sedimentation rate (ESR) measurement, is one of the common routine hematology tests that can be used as an inflammatory and infection marker.[1] In recent years, high-sensitivity C-reactive protein (hs-CRP), rheumatoid factor (RF), and interleukin-6 (IL-6) become more noticeable within clinical laboratory fields due to their higher specificity. However, physicians still request the measurement of an ESR. Formerly, ESR was integrated with other diagnostic criteria for many conditions such as rheumatoid diseases, autoimmune diseases, infections,

* **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: suwit.du@cmu.ac.th
doi: 10.12982/JAMS.2024.055
E-ISSN: 2539-6056

and tumors.[2] The authentic measurement of an erythrocyte sedimentation rate, Westergren method, was performed by determining the rate of red blood cell sedimentation by observing the level to which the cells fall in a given time interval. Conventionally, a whole blood sample mixed with a particular amount of diluent is aspirated into a specific vertical tube, straightly stand for 1 hour, and recorded the plasma length in millimeters/hour (mm./hr.).[3] However, numerous variables can change the pace at which erythrocytes settle out. Various factors, both physiological and pathological, as well as laboratory conditions, can affect ESR levels. Menstruation, pregnancy, malignancies, incorrect anticoagulant use, and elevated temperatures are among the circumstances that may elevate ESR values. In contrast, lower ESR levels may be observed in males, as well as in cases involving hemolyzed blood samples, polycythemia vera, and plasma hyper-viscosity.[4-7]

The standard Westergren method is not advisable for routine laboratory testing because of its long assay time, complicated procedure, and high risk of contamination. Over decades, automation has been continuously invented and developed for application in clinical laboratory practice and to complete its main purposes, including workload reduction, turnaround time shortening, worker safety improvement, and human error lessening. Celltac Alpha+ (MEK-1305) automated hematology and ESR analyzer can measure 20 hematological parameters, including the ESR. The automated hematology analyzer uses optical measurement of the rouleaux formation and aggregation of red blood cells that occur in the initial phase of the sedimentation phenomenon to measure ESR quickly. Rouleaux formation and aggregation of red blood cells are not equivalent to the sedimentation rate. Still, by utilizing the hematocrit (Hct) and mean corpuscular volume (MCV) measurement results, which are closely related to ESR (1-hour value), the analyzer can achieve an ESR value in a shorter amount of time that has a high correlation with the reference method. In the ESR measuring unit, light from an LED is emitted into the agitated blood, and the light that passes through is continuously measured by a light-receiving element. The rouleaux formation and aggregation of the red blood cells begin as soon as agitation ends, causing the intensity of the light passing through the blood to change over time. The waveform expressing this change in light transmission over time is called a syllectogram. At the same time, the CBC measuring unit measures HCT and MCV. The automated hematology analyzer uses a calculation method exclusive to Nihon Kohden to calculate ESR (1-hour value) based on the HCT and MCV values obtained by the CBC measuring unit and the syllabogram produced by the ESR measuring unit.[8]

To provide the method comparability between the automated method and the standard method, the International Council for Standardization in Hematology and Clinical (ICSH) and Laboratory Standards Institute (CLSI) recommended that new methodologies are needed to standardize against its own reference method known as method comparison.[9-12]

Therefore, the study's objective was to compare the Celltac Alpha+ (MEK-1305) Automated Hematology and ESR analyzer assessment of the erythrocyte sedimentation rate to the traditional Westergren technique.

Materials and methods

Blood sample collection

A total of 220 EDTA whole blood samples were all leftovers from daily routine samples of both inpatients and outpatients from Maharaj Nakorn Chiang Mai Hospital and PROMT Healthcare Center AMS CMU, Chiang Mai, Thailand. The use of de-identified blood samples was approved by the Ethics Committee of the Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand (exempted number AMSEC-65EM-011). Hemolyzed and clotted samples were excluded. Whole blood samples were collected in 3.0 mL tripotassium EDTA, processed both by manual Westergren method and by Celltac Alpha+ (MEK 1305) analyzer following the manufacturer's instructions, and examined within 4 hours of venipuncture.

Population study

Population study was classified as descriptive statistics, which is used to describe population features such as frequency, distribution patterns, central tendency, data range, standard deviation (SD), relative standard deviation (RSD), and standard error of the mean (SEM). Following the ICSH recommendations, method comparisons were further assessed in three subgroups, according to the ESR values obtained with the Westergren method, that is, normal (<20 mm.), high (20-60 mm.), and very high (>60 mm.) of the analytical range.

Precision study

B228N, B228L (commercial quality control materials), and a normal blood sample (ESR less than 20 mm.) were used to evaluate the intra-run precision. Each sample was examined 20 times with the MEK-1305 within a day. For the inter-run precision, B228N and B228L were measured once a day for 20 consecutive days. All the results were calculated to find the arithmetic mean, standard deviation, and % coefficient of variation (%CV) Then, %CV was compared to the total allowable error (%TEa).

Sample carryover assessment

Conforming to the CLSI H26-A2 guidelines, any analyzers that utilized a single capillary were required to assess the carryover between samples with high results and low results. The MEK-1305 was evaluated for carryover because it aspirates a specific volume of blood into a capillary that is used to analyze all samples, followed by a washing cycle. Samples were assessed in triplicate for both the high target value (HTV) and the low target value (LTV). % Carryover was calculated using the following formula.

Comparative study

This study performed statistical calculations using MedCalc and IBM SPSS. Three statistics were involved: Spearman's rank correlation coefficient, Passing-Bablok linear regression, and the Bland-Altman mean difference

plot for the method comparison study. Bias, accuracy, and limits of agreement were derived using the Bland-Altman plot.

Results

Population study

According to the population study, the average age of the 220 patients was 52.95 years old. The author determined that the three age groups were compatible with the ESR reference range. The author concluded that three age groups i.e., 0-50 years old, 51-85 years old, and older than 85 years old, agreed with the reference range of ESR. It showed that most patients were between the ages of 51 and 85. In addition, it also showed that most of the patients were women (60.5%) (Table 1)

Table 1. Basic statistical information of population characteristics.

Characteristics	Study population (N=220)
Mean age (years)	52.95
Age groups	
0-50 (years)	92 (41.8%)
51-85 (years)	125 (56.8%)
>85 (years)	3 (1.4%)
Gender	
Male	87 (39.5%)
Female	133 (60.5%)

Initial analysis

A total of 220 samples were processed on Celltac Alpha+ (MEK-1305) automated hematology and ESR analyzer compared to the standard Westergren method. The standard method yielded minimum, maximum, and mean ESR values of 0 mm/hr, 140 mm/hr, and 22.6 mm/hr, respectively. In contrast, the automated method yielded minimum, maximum, and mean ESR values of 1 mm/hr, 119.7 mm/hr, and 20.7 mm/hr, respectively.

Individually, the analysis revealed that ESR values in approximately 70.5% of patient specimens (N=155) were within the reference range (1 to 20 mm/hr) defined as "Normal ESR" and approximately 26.5% of patient specimens (N=65) were beyond the reference range. Results were based on the standard Westergren method. Analyzed results are shown in Table 2.

Table 2. ESR values of a sample of 220 cases categorized according to ESR (mm/hr) ranges when analyzed using the automated analyzer (MEK-1305) compared to the standard Westergren method.

ESR (mm/hr)	N (%)	Method	Min to Max	Mean (95%CI)	SD	RSD	SEM
1-20	155 (70.5)	Westergren	1-20	10.5 (9.6-11.3)	5.627	0.528	0.452
		MEK-1305	1-27	9.9 (8.9-10.9)	6.291	0.634	0.505
21-60	50 (22.7)	Westergren	21-60	36.4 (33.3-39.5)	10.921	0.299	1.545
		MEK-1305	16.7-54.3	32.8 (29.5-36.0)	11.418	0.348	1.615
>60	15 (3.8)	Westergren	70-140	95.6 (82.5-116.5)	23.709	0.248	6.1216
		MEK-1305	66-119.7	92.0 (81.1-102.9)	19.709	0.214	5.089
Total	220	Westergren	1-140	22.2 (18.9-25.4)	24.449	1.103	1.648
		MEK-1305	1-119.7	20.7 (17.6-23.8)	23.361	1.128	1.575

Precision study

Normal blood samples and commercial quality control materials (B228N and B228L), developed by Nihon Kohden, had been used to carry out a precision study. In an intra-run precision, %CV of 8.12 and 2.01 were obtained, while %CV of 7.6 and 2.9 were obtained in an inter-run precision. Moreover, intra-run precision results obtained by analyzing normal blood samples (sample 1) showed a good level of precision below 10% (CV 7.7%). It could be concluded that the variations that occurred were acceptable according to the %CV value assigned by the manufacturing company (%Total allowable error < 10%). Analyzed results are shown in Table 3.

Table 3 Intra-run and Inter-run precision of Celltac Alpha+ (MEK 1305) analyzer.

	Intra-run precision (N=20)		Inter-run precision (N=20)		
	B228N (Normal)	B228L (Abnormal)	Sample 1	B228N (Normal)	B228L (Abnormal)
Mean±SD	7.5±0.6	54.6±1.1	21.2±1.6	6.6±0.5	52.8±1.5
% CV	8.12	2.01	7.7	7.6	2.9

Sample carryover assessment

High target value samples (HTV) and low target value samples (LTV) were measured three times each by Celltac Alpha+ (MEK-1305). The study showed that no contamination happened between the process with %Carryover = 0.0%

Statistical analysis for method comparison

A total of 220 EDTA whole blood samples were processed and analyzed to compare each method statistically. As mentioned in the previous topic, data were categorized into specific parameters and intervals.

Spearman's rank correlation coefficient

The analysis indicated a positive correlation between the standard Westergren method and the automated method. All ranges (N=220) had ρ=0.916 (95%CI: 0.892 to 0.935). Additionally, the highest ρ was given in the low % Hct parameter (ρ=0.969, 95%CI: 0.938 to 0.984). The analyzed results are shown in Table 4.

Table 4 Method comparison results between ESR measurement by Westergren and Celltac Alpha+ (MEK 1305).

		Spearman's rank Correlation		Passing-Bablok linear regression					Bland-Altman	
		ρ	Equation	Slope	Intercept	Mean bias	SD	1.96SD	Upper LOA	Lower LOA
	All ranges (N=220)	0.916	Y = 0.963x - 0.815	0.963	-0.815	-1.446	-4.631	-9.076	-10.523	7.630
ESR (mm/hr)	1-20 (N=155)	0.783	Y = 1.100x - 1.300	1.100	-1.300	-0.529	-3.971	-7.784	-8.313	7.255
	21-60 (N=50)	0.891	Y = 1.023x - 4.912	1.023	-4.912	-3.65	-4.548	-8.913	-12.565	5.261
	>60 (N=15)	0.922	Y = 0.902x + 5.920	0.902	5.912	-3.5733	-7.619	-14.932	-18.506	11.359
Hct (%)	<35 (N=35)	0.969	Y = 0.977x - 1.262	0.977	-1.262	-2.171	-6.281	-12.311	-14.482	10.139
	35-45 (N=135)	0.874	Y = 0.926x + 0.111	0.926	0.111	-1.155	-4.600	-9.018	-10.172	7.863
	>45 (N=52)	0.655	Y = 0.833x - 0.350	0.833	-0.350	-1.704	-3.217	-6.305	-8.009	4.601
MCV (fL)	<80 (N=45)	0.892	Y = 0.997x - 1.661	0.997	-1.661	-2.424	-5.503	-10.785	-13.210	8.361
	≥80 (N=175)	0.919	Y = 0.958x - 0.617	0.958	-0.617	-1.195	-4.361	-8.548	-9.743	7.354

Passing-Bablok linear regression

To evaluate the relationship between the two methods, it indicated a systematic difference, which is described by a linear regression equation y = 0.963x - 0.815 at 95%CI, as represented in Table 4 and Figure 1A. A slope was used to determine if the slope equals 1; the hypothesis would only be accepted if the 95% confidence interval contained 1. The experiment revealed that the slope range was 0.933 and 1.000, which consisted of 1, indicating no statistically significant difference between the slope and 1. Next, the y-intercept was used to determine if the intercept equals 0; the hypothesis would only be accepted if the 95% confidence interval contained 0. The experiment revealed that the range of the y-intercept was -1.300 and -0.150, which does not consist of 0, indicating a statistically significant difference, which meant the y-intercept of the linear regression equation was unacceptable. The possible cause of the unacceptable y-intercept might come from the specific principle that Celltac Alpha+ The MEK-1305 used, leading to a lower result than the standard Westergren method; this is correlated to the prior research of the others.

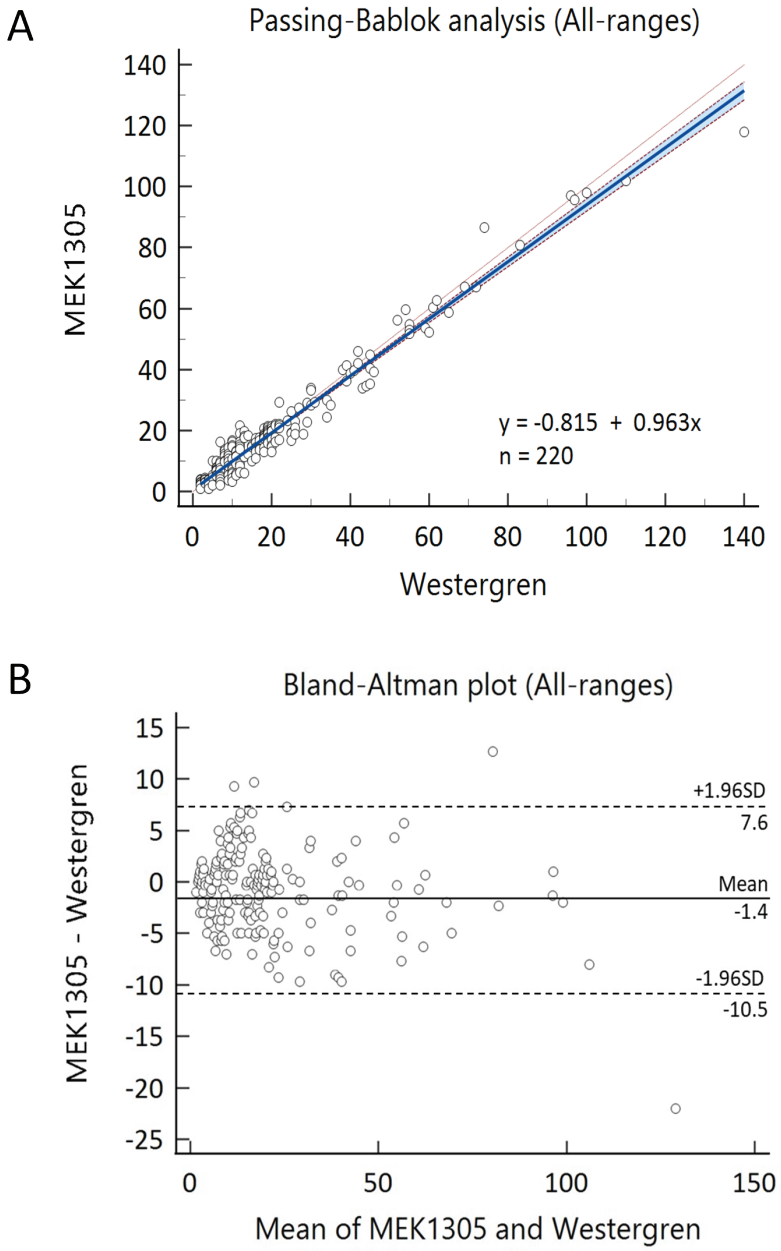


Figure 1. *Comparison between erythrocyte sedimentation rate values using Celltac Alpha+ (MEK 1305) analyzer against the Westergren method (n=220). A: Passing-Bablok scatter diagram with an equation y=-0.815 (95% CI: -1.300 to -0.150) + 0.963 (95% CI: 0.933-1.000)x, B: Bland–Altman scatter diagram with a mean bias of -1.4 (95% CI: -0.8 to 2.1), the value of 7.6 for the upper limit and -10.5 for the lower limit can be considered acceptable limits, implicating a non-significant bias based on a clinical criterion.*

Bland-Altman mean difference plot

The result represented mean bias =-1.4 (95% CI: -0.8 to 2.1), upper limit of agreement (Upper LOA) =7.6 (95% CI: 6.6 to 8.7), and lower limit of agreement (Lower LOA) =-10.5 (95% CI: -11.6 to -9.5) when measuring agreement between two qualitative measurements in 220 samples (Table 4 and Figure 1B). Additionally, negative mean biases could be observed in those parameters that are not considered normal. However, more blood samples should be collected to carry out the analysis effectively since there were not enough samples in some ranges of parameters.

Discussion

The erythrocyte sedimentation rate (ESR) is a commonly used laboratory test in clinical settings. It measures the rate at which red blood cells settle in a specific pipette within a defined interval, typically observed over 1 hour. ESR is an informative indicator for inflammatory conditions such as rheumatoid arthritis, giant cell arthritis, polymyalgia rheumatica, and other connective tissue disorders. Advancements in the study of the sedimentation phenomenon have led to the automation of ESR testing, resulting in the development of numerous automated instruments now commonly

utilized in routine clinical practice. In this study, in line with ICSH recommendations, we assessed the analytical performance of the Celltac Alpha+ (MEK-1305). This new automated ESR analyzer operates on the sedimentation principle using undiluted EDTA samples.

A nonnormal distribution resulted from a population study comparing ESR results obtained by the automated analyzer and the standard Westergren method. Of 220 EDTA whole blood samples, 60.5% were female, and 52.95% were aged between 51 and 85. The utilization of the MEK-1305 yielded a lower result (20.7 mm./hr.), whereas the standard Westergren method provided a higher result (22.2 mm./hr.). These differences occurred due to a contrasting principle between these methods. The automated analyzer utilizes the principle of measuring Rouleaux formation of red blood cells and then calculates it into the red blood cell sedimentation rate. Consequently, this may result in different values.

This study aligns with the previous research by Manoj A. Kahar and colleagues,[13] which found that automated analyzers yield lower test results than the standard Westergren method. From the initial statistical analysis of the entire sample of 220 cases, the minimum ESR value obtained from the standard Westergren method was 2 mm./hr., whereas from the automated analyzer, it was recorded as 1.0 mm./hr. This discrepancy may have arisen due to clerical errors during the operators' interpretation of the result. Additionally, it was observed that in blood samples with hematocrit less than 35% and an MCV less than 80 fL, there was a tendency toward higher than normal ESR values. Conversely, blood samples within the normal range of hematocrit and MCV also showed a tendency for ESR values within the normal range.

The precision study used normal blood samples from volunteers and quality control materials at 2 levels, namely B228N (Normal ESR level) and B228L (Abnormal ESR level). The intra-run precision study found that in B228N, the %CV was higher than the %CV of B228L (%CV, B228N=8.12%. %CV, B228L=2.01%). This aligns with previous studies by C. Cha *et al*.,[14] Nihal *et al*.,[15] and Ivana Lapic et al.[16] Due to the superior physicochemical characteristics of the quality control sample compared to blood samples from individuals, this results in a higher %CV for Sample1 when compared to B228N and B228L. This reference is based on the study conducted by Plebani and Piva.[17] The study on inter-run precision also revealed that the %CV of B228N was higher compared to the %CV of B228L. This finding aligns with previous research by A. Kahar *et al*.,[13] Mahlangu and Davids[18], Horsti *et al*.,[19] and Schapkaitz *et al*.[20] However, it's important to note that based on the study results, the %CV obtained exceeded the %TEa criterion set by the company at %TEa=10%. Moreover, the study revealed that this automated analyzer does not exhibit sample carryover, with the calculated %Carryover =0.0%. This finding aligns with the earlier study of Ryosuke Maki and colleagues in 2021.[21] The statistical analysis using Spearman's rank correlation coefficient within the entire sample of 220 cases revealed a positive correlation. Additionally, it was observed that the trend of the ρ value increased in samples with higher ESR values, lower hematocrit values, and higher MCV values. From the Passing-Bablok linear regression analysis, it was observed that the slope did not meet the criteria. This might be because the Celltac Alpha+ (MEK-1305) Automated Hematology and ESR analyzer tends to provide lower ESR values than the Westergren method. Consequently, this affects the y-intercept of the equation, which does not meet the criteria. This finding is consistent with previous studies. Moreover, when the groups were divided based on ESR values (mm/hr) into different ranges: normal range (1 to 20 mm/hr), high (21 to 60 mm/hr), and very high (>60 mm/hr), it was found that the criteria were met only in the very high ESR group. This might be attributed to a smaller number of samples in the very high ESR group, leading to this outcome. From the Bland-Altman mean difference plot analysis across the entire sample of 220 cases, it was observed that there is consistency in all parameters. However, a relatively high mean bias was noted for the ESR parameter within the range of ESR values between 21 and 60 mm/hr and for ESR values greater than 60 mm./hr. The mean bias was recorded as -3.7 (95% CI: -4.9 to -2.4) and -3.6 (95% CI: -7.8 to 0.6), respectively.

Based on the statistical analysis of the three testing statistics, this automated analyzer can be effectively utilized in clinical settings, provided that the manufacturer's guidelines are adhered to accurately.

Conclusion

Celltac Alpha+ (MEK 1305) analyzer showed good correlation with the conventional Westergren method and an acceptable bias over the entire range of ESR, exhibiting satisfactory concordance of ESR results between Celltac Alpha+ (MEK 1305) and reference Westergren method. Celltac Alpha+ (MEK 1305) analyzer offers major advantages such as use of EDTA sample, reduced sample volume, ease of performance, reduction in biohazard risk, and reliability, making it a valid substitute for reference Westergren method for ESR determination.

Acknowledgements

The author gratefully acknowledges the help extended by Narinthip Sattayarom, Nihon Kohden, Thailand, for supplying the Celltac Alpha+ (MEK 1305) analyzer and consumables for this evaluation.

References

[1] ICSH recommendations for measurement of erythrocyte sedimentation rate. International Council for Standardization in Haematology (Expert Panel on Blood Rheology). J Clin Pathol. 1993; 46(3): 198-203. doi: 10.1136/jcp.46.3.198

[2] Brigden ML. Clinical utility of the erythrocyte sedimentation rate. Am Fam Physician. 1999; 60: 1443-50. PMID: 10524488

[3] Piva E, Pajola R, Temporin V, Plebani M. A new turbidimetric standard to improve the quality assurance of the erythrocyte sedimentation rate measurement. Clin Biochem. 2007; 40: 491-5. doi:

10.1016/j.clinbiochem.2006.12.002.

[4] Kernick D, Jay AW, Rowlands S. Erythrocyte settling. Can J Physiol Pharmacol. 1974; 52(6): 1167-77. doi: 10.1139/y74-152

[5] Zhao TX, Jacobson B. Quantitative correlations among fibrinogen concentration, sedimentation rate and electrical impedance of blood. Med Biol Eng Comput. 1997; 35(3): 181-5. doi: 10.1007/BF02530035

[6] Holley L, Woodland N, Hung WT, Cordatos K, Reuben A. Influence of fibrinogen and haematocrit on erythrocyte sedimentation kinetics. Biorheology. 1999; 36(4): 287-97. PMID: 10690265.

[7] Oka S. A physical theory of erythrocyte sedimentation. Biorheology. 1985; 22(4): 315-21. doi: 10.3233/bir-1985-22404

[8] Operator's Manual of Celltac α+ Automated Hematology and ESR Analyzer Nihon Kohden, Japan.

[9] Kratz A, Plebani M, Peng M, Lee YK, McCafferty R, Machin SJ. ICSH recommendations for modified and alternate methods measuring the erythrocyte sedimentation rate. Int J Lab Hematol. 2017; 39(5): 448-57. doi: 10.1111/ijlh.12693.

[10] Clinical and Laboratory Standards Institute (CLSI). Measurement Procedure Comparison and Bias Estimation Using Patient Samples, Approved Guideline, 3rd. Ed. EP09-A3rd. Wayne, PA: CLSI; 2018.

[11] Clinical and Laboratory Standards Institute (CLSI). Evaluation of detection capability for clinical laboratory measurement procedures. Approved Guideline, 2nd Ed. EP17-A2. Wayne, PA: CLSI; 2012.

[12] Clinical and Laboratory Standards Institute (CLSI). Validation and Quality Assurance of Automated Hematology Analyzers. Approved Standard, 2nd Ed. H26-A2. Wayne, PA: CLSI; 2010.

[13] Kahar MA. Comparison of alternate erythrocyte sedimentation rate measurement by automated Celltac α+ (MEK 1305) and reference Westergren method. J Hematol Allied Sci. 2022; 2: 39-45. doi: 10.25259/JHAS_12_2022.

[14] Cha C, Cha YJ, Park C, Kim H, Cha E, Kim DH, *et al*. Erythrocyte sedimentation rate measurements by TEST 1 better reflect inflammation than do those by the Westergren method in patients with malignancy, autoimmune disease, or infection. Am J Clin Pathol. 2009; 131(2): 189-94. doi: 10.1309/AJCP0U1ASTL-RANIJ.

[15] Boğdaycioğlu N, Yilmaz FM, Sezer S, Oğuz E. Comparison of iSED and Ves-Matic Cube 200 Erythrocyte Sedimentation Rate Measurements With Westergren Method. J Clin Lab Anal. 2015; 29(5): 397-404. doi: 10.1002/jcla.21786.

[16] Lapić I, Piva E, Spolaore F, Tosato F, Pelloso M, Plebani M. Automated measurement of the erythrocyte sedimentation rate: method validation and comparison. Clin Chem Lab Med. 2019; 57(9): 1364-73. doi: 10.1515/cclm-2019-0204.

[17] Plebani M, Piva E. Erythrocyte sedimentation rate: use of fresh blood for quality control. Am J Clin Pathol. 2002; 117(4): 621-6. doi: 10.1309/QB1G-6FRR-DNWX-BKQ9.

[18] Mahlangu JN, Davids M. Three-way comparison of methods for the measurement of the erythrocyte sedimentation rate. J Clin Lab Anal. 2008; 22: 346-52. doi: 10.1002/jcla.20267.

[19] Horsti J, Rontu R, Collings A. A comparison between the StaRRsed auto-compact erythrocyte sedimentation rate instrument and the westergren method. J Clin Med Res. 2010; 2: 261-5. doi: 10.4021/jocmr476w.

[20] Schapkaitz E, RabuRabu S, Engelbrecht M. Differences in erythrocyte sedimentation rates using a modified Westergren method and an alternate method. J Clin Lab Anal. 2019; 33: e22661. doi: 10.1002/jcla.22661.

[21] Ryosuke M, Yoko T, Yamamoto T, Hiroyuki T, Tomohiro S, Esaki T, et al. Accuracy study of a novel alternate method measuring erythrocyte sedimentation rate for prototype hematology analyzer Celltac α+. Int J Lab Hematol. 2021; 43: 588-96. doi: 10.1111/ijlh.13554.

Journal of Associated Medical Sciences 2024; 57(3): 138-154

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Review article: An overview of exosomes in biology and their potential applications in regenerative medicine

Jose James[1] Janardhana PB[2] K.R. Padma[3*] K.R. Don[4]

[1]Department of Exosomes and Nanotechnology, Stellixir biotech, Bengaluru, Karnataka, India.
[2]Lihan Biologics, Bengaluru, Karnataka, India.
[3]Department of Biotechnology, Sri Padmavati Mahila Visvavidyalayam (Women's University), Tirupati, Andhra Pradesh, India.
[4]Department of Oral Pathology & Microbiology, Sree Balaji Dental College & Hospital, Bharath Institute of Higher Education & Research (BIHER) Bharath University, Chennai, Tamil Nadu, India.

ARTICLE INFO

Article history:
Received 14 January 2024
Accepted as revised 10 July 2024
Available online 15 July 2024

Keywords:
Mesenchymal Stem cells (MSCs), exosomes, immunomodulation, tissue repair, wound healing, cell differentiation, hair follicle regeneration, diabetics, regenerative medicine.

ABSTRACT

Extracellular vesicles (EVs), commonly acknowledged as Exosomes, are tiny, single-membrane, secreted organelles that range in size from 40 to 150 nm. They are noticeably abundant in various proteins, lipids, nucleic acids, and glycoconjugates and share the same structure as cells. Numerous and non-hematopoietic cell types continuously manufacture and release stable, less toxic, and biocompatible exosomes with many complex compounds (in the form of various signaling molecules, miRNA, and mRNA) in the liquid parts of the body. Exosomes help in intercellular communication/transfer of proteins, RNA, cell differentiation, immune signaling, delivering antigens, angiogenesis, and stress response. In recent studies, researchers found that Mesenchymal Stem cells (MSCs) generate exosomes that symbolize biological processes like tissue regeneration by encasing and delivering active biomolecular species to the infected/damaged cells and tissues. The most extensive research in regenerative medicine has focused on MSCs-Exosomes. Regenerative medicine plays a crucial role in restoring the damaged/lost parts of organs and tissues and aiding in wound healing. Immunomodulation and tissue repair are possible by introducing Mesenchymal Stem cell (MSC) exosomes, which have triggered remodeling reactions. They produce local anti-inflammatory and healing signals crucial for regeneration and tissue repair. The primary goal of this review is to highlight the MSCs-exosome's mechanism of action and its therapeutic uses in clinical settings. Also, it highlights new developments in employing MSCs-exosomes to treat various ailments and disorders.

Introduction

Mesenchymal stromal/stem cell (MSCs) investigations for immunomodulation and regenerative medicine are widely used in cellular treatment.[1] Mesoderm-pluripotent embryonic stem cells (MPSCs) can be discriminated into diverse cell forms.[2] The reassuring characteristics of MSCs-exosomes, such as their outstanding capacity for cell differentiation and regeneration, have generated much study attention over the past few decades. MSCs, which may be extracted from several organs, are the adult stem cells most frequently used in regenerative medicine, they have a high potential for replication in culture and can develop into different cells such as adipocytes, chondrocytes, and osteoblasts.[3] They are predominantly located in the perivascular spaces encompassing nearly every human tissue and system. Bone marrow, fatty tissues, umbilical cord, and maternal tissues are the primary sources of MSCs, as shown in Figure 1.[4.] Under this, some studies claimed that MSCs-EVs/exo, i.e., exosomes isolated from MSCs have superior curative benefits to general MSCs.[5]

∗ **Corresponding contributor.**
Author's Address: *Department of Biotechnology, Sri Padmavati Mahila Visvavidyalayam (Women's University), Tirupati, Andhra Pradesh, India.*
E-mail address: thulasipadi@gmail.com
doi: 10.12982/JAMS.2024.056
E-ISSN: 2539-6056

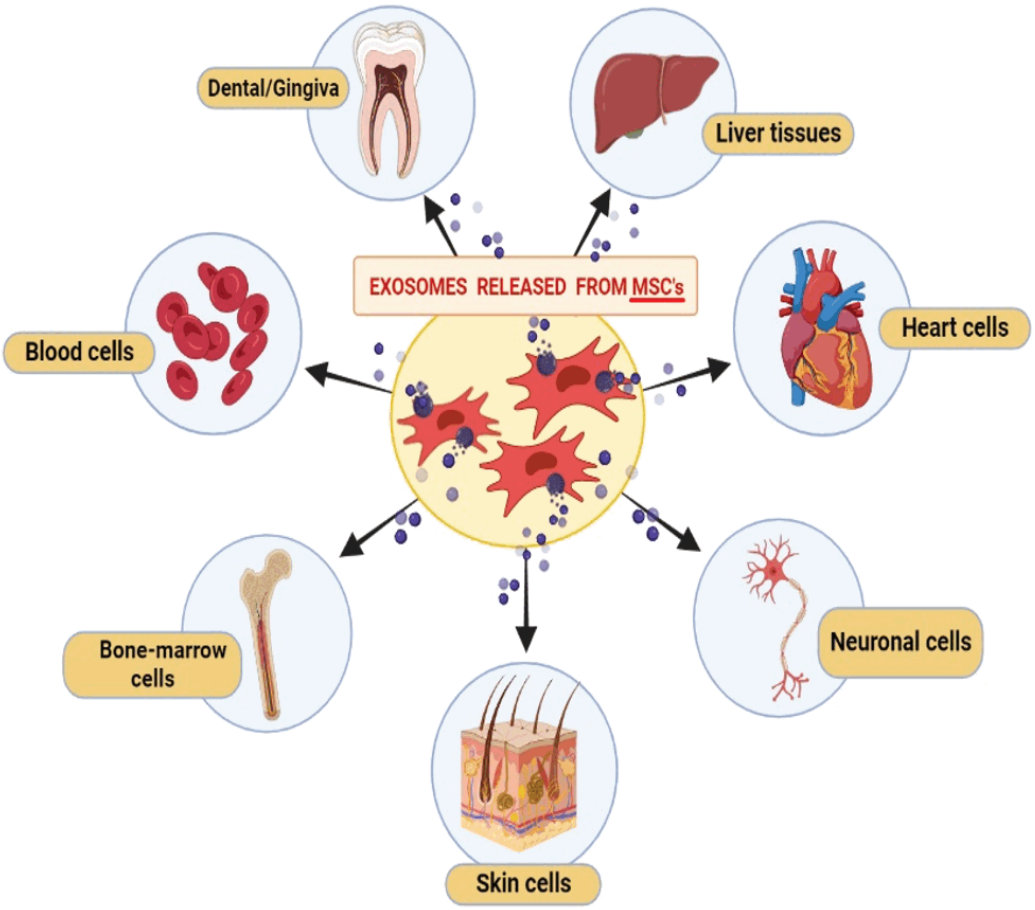


Figure 1. *Mesenchymal stem cells (MSCs) termed exosomes isolated from diverse organs/ tissues (bone marrow, adipose tissue, dental pulp, umbilical cord blood, placenta, liver, and neuronal cells).*

Since MSCs have receptors for various factors expressed on their outermost cell layer, they can go to the site of cancer or inflammation and release inflammatory factors, including IL-6, IL-8, and MCP-1, into the environment of such inflamed areas. These elements support MSCs' instructed movement.[6,7] Exosomes produced from MSCs (MSCs-Exo) have been the subject of numerous studies, and it has been found that these exosomes share many of the same processes as MSCs, including the ability to repair tissue damage, suppress inflammatory responses, advance tumors, and promote angiogenesis.[8] Extracellular vesicles (EVs) or MSCs-Exos, which are released by the majority of MSCs cells, as well as secretory vesicles, which are produced by some MSCs cells and allow the vesicular transport of cargo like neurotransmitters or hormones, are among the membrane vesicles that MSCs cells can produce.[9] It was initially believed that the production of EVs served as a mechanism for the cell to get rid of waste products in the form of "platelet dust/cellular debris".[10] The ability of EVs to carry cargo between cells, such as DNA, lipids, and proteins, is the main area of interest in this subject because it has become clear that EVs are more than merely trash carriers.

The EVs are crucial for maintaining or developing cell pathologies as cellular communication agents. EVs can be split into two significant groups according to their origin: Micro- and exo-vesicles (MVs).[11] According to the International Society for Extracellular Vesicles (ISEV), extracellular vesicles might fall into 3 categories: exosomes, microvesicles, and apoptotic structures. These groups' location, size, and substance are crucial.[12,13] The smallest of these are exosomes, with a size assortment of ~ (40-150) nm on the nanoscale [14]. Furthermore, their biogenesis is another distinguishing trait of exosomes that sets them apart from other EVs. Exosomes have an endosomal origin in contrast to microvesicles, which form directly from the plasma membrane, and apoptotic bodies, which form during apoptosis (Figure 2).[15] Exosomes are formed by the fusion of plasma membrane with multivesicular bodies (MVB).[16] Multivesicular bodies (MVB), a type of internal multivesicular compartment, are created when the endosomal membrane invaginates.[17,18] Exosomes have been characterized as having a "cup-shaped" structure with varying sucrose densities between 1.13 and 1.19 g/mL.[19]

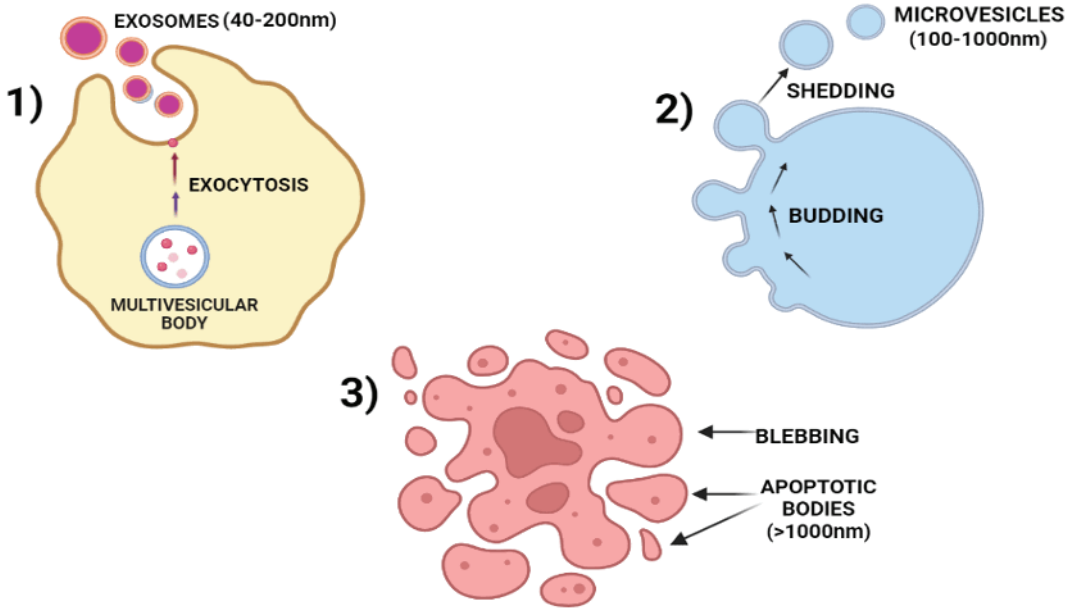


***Figure 2**. Extracellular vesicles (EVs) classification.*

A few of the cells that produce exosomes are immune cells, platelets, smooth muscle cells, and endothelial cells. When the extracorporeal membrane of the polyvesicles coalesces with the plasma membrane. Exosomes are produced inside cells and released into the extracellular matrix by lysosomal polyvesicles.[20] Exosomes, which are generally released into receptor cells by host cells, regulate the biological functions of receptor cells by supplying substances that affect how receptor cells behave. These substances include proteins, lipids, nucleic acids (NAs), and other added substances. These outcomes sustain the use of exosomes as biomarkers and in immunotherapeutic strategies. Our study presents a brief overview of the function of exosomes in connection to their origin and roles, and it addresses the possible use of exosomes to treat various disorders, such as diabetes, immunomodulation, wound healing, and hair follicle growth.[21]

Three distinctive classes of EVs are portrayed above where: 1) Exosomes, which are 40-200 nm in diameter, are produced by endocytic pathways and are let out into extracellular spaces by a process called exocytosis; 2) Microvesicles are irregular shaped bodies that are of diameter size 100-1000 nm and are generated by budding from the cytoplasmic membrane; and 3) Cells that undergo apoptosis produce apoptotic bodies which are of diameter size >1000 nm by the process of cell blebbing.[22] In Table 1 depicts the diverse forms of exosomes.

Table 1. Depicts the origin, size, markers and different types of exosomes

Different types of vesicles	Origin	Size (nm)	Markers
Microvesicles	Plasma membrane and many other cell types	20-1000	Broad variety of non-specific markers such as selectin, integrins and CD40 ligand[23,24]
Apoptotic bodies	Cell membrane from Endoplasmic reticulum	1000-5000	Phosphatidylserine, DNA & its products, Histones[25,26]
Membrane fragments	Epithelial cell membrane	50-80	Prominin-1 (CD133) [27]
Exosomes	Endosomes from various cell types	40-100	other endosomal-related indicators, such as lipid rafts, Tetraspanins, flotillin, Alix, TSG101, sphingomyelin, and Rab5b. All markers, however, specifically target exosomes. [28]

Discovery of exosomes

Exosomes were once believed to be apoptotic bodies for the minimally disruptive discharge of cellular debris from cell injury or byproducts of cell homeostasis.[29] Harding *et al.*,[30,31] and Pan *et al*.,[32] were the two scientists who first described exosomes in their research papers on two separately different animal studies, and this was published a week apart in 1983. In their research, scientists noticed how sheep Transferrin Receptors (TfRs) were released from the plasma membrane and entered

growing human reticulocytes. It was discovered that transferrin receptors interact with cells and are then packed into tiny (~50 nm) vesicles.[33,34] Schirrmacher and Barz, a year after this event, noticed that tumor-derived exosomes (TDEs) had antigens that were comparable to those of the matching tumor cells.[35] These vesicles were subsequently named "exosomes" in 1987 by Johnstone *et al*.,[36] which were initially believed to be on their way to lysosomes for destruction before being secreted into the extracellular area by maturing blood reticulocytes.[37]

Extensive study and research on exosomes increased due to the development of more sophisticated approaches to understand functional biology clearly. They extended therapeutic applications of exosomes, such as techniques for tissue regeneration, immunomodulation, and cell differentiation [38]. The initial findings from clinical tests on exosome-based immunizations were published in the mid-2000s.[39] In-vitro conditions are a viable therapeutic option for several illnesses with tumor-specific biomarkers. Nevertheless, MSCs-Exos applicability in orthopedics, dentistry, and COVID-19 therapy, as well as liposome-mediated drug delivery in cardiovascular disorders and cancer using tumor-specific biomarkers.[40-43]

Structure and composition of exosomes

Following diverse proteomics techniques like mass spectrometry, exosomes are said to contain more than 4,000 distinct types of biomolecules, involving particular groups of lipids, proteins, nucleic acids, mitochondrial DNA (mtDNA), microRNAs (miRNAs), mRNAs such as transfer RNA (tRNA), long noncoding RNAs (lncRNAs), piwi-interacting proteins RNAs (piRNAs), cytokines, transcription factor receptors, and other bioactive compounds are among the different kinds of RNA. Exosomes from progenitor cells transport these RNAs to target cells, where they can perform specific functional tasks. Exosomes vary in constituent parts due to various physiological and pathological circumstances and specific cell types; the cell of origin significantly influences the exosomal contents.[44-49]

Exosome protein components are classified into two categories; one category consists of elements that are involved in the creation, release, and exosome biogenesis of intraluminal vesicles (ILVs) regularly, like programmed cell death 6 interacting protein (PDCD6IP or ALIX), and another category made up of the tumor susceptibility gene 101 (eg. tetraspanins-8 (TSPAN 8), Vacuolar protein sorting-associated protein 4 (VPS4), TSG101, SD106), whose release is mainly reliant on proteins like Rab GTPases.[50-53] Additionally, exosomes contain non-specific proteins like cytoskeleton proteins (such as myosin, actin, and tubulin), cell-specific antigen-presenting components like CD45 and Major Histocompatibility Complex proteins (such as MHC I and MHC II), which are typically seen in donor cells, heat shock proteins (Hsp) (such as Hsp70, Hsp90, Hsp110), membrane integrating transfer proteins (such as flotillins, annexins, and Rab), ESCRT, and CD63.[54,55] Compounds such as sphingomyelin, cholesterol, phosphatidylserine, glycosphingolipids, and ceramide, which impact signaling, exosome release, structure, and cargo separation, are also seen as abundant in exosomes. These lipid components are stable and necessary for the preservation of exosome structure, biogenesis of exosome, and homeostatic control of the target cell.[56] Additionally, these proangiogenic proteins, like basic fibroblast growth factor (bFGF), interleukin 6 (IL-6), monocyte chemokine protein-1 (MCP-1), migration-promoting chemokines, vascular endothelial growth factor (VEGF), towards the inflammation site, proteins involved with the mitochondria, proteasomes, and endosomal reticulum, are strongly expressed in MSCs-derived microvesicles (MVs) (Figure 3).[57]

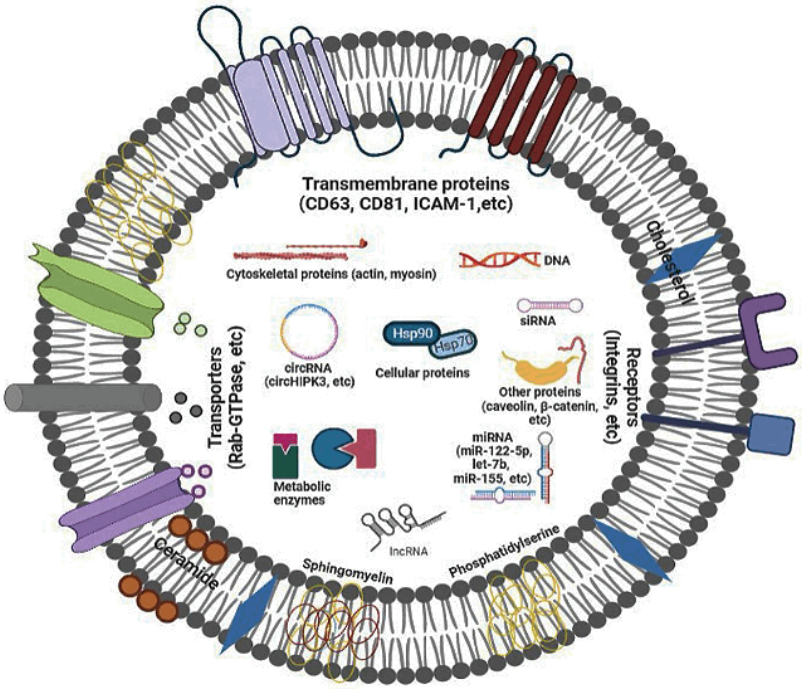


Figure 3. *Structure and composition of exosome.*

Exosome biogenesis, release, and uptake

Initiation, endocytosis, multivesicular body growth, and secretion are the four stages of exosome maturation.[58] When the plasma membrane invaginates, early endosomes are created that include membrane protein and multivesicular bodies (MVBs), which are exosomes. Sometimes, it is also referred to as late endosomal structures comprising several ILVs with a changing subcellular structure.[59] Finally, MVBs are fused with the plasma membrane to release exosomes into the extracellular environment or carried to lysosomes to degrade all transported compounds. They are also transferred to the trans-Golgi network (TGN) for endosome recycling.[60] Early endosomes (EEs), created by merging endocytic vesicles, are the first stage in biogenesis.[61] Later, the formed EEs are packaged into late endosomes (LEs) or return the cargo proteins outside the cell.[62] To move proteins, the packing of proteins inside multivesicular bodies and intraluminal vesicles (ILVs) relies on the endosomal-sorting complex (ESCRT). Typically 4 ESCRT protein structures such as ESCRT0, ESCRTI, ESCRTII, and ESCRTIII, in conjunction with AAA, ATPase, Tsp101, ALIX, and Vps4 accessory protein complex that relies on ubiquitin protein for further mechanism or go through ubiquitin independent mechanism of action for selection of required cargo, and the production of target exosomes.[63-65] Both the Vps4 complex and the ESCRT-III complex work together to cause vesicle neck scission; ESCRT-III multifaceted dissociation and regeneration occur. The multifaceted ESCRT-0 also arranges ubiquitinated cargo proteins into the lipid area. ESCRT-I, as well as ESCRT-II, drive the deformation of the membrane to develop the membrane-stable neck.[66] The first stage is the creation of ESCRT-0 complex, a heterodimer made up of the ubiquitylated cargo-identifying proteins STAM1/2 and HRS.[67] The FYVE domain covers 3 retained components such as C-terminal RVC motifs, which form phosphatidylinositol 3-phosphate (PtdIns-3) binding site, N-terminal WxxD, the central RR/KHHCR, cytoplasmic protein HRS, and early endosome antigen 1 (EEA1) results in formation of endosomes,[68,69] which connects to phosphatidylinositol 3-phosphate (PtdIns3P). This lipid is plentiful on the surface of pre-MVB endosomes.[70] Then, HRS employs Clathrin,71, who assists in the corralling and ubiquitylated freight congregating.[72] At the site of ILV creation, ESCRT-0 interacts with ESCRT-I and ESCRT-II. Since both contain ubiquitin-interaction domains, a sorting domain with elevated avidity for ubiquitylated cargo must be created.[73-76] ESCRT-III, which promotes membrane deformation along with tightening of the collar through consequent intussusception,[77] further is also recruited at that precise moment.[78,79] [Then, ubiquitin is eliminated from the payload by Deubiquitinating enzymes (DUBs),[80] and the ESCRT complex is dissociated so that its constituent parts can be repurposed by the ATPase VPS4 and co-factor VTA1.[81] Theoretically, suppression of any of these factors should prevent the production of ESCRT-dependent exosomes, yet without having an impact on other procedures like lysosomal targeting.[82]

RNA entry into exosomes seems to be lipid-mediated, unlike proteins organized by ESCRT that rely on self-regulating lipid and cargo locations. Nucleotide sequences demonstrate the increased attraction for the phospholipid bilayer and largely depend on elements like lipid structure (lipid rafts), hydrophobic alterations, and sphingosine at an average proportion in rafted membranes (Table 2).[83] The plasma membrane's fatty rafts are regions that are abundant in proteins that are glycosylphosphatidylinositol (GPI)-anchored, sphingolipids, and cholesterol. ILVs are formed spontaneously by the budding-in process while ceramide, lysophospholipid, and glycosphingolipid molecules accumulate on the limiting membrane. In the presence of ceramidase and sphingosine kinase, the ceramide is transformed to sphingosine and sphingosine-1-phosphate (S1P), and continual stimulation of sphingosine-1-phosphate receptors on the limiting membrane facilitates the organization of tetraspanin into ILVs.[84] Tetraspanins organize membrane microdomains, defined as tetraspanin-enriched microdomains (TEMs) with many transmembrane and cytosolic signaling proteins. Four transmembrane domains set tetraspanins apart as a protein superfamily of transmembrane proteins linked to cell surfaces.[85] According to a 2009 study by Stuffers, Susanne *et al*., the absence of the ESCRT tool in mammalian cells did not prevent the development of MVB vesicles. Still, it did result in abnormalities in the amount and size of ILVs and in the cargo-sorting into ILVs. This suggests that exosome biogenesis may be a synchronized process with both ESCRT-dependent and ESCRT-independent pathways (Figure 4).[86]

The first step in producing and generating exosomes is the creation of early endosomes, which include RNA and cytosolic proteins. These early endosomes later develop into active subcellular structures called MVBs (Multivesicular Bodies). MVBs can then be destroyed by lysosome fusion or released by plasma membranes to create exosomes. Lastly, there are numerous methods by which exosomes deliver their cargo, which includes proteins, DNA, and microRNA, to the target cell.

Table 2. The process of synchronization of exosomes and their release.

ESCRT-Dependent exosome

Protein	Cell lines used for *In vitro* studies	ESCRT-dependent exosome proteins used
Syndecan	MCF-7	HSP70, Alix, CD63 (membrane cargo binding to syntenin-1)[87,88]
Hepatocyte growth factor receptor tyrosine kinase substrate (HRS) an endosomal protein	HeLa-CIITA, DCs	VPS4B, MHC-II, CD63, Tsg101[89-91]
Tsg101 (VPS23)	MDCK, MCF-7, HeLa-CIITA, DCs	CD81, MHC-II, CD63, syndecan-1, ALIX, HSC70[92]
Syntenin	MCF-7	HSP70, CD63[93]
VPS4	MCF-7, DCs, HeLa CIITA	syndecan-1, HSC70[94]
Alix	MCF-7, DCs, HeLa, CIITA	syndecan-1, TSG101, RAB5, HRS, HSC70[95]
CHMP4C (SNF7C)	HeLa-CIITA	CD81, HSC70, CD63, MHC-II[96]
CHMP4B (SNF7B)	HeLa-CIITA	HRS, TSG101, RAB5[97]
STAM1	HeLa-CIITA	MHC-II, HSC70, CD63, CD81[98]

ESCRT-independent exosomes

Protein	Cell lines used for *In vitro* studies	ESCRT-Independent exosome proteins used
RAB31	HeLa, HEK-293T	Tsg101, CD81, Alix, CD9, CD63[99]
nSMase2	HEp-2, Oli-neu	Tsg101, Hrs, PLP, Alix, Syntenin[100,101]
DGKα	J-HM1–2.2	B-Actin, Fasl, CD63[102]
PLD2	RBL-2H3, MCF-7	SDC1CTF, CD63[103,104]
CD9	BMDCs, HEK293	Flotillin-1, β-catenin[105]
CD63	DG-75, HEK293, HK1, Rat1, MNT-1, HeLa	Calnexin, CD81, HSC70[106]
CD82	HEK293	β- Catenin[107]

Release of exosome

Protein	Cell lines used for In Vitro studies	Release of Exosome
Tetherin	HeLa	ALIX, TSG101, CD63[108]
RAB11	Drosophila S2, K562	HSC70, Evi (WB), Lyn, Transferrin receptor[109,110]
YKT6	A549	Tsg101[111]
RalA, RalB	4T1	HSC70, ALIX, TSG101, CD63[112]
VAMP7	K562	Acetylcholinesterase activity[113]
RAB27a/b	Human peripheral blood, HeLa-CIITA	Tsg101, Hsc70, Hsp70, CD63[114]

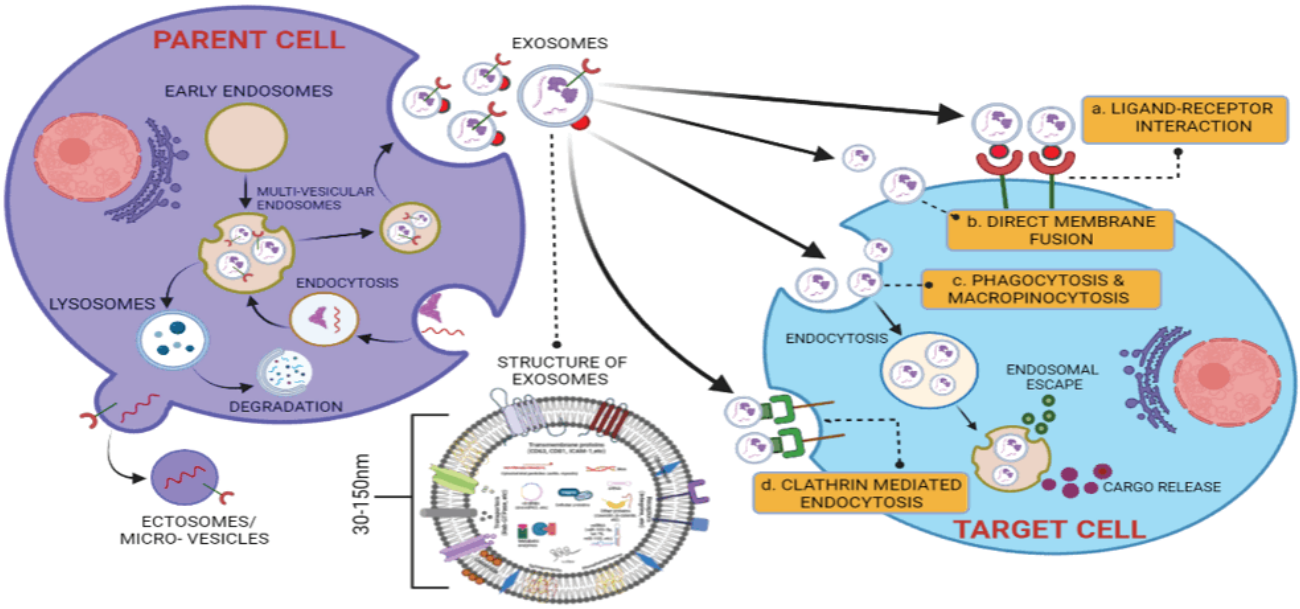


***Figure 4**. Exosome biogenesis, release, and uptake.*

Uptake of exosome

Exosomes communicate with target cells via three different mechanisms: (a) receptor-ligand contact, (b) direct membrane fusion, and (c) endocytosis using phagocytosis. Additionally, several proteins function as ligands to stimulate the absorption of the exosome, such as ICAM-1 for Tim 1/4 for B-cells and cells that present antigens (APCs).[115-117] Exosomes perform an integral part in pathological and physiological methods based on their capability to communicate with target cells once discharged outside the cell and transmit proteins, lipids, and nucleic acids. Even though numerous techniques have been proposed.[118-120]

Applications of exosomes

Immunomodulatory activity of exosomes

MSCs are extremely important in the field of regenerative medicine and are found to interfere with numerous immune response mechanisms and exhibit immunomodulatory functions. MSCs have broad immunoregulatory capabilities through interactions with immune cells in both the adaptive and innate immune systems, resulting in immunosuppression of diverse effector functions.[121] MSCs regulate immunomodulation by reducing B-cell proliferation and activation, hindering NK cell cytotoxicity and proliferation, decreasing dendritic cell maturation and suppressing T-cells, and also by facilitating the generation of regulatory T-cells through cell-cell contact or with the help of several soluble factors such as prostaglandin E2 (PGE2), hepatocyte growth factor (HGF), transforming growth factor-β1 (TGF-β1), etc.[122] Exosomes and microvesicles generated from MSCs have been found to have similar immunosuppressive properties.[123]

Zhang *et al*. noticed that MSC-derived exosomes persuaded the phenotype of monocytes to M2-like promoted T-cells, which further evolved into regulatory T-cells.[124] Nonetheless, lowered immunological actions *in vivo* augment the viability of mice allogeneic skin grafts.[125] Morrison *et al*. demonstrated that MSCs-derived EVs promote a highly phagocytic and anti-inflammatory macrophage phenotype through mitochondrial delivery.[126] Responses via T-cell mediated immunomodulation is a powerful strategy for controlling autoimmune and inflammatory disorders. *In vitro* experiments with T-cells treated with MSCs-derived EVs revealed a significant reduction in T-cell-driven proliferation and decreased production of specific proteins, including IFN-γ and TNF-α.[127] Animal models were utilized to investigate the immunomodulatory role of MSCs-derived EVs, with results revealing immunological activity and changes in the production of pro-inflammatory and anti-inflammatory cytokines.[128] The impact of immunomodulation by the exosomes produced from MSCs in animal models of antigen-driven tissue damage exhibited a variety of effects, including a reduction in the numbers of synovial joint lymphocytes and reduced expression of TNF-α mRNA in the synovial joints increased in survival of induced lung injury as well as an increase in regulatory T-cells after concanavalin-A-induced liver damage.[129-131] According to these findings, MSC-derived EVs preserve the biological functions performed by the mother cells and can be used as an immunosuppressive tool.[132]

MSCs have recently been proposed as a promising treatment for SARS-CoV-2. MSCs suppress viral infections by releasing certain cytokines; these features are inherent in MSCs even before they are separated from their parent tissue. As a result, when these MSCs and their exosomes (MSCs-Exo) are transplanted into a patient with proven SARS-CoV-2 infection, they are likely to survive.[133]

Exosomes in hair follicle regeneration

The skin, the biggest organ, is a vital bodily barrier, playing a major role in detecting impulses and defending against infections and other environmental substances.

The skin appendices also contain hair follicles (HF) and glands; these hair follicles originate from the first two layers of skin, epidermis and dermis, respectively. The HF is generally made of layers of cells within the layers of the skin.[134] The cycle of HF development occurs continuously throughout the life of the organism and occurs in phases, which are telogen (rest phase), anogen (growth phase), and catagen (regression phase). Epithelial, melanocyte, and mesenchymal stem cells (MSCs), which self-renew, differentiate, control hair growth, and uphold skin homeostasis, are among the numerous skin stem cell groups found in the bulge section of the HF.[135] The skin has drawn much interest as a potential target for regeneration treatment owing to the improved access and knowledge about the localization action of skin stem cells, in general, the skin, and the HF in specific.[136,137]

Wnt factors are well-known primary controllers of hair growth and HF morphogenesis.[138] Wnt ligands are critical in de novo hair creation and are induced by adult skin wounds. Active Wnt factors, on the contrary, have been revealed as molecules that are secreted by the exosomes that can be both housed within and released from these vesicles.[139,140] Wnt4 and Wnt11, obtained from the MSCs of human umbilical cord exosomes, have been proven to promote the revitalization of a cutaneous layer of skin in a burnt skin rat model. Wnt4 improved the skin's angiogenesis and Wnt/- β-catenin signaling processes.[141,142] Additionally, it was discovered that different subsets of vesicles containing Wnt could be isolated in a specific pattern in epithelial cells of the polarised type. Moreover, a mechanism for the discharge of Wnt3a and Wnt11 from MDCK cells has also been shown to be crucial for the structural organization of the HF and interfollicular epidermis.[143] These findings show that enhanced expression of Wnt3a and Wnt5a was linked to stimulation of hair formation in mouse skin treated with intradermally injected EVs generated from MSCs.[144] When injected into the skin of mice, exosomes derived from human Dermal Papilla (DP) cells were found to extend the anagen phase of the hair cycle by promoting the production of β-catenin and Sonic Hedgehog (Shh). The Hedgehog pathway is a crucial pathway that plays a major role in homeostasis, repair, and the development of the HF epidermis and the maintenance of HF bulge stem cells.[145] MicroRNAs (miRNAs) can also be transported in EVs, and it has been found that these molecules also have a major role to play in the regulation of skin and HF development through Wnt signaling modulation.[146] MiR-181c, detected in the exosomes of MSCs from the human umbilical cord, has been proven in studies to be an important factor in lowering inflammation brought on by burn in a rat model.[147] Additionally, synovium-MSCs release exosomes that overexpress miR-126-3p, which has been shown to trigger increased P-AKT and ERK1/2 production in HMEC-1 endothelial cells and support diabetic rat skin wound healing.[148]

Subcutaneous administration with conditioned media produced from MSCs derived from human amniotic fluid improved hair regrowth and expedited wound healing in a wound model in rats.[149] Intradermally injected MSCs-EVs improved the telogen-to-anagen transition in a mouse model.[150] Exosomes from human DP were injected into mice's skin to promote hair growth, possibly simulating the paracrine effects that DP cells naturally have on epithelial cells. The outer root sheath cells of epithelial hair follicles taken from human scalps and cultured with DP exosomes showed elevated Shh and β -catenin levels in the treated skin. In response to exosomes, these findings repeatedly showed how the Wnt/ β -catenin and Shh pathways regulate hair growth (Figure 5).[151]

EVs are natural hair cycle regulators and potential delivery systems for enhancing hair and skin regeneration. The precise physiological significance and contribution of exosomes to the HF cycle in vivo and the therapeutic implications of using exosomes to enhance hair growth in clinics require further investigation.

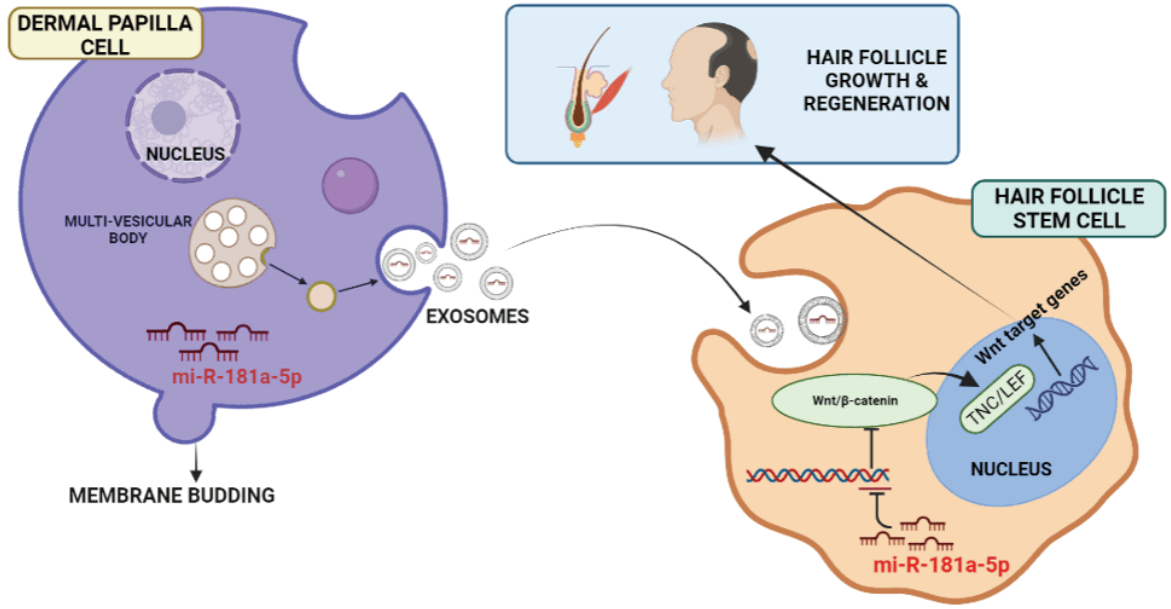


***Figure 5**. Exosomes used in Hair follicle regeneration.*

Exosomes in wound healing

Mesenchymal stromal cells (MSCs) and their exosomes have recently emerged as game-changing tools in tissue engineering and regenerative medicine. Exosomes are generally known to transport functional cargos like miRNA, cytokine, growth factor, etc., from MSCs to the respective target cells, thus influencing the biological processes in the recipient skin cells like proliferation, migration, and ECM component secretion (e.g., collagen, fibroblasts, and keratinocytes).[152] Wound healing is a significant difficulty in certain cases like pathological disorders like diabetic foot ulcers (DFUs), plastic surgery, and others. Persistent wounds cause significant patient morbidity, negatively impacting patient quality of life and exacerbating pain, stress, and despair.[153,154] The current therapeutic procedures are prohibitively expensive, ineffective, and time-demanding, and more than half of infected wounds display strong resistance to topical therapy, which cannot reduce scarring.[155] Mesenchymal stromal cells (MSCs) have emerged as an essential strategy for restoring skin wound healing. MSCs show their therapeutic efficacy primarily through numerous actions on numerous cell types and at all stages of wound healing, from hemostasis through remodeling.[156] Exosomes generated from MSCs have biological functions like the parental cells and can thus help tissue regeneration by moving their contents to neighboring cells.[157] Although clinical trials show that MSCs-based treatments are safe, feasible, and helpful, the limited sample size and absence of long-term follow-up make these trials unsatisfactory.

The cell and metabolic processes throughout the wound repair process are classified into four major phases: hemostasis, inflammatory, proliferative, and remodeling (or maturation phase).[158] As previously noted, MSC-derived exosomes (MSCs-exosomes) will have resource cell characteristics, encouraging tissue regeneration and self-healing of cells, reestablishing tissue homeostasis, and accelerating wound healing in damaged regions.[159] Emerging data suggests that MSCs produce bioactive substances to target tissues and cells via endocrine and paracrine pathways, lowering wound inflammation and enhancing tissue healing.[160] Exosomes produced by various MSCs can attenuate the inflammation by downregulating proinflammatory enzymes such as COX-2, inducible nitric oxide synthase, and chemokines such as chemoattractant protein (MCP)-1, TNF-α and IL-1β. Also, MSCs-exosomes can induce anti-inflammatory cytokine production called IL-10, which is known to be significant in regulating inflammation of cutaneous wounds and scar formation in numerous disease types.[161] Cell proliferation and re-epithelialization of the skin are critical for cutaneous regeneration. Skin fibroblasts are involved in the contraction of wounds, deposition of extracellular, remodeling of tissues, and other aspects of skin tissue healing and regeneration.[162]

It has been demonstrated that MSCs-exosomes control the expression of the genes linked with growth factors, which in turn controls the production and fibroblast migration. Therefore, it promotes the development of granulation tissue and collagen, facilitating structural support for wound healing.[163] Exosomes derived from Human fibrocytes include miRNAs and proteins along with various biological functions, and these exosomes improve the healing of wounds in diabetic rat models by promoting skin migration and proliferation of cells.[164] Studies also have demonstrated that, when transplanted to wound sites, exosomes derived from human amniotic epithelial cells (hAEC-exosomes) were shown to fasten the process of re-epithelialization and wound closure.[165]

One of the mesenchymal stem cells (MSCs) is adipose-derived stem cells (ADSC), which have a variety of origins, are simple to isolate and amplify, and are less immunogenic. It is commonly recognized that ADSC can accelerate wound healing by controlling the activity of several effector cells involved in the healing process. Exosomes that are derived from ADSCs (ADSC-exosomes) have also shown that they can control the synthesis of collagen at various stages of the wound healing process by simply speeding up the healing of a wound by an increase in type I and type III collagen production during the early stage and preventing the synthesis of collagen in a late stage, and thus decreasing the formation of scar.[166] *In vivo*, MSCs administration using the scaffolds of collagen hybrid improved the deposition of collagen as well as angiogenesis in diabetic wound healing.[40] Polarized macrophages are formed into two different phenotypes, M1 (proinflammatory) or M2, in response to activation cues (anti-inflammatory). Evidence suggests that macrophages, which are M2, can express mediators involved in the inflammation resolution and remodeling of tissue and hence enhance healing of the wound (Figure 6).[167]

MSCs' differentiation capacities and possible additional qualities, such as inducing the release of anti-inflammatory and pro-angiogenic mediators, emphasize their importance in wound healing and skin regeneration.

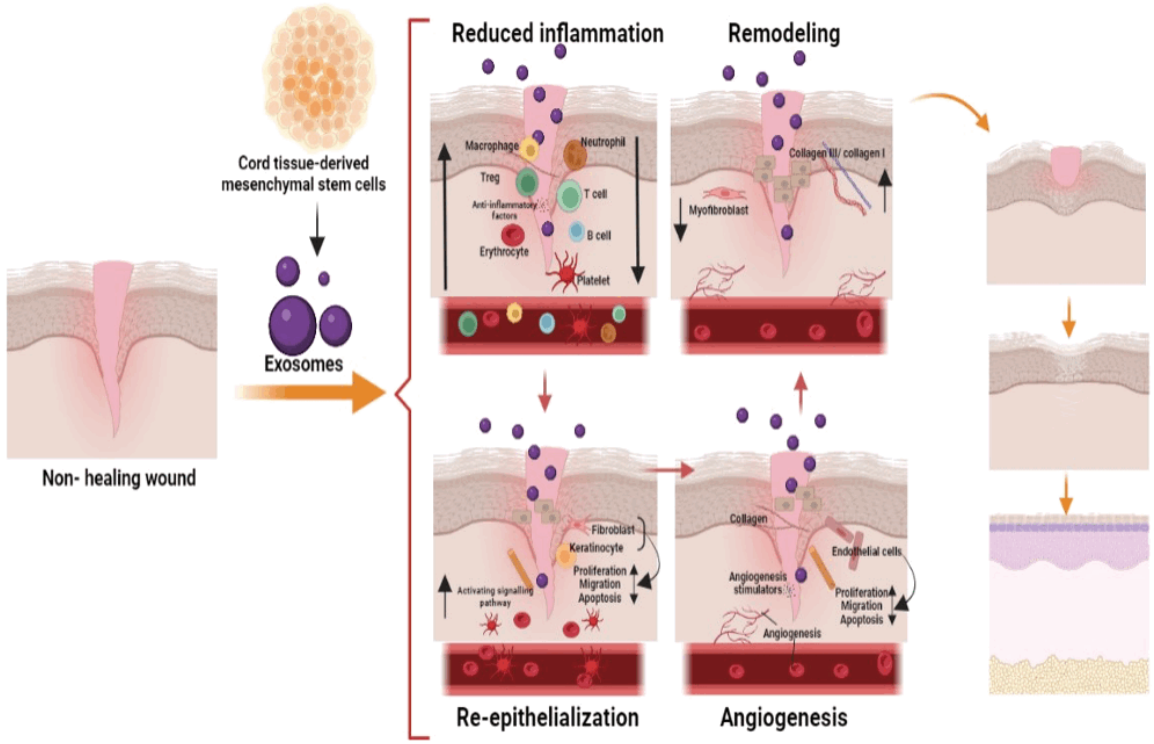


***Figure 6**. Exosomes in wound healing.*

Exosomes for the treatment of diabetes

A global metabolic disorder called diabetes mellitus (DM) is marked by a lack of insulin production, increasing blood glucose levels. Diabetes is frequently associated with several macrovascular problems, including coronary heart disease, macrovascular arteriosclerosis, microvascular disease, and hypertension. MSCs-EVs have been shown to offer outstanding healing effects when it comes to treating various organ damages as the MSCs-EVs have demonstrated sound healing effects in a variety of tissue injuries, including cardiovascular, liver, and skin wounds, which include processes of angiogenesis, cell proliferation regulation, and immunological modulation.[168]

Type 1 and type 2 are the two major types of diabetes, and both are associated with some degree of destruction of pancreatic islet cells. Diabetic patients may benefit from a potential novel medication that promotes the regeneration of cells of islets and improves the susceptibility of targeted insulin tissues.[169] Deficiency of Insulin is caused by type 1 diabetes, which destroys autoimmune tissues, particularly in the islet cells of patients. Furthermore, it has been shown that MSCs-EVs have the exclusive capacity to manage immune repair in pancreatic islet tissue to treat diabetes. MSCs-Exos can chemo-attract pancreatic tissue and activate the pancreatic and duodenal homeobox 1 pathway, further stimulating pancreatic β-cell regeneration and insulin production.[170] Many research reports have revealed that MSCs-Exos increase the regulatory T-cell levels and factors like IL-4. IL-10 and transforming growth factor (TGF). Thus, it ultimately improves diabetic mice's autoimmune response and islet regeneration. Type 2 diabetes is caused by the reduced insulin sensitivity of the peripheral tissues, which is also associated with insulin secretion defectiveness by pancreatic beta cells. The expression of muscle transporter 4 results in the reversal of peripheral insulin resistance, as well as the reduction of islet cell apoptosis and the restoration of insulin secretion function, according to studies showing the ability of MSCs-Exos to reinstate the phosphorylation protein kinase B and insulin receptor substrate 1.

As mentioned, a patient suffering from diabetes is generally associated with various other complications, one of which is diabetes ulcers, which majorly occur on the feet of patients, also called diabetic feet. MSCs-Exos improves the polarization ratio of M2/M1 (macrophage), which reduces inflammation and accelerates healing in diabetic ulcer wounds. The extent of wound vascularization impacts the wound's healing pace and remodeling. Also, as mentioned, MSCs-Exos are abundant in therapeutic noncoding RNAs and growth factors, and they may efficiently enhance the vascularization of wound skin while being more stable and safer than cell treatment. MSCs-Exos have also been shown to improve renal function, restore podocyte function, and postpone renal fibrosis.[170]

MSCs-Exos include a wide range of proteins for repair, growth cytokines, and noncoding RNAs. These have therapeutic effects that can improve organ repair done by diabetes, and their consequences aid in controlling vascularization, anti-apoptotic, and inflammation processes. MSCs-Exos may be a successful therapy method for diabetes and its associated consequences. Diabetes is frequently associated with several macrovascular problems, including hypertension, coronary heart disease, macrovascular, microvascular, and arteriosclerosis issues. MSCs-Exos have been shown to offer outstanding healing effects when it comes to treating various organ damages

as the MSCs-Exos have demonstrated good healing effects in a variety of tissue injuries, including cardiovascular, liver, and wounds on the skin, which include processes of angiogenesis, cell proliferation and regulation along with immunological modulation.

Conclusion

Therefore, stem cell therapies for various illnesses are increasingly being considered in tandem with the investigation of multiple types of stem and progenitor cells. The field of transplanting human embryonic or post-natal pluripotent stem cell-derived cells is rapidly expanding, and there have been encouraging outcomes in the grafting of dopaminergic neurons obtained from embryonic stem cells. This is because most scenarios currently cannot be imagined utilizing lineage-specific, tissue-resident natural stem cells. Does the future hold promise for MSCs as therapeutic cells? MSC treatments have been effective in maintaining patients' health. However, the safety of ES- or iPS-cell-based therapy has not yet been established. Exosome research and interest in exosome roles in disease pathology and possible treatments have led to exponential growth in the exosome field; however, inconsistent exosome collection, isolation, and analysis methods have posed a major obstacle to the field's quick progress. The International Society for Extracellular Vesicles (ISEV) has released a policy statement addressing these concerns and providing recommendations to investigators to avoid discrepancies in exosome and EV research.

Nonetheless, MSC therapies have been effective in maintaining patients' health. Examining the mechanism of exosomes in treating diseases is the primary connection to future clinical research. Exosomes are a cutting-edge therapeutic concept with unique advantages. MSC exosomes are signaling molecules that work similarly to MSCs but have a more robust membrane structure. Exosomes produced from MSCs are more immunogenic and well-tolerated than whole-cell therapy. There are more options for treating illnesses thanks to these advantages. There are several limitations to exosome manufacturing and purification, and more studies are required to determine the exact nature of their therapeutic benefits. Exosomes need to clear up a few things before they may be used in a clinical context. Detailed characterization will be required to determine the different exosome subpopulations. Exosomes have a wealth of evidence supporting their potential as novel therapeutic agents, but this needs to be clinically confirmed.

References

[1] Jacques G, et Luc Sensébé. Mesenchymal stromal cells: Clinical challenges and therapeutic Opportunities. Cell Stem Cell. 2018; 22(6): 824-33. doi.org/10.1016/j.stem.2018.05.004.

[2] Tiziana S, *et al*. Clinical trials with mesenchymal stem cells: An update. Cell Transplantation. 2016; 25(5): 829-48. doi.org/10.3727/096368915X689622.

[3] Hass R, Kasper C, Bohm S, Jacobs R. Different populations and sources of human mesenchymal stem cells (MSCs): a comparison of adult and neonatal tissue-derived MSCs. Cell Commun Signal. 2011; 9 (1): 12.

[4] Maqsood M, *et al*. Adult mesenchymal stem cells and their exosomes: Sources, characteristics, and application in regenerative medicine. Life Sciences. 2020; 256: 118002. doi.org/10.1016/j.lfs.2020.118002.

[5] Yuanxia Z, *et al*. Mesenchymal stem cell-derived extracellular vesicles/exosome: A promising therapeutic strategy for Intracerebral hemorrhage. Regen Ther. 2023; 22: 181-90. doi.org/10.1016/j.reth.2023.01.006.

[6] He L, *et al*. Bone marrow mesenchymal stem cell-derived exosomes protect cartilage damage and relieve knee osteoarthritis pain in a rat model of osteoarthritis. Stem Cell Res Ther. 2020; 11(1): 276. doi: 10.1186/s13287-020-01781-w.

[7] Li T, *et al*. Mesenchymal stem cell-derived exosomal microRNA-3940-5p inhibits colorectal cancer metastasis by targeting integrin α6. Dig Dis Sci. 2021; 66(6): 1916-27. doi: 10.1007/s10620-020-06458-1.

[8] Gillian C, *et al*. Exosomes and other extracellular vesicles: The new communicators in parasite infections. Trends Parasitol. 2015; 31(10): 477-89. doi.org/10.1016/j.pt.2015.06.009.

[9] Peter W. The nature and significance of platelet products in human plasma. British J Haematol. 1967; 13(3): 269-88. doi.org/10.1111/j.1365-2141.1967.tb08741.x.

[10] Johnstone RM, *et al*. Vesicle formation during reticulocyte maturation. Association of plasma membrane activities with released vesicles (exosomes). J Biol Chem. 1987; 262(19): 9412-20. doi.org/10.1016/S0021-9258(18)48095-7.

[11] Abbaszadeh H, *et al*. Human umbilical cord mesenchymal stem cell-derived extracellular vesicles: a novel therapeutic paradigm. J Cell Physiol. 2020; 235(2): 706-17. doi: 10.1002/jcp.29004.

[12] Théry C, Witwer KW, Aikawa E, Alcaraz MJ, Anderson JD, Andriantsitohaina R, *et al*. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): A position statement of the International Society for Extracellular Vesicles and Update of the MISEV2014 Guidelines. J Extracell Vesicles. 2018: 23; 7(1): 1535750. doi: 10.1080/20013078.2018.1535750.

[13] Lötvall J, Hill AF, Hochberg F *et al*. Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the International Society for Extracellular Vesicles. J Extracell Vesicles. 2014; 22(3): 26913. doi.org/10.3402/jev.v3.26913.

[14] Pap E, *et al*. Highlights of a New Type of Intercellular Communication: Microvesicle-Based Information Transfer. Inflamm Res. 2009; 58(1): 1-8. doi.org/10.1007/s00011-008-8210-7.

[15] Sonam G, *et al*. The exosome journey: From biogenesis to uptake and intracellular signaling. Cell Commun Signal. 2021; 19(1): 47. doi.org/10.1186/s12964-021-00730-1.

[16] Szabo G, Momen-Heravi F. Extracellular vesicles and exosomes: biology and pathobiology. The Liver:

Biology and pathobiology. Wiley, New York, 2020, pp 1022-7. doi.org/10.1002/9781119436812.ch78

[17] Antonelou MH, Seghatchian J. Update on extracellular vesicles inside red blood cell storage units: adjust the sails closer to the new wind. Transfus Apher Sci. 2016; 55(1): 92-104. doi: 10.1016/j.transci. 2016.07.016.

[18] Michelle M.and Leonard J. FedExosomes: Engineering therapeutic biological nanoparticles that truly deliver. Pharmaceuticals. 2013; 6(5): 659-80. doi.org/10.3390/ph6050659.

[19] Pluchino S. and Smith JA. Explicating exosomes: reclassifying the rising stars of intercellular communication. Cell. 2019; 177(2): 225-7. doi: 10.1016/j.cell.2019.03.020.

[20] Mathivanan S, Fahner CJ, Reid GE, and Simpson RJ. ExoCarta. Database of exosomal proteins, RNA and lipids. Nucleic Acids Res. 2012; 40(D1): D1241-4, 2012. doi: 10.1093/nar/gkr828.

[21] Simpson RJ, Lim JWE, Moritz RL, and Mathivanan S. Exosomes: proteomic insights and diagnostic potential. Expert Rev Proteomics. 2014; 6(3): 267-83. doi: 10.1586/epr.09.17.

[22] Munir H, Ward LSC, McGettrick HM. Mesenchymal stem cells as endogenous regulators of inflammation in stromal immunology. Adv Exp Med Biol. 2018; 1060: 73-98. doi: 10.1007/978-3-319-78127-3_5.

[23] Tanaka Y, Kamohara H, Kinoshita K, *et al*. Clinical impact of serum exosomal microRNA-21 as a clinical biomarker in human esophageal squamous cell carcinoma. Cancer. 2013; 119(6): 1159-67. doi: 10.1002/cncr.27895.

[24] Raposo G, Stoorvogel W. Extracellular vesicles: exosomes, microvesicles, and friends. J Cell Biol. 2013; 200(4): 373-83. doi: 10.1083/jcb.201211138.

[25] Edwin van der Pol, Böing AN, Harrison P, Sturk A, Nieuwland R. Classification, functions, and clinical relevance of extracellular vesicles. Pharmacological Reviews. 2012; 64(3): 676-705. doi: 10.1124/pr.112.005983.

[26] Lee TH, D'Asti E, Magnus N, Al-Nedawi K, Meehan B, and Rak J. Microvesicles as mediators of intercellular communication in cancer—the emerging science of cellular 'debris'. Semin Immunopathol. 2011; 33(5): 455-67. doi 10.1007/s00281-011-0250-3 10.

[27] Corbeil D, Marzesco AM, Wilsch-Brauninger M, Huttner WB. (2010). The intriguing links between prominin-1 (CD133), cholesterol-based membrane microdomains, remodeling of apical plasma membrane protrusions, extracellular membrane particles, and (neuro)epithelial cell differentiation. FEBS Letters. 2010; 584(9): 1659-64. doi: 10.1016/j.febslet.2010.01.050.

[28] Lai RC, Chen TS, Lim SK. Mesenchymal stem cell exosome: a novel stem cell-based therapy for cardiovascular disease. Regen Med. 2011; 6(4): 481-92. doi: 10.2217/ rme.11.35.

[29] Yuan Z, *et al*. Exosomes: Biogenesis, biologic function and clinical potential. Cell Biosci. 2019; 9(1): 19. doi.org/10.1186/s13578-019-0282-2.

[30] Clifford V. Harding, John E. Heuser, Philip D. Stahl. Exosomes: Looking back three decades and into the future. J Cell Biol. 2013; 200(4): 367-71. doi: 10.1083/jcb.201212113.

[31] Harding C, Heuser J, Stahl P. Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes. J Cell Biol. 1983; 97, 329-39. doi: 10.1083/jcb.97.2.329.

[32] Pan BT, Johnstone RM. (1983). Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro: Selective externalization of the receptor. Cell. 1983; 33: 967-78. doi: 10.1016/0092-8674(83)90040-5.

[33] Kalra H, Drummen GP, Mathivanan S. Focus on extracellular vesicles: Introducing the next small big thing. International Journal of Molecular Sciences. 2016; 17: 1-30. doi: 10.3390/ijms17020170.

[34] Volker S, Barz D. Characterization of Cellular and Extracellular Plasmamembrane Vesicles from a Low Metastatic Lymphoma (Eb) and Its High Metastatic Variant (ESb): Inhibitory Capacity in Cell-Cell Interaction Systems. Biochimica et Biophysica Acta (BBA)-Biomembranes. 1986; 860(2): 236-42. doi.org/10.1016/0005-2736(86)90519-5.

[35] Pan, B. T., & Johnstone, R. M. (1983). Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro: Selective externalization of the receptor. *Cell, 33,* 967–978.

[36] Johnstone RM, Mathew A, Mason AB, Teng K. Exosome formation during maturation of mammalian and avian reticulocytes: Evidence that exosome release is a major route for externalization of obsolete membrane proteins. Journal of Cellular Physiology. 1991; 147: 27-36. doi.org/10.1002/jcp.1041470105.

[37] Yin K, Wan, S, Zhao RC. Exosomes from mesenchymal stem/stromal cells: A new therapeutic paradigm. Biomark Res. 2019; 7: 1-8. doi.org/10.1186/s40364-019-0159-x.

[38] Harrell CR, *et al*. Mesenchymal stem cell-derived exosomes and other extracellular vesicles as new remedies in the therapy of inflammatory diseases. Cells. 2019; 8(12): 1605. doi: 10.3390/cells8121605.

[39] Lamparski HG, Metha-Damani A, Yao J-Y, Patel S, Hsu D-H, Ruegg C, *et al*. Production and characterization of clinical grade exosomes derived from dendritic cells. J Immunol Methods (2002) 270(2):211–26. doi: 10.1016/s0022-1759(02)00330-7.

[40] Min L, Zhu S, Chen L, Liu X, Wei R, Zhao L, *et al*. Evaluation of circulating small extracellular vesicles derived miRNAs as biomarkers of early colon cancer: A comparison with plasma total miRNAs. J Extracel Vesicles. 2019; 8: 1643670. doi:10.1080/20013078.2019.1643670.

[41] Wennink JWH, Liu Y, Mäkinen PI, *et al*. Macrophage selective photodynamic therapy by meta-tetra (hydroxyphenyl)chlorin loaded polymeric micelles: a possible treatment for cardiovascular diseases. Eur J Pharm Sci. 2017; 107 :112-125. doi:10.1016/j.ejps.2017.06.038.

[42] Pala RR, Anju VT, Dyavaiah M, Busi S, Nauli SM. Nanoparticle-mediated drug delivery for the treatment

of cardiovascular diseases. Int J Nanomedicine. 2020: 15: 3741-69. doi: 10.2147/IJN.S250872

[43] Wang R, Ding Q, Yaqoob U, et al. Exosome adherence and internalization by hepatic stellate cells triggers Sphingosine 1-phosphate-dependent migration. J Biol Chem. 2015; 290: 30684-96. doi: 10.1074/jbc.M115.671735.

[44] Tickner JA, Urquhart AJ, Stephenson SA, Richard DJ, O'Byrne KJ. Functions and therapeutic roles of exosomes in cancer. Front Oncol. 2014; 4: 127. doi: 10.3389/fonc.2014.00127.

[45] Frydrychowicz Kolecka-Bednarczyk MA, Madejczyk M, Yasar MS, Dworacki G. Exosomes-structure, biogenesis and biological role in non-small-cell lung cancer. Scand J Immunol. 2015; 81(1): 2-10. doi: 10.1111/sji.12247.

[46] Subra C, Laulagnier K, Perret B, Record M. Exosome lipidomics unravels lipid sorting at the level of multivesicular bodies. Biochimie. 2007; 89: 205-12. doi: 10.1016/j.biochi.2006.10.014.

[47] Pfeffer SR. Two Rabs for exosome release. Nat Cell Biol. 2010; 12(1): 3-4. doi: 10.1038/ncb0110-3.

[48] Nabhan JF, Hu R, Oh RS, Cohen SN, Lu Q, Formation and release of arrestin domain-containing protein 1- mediated microvesicles (ARMMs) at plasma membrane by recruitment of TSG101 protein. Proc Natl Acad Sci. 2012; 109(11): 4146-51. doi: 10.1073/pnas.1200448109.

[49] Lou G, Chen GZ, Zheng M, Liu Y. Mesenchymal stem cell-derived exosomes as a new therapeutic strategy for liver diseases. Ex Mo Med. 2017; 49(6): e346.

[50] Schey KL, Luther JM, Rose KL. Proteomics characterization of exosome cargo. Methods. 2015; 87: 75-82. doi: 10.1016/j.ymeth.2015.03.018.

[51] Tai YL, Chen KC, Hsieh JT, Shen TL. Exosomes in cancer development and clinical applications. Cancer Sci. 2018; 109(8): 2364-74. doi: 10.1111/cas.13697

[52] Zhang H, Freitas D, Kim HS, *et al*. Identification of distinct nanoparticles and subsets of extracellular vesicles by asymmetric flow field-flow fractionation. Nat Cell Biol. 2018; 20(3): 332-43. doi: 10.1038/s41556-018-0040-4.

[53] Cosenza S, Toupet K, Maumus M, *et al*. Mesenchymal stem cells-derived exosomes are more immunosuppressive than microparticles in inflammatory arthritis. Theranostics. 2018; 8(5): 1399-410. doi: 10.7150/thno.21072.

[54] K. Yin, S. Wang, and R. C. Zhao, "Exosomes from mesenchymal stem/stromal cells: a new therapeutic paradigm," Biomarker Research. 2019; 7(1): 8.

[55] Xie Y, Dang W, Zhang S, *et al*. The role of exosomal noncoding RNAs in cancer. Mol Cancer. 2019; 18: 37.

[56] Li SP, Lin ZX, Jiang XY, Yu XY. Exosomal cargo-loading and synthetic exosome-mimics as potential therapeutic tools. Acta Pharmacol Sin. 2018; 39: 542-51.

[57] Fan XL, Zhang Y, Li X, Fu QL. Mechanisms underlying the protective effects of mesenchymal stem cell-based therapy. CMLS. 2020; 77(14): 2771-94. doi: 10.1007/s00018-020-03454-6.

[58] Mikael S, Raposo G. Exosomes - Vesicular carriers for intercellular communication. Curr Opin Cell Biol. 2009; 21(4): 575-81. doi.org/10.1016/j.ceb.2009.03.007.

[59] Williams RL, Urbé S. The emerging shape of the ESCRT machinery. Nat Rev Mol Cell Biol. 2007; 8(5): 355-68. doi.org/10.1038/nrm2162.

[60] Jatta H, Helenius A. Endosome maturation: Endosome maturation. EMBO J. 2011; 30(17): 3481-500. doi.org/10.1038/emboj.2011.286.

[61] Mashouri L, Yousefi H, Aref AR, Ahadi AM, Molaei F, and Alahari SK. Exosomes: composition, biogenesis, and mechanisms in cancer metastasis and drug resistance. Mol Cancer. 2019; 18: 75.

[62] Harding C, Heuser J, Stahl P. Endocytosis and intracellular processing of transferrin and colloidal gold-transferrin in rat reticulocytes: demonstration of a pathway for receptor shedding. Eur J Cell Biol. 1984; 35: 256-63

[63] Henne WM, Buchkovich NJ, Emr SD. The ESCRT pathway. Dev. Cell. 2011; 21: 77-91. doi: 10.1016/j.devcel.2011.05.015.

[64] Hirano S, Kawasaki M, Ura H, Kato R, Raiborg C, Stenmark H, *et al*. Double-sided ubiquitin binding of Hrs-UIM in endosomal protein sorting. Nat Struct Mol Biol 2006; 13: 272-7. doi: 10.1038/nsmb1051.

[65] Babst M, Sato TK, Banta LM, Emr SD. Endosomal transport function in yeast requires a novel AAA-type ATPase, Vps4p. EMBO J. 1997; 16: 1820-31. doi: 10.1093/emboj/16.8.1820.

[66] Thomas J, Fürthauer M. Biogenesis and function of ESCRT-dependent extracellular vesicles. Semin Cell Dev Biol. 2018; 74: 66-77. doi.org/10.1016/j.semcdb.2017.08.022.

[67] Katzmann DJ, Babst M, Emr SD. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. Cell. 2001; 106: 145-55. doi: 10.1016/s0092-8674(01)00434-2.

[68] Misra S. Hurley JH. Crystal structure of a phosphatidylinositol 3-phosphate-specific membrane-targeting motif, the FYVE domain of Vps27p. Cell. 1999; 97: 657-66. doi: 10.1016/s0092-8674(00)80776-x.

[69] Raiborg C, Bremnes B, Mehlum A, Gillooly DJ, D'Arrigo A, Stang E, *et al*. FYVE and coiled-coil domains determine the specific localization of Hrs to early endosomes. J Cell Sci. 2001b; 114: 2255-63. doi: 10.1242/jcs.114.12.2255.

[70] Stahelin RV, Long F, Diraviyam K, Bruzik KS, Murray D, Cho W. Phosphatidylinositol 3-phosphate induces the membrane penetration of the FYVE domains of Vps27p and Hrs. J Biol Chem. 2002; 277: 26379-88. doi: 10.1074/jbc.M201106200.

[71] Raiborg C, Bache KG, Mehlum A, Stang E, Stenmark H. Hrs recruits clathrin to early endosomes. EMBO J. 2001a; 20: 5008-21. doi: 10.1093/emboj/20.17.5008.

[72] Raiborg C, Bache KG, Gillooly DJ, Madshus IH, Stang E, Stenmark H. Hrs sorts ubiquitinated proteins into clathrin-coated microdomains of early endosomes. Nat. Cell Biol. 2002; 4: 394-8. doi: 10.1038/ncb791

[73] Raiborg C, Wesche J, Malerod L, Stenmark H. Flat clathrin coats on endosomes mediate degradative protein sorting by scaffolding Hrs in dynamic microdomains. J Cell Sci.2006; 119: 2414-24. doi: 10.1242/jcs.02978.

[74] Hurley JH. The ESCRT Complexes. Crit Rev Biochem Mol Biol. 2010; 45(6): 463-87. doi.org/10.3109/10409238.2010.502516.

[75] Raiborg C, Stenmark H. The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins. Nature. 2009; 458(7237): 445-52. doi.org/10.1038/nature07961.

[76] Babst M, Katzmann DJ, Estepa-Sabal EJ, Meerloo T, Emr SD. Escrt-III: an endosome-associated heterooligomeric protein complex required for mvb sorting. Dev Cell. 2002; 3: 271-82. doi: 10.1016/s1534-5807(02)00220-4.

[77] Chiaruttini N, Redondo-Morata L, Colom A, Humbert F, Lenz M, Scheuring S. Relaxation of loaded ESCRT-III spiral springs drives membrane deformation. Cell. 2015; 163: 866-79. doi: 10.1016/j.cell.2015.10.017.

[78] Henne WM, Stenmark H, Emr SD. Molecular mechanisms of the membrane sculpting ESCRT pathway. Cold Spring Harb Perspect Biol. 2013; 5: a016766. doi: 10.1101/cshperspect.a016766

[79] Agromayor M, Martin-Serrano J. Interaction of AMSH with ESCRT-III and deubiquitination of endosomal cargo. J Biol Chem. 2006; 281: 23083-91.

[80] Hurley JH. ESCRT complexes and the biogenesis of multivesicular bodies. Curr Opin Cell Biol. 2008; 20: 4-11. doi: 10.1016/j.ceb.2007.12.002.

[81] Dumas JJ, Merithew E, Sudharshan E, Rajamani D, Hayes S, Lawe D, *et al*. Multivalent endosome targeting by homodimeric EEA1. Mol Cell. 2001; 8(5): 947-58. doi.org/10.1016/S1097-2765(01)00385-9.

[82] Janas T, *et al*. Mechanisms of RNA Loading into Exosomes. FEBS Lett. 2015; 589: 1391-8. doi.org/10.1016/j.febslet.2015.04.036.

[83] Santosh P, et al. Regulation of exosome release by glycosphingolipids and flotillins. FEBS J. 2014; 281: 2214-27. doi.org/10.1111/febs.12775.

[84] Hemler ME. Tetraspanin proteins mediate cellular penetration, invasion, and fusion events and define a novel type of membrane microdomain. Annu Rev Cell Dev Biol. 2003; 19: 397-422. doi: 10.1146/annurev.cellbio.19.111301.153609.

[85] Kajimoto T, Okada T, Miya S, *et al*. Ongoing activation of sphingosine 1-phosphate receptors mediates maturation of exosomal multivesicular endosomes. Nat Commun. 2013; 4: 2712. doi: 10.1038/ncomms3712.

[86] Stuffers S, *et al*. Multivesicular endosome biogenesis in the absence of ESCRTs. Traffic. 2009; 10(7): 925-37. doi.org/10.1111/j.1600-0854.2009.00920.x.

[87] Baietti MF, Zhang Z, Mortier E, Melchior A, Degeest G, Geeraerts A, *et al*. Syndecan-syntenin-ALIX regulates the biogenesis of exosomes. Nat Cell Biol. 2012; 14: 677-85.

[88] Thompson CA, Purushothaman A, Ramani VC, Vlodavsky I, Sanderson RD. Heparanase regulates secretion, composition, and function of tumor cell-derived exosomes. J Biol Chem 2013; 288: 10093-9. doi: 10.1074/jbc.C112.444562.

[89] Colombo M, Moita C, Van Niel G, Kowal J, Vigneron J, Benaroch P, *et al*. Analysis of ESCRT functions in exosome biogenesis, composition and secretion highlights the heterogeneity of extracellular vesicles. J. Cell Sci. 2013; 126: 5553-65. doi: 10.1242/jcs.128868.

[90] Tamai K, Tanaka N, Nakano T, Kakazu E, Kondo Y, Inoue J, *et al*. Exosome secretion of dendritic cells is regulated by Hrs, an ESCRT-0 protein. Biochem. Biophys Res. Commun. 2010; 399: 384-90. doi: 10.1016/j.bbrc.2010.07.083.

[91] Razi M, Futter CE. Distinct roles for Tsg101 and Hrs in multivesicular body formation and inward vesiculation. Mol Biol Cell. 2006; 17: 3469-83. doi: 10.1091/mbc.e05-11-1054.

[92] Baietti M.F., Zhang Z, Mortier E, Melchior A, Degeest G, Geeraerts A, Ivarsson Y, Depoortere F, Coomans C, Vermeiren E, *et al*.(2012). Syndecansyntenin-ALIX regulates the biogenesis of exosomes. Nat. Cell Biol. 2012, 14, 677-85

[93] Banfer S, Schneider D, Dewes J, Strauss MT, Freibert SA, Heimerl T, *et al*. Molecular mechanism to recruit galectin-3 into multivesicular bodies for polarized exosomal secretion. Proc Natl Acad Sci USA. 2018; 115: E4396-E4405. doi: 10.1073/pnas.1718921115.

[94] Hanson PI, Cashikar A. Multivesicular body morphogenesis. Annu Rev Cell Dev Biol. 2012; 28: 337-62. doi: 10.1146/annurev-cellbio-092910-154152.

[95] Larios J, Mercier V, Roux A, Gruenberg J. ALIX- and ESCRT-III dependent sorting of tetraspanins to exosomes. J Cell Biol. 2020, 219(3): e201904113. doi.org/10.1083/jcb.201904113.

[96] Campsteijn C, Vietri M, Stenmark H. Novel ESCRT functions in cell biology: Spiraling out of control? Curr Opin Cell Biol. 2016; 41: 1-8. doi.org/10.1016/j.ceb.2016.03.008.

[97] Petsalaki E, Zachos G. Clks 1, 2 and 4 prevent chromatin breakage by regulating the Aurora B-dependent abscission checkpoint. Nat Commun. 2016; 7: 11451. doi.org/10.1038/ncomms11451.

[98] Wei D, Zhan W, Gao Y, Huang L, Gong R, Wang W, *et al*. RAB31 marks and controls an ESCRT-independent exosome pathway. Cell Res. 2021; 31: 157-77.

[99] Stuffers S, Wegner CS, Stenmark H, Brech A. Multivesicular endosome biogenesis in the absence of ESCRTs. Traffic. 2009; 10: 925-37. doi: 10.1111/j.1600-0854.2009.00920.x.

[100] Trajkovic K, Hsu C, Chiantia S, Rajendran L, Wenzel D, Wieland F, *et al*. Ceramide triggers budding of exosome vesicles into multivesicular endosomes. Science. 2008; 29; 319(5867): 1244-7.doi: 10.1126/science.1153124.

[101] Mazzeo C, Calvo V, Alonso R, Merida I, Izquierdo M. Protein kinase D1/2 is involved in the maturation of multivesicular bodies and secretion of exosomes in

T and B lymphocytes. Cell Death Differ. 2016: 23, 99-109. doi: 10.1038/cdd.2015.72.

[102] Laulagnier K, Grand D, Dujardin A, Hamdi S, Vincent-Schneider H, Lankar D, *et al*. PLD2 is enriched on exosomes and its activity is correlated to the release of exosomes. FEBS Lett. 2004; 572: 11-14. doi: 10.1016/j.febslet.2004.06.082.

[103] Ghossoub R, Lembo F, Rubio A, Gaillard CB, Bouchet J, Vitale N, *et al*. Syntenin-ALIX exosome biogenesis and budding into multivesicular bodies are controlled by ARF6 and PLD2. Nat Commun. 2014; 5: 3477. doi: 10.1038/ncomms447.

[104] Chairoungdua A, Smith DL, Pochard P, Hull M, Caplan MJ. Exosome release of beta-catenin: A novel mechanism that antagonizes Wnt signaling. J Cell Biol. 2010; 190: 1079-91. doi: 10.1083/jcb.201002049.

[105] Van Niel G, Charrin S, Simoes S, Romao M, Rochin L, Saftig P, *et al*. The tetraspanin CD63 regulates ESCRT independent and -dependent endosomal sorting during melanogenesis. Dev Cell. 2011; 21: 708-21.

[106] Hurwitz SN, Nkosi D, Conlon MM, York SB, Liu X, Tremblay DC, *et al*. CD63 regulates Epstein-Barr virus LMP1 exosomal packaging, enhancement of vesicle production, and noncanonical NF-kappaB signaling. J Virol. 2017; 91(5): e02251-16. doi: 10.1128/JVI.02251-16.

[107] Edgar JR, Manna PT, Nishimura S, Banting G, Robinson MS. Tetherin is an exosomal tether. Elife 2016;5: e17180. doi.org/10.7554/eLife.17180.

[108] Savina A, Vidal M, Colombo MI. The exosome pathway in K562 cells is regulated by Rab11. J Cell Sci. 2002; 115: 2505-15. doi: 10.1242/jcs.115.12.2505.

[109] Savina A, Fader CM, Damiani MT, Colombo MI. Rab11 promotes docking and fusion of multivesicular bodies in a calcium-dependent manner. Traffic. 2005; 6: 131-43. doi: 10.1111/j.1600-0854.2004.00257.x.

[110] Messenger SW, Woo SS, Sun ZZ, Martin TFJ. A Ca^{2+}-stimulated exosome release pathway in cancer cells is regulated by Munc13-4. J Cell Biol. 2019; 218: 1422. doi: 10.1083/jcb.201710132.

[111] Ruiz-Martinez M, Navarro A, Marrades RM, Vinolas N, Santasusagna S, Munoz C, et al. YKT6 expression, exosome release, and survival in non-small cell lung cancer. Oncotarget. 2016; 7: 51515-24. doi: 10.18632/oncotarget.9862.

[112] Hyenne V, Apaydin A, Rodriguez D, Spiegelhalter C, Hoff-Yoessle S, Diem M, et al. RAL-1 controls multivesicular body biogenesis and exosome secretion. J Cell Biol. 2015; 211: 27-37. doi: 10.1083/jcb.201504136.

[113] Fader CM, Sanchez DG, Mestre MB, Colombo MI. TIVAMP/VAMP7 and VAMP3/cellubrevin: Two v-SNARE proteins involved in specific steps of the autophagy/multivesicular body pathways. Biochim Biophys Acta. 2009; 1793: 1901-16. doi: 10.1016/j.bbamcr.2009.09.011.

[114] Ostrowski M, Carmo NB, Krumeich, S Fange I, Raposo G, Savina A, *et al*. Rab27a and Rab27b control different steps of the exosome secretion pathway. Nat Cell Biol. 2010; 12: 19-30. doi: 10.1038/ncb2000.

[115] Zhao R, Chen X, Song H, Bie B, Zhang B. Dual role of MSCs-derived exosomes in tumor development. Stem Cells Intl. 2020; 2020: 88447302020 doi: 10.1155/2020/8844730.

[116] Tschuschke M, Kocherova I, Bryja A, Mozdziak P, Volponi AA, Janowicz K, *et al*. Inclusion biogenesis, methods of isolation and clinical application of human cellular exosomes. J Clin Med. 2020; 9(2): 436. doi.org/10.3390/jcm9020436.

[117] Miyanishi M, Tada K, Koike M, Uchiyama Y, Kitamura T, Nagata S. Identification of Tim4 as a phosphatidylserine receptor. Nature. 2007; 450(7168); 435-9. doi: 10.1038/nature06307.

[118] Simons M, Raposo G. Exosomes-vesicular carriers for intercellular communication. Curr Opin Cell Biol. 2009; 21: 575-81. doi.org/10.1016/j.ceb.2009.03.007.

[119] Mathivanan S, Lim JW, Tauro BJ, Ji H, Moritz RL, Simpson RJ. Proteomics analysis of A33 immunoaffinity-purified exosomes released from the human colon tumor cell line LIM1215 reveals a tissue-specific protein signature. Mol Cell Proteomics. 2010; 9(2): 197-208. doi: 10.1074/mcp.M900152-MCP200

[120] Feng D, Zhao WL, Ye YY, Bai XC, Liu RQ, Chang LF, *et al*. Cellular internalization of exosomes occurs through phagocytosis. Traffic. 2010; 11: 675-87. doi.org/10.1111/j.1600-0854.2010.01041.

[121] Wang M, Yuan Q, Xie L. Mesenchymal stem cell-based immunomodulation: properties and clinical application. Stem Cells Int. 2018; 2018: 3057624. doi: 10.1155/2018/3057624.

[122] Gao F, Chiu SM, Motan DA, Zhang Z, Chen L, Ji HL, *et al*. Mesenchymal stem cells and immunomodulation: current status and future prospects. Cell Death Dis. 2016; 7: e2062. doi: 10.1038/cddis.2015.327.

[123] Gomzikova MO, James V, Rizvanov AA. Therapeutic application of mesenchymal stem cells derived extracellular vesicles for immunomodulation. Front Immunol. 2019; 10: article 2663. doi.org/10.3389/fimmu.2019.02663

[124] Zhang B, Yin Y, Lai RC, Tan SS, Choo AB, Lim SK. Mesenchymal stem cells secrete immunologically active exosomes. Stem Cells Dev. 2014; 23: 123-44. doi: 10.1089/scd.2013.0479.

[125] Morrison TJ, Jackson MV, Cunningham EK, Kissenpfennig A, McAuley DF, O'Kane CM, *et al*. Mesenchymal stromal cells modulate macrophages in clinically relevant lung injury models by extracellular vesicle mitochondrial transfer. Am J Respir Crit Care Med. 2017; 196:1275-86. doi: 10.1164/rccm.201701-0170OC.

[126] Van den Akker F, Vrijsen KR, Deddens JC, Buikema JW, Mokry M, van Laake LW, *et al*. Suppression of T cells by mesenchymal and cardiac progenitor cells is partly mediated via extracellular vesicles. Heliyon. 2018; 4: e00642. doi: 10.1016/j.heliyon.2018.e00642.

[127] Casado JG, Blazquez R, Vela FJ, Alvarez V, Tarazona R, Sanchez-Margallo FM. Mesenchymal stem cell-derived exosomes: immunomodulatory evaluation

in an antigen-induced synovitis porcine model. Front Vet Sci. 2017; 4: 39. doi: 10.3389/fvets.2017.00039.

[128] Monsel A, Zhu YG, Gennai S, Hao Q, Hu S, Rouby JJ, *et al*. Therapeutic effects of human mesenchymal stem cell-derived microvesicles in severe pneumonia in mice. Am J Respir Crit Care Med. 2015; 192: 324-6. doi: 10.1164/rccm.201410-1765OC.

[129] Shigemoto-Kuroda T, Oh JY, Kim DK, Jeong HJ, Park SY, Lee HJ, *et al*. MSCs-derived extracellular vesicles attenuate immune responses in two autoimmune murine models: type 1 diabetes and uveoretinitis. Stem Cell Rep. 2017; 8: 1214-25. doi: 10.1016/j.stemcr.2017.04.008.

[130] Alzahrani FA, Saadeldin IM, Ahmad A, Kumar D, Azhar EI, Siddiqui AJ, *et al*. The potential use of mesenchymal stem cells and their derived exosomes as immunomodulatory agents for COVID-19 patients. Stem Cells Int. 2020; 2020: 8835986. doi: 10.1155/2020/8835986

[131] Carrasco E, Soto-Heredero G, Mittelbrunn M. The role of extracellular vesicles in cutaneous remodeling and hair follicle dynamics. Int J Mol Sci. 2019; 20(11): 2758. doi: 10.3390/ijms20112758.Solanas G,

[132] Solanas G, Benitah SA. Regenerating the Skin: A task for the heterogeneous stem cell pool and surrounding niche. Nat Rev Mol Cell Biol. 2013; 14: 737-48. doi: 10.1038/nrm3675.

[133] Wang B, Liu XM, Liu ZN, Wang Y, Han X, Lian AB, *et al*. Human hair follicle-derived mesenchymal stem cells: Isolation, expansion, and differentiation. World J Stem Cells. 2020; 12(6): 462-70. doi: 10.4252/wjsc.v12.i6.462.

[134] Millar SE. Molecular mechanisms regulating hair follicle development. J Investig Dermatol. 2002; 118: 216-25. doi: 10.1046/j.0022-202x.2001.01670.x.

[135] Gross JC, Chaudhary V, Bartscherer K, Boutros M. Active Wnt proteins are secreted on exosomes. Nat Cell Biol. 2012; 14: 1036-45. doi: 10.1038/ncb2574.

[136] McBride JD, Rodriguez-Menocal L, Guzman W, Candanedo A, Garcia-Contreras M, Badiavas EV. Bone marrow mesenchymal stem cell-derived CD63+ exosomes transport Wnt3a exteriorly and enhance dermal fibroblast proliferation, migration, and angiogenesis *In vitro*. Stem Cells Dev. 2017; 26: 1384-98. doi: 10.1089/scd.2017.0087.

[137] Zhang B, Wang M, Gong A, Zhang X, Wu X, Zhu Y, *et al*. HucMScs-exosome mediated-Wnt4 signaling is required for cutaneous wound healing. Stem Cells .2015; 33: 2158-68. doi: 10.1002/stem.1771.

[138] Zhang B, Wu X, Zhang X, Sun Y, Yan Y, Shi H, *et al*. Human umbilical cord mesenchymal stem cell exosomes enhance angiogenesis through the Wnt4/β-Catenin pathway. Stem Cells Transl Med. 2015; 4: 513-22. doi: 10.5966/sctm.2014-0267.

[139] Chen Q, Takada R, Noda C, Kobayashi S, Takada S. Different populations of Wnt-containing vesicles are individually released from polarized epithelial cells. Sci Rep. 2016; 6: 35562. doi: 10.1038/srep35562.

[140] Rajendran RL, Gangadaran P, Bak SS, Oh JM, Kalimuthu S, Lee HW, *et al*. Extracellular vesicles derived from MSCs activates dermal papilla cell *in vitro* and promotes hair follicle conversion from Telogen to Anagen in mice. Sci Rep. 2017; 7: 155604. doi: 10.1038/s41598-017-15505-3.

[141] Zhou L, Wang H, Jing J, Yu L, Wu X, Lu Z. Regulation of hair follicle development by exosomes derived from dermal papilla cells. Biochem Biophys Res Commun. 2018; 500(2): 325-32. doi: 10.1016/j.bbrc.2018.04.067.

[142] Ahmed MI., Alam M, Emelianov VU, Poterlowicz K, Patel A, Sharov AA, *et al*. MicroRNA-214 controls skin and hair follicle development by modulating the activity of the Wnt pathway. J Cell Biol. 2014; 207: 549-67. doi: 10.1083/jcb.201404001.

[143] Li X, Liu L, Yang J, Yu Y, Chai J, Wang L, *et al*. Exosome derived from human umbilical cord mesenchymal stem cell mediates MiR-181c attenuating burn-induced excessive inflammation. EBioMedicine. 2016; 8: 72-82. doi: 10.1016/j.ebiom.2016.04.030.

[144] Tao SC, Guo SC, Li M, Ke QF, Guo YP, Zhang CQ. Chitosan wound dressings incorporating exosomes derived from microRNA-126-overexpressing synovium mesenchymal stem cells provide sustained release of exosomes and heal full-thickness skin defects in a diabetic rat model. Stem Cells Transl Med. 2017; 6: 736-47. doi: 10.5966/sctm.2016-0275.

[145] Dong L, Hao H, Xia L, Liu J, Ti D, Tong C, *et al*. Treatment of MSCs with Wnt1a-conditioned medium activates DP cells and promotes hair follicle regrowth. Sci Rep. 2014; 4: 5432. doi: 10.1038/srep05432.

[146] Gross JC, Chaudhary V, Bartscherer K, Boutros M. Active Wnt Proteins Are Secreted on Exosomes. Nat Cell Biol. 2012; 14: 1036-45.

[147] Bian D, Wu Y, Song G, Azizi R, Zamani A. The application of mesenchymal stromal cells (MSCs) and their derivative exosome in Skin wound healing: A comprehensive review. Stem Cell Res Ther. 2022; 13(1): 24. doi.org/10.1186/s13287-021-02697-9.

[148] Tottoli EM, Dorati R, Genta I, Chiesa E, Pisani S, Conti B. Skin wound healing process and new emerging technologies for skin wound care and regeneration. Pharmaceutics. 2020; 12(8): 735. doi: 10.3390/pharmaceutics12080735.

[149] Frykberg Robert G. Challenges in the treatment of chronic wounds. Adv Wound Care. 2015; 4(9): 560-82. doi: 10.1089/wound.2015.0635.

[150] Hocking AM, Gibran NS. Mesenchymal stem cells: paracrine signaling and differentiation during cutaneous wound repair. Exp Cell Res. 2010; 316(14): 2213-9. doi: 10.1016/j.yexcr.2010.05.009

[151] Rani S, Ritter T. The exosome-A naturally secreted nanoparticle and its application to wound healing. Adv Mater. 2016; 28(27): 5542-52. doi: 10.1002/adma.201504009.

[152] Stephens P, Thomas DW. The cellular proliferative phase of the wound repair process. J Wound Care. 2002; 11(7): 253-61. doi: 10.12968/jowc.2002.11.7.26421.

[153] Lai RC, Yeo RW, Lim SK. Mesenchymal stem cell

exosomes. Semin Cell Dev Biol. 2015; 40: 82-8. doi: 10.1016/j.semcdb.2015.03.00125765629.

[154] Hu P, Yang Q, Wang Q, Shi C, Wang D, Armato U, *et al*. Mesenchymal stromal cells-exosomes: A promising cell-free therapeutic tool for wound healing and cutaneous regeneration. Burns Trauma. 2019; 7: 38. doi: 10.1186/s41038-019-0178-8

[155] Yang J, Liu XX, Fan H, Tang Q, Shou ZX, Zuo DM, *et al*. Extracellular vesicles derived from bone marrow mesenchymal stem cells protect against experimental colitis via attenuating colon inflammation, oxidative stress and apoptosis. PLoS One. 2015; 10: e014055 14607447. doi: 10.1371/journal.pone.0140551

[156] Hu L, Wang J, Zhou X, Xiong Z, Zhao J, Yu R, *et al*. Exosomes derived from human adipose mensenchymal stem cells accelerates cutaneous wound healing via optimizing the characteristics of fibroblasts. Sci Rep. 2016; 6: 32993. doi: 10.1038/srep32993.

[157] Geiger A, Walker A, Nissen E. Human fibrocyte-derived exosomes accelerate wound healing in genetically diabetic mice. Biochem Biophys Res Commun. 2015; 467: 303-9. doi: 10.1016/j.bbrc.2015.09.166264 54169.

[158] Zhao B, Zhang Y, Han S, Zhang W, Zhou Q, Guan H *et al*. Exosomes derived from human amniotic epithelial cells accelerate wound healing and inhibit scar formation. J Mol Histol. 2017; 48: 121-32. doi: 10. 1007/s10735-017-9711-x28229263.

[159] Kou X, Xingtian Xu X, Chen C, Sanmillan ML, Cai T, Zhou Y. The Fas/Fap-1/Cav-1 complex regulates IL-1RA secretion in mesenchymal stem cells to accelerate wound healing. Science Translational Medicine. 2018; 10(432): eaai8524. doi: 10.1126/scitranslmed.aai8524.

[160] Marofi F, Alexandrovna KI, Margiana R, Bahramali M, Suksatan W, Abdelbasset WK, et al. MSCs and their exosomes: A rapidly evolving approach in the context of cutaneous wound therapy. Stem Cell Res Ther. 2012; 12(1): 597. doi.org/10.1186/s13287-021-02662-6.

[161] He X, Dong Z, Cao Y, Wang H, Liu S, Liao L, *et al*. MSCs-derived exosome promotes M2 polarization and enhances cutaneous wound healing. Stem Cells Int. 2019; 2019: 7132708 doi: 10.1155/2019/7132708.

[162] Xiong J, Hu H, Guo R, Wang H, Jiang H. Mesenchymal stem cell exosomes as a new strategy for the treatment of diabetes complications. Front Endocrinol. 2021; 12: 646233. doi.org/10.3389/fendo.2021.646233.

[163] Guay C, Regazzi R. Exosomes as new players in metabolic organ cross-talk. Diabetes Obes Metab. 2017; 19 Suppl 1: 137-46. doi: 10.1111/dom.13027.

[164] Mahdipour E, Salmasi Z, Sabeti N. Potential of stem cell-derived exosomes to regenerate β Islets through Pdx-1 dependent mechanism in a rat Model of type 1 diabetes. J Cell Physiol. 2019; 234: 20310-21. doi: 10.1002/jcp.28631.

[165] Nojehdehi S, Soudi S, Hesampour A, Rasouli S, Soleimani M, Hashemi SM. Immunomodulatory effects of mesenchymal stem cell-derived exosomes on experimental type-1 autoimmune diabetes. J Cell Biochem. 2018; 119: 9433-43. doi: 10.1002/jcb.27260.

[166] Sabry D, Marzouk S, Zakaria R, Ibrahim HA, Samir M. The effect of exosomes derived from mesenchymal stem cells in the treatment of induced type 1 diabetes mellitus in rats. Biotechnol Lett. 2020; 42: 1597-610. doi: 10.1007/s10529-020-02908-y.

[167] Sun Y, Shi H, Yin S, Ji C, Zhang X, Zhang B, *et al*. Human mesenchymal stem cell derived exosomes alleviate type 2 diabetes mellitus by reversing peripheral insulin resistance and relieving β-cell destruction. ACS Nano. 2018; 12: 7613-28. doi: 10.1021/acsnano.7b07643.

[168] Guay C, Regazzi R. Exosomes as new players in metabolic organ cross-talk. Diabetes Obes Metab. 2017; 19 Suppl 1: 137-46. doi: 10.1111/dom.13027.

[169] Dalirfardouei R, Jamialahmadi K, Jafarian AH, Mahdipour E. Promising effects of exosomes isolated from menstrual blood-derived mesenchymal stem cell on wound-healing process in diabetic mouse model. J Tissue Eng Regen Med. 2019; 13: 555-68. doi: 10.1002/term.2799.

[170] Rackov G, Garcia-Romero N, Esteban-Rubio S, Carrión-Navarro J, Belda-Iniesta C, Ayuso-Sacido A. vesicle-mediated control of cell function: The role of extracellular matrix and microenvironment. Front Physiol. 2018; 9: 651. doi: 10.3389/fphys.2018.00651.

Journal of Associated Medical Sciences 2024; 57(3): 155-165

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Exploring machine learning approaches for early diabetes risk prediction: A comprehensive examination of health indicators and models

Nihar Ranjan Panda[1*] Jatindra Nath Mohanty[2] Ruchi Bhuyan[1] Prasanta Kumar Raut[3] Manulata[4]

[1]Department of Medical Research, IMS & SUM Hospital, SOA Deemed to be university, Bhubaneswar, Odisha, India.
[2]School of Basic and Applied Science, Centurian University of Technology & Management, Bhubaneswar, Odisha, India.
[3]Trident academy of technology, Bhubaneswar, Odisha, India.
[4]All India Institute of Medical Science, Bibinagar, India.

ARTICLE INFO

Article history:
Received 31 May 2024
Accepted as revised 21 July 2024
Available online 23 July 2024

Keywords:
Health indicators, diabetes, classification, decision tree, random forest.

ABSTRACT

Background: The increased prevalence of morbidity and mortality associated with Type 2 diabetes is due to changing lifestyles, demanding improved disease management measures. To tackle this, scientists are increasingly looking to technological advances, notably machine learning, for illness prevention and management, particularly in non-communicable diseases. The emphasis is on establishing an early detection system to identify Type 2 diabetes risk factors, enabling prompt treatments and preventative steps to reduce the disease's rising prevalence.

Materials and methods: The research aimed to assess the association of diabetes class with health indicators. Five machine learning models were employed with cross-validation techniques to predict early diabetes risk. The performance matrices of the models were evaluated and compared with the existing work.

Results: In multivariate analysis, we found polyuria (β=3.492; Aor=32.872; 95% CI=11.09,97.35; p<0.001), polydipsia (β=-4.100; Aor=60.378; 95%CI=18.28,199.37; p<0.001), polyphagia (β=1.181; Aor=3.25; 95%CI=1.23,8.57; p=0.017), genital thrush (β=1.08; Aor=2.96; 95%CI=1.26,7.53; p=0.023), irritability (β=2.28; Aor=9.82; 95%CI=3.41,28.26; p<0.001), and partial paresis (β=1.2406; Aor=3.45; 95% CI=1.35,8.79; p=0.009) are the potential health risk indicators for positive diabetes class.

Conclusion: Using an interpretable feature learning approach for early diabetes prediction improves the use of global health data. This method forecasts hazards correctly and gives insights into influential aspects. As a result, a more proactive healthcare strategy is implemented, allowing for more prompt treatments and encouraging a more hopeful future by improving patient outcomes and lowering the total burden of diabetes on individuals and healthcare systems.

Introduction

Diabetes mellitus is a chronic condition that affects people of all ages. It can cause blindness, renal failure, amputation, heart failure, and stroke, among other serious effects.[1] After we eat, our bodies convert calories into glucose, and the insulin produced by the pancreas nettles the cells to absorb glucose. The pancreas would not produce enough insulin, or the body would not respond efficiently to the insulin produced, leading to increased glucose levels in the patient. The World Health Organization says that diabetic disease was the ninth most significant cause of death on the globe in 2019 and was number six in upper-middle-income countries. Diabetes is

∗ **Corresponding contributor.**
Author's Address: *Department of Medical Research, IMS & SUM Hospital, SOA Deemed to be university, Bhubaneswar, Odisha, India.*
E-mail address: niharpanda1994@gmail.com
doi: 10.12982/JAMS.2024.057
E-ISSN: 2539-6056

classified into three types: T1, T2, and GDM. Presence of Type 1 diabetes arises when the pancreas does not do its job correctly in producing enough insulin, and it frequently strikes kids, and teenagers.[2] Common symptoms include Excessive thirst, dry mouth, weight loss, impaired eyesight, and frequent urination. Individuals with Type 1 diabetes are more likely to develop heart disease. Type 2 diabetes is distinguished by cells that fail to react effectively to insulin, resulting in insulin resistance.[3-5] Type 2 diabetes accounts for approximately 90% of all diabetes cases worldwide. While less severe than Type 1, it can cause health problems, particularly in the kidneys, nerves, and eyes. Type 2 diabetes, which was previously only observed in adults, is increasingly impacting youngsters. Gestational diabetes, the third primary form, occurs in pregnant women who have no history of diabetes and causes blood sugar levels to rise.[6,7] It usually goes away after delivery, but up to 10% of cases might evolve into Type 2 diabetes later in life. Babies born to moms who have gestational diabetes face problems. Diabetes prevalence is highest in China and India.

Related work

Isfafuzzaman *et al.* used a semi-supervised model using extreme gradient boosting (XGBoost) to predict the insulin characteristics of the private dataset.[8] SMOTE and ADASYN techniques were used to address the class imbalance issue. Machine learning classification methods such as decision tree, SVM, Random Forest, Logistic Regression, KNN, and different ensemble approaches were employed to establish the best prediction outcomes. With 81% accuracy, 0.81 F1 coefficients, and an AUC of 0.84, the XGBoost classifier with the ADASYN technique produced the best results.B. Shamreen*et et al*. tested various machine learning classifiers for predicting Type 2 diabetes mellitus, including logistic regression, XGBoost, gradient boosting, decision trees, ExtraTrees, random forest, and Light Gradient Boosting Machine (LGBM). LGBM had the best accuracy among these classifiers, reaching 95.20% and outperforming the other methods.[9] Aishwarya *et al.* For improved diabetes categorization, research presents a diabetes prediction model that adds external parameters and regular criteria such as glucose, BMI, age, insulin, etc. When compared to the previous dataset, the new dataset improves classification accuracy. A pipeline model is also enforced to boost classification accuracy for diabetes prediction. The research describes using clustering, especially K-means clustering, on the dataset to categorize each patient as diabetes or non-diabetic. Clustering was performed using highly associated characteristics, such as glucose and age. Sandip *et al*. provided a diabetes prediction model that uses a variety of machine learning algorithms, including Logistic Regression, SVM, Nave Bayes, Random Forest, XGBoost, LightGBM, CatBoost, Adaboost, and Bagging.[10] CatBoost is the most successful ensemble approach tested, attaining a high accuracy rate of 95.4%. CatBoost also surpasses XGBoost, with a better AUC-ROC score of 0.99 vs. XGBoost's accuracy rate of 94.3% and AUC-ROC score of 0.98. Quan *et al.* used decision trees, random forests, and neural network approaches to predict diabetes mellitus, with random forests attaining the most remarkable accuracy (ACC=0.8084) when all characteristics were applied.[11] The research also utilized mRMR to pick characteristics and discovered that the first five variables (height, HDL, fasting glucose, breath, and LDL) predicted diabetes well using the Luzhou dataset. The first three characteristics (glucose, 2-hour serum insulin, and age) were chosen for the Pima Indians dataset.

The research aimed to evaluate and compare the effectiveness of five machine-learning models in predicting diabetes using lifestyle data from the NHANES database. The following are the study's primary conclusions and information:

1. **Comparison of model**
   - The researchers examined the predictive abilities of CATBoost, XGBoost, RF, LR, and SVM.
   - In terms of predictive performance, CATBoost beat the other models.
   - The models were most likely assessed using sensitivity, accuracy, precision, AUC, and ROC measures.
2. **Performance of CATboost**
   - CATBoost produced an AUC (Area Under the Curve) of 0.83 and an accuracy of 82.1%.
   - According to these criteria, CATBoost displayed a high degree of accuracy and discrimination in discriminating between those with and without diabetes.
3. **Contributing factors**
   - The study found that calorie, carbohydrate, and fat intake levels were the most important predictors of diabetes patients.
   - This means that the machine-learning algorithms, notably CATBoost, used nutritional data to produce accurate predictions.

Materials and methods

This work is based on the ''Early-Stage Diabetes Risk Prediction'' dataset from the University of California, Irvine (UCI) machine learning repository.[29] This dataset contains details of 520 people who report diabetes-related symptoms. It gives us information about people, including features that can cause diabetes to develop. This section describes the method of identifying diabetes using a machine learning technique. The prediction algorithm is initially built using a publicly accessible diabetes dataset. EDA is used to find significant patterns among parameters in discriminating between diabetes and non-diabetic patients. Basic preprocessing activities, such as addressing missing values, numerical representation of categorical data, and normalization, were carried out. After that, the dataset is partitioned into training and test sets, with the training set further subdivided into training and validation subsets. Feature selection minimizes dimensionality, using insights from EDA and an additional trees classifier to find key features. Several machine learning techniques are used to build the prediction model, which is then validated on the validation set. Various measures are used to evaluate performance, and model parameters are fine-tuned to

improve efficacy. The resulting model is then applied to the test set for additional assessment, concluding the development and evaluation of a machine learning-based diabetes detection system. The detailed roadmap of the work is shown in Figure 1. A logistic regression model was used to find potential risk factors for diabetes.[31]

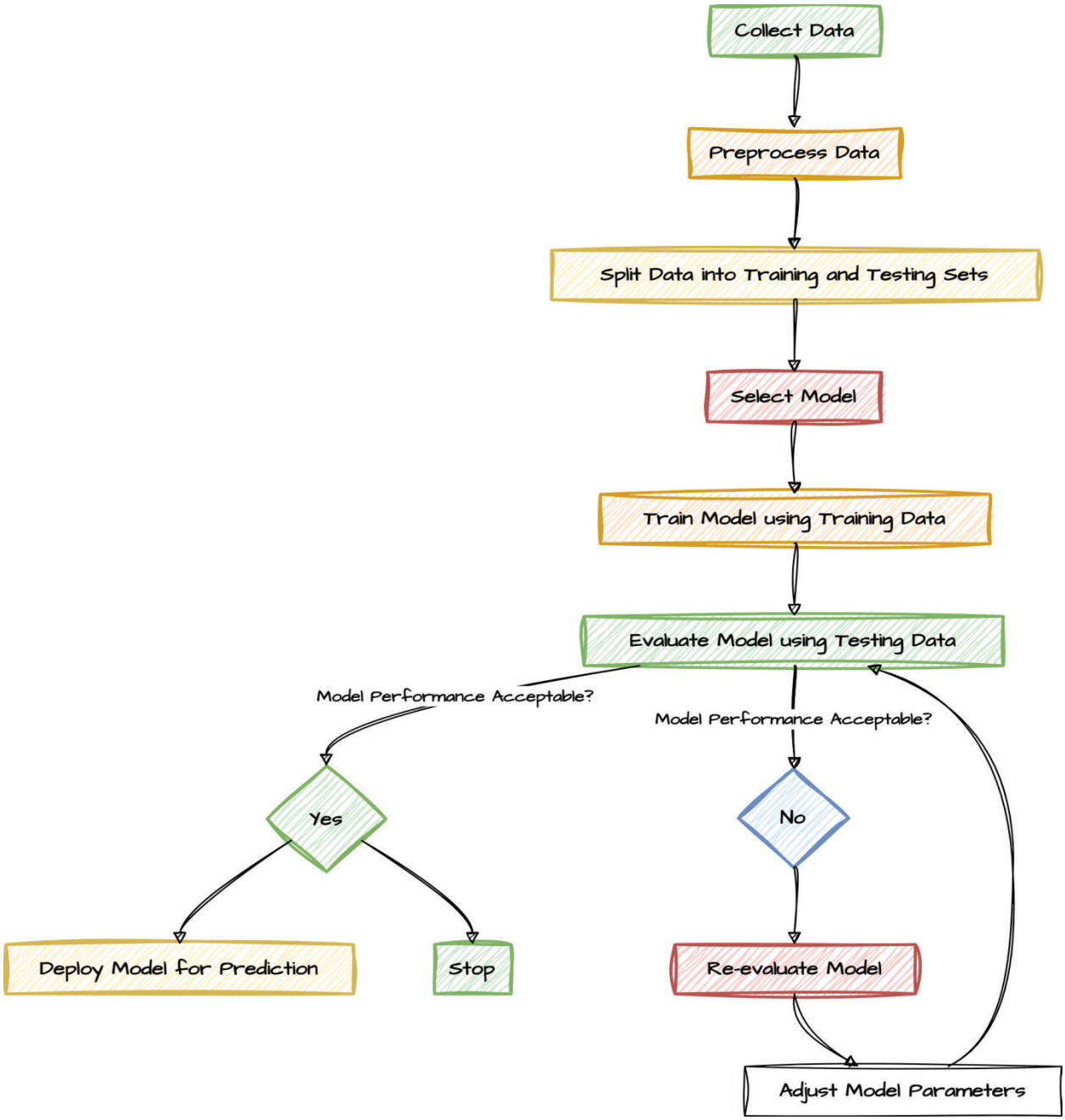


***Figure 1**. A roadmap for building a classification model for predicting type 2 diabetes.*

Statistical analysis and machine learning models

The data were expressed in frequency (%) for categorical variables and mean±SD for continuous variables. A chi-square test was applied to find the association between diabetes class and other parameters. For the mean difference between the two classes, we performed the test. The significance level was 0.05. After conducting an association between diabetes class and other parameters, we used the multivariate logistic regression method to find the most significant risk factors. The factors found to be substantial in the Univariate analysis were run in the final multivariate analysis. To predict early diabetes risk, we used five machine learning models. All the models were tested by their performance matrices.

Data processing and Eda

The effectiveness of a machine learning model is dependent on precise data preparation. It is critical for best model performance to analyze the dataset for redundancy, missing, or irrelevant data. The primary goal is to fine-tune the dataset for training and testing. Using

visualizations, such as charts, may help you understand how attributes affect the target variable. This astute data-driven strategy improves the model's capacity to generalize and uncover important patterns, resulting in robust prediction outcomes in cases such as diabetes diagnosis. The association between diabetes class and some health indicators is shown in Figure 2.

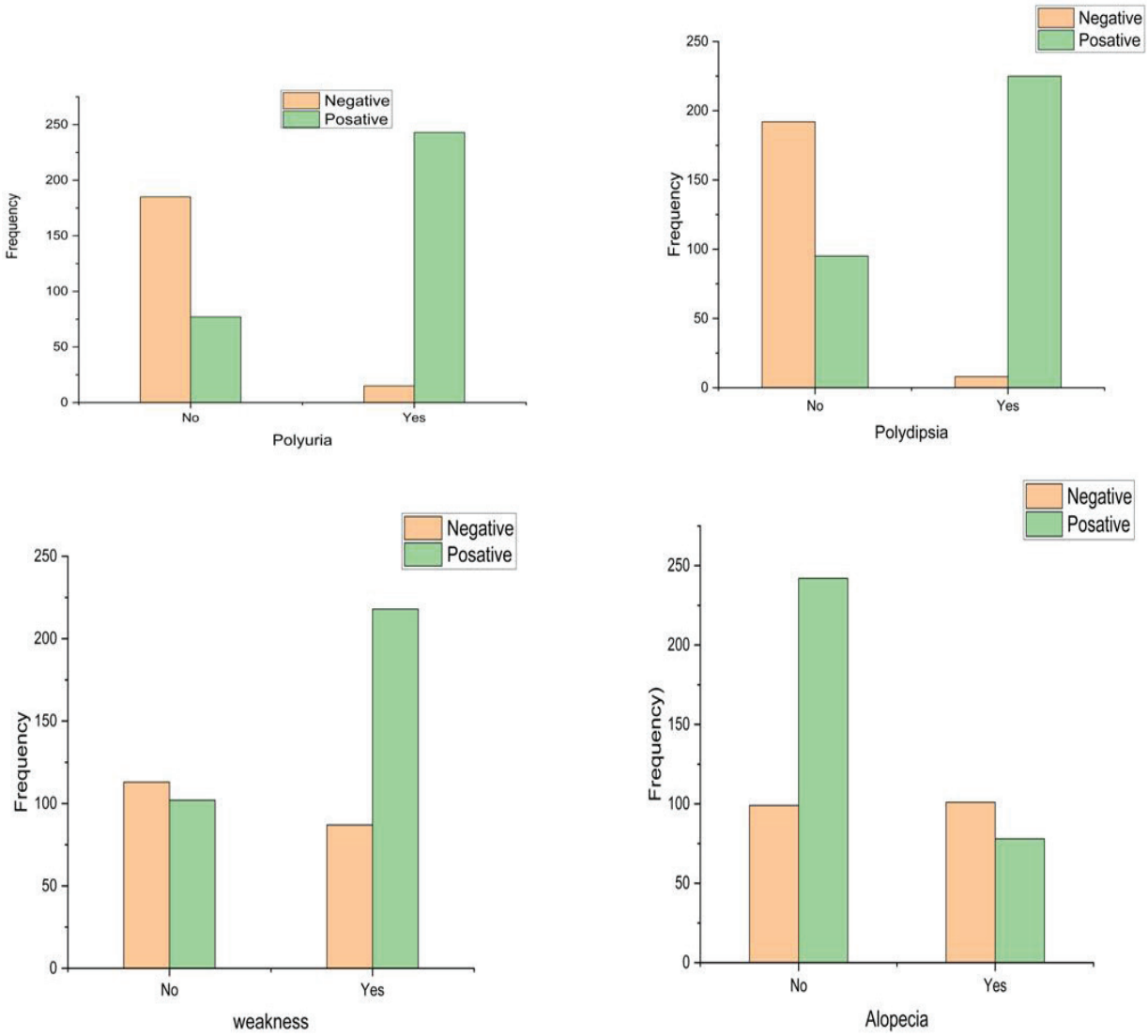


***Figure 2**. Association between diabetes classes with some health indicators.*

Boosting classification

Combining numerous weak learners into a robust ensemble model improves prediction performance.[12] Five boosting methods, including the Gradient Boosting Machine (GBM), were used in this study. GBM combines predictions from numerous decision trees to create powerful learners. For optimum splits, each tree's nodes employ unique feature subsets. Sequential trees rectify prior faults, constantly enhancing the model. GBM performance is further enhanced by hyperparameter optimization, resulting in a considerable improvement. GBM can generate accurate predictions thanks to this iterative learning process that leverages the collective wisdom of numerous trees, making it a valuable tool in machine learning applications.

Decision tree

A Decision Tree is a supervised learning technique with a tree-like structure that starts at the root node and branches out.[13] It makes judgments based on input features and is mainly used for categorization. The tree's leaf nodes indicate outcomes, while inside nodes include dataset attributes and branching decision criteria. The information gain, calculated for each characteristic at each node, directs feature selection, optimizing predicted accuracy. Grid Search CV and Randomized Search CV are used in hyperparameter optimization. The minimum leaf sample size (10), maximum depth (6), and the criteria ('gin) are all fine-tuned. This careful optimization procedure refines the structure of the Decision Tree, ensuring that it efficiently captures patterns and correlations within the data, boosting its prediction skills in a supervised learning setting.[14]

Random forest

Random Forest is an ensemble learning approach that builds a 'forest' of decision trees, each of which has been trained using the Bootstrap Aggregating (bagging) methodology.[15] The primary objective is to improve individual Decision Tree performance by minimizing variation and enhancing model unpredictability. Random Forest reduces overfitting and provides a more robust, generalizable model by developing many trees and using only random subsets of features at each split. Instead of focusing on the most crucial feature during node division, the method chooses the best feature from a random subset, boosting tree variety. Configuring parameters such as min_samples_split, min_samples_leaf, max_depth, max_features, n_estimators, and Bootstrap is part of the offered hyperparameter tuning. It is critical to fine-tune these hyperparameters to optimize the model's performance. However, the cost of computing is evident since the Randomized Search CV technique, used for hyperparameter tuning, takes a long time to execute due to its broad parameter space search. Regardless of the computational requirement, Random Forest is a robust and frequently used machine learning technique that boosts forecast accuracy and manages complicated datasets via its ensemble of varied decision trees.[16]

KNN classification

K-Nearest Neighbours (KNN) is a lazy learning algorithm that differs from eager learners, such as Random Forest, because it does not have a dedicated training phase.[17] KNN is built on the notion of 'feature similarity,' it works by categorizing incoming data points based on their closeness to existing training data. A Minkowski metric was used to measure distances during the model-building process, and a leaf size of 20 was chosen. The best accuracy was obtained after extensive parameter optimization using the Euclidean metric and setting the number of neighbors to four. This means that during the classification step, the algorithm identifies the class of a new data point in the feature space by considering the classes of its four nearest neighbors. KNN's adaptability makes it useful for various applications.[30]

Support vector classification

The Support Vector Machine (SVM) is a popular supervised machine learning technique representing each data point in n dimensions.[18] The data characteristics function as coordinates in this space in this method. SVM performs classification by determining which hyperplane best divides the two classes. A margin, defined as the distance between the decision border and the nearest points of each class, is determined to identify the best hyperplane. SVM chooses the hyperplane with the most significant margin. In other circumstances, however, precise class prediction precedes maximizing the margin. The selection of hyperparameters is critical for the overall accuracy of SVM since they have a major influence on its performance. Randomized Search CV, a method that effectively searches the parameter space, was used to optimize the hyperparameter.[19]

Results

Table 1 describes the association between diabetes class and health indicators. The table shows that the positive diabetes class in females is higher (54.1%) than in males (45.9%). Which is statistically significant (p<0.001). The number of polyuria cases is relatively higher, 243 (75.9%), in the positive diabetes class, compared to the negative diabetes class, 15 (7.5%). We found that polydipsia is strongly associated with diabetes class (p<0.001). It can be observed that the polydipsia cases are much higher, 225 (70.3%) as compared to negative diabetes class (4%). Similarly, sudden weight loss, weakness, polyphagia, visual blurring, irritability, partial paresis, and alopecia were found to be highly significant (p<0.001) with diabetes class. The proportion of all these symptoms is found to be higher in positive diabetes classes. However, delayed healing(p=0.284), obesity, and itching (p=0.760) were not associated with the diabetes class. The mean age of the positive diabetes class was 49.1±12.1, and the negative diabetes class was 46.4±12.1 (p=0.013).

Table 1. Features description of the data set.

Features	Levels	Negative (N=200)	Positive (N=320)	*p*
Sex	Female	19 (9.5)	173 (54.1)	<0.001
	Male	181 (90.5)	147 (45.9)	
Polyuria	No	185 (92.5)	77 (24.1)	<0.001
	Yes	15 (7.5)	243 (75.9)	
Polydipsia	No	192 (96)	95 (29.7)	<0.001
	Yes	8 (4)	225 (70.3)	
Sudden weight loss	No	171 (85.5)	132 (41.3)	<0.001
	Yes	29 (14.5)	188 (58.8)	
Weakness	No	113 (56.5)	102 (31.9)	<0.001
	Yes	87 (43.5)	218 (68.1)	
Polyphagia	No	152 (76.0)	131 (40.9)	<0.001
	Yes	48 (24.0)	189 (59.1)	
Genital thrush	No	167 (83.5)	237 (74.1)	0.012
	Yes	33 (16.5)	83 (25.9)	
Visual blurring	No	142 (71)	145 (45.3)	<0.001
	Yes	58 (29)	175 (54.7)	
Itching	No	101 (50.5)	166 (51.9)	0.760
	Yes	99 (49.5)	154 (48.9)	
Irritability	No	184 (92)	210 (65.6)	<0.001
	Yes	16 (8)	110 (34.4)	
Delayed healing	No	114 (57)	167 (52.2)	0.284
	Yes	86 (43)	153 (47.8)	
Partial paresis	No	168 (84)	128 (40)	<0.001
	Yes	32 (16)	192 (60)	
Muscle stiffness	No	140 (70)	185 (57.83)	0.005
	Yes	60 (30)	135 (42.2)	
Alopecia	No	99 (49.5)	242 (75.6)	<0.001
	Yes	101 (51.5)	78 (24.4)	
Obesity	No	173 (86.5)	259 (80.9)	1.000
	Yes	27 (13.5)	61 (19.1)	
Age		46.4±12.1	49.1±12.1	0.013

Table 2 describes the multivariate analysis among diabetes class and health indicators. We removed itching, delayed healing, and obesity from the analysis since they were found to be nonsignificant in the Univariate analysis. In multivariate analysis, we found polyuria (β=3.492; Aor=32.872; 95%CI=11.09,97.35; p<0.001), polydipsia (β=-4.100; Aor=60.378; 95%CI=18.28,199.37; p<0.001), polyphagia (β=1.181; Aor=3.25; 95%CI=1.23,8.57; p=0.017), genital thrush (β=1.08; Aor=2.96; 95%CI=1.26,7.53; p=0.023), irritability (β=2.28; Aor=9.82; 95%CI=3.41,28.26; p<0.001), and partial paresis (β=1.2406; Aor=3.45; 95%CI=1.35,8.79; p=0.009) are the potential health risk indicators for positive diabetes class. Apart from these, we found gender and age are also associated (p<0.05) in positive diabetes class in multivariate analysis.

Table 2 Multivariate analysis for potential risk health indicators.

	Model coefficients - class					95% Confidence interval	
Predictor	**Estimate**	**SE**	**Z**	***p***	**Odds ratio**	**Lower**	**Upper**
Intercept	2.4927	0.8838	2.8205	0.005	12.0939	2.13936	68.3679
Gender							
Male- Female	-3.7962	0.5170	-7.3433	< 0.001	0.0225	0.00815	0.0619
Polyuria							
Yes-No	3.4926	0.5540	6.3048	**<0 .001**	**32.8720**	**11.09928**	97.3547
Polydipsia							
Yes-No	4.1006	0.6095	6.7281	**<0 .001**	**60.3782**	**18.28469**	199.3756
Sudden weight loss							
Yes-No	0.5206	0.5023	1.0365	0.300	1.6831	0.62885	4.5048
Weakness							
Yes-No	0.0438	0.4723	0.0928	0.926	1.0448	0.41400	2.6367
Polyphagia							
Yes-No	1.1814	0.4939	2.3920	**0.017**	**3.2588**	**1.23783**	8.5794
Genital thrush							
Yes-No	1.0855	0.4766	2.2779	**0.023**	**2.9610**	**1.16357**	7.5350
Visual blurring							
Yes-No	-0.1766	0.5283	0.3344	0.738	0.8381	0.29760	2.3602
Irritability							
Yes-No	2.2851	0.5392	4.2383	**<0.001**	**9.8265**	**3.41564**	28.2699
Partial paresis							
Yes-No	1.2406	0.4763	2.6046	**0.009**	**3.4578**	**1.35943**	8.7953
Muscle stiffness							
Yes-No	-0.6180	0.4992	1.2378	0.216	0.5390	0.20261	1.4341
Alopecia							
Yes-No	-0.5610	0.5044	1.1123	0.266	0.5706	0.21234	1.5335
Age	-0.0552	0.0207	2.6597	0.008	0.9463	0.90863	0.9856

Table 3 describes the classification performances of various machine learning techniques using the data set. The k-fold cross-validation technique was used for all the models. In machine learning, greater k values in cross-validation frequently result in higher accuracy but can also lead to overfitting. Leave-one-out cross-validation is appropriate for small datasets (often less than 100 occurrences) to maximize data utilization, but it can be computationally costly. Holdout validation is commonly recommended for large datasets since it reduces training time. However, this essay seeks to refute that assertion. Despite the increased time investment, it contends that adopting k-fold cross-validation for massive datasets yields significant accuracy benefits. The purpose is to show through results that the time cost of employing k-fold validation over holdout validation is justified, especially when the value of k is kept low enough to provide adequate classification quality. Among all the models, KNN performed the best accuracy (0.962), with an AUC(0.969) and F1 score (0.962). The Roc curves of all the machine learning models are shown in Figure 3.

Table 3. Performance matrices for classification model to predict diabetes class.

Evaluation Metrics	Boosting	DT	KNN	RF	SVM
Accuracy	0.923	0.920	0.962	0.923	0.913
Precision (Positive predictive value)	0.932	0.923	0.965	0.927	0.913
Recall (True positive rate)	0.923	0.923	0.962	0.923	0.913
False positive rate	0.065	0.085	0.031	0.072	0.099
False discovery rate	0.091	0.085	0.047	0.086	0.093
F1 score	0.924	0.923	0.962	0.924	0.913
Matthews Correlation Coefficient	0.843	0.83	0.923	0.842	0.808
Area Under Curve (AUC)	0.973	0.915	0.969	0.977	0.901
Negative predictive value	0.909	0.915	0.953	0.914	0.907
True negative rate	0.935	0.915	0.969	0.928	0.901
False negative rate	0.065	0.085	0.031	0.072	0.099
False omission rate	0.091	0.085	0.047	0.086	0.093
Threat score	4.556	4	10.063	4.271	3.478
Statistical parity	1	1	1	1	1

Note: *DT: decision tree, KNN: K nearest neighbor, RF: random forest, SVM: support vector machine.*

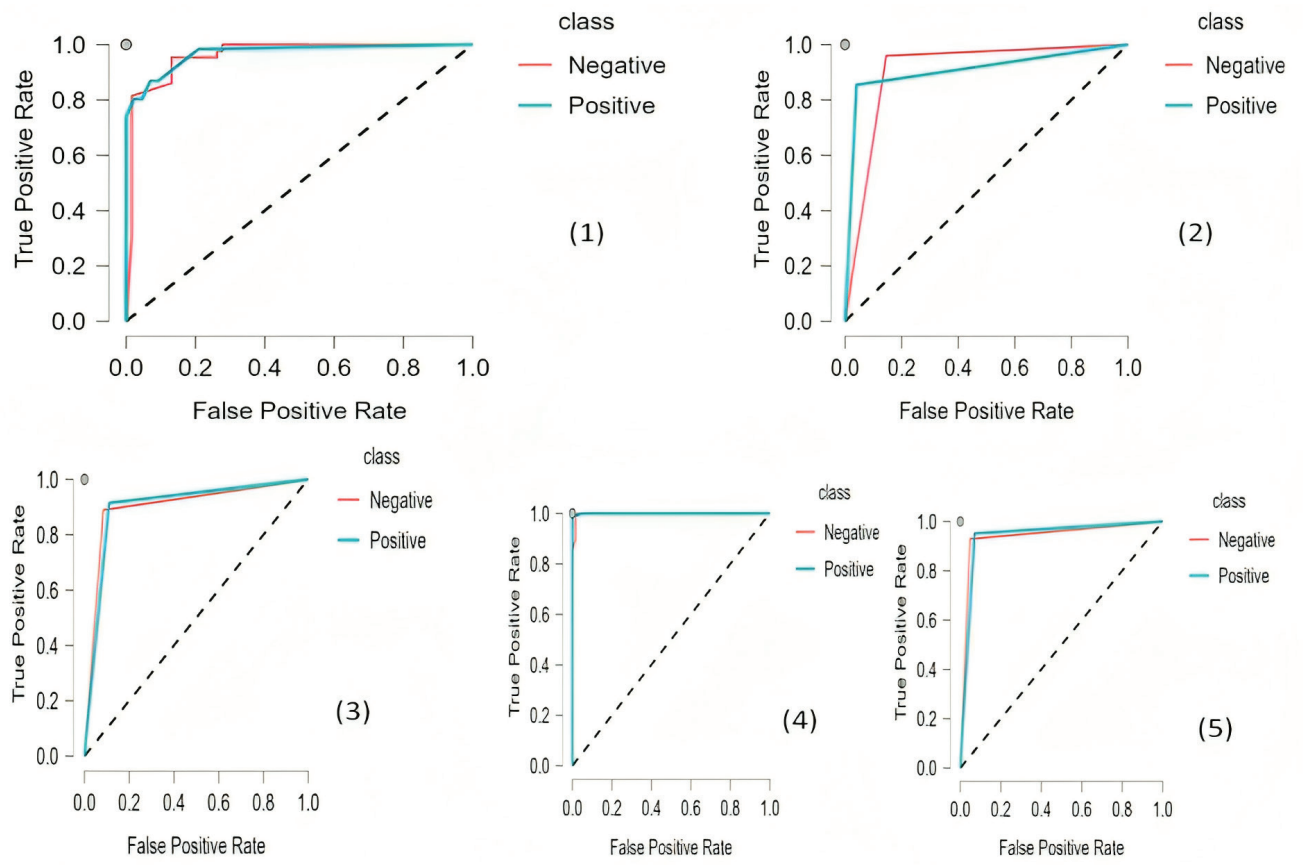


Figure 3. *ROC curves for all the machine learning models.*

Similarly, random forest and boosting algorithms were performed with the same accuracy (0.923). The KNN model beats the others in terms of accuracy, with a score of 0.962, followed by Boosting and Random Forest, with scores of 0.923. Decision Tree and Support Vector Machine (SVM) have lower accuracies of 0.920 and 0.913, respectively. KNN has the best precision and recall and the lowest false positive rate. However, Random Forest has the most significant AUC (0.977), suggesting higher overall performance in classification tasks. Despite its excellent accuracy, KNN has a significantly higher threat score than the other models, indicating that it may be more prone to mistakes. In our research, Figure 4 shows classification accuracy and other related curves of the model.

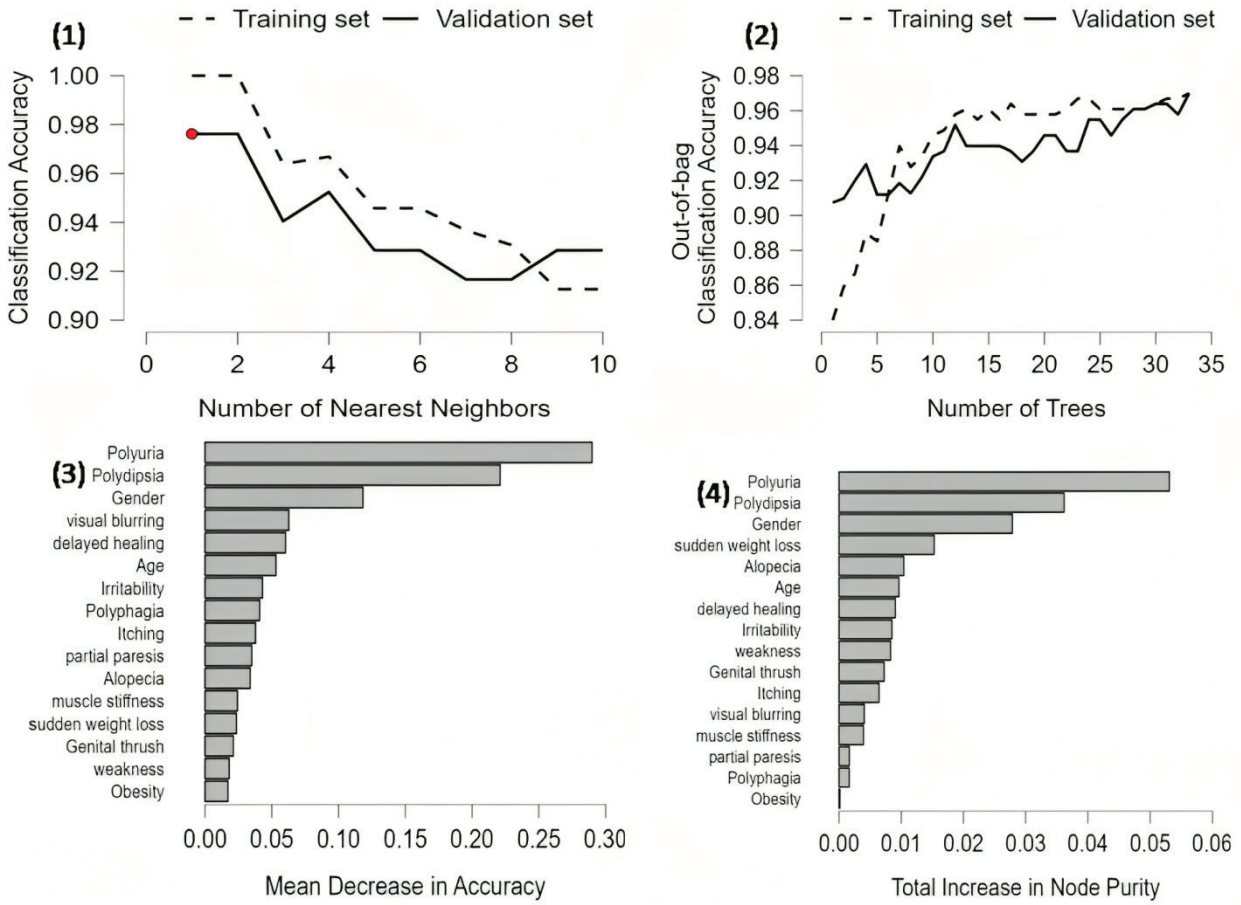


Figure 4*. Classification accuracy and other related curves of the model.*

Discussion

Early detection of any disease facilitates prompt decision-making regarding the condition, mitigates the exacerbation of complications, and conserves both time and money. Machine learning plays a significant role in diabetes prediction by leveraging various data sources to develop predictive models. The study thoroughly assesses the efficacy of different machine-learning techniques in classifying patients as either diabetic or non-diabetic in the early stages. Various metrics such as accuracy, sensitivity, specificity, the ROC curve (AUC-ROC), and precision-recall curve are used to evaluate here the performance of diabetes prediction models. We found that the positive diabetes class in females is higher (54.1%) as compared to males (45.9%), which aligns with a previous study[20,21] but contrasts with the Olufemi *et al*., where gender predictor variable indicates males are more likely to have diabetes than female in their logistic regression model in prediction of diabetes across the US.[23]

The incidence of polyuria cases in our study is found to be significantly elevated, comprising 243 (75.9%) instances within the positive diabetes class, in contrast to the negative diabetes class. In a recent study, Olufemi *et al*. utilized a logistic regression model to predict early diabetes prevalence across the United States. Their findings revealed that polyuria and polydipsia contributed most significantly to predicting the "Positive" class, as evidenced by their parameter values and odds ratios.[23] Various other data mining approaches have been employed to predict diabetes. Using the decision tree algorithm (C4.5), reports indicate that polydipsia is the most influential parameter for diabetes prediction. The performance results demonstrate a notable accuracy rate of 90.38%, suggesting the effectiveness of this algorithmic model.[22] Our comparative machine learning model also found a significant association between polydipsia and the diabetes class (p<0.001). Furthermore, we observed a substantial prevalence of polydipsia cases, accounting for

225 (70.3%) instances, compared to only 8 (4%) cases in the negative diabetes class.

Our study revealed that sudden weight loss, weakness, polyphagia, visual blurring, irritability, partial paresis, and alopecia were significantly associated with the diabetes class (p<0.001). These findings align with the research conducted by Dritsas and Trigka, who also identified polyphagia, irritability, alopecia, visual blurring, and weakness as prominent features correlated with diabetes. In contrast, other features showed negligible correlation (rank <0.2).[24] Utilizing various machine-learning models, they aimed to pinpoint individuals at risk of diabetes based on specific risk factors such as weight loss, weakness, polyphagia, visual blurring, and irritability.[24] Delayed healing (p=0.284), obesity, and itching (p=0.760) were determined to have no significant association with the diabetes class in our study. However, a previous study by Dritsas and Trigka found that diabetes prevalence was notably linked to delayed healing and visual blurring features, with 50% of diagnosed individuals exhibiting these symptoms.[24] Similarly, in our study, the mean age of the positive diabetes class was 49.1±12.1, while for the negative diabetes class, it was 46.4±12.1 (p=0.013). These findings closely resemble those of an interventional study conducted by Wicaksana *et al*. where the average age of diabetic patient participants was reported to be 55.13 years.[25]

In multivariate analysis among diabetes class and health indicators, we also analyzed the indicators other than polydipsia, polyuria, polyphagia, visual blurring, irritability, and alopecia like genital thrush, irritability, and Partial paresis are the potential health risk indicators for a positive diabetes class. Previous studies also obtained similar findings.[26,27] Our comparative analysis found that k-fold validation offers justified time costs compared to holdout validation, mainly when k is maintained at a low level to ensure high-quality classification. Among the various models examined, KNN demonstrated the highest accuracy (0.962), along with notable AUC (0.969) and F1 score (0.962). These findings are consistent with the findings reported in the previous study by Ghosh et al., where they performed another comparative analysis of different machine learning tools in detecting diabetes in their experiments.[28]

Limitations

Although the study identified several possible risk factors and symptoms for early-stage diabetes, certain limitations remain. Since the data set is publicly available, we have only considered a few variables. Further elements or variables might be included to strengthen the study. The collection of data was of a limited size. Large data sets may be used for additional studies to create more precise machine-learning models. Another similar flaw in this research is that internal validation models were used to validate machine learning models by dividing the data into an 80:20 ratio. For generalization to apply to other populations, external validation is necessary. Our future work will focus on external validation of the machine learning models to predict early diabetes risk.

Conclusion

In conclusion, we identified several potential health risk indicators for the positive diabetes class, including polyuria, polydipsia, polyphagia, genital thrush, irritability, and partial paresis. The advancement of diagnostic techniques for diabetes, emphasizing a recent trend towards technology-driven approaches, notably those employing Artificial Intelligence (AI), greatly empowers researchers and healthcare professionals in managing the disease. This shift presents opportunities for early identification of individuals at elevated risk of developing diabetes. Our research indicates that machine learning methods utilizing AI hold promise for accurately predicting diabetes. These models in healthcare effectively mitigate human-induced observation errors, resulting in enhanced outcomes. Timely detection facilitates swift and efficient treatment, potentially lowering the burden of morbidity and mortality linked to the disease.

Funding

No funding was provided for this research.

Conflict of interest

There is no potential conflict of interest.

Ethical Approval

Not applicable.

Acknowledgements

We wish to acknowledge the founder of Siksha O Anusandhan University and the chairman, Prof Dr Manoj Ranjan Nayak, for providing all the research facilities.

References

[1] Caughey GE, Roughead EE, Vitry AI, McDermott RA, Shakib S, Gilbert AL. Comorbidity in the elderly with diabetes: Identification of areas of potential treatment conflicts. Diabetes Res Clin Pract. 2010; 87(3): 385-93. doi: 10.1016/j.diabres.2009.10.019.

[2] American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2014; 37(Supp1): S81-90. doi10.2337/dc14-S081

[3] DeFronzo RA, Ferrannini E, Groop L, Henry RR, Herman WH, Holst JJ, Hu FB, Kahn CR, Raz I, Shulman GI, Simonson DC. Type 2 diabetes mellitus. Nat Rev Dis Primers. 2015; 1(1): 1-22. doi: 10.1038/nrdp.2015.19.

[4] Olokoba AB, Obateru OA, Olokoba LB. Type 2 diabetes mellitus: a review of current trends. Oman Med J. 2012; 27(4): 269. doi: 10.5001/omj.2012.68.

[5] Ginter E, Simko V. Type 2 diabetes mellitus, pandemic in 21st century. Diabetes: an old disease, a new insight. 2013: 42-50. doi: 10.1007/978-1-4614-5441-0_6.

[6] Buchanan TA, Xiang AH. Gestational diabetes mellitus. J Clin Invest. 2005; 115(3): 485-91. doi: 10.1172/JCI24531.

[7] Plows JF, Stanley JL, Baker PN, Reynolds CM, Vickers MH. The pathophysiology of gestational diabetes

mellitus. Int J Mol Sci. 2018; 19(11): 3342. doi: 10.3390/ijms19113342

[8] Tasin I, Nabil TU, Islam S, Khan R. Diabetes prediction using machine learning and explainable AI techniques. Healthc Technol Lett. 2023; 10(1-2): 1-10. doi: 10.1049/htl2.12039

[9] Ahamed BS, Arya MS, Nancy V AO. Prediction of type-2 diabetes mellitus disease using machine learning classifiers and techniques. Front Comput Sci. 2022; 4: 835242. doi.org/10.3389/fcomp.2022.835242

[10] Modak, S.K.S., Jha, V.K. Diabetes prediction model using machine learning techniques. Multimed Tools Appl. 2023; 83: 38523-49. doi: 10.1007/s11042-023-16745-4

[11] Zou Q, Qu K, Luo Y, Yin D, Ju Y, Tang H. Predicting diabetes mellitus with machine learning techniques. Front Genet. 2018; 9: 515. doi.org/10.3389/fgene.2018.00515

[12] Sutton CD. Classification and regression trees, bagging, and boosting. Handbook of statistics. 2005; 24: 303-29. doi:10.1016/S0169-7161(04)24011-1

[13] Song YY, Ying LU. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry. 2015; 27(2): 130. doi: 10.11919/j.issn.1002-0829.215044

[14] Panda NR, Pati JK, Pati T, Satpathy S, Bhuyan R. Comparison of artificial neural network and decision tree methods for predicting the maternal outcome in a tertiary care hospital in Odisha, India. Nat J Community Med. 2022; 13(11): 821-7. doi.org/10.55489/njcm.131120222262

[15] Speiser JL, Miller ME, Tooze J, Ip E. A comparison of random forest variable selection methods for classification prediction modeling. Expert Syst Appl. 2019; 134: 93-101. doi.org/10.1016/j.eswa.2019.05.028

[16] Panda NR, Mahanta KL, Pati JK, Varanasi PR, Bhuyan R. Comparison of Some Prediction Models and their Relevance in the Clinical Research. Int J Stats Med Res. 2023; 12: 12-9. doi.org/10.6000/1929-6029.2023.12.02

[17] Mucherino A, Papajorgji PJ, Pardalos PM, Mucherino A, Papajorgji PJ, Pardalos PM. K-nearest neighbor classification. Data Mining in Agr. 2009: 83-106. doi: 10.1007/978-0-387-88615-2_4

[18] Yu H, Kim S. SVM Tutorial-Classification, Regression and Ranking. Handbook of Nat comp. 2012; 1: 479-506. doi.org/10.1007/978-3-540-92910-9_15

[19] Patle A, Chouhan DS. SVM kernel functions for classification. In: 2013 International conference on advances in technology and engineering (ICATE). 2013 Jan 23 (pp.1-9), IEEE. doi. 10.1109/ICAdTE.2013.6524743

[20] Charoensakulchai S, Usawachoke S, Kongbangpor W, Thanavirun P, Mitsiriswat A, Pinijnai O, Kaensingh S, Chaiyakham N, Chamnanmont C, Ninnakala N, Hiri-o-Tappa P. Prevalence and associated factors influencing depression in older adults living in rural Thailand: A cross-sectional study. Geriatr Gerontol Int. 2019; 19(12): 1248-53. doi: 10.1111/ggi.13804.

[21] Gao M, Jebb SA, Aveyard P, Ambrosini GL, Perez-Cornago A, Papier K, Carter J, Piernas C. Associations between dietary patterns and incident type 2 diabetes: prospective cohort study of 120,343 UK biobank participants. Diabetes Care. 2022; 45(6): 1315-25. doi: 10.2337/dc21-2258.

[22] Permana BA, Ahmad R, Bahtiar H, Sudianto A, Gunawan I. Classification of diabetes disease using decision tree algorithm (C4. 5). In: Journal of Physics: Conference Series 2021; 1869 (1): 012082, IOP Publishing. doi: 10.1088/1742-6596/1869/1/012082

[23] Olufemi I, Obunadike C, Adefabi A, Abimbola D. Application of Logistic Regression Model in Prediction of Early Diabetes Across United States. Int J Sci Manag Res. 2023; 6(05): 34-48. doi: 10.0130/2023230563

[24] Dritsas E, Trigka M. Data-driven machine-learning methods for diabetes risk prediction. Sensors. 2022; 22(14): 5304. doi.org/10.3390/s22145304

[25] Wicaksana AL, Apriliyasari RW, Tsai PS. Effect of self-help interventions on psychological, glycemic, and behavioral outcomes in patients with diabetes: A meta-analysis of randomized controlled trials. Int J Nurs Stud. 2024; 149: 104626. doi.org/10.1016/j.ijnurstu.2023.104626

[26] AtınçYılmaz. PREDICTION OF TYPE 2 DIABETES MELLITUS USING FEATURE SELECTION-BASED MACHINE LEARNING ALGORITHMS. Health Prob Civil. 2022; 16(2): 128-39. doi.org/10.5114/hpc.2022.114541

[27] Sabejon JA, Rejas JB, Lumacad GS, Zarate RL, Mendez EA, Tinoy FM. XGBoost–Based Analysis of the Early-Stage Diabetes Risk Dataset. In: 2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT) 2023 Jun 9 (pp. 19-24). IEEE. doi: 10.1109/APSIT58554.2023.10201658.

[28] Ghosh P, Azam S, Karim A, Hassan M, Roy K, Jonkman M. A comparative study of different machine learning tools in detecting diabetes. Procedia Comput Sci. 2021; 192: 467-77. doi.org/10.1016/j.procs.2021.08.048

[29] Early-stage diabetes risk prediction dataset, 2020, doi.org/10.24432/C5VG8H, UCI Machine Learning Repository.

[30] Panda NR, Mahanta KL, Pati JK, Pati T. Development and Validation of Prediction Model for Neonatal Intensive Care Unit (NICU) Admission Using Machine Learning and Multivariate Statistical Approach. J Obstet Gynaecol India. 2024; 74(3): 1-9. doi: 10.1007/s13224-024-02009-0

[31] Panda NR. A review on logistic regression in medical research. Nat J Community Med. 2022; 13(04): 265-70. doi.org/10.55489/njcm.134202222

Journal of Associated Medical Sciences 2024; 57(3): 166-176

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

The influence of 10 kVp and 15% rule applications on patient dose and image quality in extremities radiography: A phantom study

Thunyarat Chusin[1,2] Waraporn Sudchai[3] Napassorn Jitnarin[1] Sirinya Sridaungpang[1] Sudarat Aree[1] Pachuen Potup[1,4*]

[1]Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.
[2]Interdisciplinary Health and Data Sciences Research Unit, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.
[3]Thailand Institute on Nuclear Technology, Nakornnayok Province, Thailand.
[4]Cellular and Molecular Immunology Research Unit, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.

ARTICLE INFO

Article history:
Received 10 May 2024
Accepted as revised 24 July 2024
Available online 26 July 2024

Keywords:
Digital radiography, extremity X-ray imaging, image quality, radiation dose.

ABSTRACT

Background: Optimizing image quality and radiation dose is crucial in general radiography, adhering to the As Low As Reasonably Achievable (ALARA) principle.

Objective: This study aimed to evaluate the impact of applying the 10 kVp and 15% rules on image quality and patient dose in extremity X-ray imaging using both computed radiography (CR) and digital radiography (DR) systems.

Materials and methods: X-ray imaging of hand, elbow, knee, and foot phantoms was performed using three different exposure techniques on both CR and DR systems. These techniques included the standard technique (ST) based on the established guidelines of the imaging systems, increased 10 kVp with a 50% mAs reduction from ST (10 kVp rule), and increased 15% kVp with a 50% mAs reduction from ST (15% rule). The entrance skin dose (ESD) was measured using nanoDot™placed on the phantom's surface. The physical image qualities in contrast-to-noise ratio (CNR) and figure of merit (FOM) were utilized to assess the balance between image quality and radiation doses.

Results: The ESD was reduced by an average of -16% and -25% when applying the 10 kVp and 15% rules for all extremity imaging. This reduction decreased image CNR by -18% and -12%, respectively. There was no significant difference in CNR between the 15% and 10 kVp rule techniques for all extremity examinations in both CR and DR systems (p>0.05). Meanwhile, the exposure and deviation indexes remained within the established guidelines for CR and DR systems. However, the FOM values tended to be greater with the 15% rule technique than other techniques.

Conclusion: The ESD reduction was observed when applying the 10 kVp and 15% rules for all extremity imaging, both in CR and DR systems, with a slight degradation in image quality. The 15% rule represents the best option for optimization of image quality and patient dose based on the FOM results.

Introduction

Plain radiography is widely used for diagnosis, treatment, and monitoring of diseases, representing the majority of radiographic imaging procedures[1]. Despite patients receiving a lower radiation dose from extremity radiography than other procedures such as abdominal and spine radiography,[1] the optimization of image quality while minimizing radiation exposure remains paramount. This optimization aligns with the As Low As Reasonably Achievable (ALARA) principle, emphasizing the continuous effort to enhance patient safety.

∗ **Corresponding contributor.**
Author's Address: *Department of Radiological Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok Province, Thailand.*
E-mail address: pachuenp@nu.ac.th
doi: 10.12982/JAMS.2024.058
E-ISSN: 2539-6056

However, a notable unintended consequence of computed radiography (CR) and digital radiography (DR) systems is increased patient radiation exposure, known as dose creep.[2-4] This issue arises because of various factors, including a lack of user experience or training in optimizing image quality while managing radiation doses across technologies from different manufacturers.[4] A primary concern is that underexposed images, which result in image noise, are often deemed more problematic than overexposed images, which can be adjusted in post-processing.[5] To address the issue of dose creep, the International Electrotechnical Commission (IEC) and the American Association of Physicists in Medicine (AAPM) have introduced the use of the exposure index (EI) and deviation index (DI) to provide feedback to help radiologic technologists (RTs) assess whether an image has been captured with an optimal balance of image quality and radiation dose.[5]

The 15% rule is based on the principle that increasing the tube potential (kVp) by 15% allows a 50% reduction in the tube current-time product (mAs) to maintain the same exposure level.[6,7] Such adjustments can reduce the patient's skin dose without sacrificing diagnostic image quality. This effect occurs because higher kVp settings produce more penetrating X-rays, less absorbed by the patient. Additionally, the RTs can compensate for the reduced contrast due to increased scatter radiation at higher kVp settings through image post-processing in the CR and DR systems. On the other hand, the 10 kVp rule follows a similar principle but simplifies the approach by halving the mAs for each 10 kVp increment.[6,7] This method provides RTs with an easier way to calculate technical factors, bypassing the complex calculations required by the 15% rule and promoting its application in clinical practices.

Many studies reported patient dose reduction by increasing the kVp without compromising image quality.[8–13] In 2020, Coffey *et a.l* found that using the 10 kVp rule for shoulder and hand radiography resulted in lower entrance skin dose (ESD) than the 15% rule in the GE Definium 8000 DR system.[10] The exposure indicator for the detector remained within the optimal range for all exposures under both rules. They suggested the 10 kVp rule as a viable alternative to the 15% rule and recommended more extensive research on various DR systems and types of examinations. In 2024, Wenman *et al*. confirmed that the 10 kVp rule for knee X-rays in the AGFA DX-D 40 DR system lowered the ESD without compromising image quality compared to standard techniques.[13] They recommended further studies on broader kVp ranges.

This study aimed to evaluate the effectiveness of the 10 kVp and the 15% rules in reducing patient dose while maintaining the quality of hand, elbow, knee, and foot radiographic images in specific CR and DR systems.

Materials and methods

Extremities imaging

The study employed the IMAGE-X 50 X-ray machine (HYUN DAI MEDICAL X-RAY CO., LTD.) in conjunction with the DirectView Vita CR system (Carestream Health, Inc.) and the VIVIX-S 1417W DR system (Vieworks Co., Ltd.) for imaging purposes. The imaging system underwent and passed the annual quality control check conducted by the Department of Medical Sciences at the Ministry of Public Health (DMSMOPH). X-ray imaging of anthropomorphic hand, elbow, knee, and foot phantoms (Radiology Support Devices, Inc.) was triple-performed using both the imaging plate (IP) of the CR system and the flat panel detector (FPD) of the DR system. Specifically, the 18×24 cm IP was used for the hand and foot X-rays, while the 24×30 cm IP was chosen for the elbow and knee X-rays. The 35×43 cm FPD was employed for all extremity X-rays. The distance from the X-ray source to the IP and FPD was fixed at 100 cm in this study. The collimation was restricted to 18×24 cm for hand and foot X-rays and 24×30 cm for elbow and knee X-rays using CR and DR systems. The imaging techniques used for these X-rays are detailed in Table 1 for CR system and Table 2 for DR system. The kVp and mAs values for the standard technique (ST), along with the 10 kVp and 15% rules, were adjusted to meet the image quality standards recommended by the CR and DR system manufacturers.[14,15] This adjustment targeted the optimal ranges of the EI and DI, which are 1,850-2,150 for CR system and -2 to 2 for DR system, respectively.[14,15] The obtained EI and DI values for the ST, 10 kVp, and 15% rules fell within these optimal ranges. The grid and added filtration were not applied in all extremity imaging.

Table 1. The imaging techniques used for extremity X-rays with CR system.

Body Parts/Projections		Exposure parameters		
		ST	15% rule	10 kVp rule
Hand	PA	50 kVp, 2.5 mAs	58 kVp, 1.3 mAs	60 kVp, 1.3 mAs
	PA Oblique	50 kVp, 3.2 mAs	58 kVp, 1.6 mAs	60 kVp, 1.6 mAs
Elbow	AP	60 kVp, 2.0 mAs	69 kVp, 1.0 mAs	70 kVp, 1.0 mAs
	Lateral	60 kVp, 2.5 mAs	69 kVp, 1.3 mAs	70 kVp, 1.3 mAs
Knee	AP	55 kVp, 10 mAs	63 kVp, 5.0 mAs	65 kVp, 5.0 mAs
	Lateral	55 kVp, 10 mAs	63 kVp, 5.0 mAs	65 kVp, 5.0 mAs
Foot	AP	50 kVp, 6.3 mAs	58 kVp, 3.2 mAs	60 kVp, 3.2 mAs
	AP Oblique	50 kVp, 5.0 mAs	58 kVp, 2.5 mAs	60 kVp, 2.5 mAs

Table 2. The imaging techniques used for extremity X-rays with DR system.

Body Parts/Projections		Exposure parameters		
		ST	15% rule	10 kVp rule
Hand	PA	50 kVp, 2.0 mAs	58 kVp, 1.0 mAs	60 kVp, 1.0 mAs
	PA Oblique	50 kVp, 2.5 mAs	58 kVp, 1.3 mAs	60 kVp, 1.3 mAs
Elbow	AP	60 kVp, 1.6 mAs	69 kVp, 0.8 mAs	70 kVp, 0.8 mAs
	Lateral	60 kVp, 2.0 mAs	69 kVp, 1.0 mAs	70 kVp, 1.0 mAs
Knee	AP	55 kVp, 8.0 mAs	63 kVp, 4.0 mAs	65 kVp, 4.0 mAs
	Lateral	55 kVp, 6.3 mAs	63 kVp, 3.2 mAs	65 kVp, 3.2 mAs
Foot	AP	50 kVp, 5.0 mAs	58 kVp, 2.5 mAs	60 kVp, 2.5 mAs
	AP Oblique	50 kVp, 5.0 mAs	58 kVp, 2.5 mAs	60 kVp, 2.5 mAs

Dose assessment

NanoDot™ is a type of small optically stimulated luminescence dosimeter (OSLD). The nanoDots™were calibrated against an ionization chamber, model 10X6-6 (Radcal Corporation, Monrovia, CA), based on the beam quality used for all extremity imaging. The ESD was measured three times, with a single nanoDot™ (Landauer, Glenwood, IL) placed at the center of the X-ray beam on the extremity phantom's surface for each measurement, as shown in Figure 1. The field sizes of the X-ray beam were 18×24 cm for hand and foot X-rays, and 24×30 cm for elbow and knee X-rays. Three nanoDots™were used to control for background radiation. Dose values for both the exposed (M) and control (M_{BG}) nanoDots™ were read three times and then reported using the microSTAR reader (Landauer, Glenwood, IL). Consequently, the ESDs were calculated using Equation 1.

$$ESD = (M - M_{BG}) \times CF \quad (1)$$

where CF represents the calibration factor for each beam quality used in all extremity imaging.

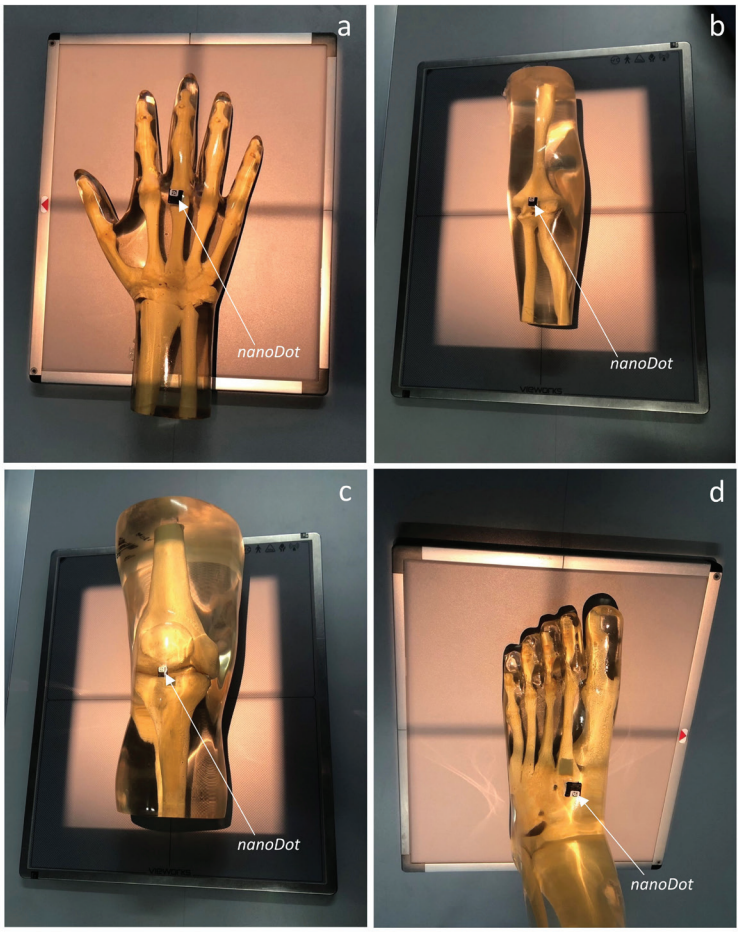


Figure 1*. Placement of a: hand, b: elbow, c: knee, and d: foot phantoms equipped with nanoDots™ for measuring the entrance skin dose (ESD) in computed radiography (CR) and digital radiography (DR) systems.*

Image quality evaluation

To evaluate the quality of 144 extremity images (four organs, two projections, three exposure techniques, X-ray repetition three times, and two imaging systems), this study utilized the contrast-to-noise ratio (CNR), calculated using Equation 2. We identified areas and measured soft tissue and bone pixel values using Image J software (National Institutes of Health, Bethesda, MD) for the CNR calculations. This process included quantifying noise as the standard deviation in soft tissue areas. Three regions of interest (ROIs) per image were selected to analyze both bone and soft tissue, aiming to calculate the mean pixel values (signal) and mean standard deviation (noise), as shown in Figure 2.

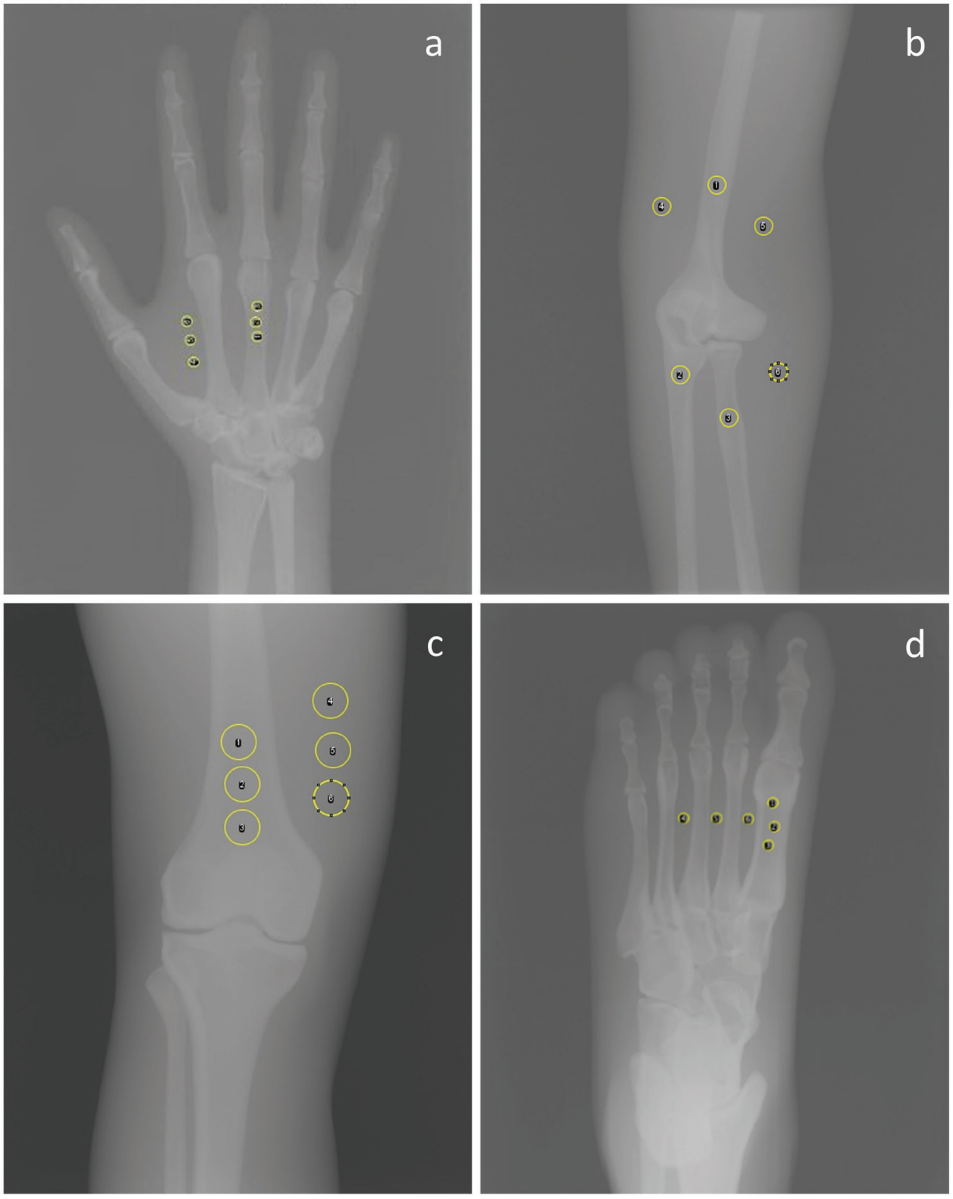


Figure 2. *Selection of the region of interest (ROI) for determining the pixel value and standard deviation in non-processed images of (a) hand, (b) elbow, (c) knee, and (d) foot.*

$$CNR = \frac{PV_{bone} - PV_{soft\ tissue}}{SD_{soft\ tissue}} \quad (2)$$

where PV_{bone} and $PV_{soft\ tissue}$ represent the mean pixel values in ROIs of bone and soft tissue areas in an extremity image.

$SD_{soft\ tissue}$ represents the mean standard deviation of pixel values in ROIs of bone and soft tissue areas in an extremity image.

Finally, the figure of merit (FOM) was calculated using Equation 3 to compare the CNR independently of ESD.[16,17]

$$FOM = \frac{CNR^2}{ESD} \quad (3)$$

Where *CNR* is the mean of contrast to noise ratio from three images for each X-ray examination.

ESD is the mean of entrance skin dose from three-time measurements for each X-ray examination.

Statistical analysis

The statistical significance of variations in ESD and CNR values was evaluated using either the one-way ANOVA test (at a 95% confidence level) or the Mann-Whitney U test, based on whether the variables were normally distributed, as verified by the Shapiro-Wilk test.

Results

The DR system tended to achieve optimal image quality with lower mAs values (radiation doses) than the CR system, as described in Tables 1 and 2. The hand, elbow, knee, and foot X-ray images were acquired using CR and DR systems with EI and DI values corresponding to three exposure techniques illustrated in Table 3. Both the 10 kVp rule and the 15% rule techniques resulted in increased doses to the image receptors, as indicated by the higher EI and DI values for these techniques compared to the ST. Meanwhile, the mean ESDs derived from different exposure techniques are shown in Figures 3 and 4. In a comparison among the ST, the 10 kVp rule, and the 15% rule techniques, the ST yielded the highest ESD (0.046±0.002–0.39±0.008 mGy), with the ESD from the 10 kVp rule (0.040±0.003–0.311±0.005 mGy) surpassing that from the 15% rule (0.036±0.002–0.276±0.005 mGy) in both systems. Compared to the ST, the mean of ESD reductions resulting from the application of the 10 kVp and 15% rules were found to be -15% (ranging from -6% to -27%) and -24% (ranging from -8% to -48%) for CR, and -18% (ranging from -10% to -23%) and -26% (ranging from -22% to -33%) for DR, respectively. Simultaneously, the EI and DI values were maintained within the optimal ranges prescribed by the CR and DR system manufacturers, as shown in Table 3. The calibration factors for the nanoDot™ ranged from 0.820 to 0.915. The coefficient of variation (CV) associated with ESD measurements using the nanoDot™ system was observed to be within ±10% across all beam qualities employed in extremity imaging. Subsequently, a variation in the ESD was noted in foot oblique X-rays, despite applying identical exposure techniques for CR and DR systems. We found a significant difference in ESDs between the ST and the 10 kVp rule techniques and between the ST and the 15% rule techniques for all extremity examinations in both CR and DR systems (p<0.05), except for AP and oblique foot X-ray using CR. However, there was no significant difference in ESDs between the 10 kVp rule and the 15% rule techniques.

Table 3. The exposure index (EI) and deviation index (DI) were obtained using the standard (ST), the 10 kVp rule, and the 15% rule techniques in computed radiography (CR) and digital radiography (DR) systems.

Body Parts/Projections		Mean EI values for CR			Mean DI values for DR		
		ST	15% rule	10 kVp rule	ST	15% rule	10 kVp rule
Hand	PA	1922	1948	2012	0.5	0.8	1.2
	PA Oblique	1999	2005	2050	0.3	0.4	1.0
Elbow	AP	1982	1983	1990	-0.1	-0.1	0.0
	Lateral	2039	2044	2075	0.1	0.2	0.4
Knee	AP	1973	2012	2037	-0.4	-0.2	0.5
	Lateral	2028	2076	2132	0.7	1.0	1.4
Foot	AP	1980	2012	2040	0.1	0.4	0.8
	AP Oblique	1959	1997	2009	0.1	0.3	0.7

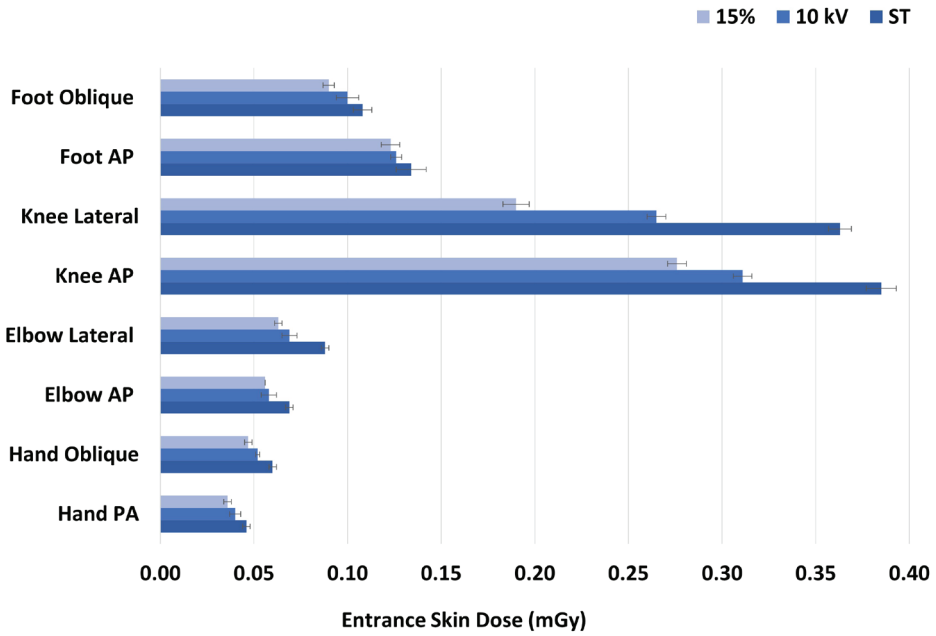


Figure 3. *The mean entrance skin dose was obtained using three different exposure techniques in computed radiography (CR).*

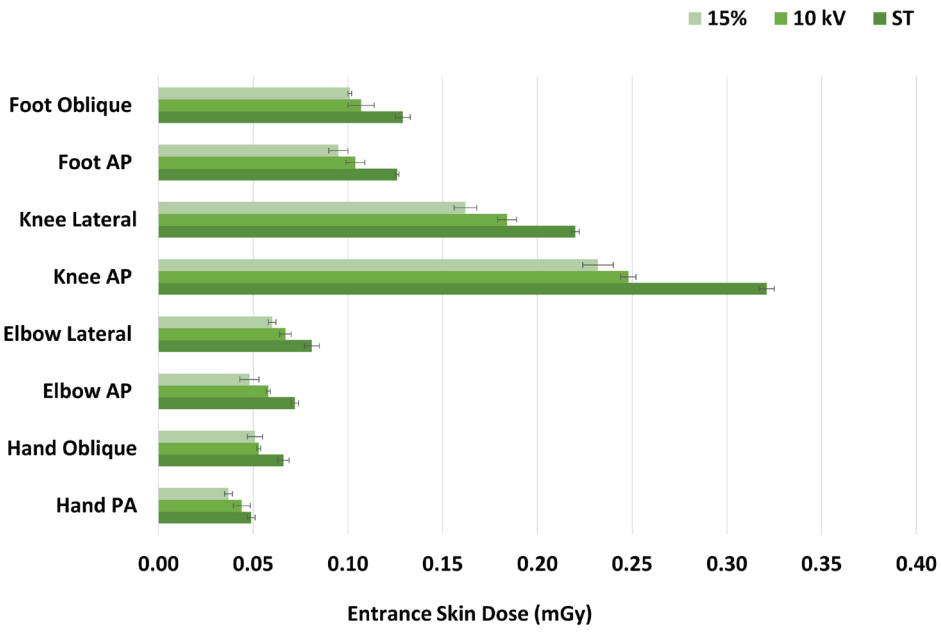


Figure 4. *The mean entrance skin dose was obtained using three different exposure techniques in digital radiography (DR).*

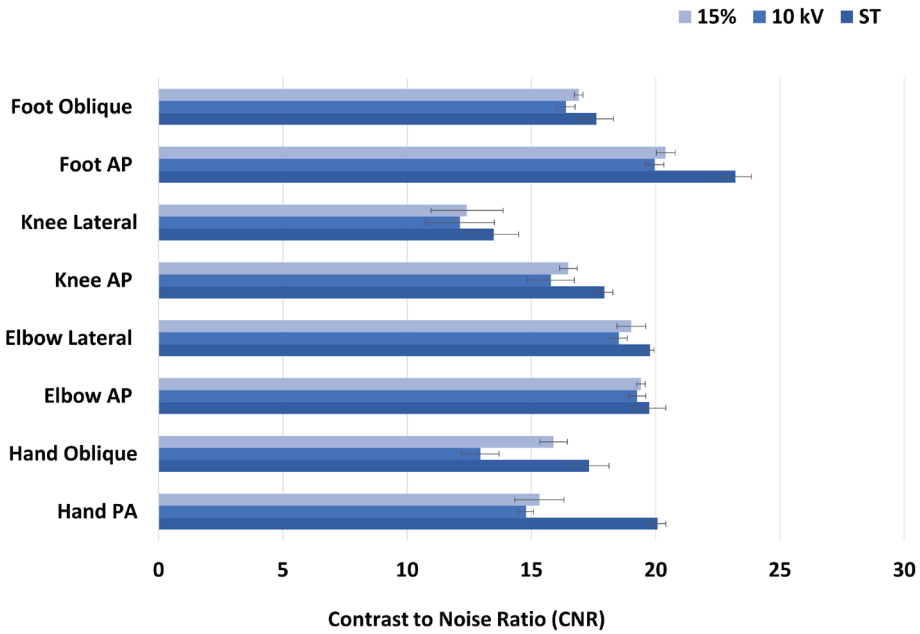


***Figure 5**. The mean contrast-to-noise ratio (CNR) was obtained using three different exposure techniques in computed radiography (CR).*

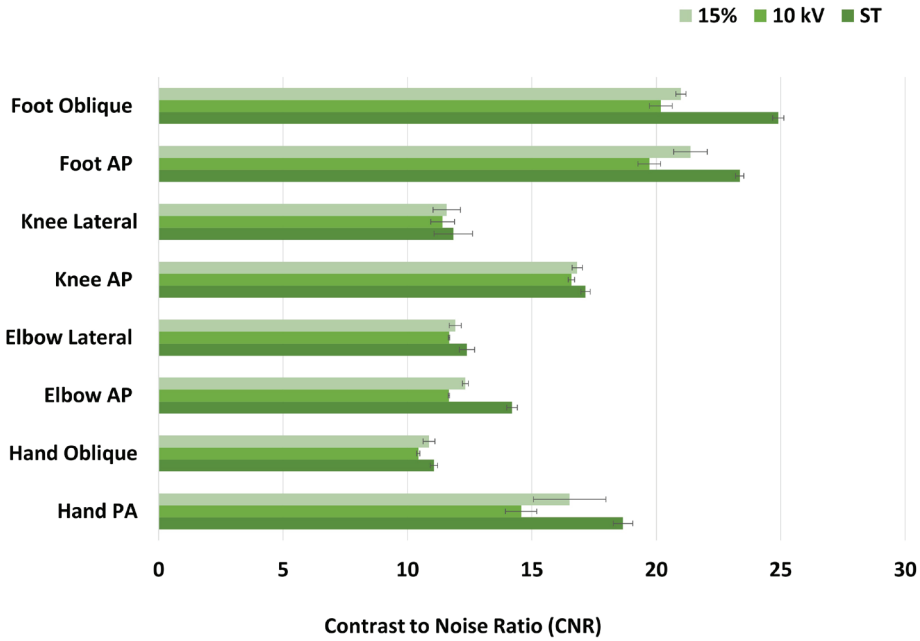


***Figure 6**. The mean contrast-to-noise ratio (CNR) was obtained using three different exposure techniques in digital radiography (DR).*

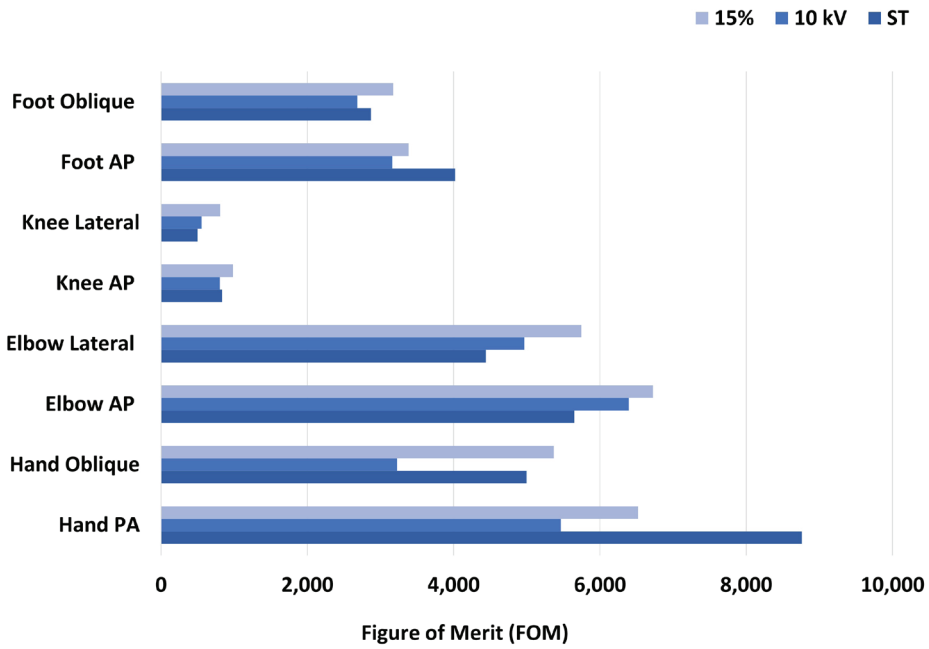


***Figure 7**. The figure of merit (FOM) was obtained using three different exposure techniques in computed radiography (CR).*

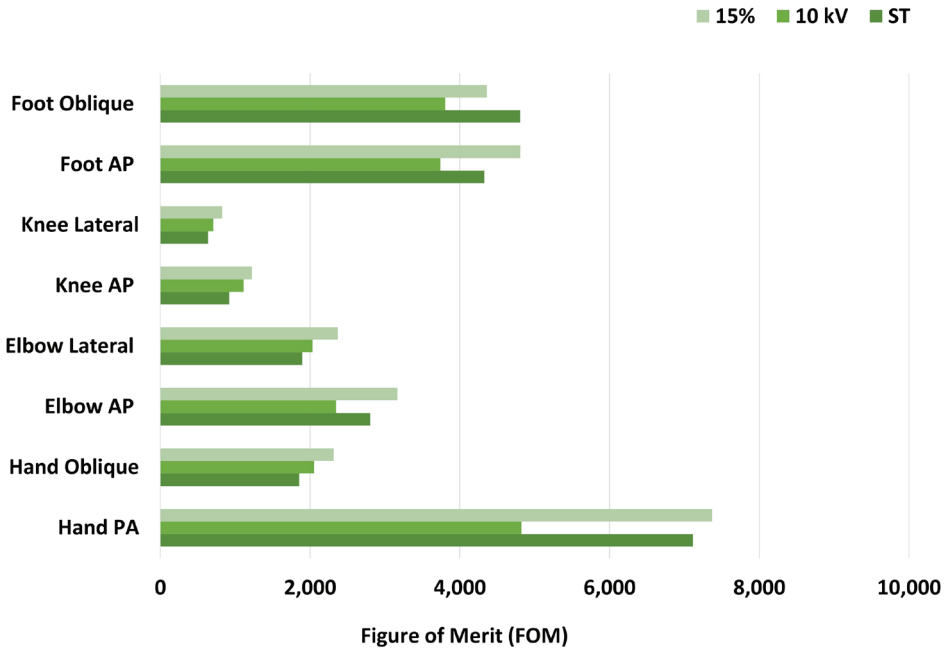


***Figure 8**. The figure of merit (FOM) was obtained using three different exposure techniques in digital radiography (DR).*

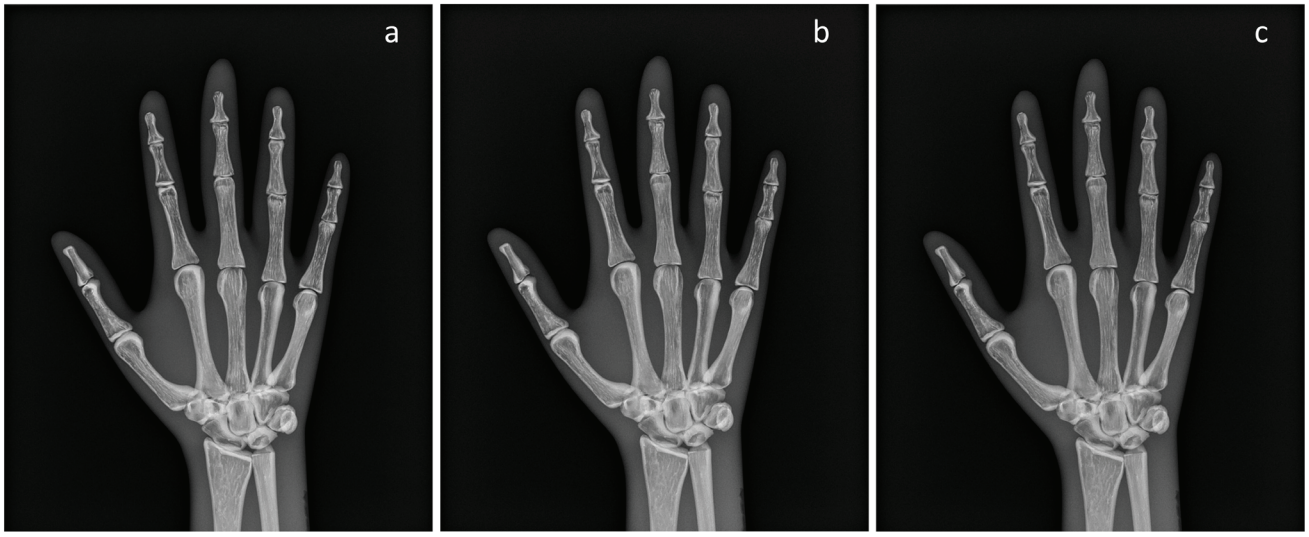


***Figure 9**. Pre-processed radiographic images of PA hand acquired using a: the standard (ST), b: the 15% rule, and c: the 10 kVp rule techniques.*

The CNR values corresponding to three distinct exposure techniques are illustrated in Figure 5 for CR and Figure 6 for DR. Comparative analysis revealed that the CR system demonstrated superior CNR values relative to the DR system, indicating a consistent correlation with the ESD outcomes. Despite employing identical exposure techniques, observations showed that the DR system demonstrated superior CNR values in foot oblique X-ray imaging compared to the CR system. The ST achieved better CNR values than those observed by applying the 10 kVp and 15% rules. Specifically, the 15% rule yielded higher CNR values than the 10 kVp rule. These findings were consistent across all CR and DR systems extremity examinations. Relative to the ST, the mean CNR reductions associated with applying the 10 kVp and 15% rules were -19% and -11% for CR, and -18% and -13% for DR, respectively. Despite the observed reductions in CNR, the critical metrics used to evaluate the optimization of radiation dose and image quality, including the EI and DI, remained within the established guidelines for CR and DR systems, as shown in Table 3. The FOM results are presented in Figure 7 for CR and Figure 8 for DR. The 15% rule technique demonstrated a superior FOM compared to other techniques, except PA hand and AP foot X-rays using CR and oblique foot X-rays using DR. The statistical tests reveal no significant difference in CNR between the 15% rule and the 10 kVp rule techniques for all extremity examinations in both CR and DR systems. However, significant differences in CNR were observed between the ST and the 10 kVp rule techniques, as well as between the ST and the 15% rule techniques for PA hand (CR and DR), oblique hand (CR), AP foot (CR and DR), and oblique foot (DR), with $p < 0.05$. An example of radiographic images of PA hand acquired with automatic pre-processing by the DR system for the ST, the 10 kVp rule, and 15% rule is shown in Figure 9.

Discussion

This study assesses the impact of implementing the 10 kVp and 15% rules as alternatives to the standard technique on the ESD and the resulting image quality in extremity radiography using CR and DR systems. Our findings reveal a notable reduction in ESD by approximately -25% when applying the 15% rule across all types of extremity imaging in both CR and DR systems. This is accompanied by a slight degradation in CNR by approximately -12.5% compared to the ST. Additionally, the FOM values suggest that the 15% rule provides superior optimal images.

Optimizing radiation dose without sacrificing image quality is a critical consideration in radiographic practices. The radiation dose reduction by applying 10 kVp and 15% rules in our study agrees with previous studies,[10,13] suggesting that higher kVp settings can reduce patient dose while maintaining acceptable image quality.[9] Notably, while the 10 kVp rule also reduced dose, its impact on CNR was less favorable than the 15% rule. This difference emphasizes the importance of selecting appropriate exposure parameters for typical examinations that balance patient safety with diagnostic efficacy. The percentage of ESD reduction in knee X-rays using the 10 kVp rule was lower in this study (ranging from -16% to -27%) compared to the study by Wenman *et al*. (ranging from -32% to -34%).[13] However, the ESD reduction in hand X-rays using the 10 kVp and 15% rules was higher in this study (ranging from -10% to -24%) compared to the study by Coffey *et al.* (ranging from -8% to -10%). These differences can be attributed to the variations in ESD values, X-ray imaging, and dose measurement systems.[10]

The kV is a critical parameter influencing image contrast, wherein techniques employing lower kVp settings enhance contrast through a heightened probability of photoelectric absorption and a reduced likelihood of Compton scattering. These observations align with our findings, wherein applying the 15% rule technique yielded images with superior CNR values compared to those derived using the 10 kVp rule. Brindhaban *et al*. found reductions in CNR values of -4% and -22% for the Fuji FCR and the Kodak Direct View CR systems due to higher kVp

settings when applying the 15% rule for lumbar spine imaging.[11] This study used a low kVp range (50-60 kVp) for extremity X-ray imaging, which resulted in a higher kVp setting for the 10 kVp rule than for the 15% rule. Therefore, the results of CNR and ESD in this study may differ from those in exams employing a high kVp range, such as shoulder, abdomen, pelvis, and spine radiography, where the 15% rule would yield a higher kVp setting than the 10 kVp rule. For example, Coffey *et al*. study found that applying the 10 kVp rule (85kV, 2.5 mAs) instead of preset exposure technique (75kV, 5 mAs) for shoulder radiography resulted in lower ESD than the 15% rule (86kV, 2.5 mAs).[10] This leads to the need for more studies to determine whether the 10 kVp rule is superior in reducing ESD while maintaining the quality of diagnostic images compared to the 15% rule.

Limitation

There are several limitations in this study. The ROI sizes varied according to the bone and soft tissue areas in each organ. This study primarily focused on the quantitative aspects of image quality. However, the most effective diagnosis should also consider qualitative factors such as the visual grading analysis (VGA) of clinically relevant anatomical structures. Our research focused on a phantom study, which may not fully capture the complexity of human anatomy and its impact on radiographic images. Future research could focus on applying these rules in a clinical setting, incorporating various anatomical areas and patient demographics to validate our findings.

Conclusion

Applying the 10 kVp and 15% rules in extremity radiographic imaging using CR and DR systems can reduce the ESD by an average of -16% and -25%, respectively. This reduction in ESD was accompanied by a modest decrease in CNR of -18% and -12%, indicating a slight compromise in image quality. However, the EI and DI remained within acceptable ranges. The FOM values were higher for the 15% rule than for the ST and the 10 kVp rule. Therefore, the 15% rule is recommended for optimizing patient dose and image quality in extremity radiography.

Conflict of interest

This work has no conflicts of interest.

Funding

This work was supported by the Thailand Institute of Nuclear Technology (Public Organization). Grant No. R2566A035

References

[1] Mettler FA, Mahesh M, Bhargavan-Chatfield M, Chambers CE, Elee JG, Frush DP, *et al*. Patient exposure from radiologic and nuclear medicine procedures in the United States: Procedure volume and effective dose for the period 2006-2016. Radiology. 2020; 295(2): 418-27. doi: 10.1148/radiol.221263.

[2] Lewis S, Pieterse T, Lawrence H. Retrospective evaluation of exposure indicators: a pilot study of exposure technique in digital radiography. J Med Radiat Sci. 2019; 66(1): 38-43. doi: 10.1002/jmrs.317.

[3] Gibson DJ, Davidson RA. Exposure creep in computed radiography. A longitudinal study. Acad Radiol. 2012; 19(4): 458-62. doi: 10.1016/j.acra.2011.12.003.

[4] Compagnone G, Baleni CM, Pagan L, Calzolaio FL, Barozzi L, Bergamini C. Comparison of radiation doses to patients undergoing standard radiographic examinations with conventional scree-film radiography, computed radiography and direct digital radiography. Br J Radiol. 2006; 79(947): 899-904. doi: 10.1259/bjr/57138583.

[5] Seibert JA, Morin RL. The standardized exposure index for digital radiography: An opportunity for optimization of radiation dose to the pediatric population. Pediatr Radiol. 2011; 41(5): 573-81. doi: 10.1007/s00247-010-1954-6.

[6] Ching W, Robinson J, Mcentee M. Patient-based radiographic exposure factor selection: A systematic review. J Med Radiat Sci. 2014; 61: 176-90. doi: 10.1002/jmrs.66.

[7] Ching W, Robinson JW. DigiBit: A system for adjusting radiographic exposure factors in the digital era. Radio Technol. 2015; 86(6): 614-22. PMID: 26199434.

[8] Martin CJ, Darragh CL, McKenzie GA, Bayliss AP. Implementation of a programme for reduction of radiographic doses and results achieved through increases in tube potential. Br J Radiol. 1993; 66(783): 228-33. doi: 10.1259/0007-1285-66-783-228.

[9] Ramanaidu S, Sta Maria RB, Ng KH, George J, Kumar G. Evaluation of radiation dose and image quality following changes to tube potential (kVp) in conventional paediatric chest radiography. Biomed Imaging Interv J. 2006; 2(3): e35. doi: 10.2349/biij.2.3.e35

[10] Coffey H, Chanopensiri V, Ly B, Nguyen D. Comparing 10 kVp and 15% Rules in Extremity Radiography. Radiol Technol. 2020; 91(6): 516-24. PMID: 32606229.

[11] Brindhaban A, Al Khalifah K, Al Wathiqi G, Al Ostath H. Effect of x-ray tube potential on image quality and patient dose for lumbar spine computed radiography examinations. Australas Phys Eng Sci Med. 2005; 28(4): 216-22. doi: 10.1007/BF03178721.

[12] Peacock NE, Steward AL, Riley PJ. An evaluation of the effect of tube potential on clinical image quality using direct digital detectors for pelvis and lumbar spine radiographs. J Med Radiat Sci. 2020; 67(4): 260-68. doi: 10.1002/jmrs.403.

[13] Wenman A, Lockwood P. Comparing the standard knee X-ray exposure factor, 10 kV rule, and modified 10 kV rule techniques in digital radiography to reduce patient radiation dose without loss of image quality. Radiography. 2024; 30(2): 574-81. doi: 10.1016/j.radi.2024.01.013.

[14] Moore QT, Don S, Goske MJ, Strauss KJ, Cohen M, Herrmann T, et al. Image gently: using exposure indicators to improve pediatric digital radiography. Radiol Technol. 2012; 84(1): 93-9. PMID: 22988269.

[15] Vieworks. VXvue Operation Manual [Internet]. [cited 2024 Mar 21]. Available from: https://wosc.info/wp-content/uploads/EquipmentManuals/VXVUE/VXVUE-Operation-Manual.V86b21_EN-Manual.pdf

[16] Freitas MB, Pimentel RB, Braga LF, Salido FSA, Neves RFCA, Medeiros RB. Patient dose optimization for computed radiography using physical and observer-based measurements as image quality metrics. Radiat Phys Chem. 2020; 172: 108768. doi: 10.1016/j.radphyschem.2020.108768.

[17] Nocetti D, Villalobos K, Marín N, Monardes M, Tapia B, Toledo MI, et al. Radiation dose reduction and image quality evaluation for lateral lumbar spine projection. Heliyon. 2023; 9(9): e19509. doi: 10.1016/j.heliyon.2023.e19509.

Journal of Associated Medical Sciences 2024; 57(3): 177-183

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

Muscle thickness of lateral abdominal muscles in sitting and standing positions during abdominal drawing-in maneuver and abdominal bracing among healthy adults

Aisha Tukur Mohammed[1,4] Hathaichanok Petchsont[1] Patraporn Sitilertpisan[1] Sompong Sriburee[2] Leonard Joseph[3] Aatit Paungmali[1*]

[1]Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[3]Centre for Regenerative Medicine and Devices, School of Sports and Health Sciences, University of Brighton, Eastbourne, United Kingdom.
[4]Department of Physiotherapy, University of Maiduguri, Borno state, Nigeria.

ARTICLE INFO

Article history:
Received 17 April 2024
Accepted as revised 25 July 2024
Available online 28 July 2024

Keywords:
Lateral abdominal muscles, abdominal drawing-in maneuver, abdominal bracing, real-time ultrasound imaging.

ABSTRACT

Background: Spinal stabilization exercises, such as abdominal drawing maneuvers (ADIM) and abdominal bracing (AB), are crucial in alleviating low back pain, especially in functional positions such as sitting and standing. However, few studies have investigated the contraction patterns of the lateral abdominal muscles during ADIM and AB in these daily living positions (i.e., sitting and standing).

Objective: The study aimed to investigate the thickness of transverse abdominis (TrA), internal abdominal oblique (IO), and external abdominal oblique (EO) during ADIM and AB between sitting and standing positions.

Materials and methods: Twenty-four healthy adults (12 males, 12 females) without low back pain or other orthopedic conditions were assessed in sitting and standing positions while performing ADIM and AB. Real-time ultrasound imaging was used to measure the thickness of the TrA, IO, and EO during the ADIM and AB.

Results: Muscle thickness of TrA and IO was significantly greater during ADIM when compared to AB ($p<0.05$). TrA thickness during AB was significantly greater in sitting than standing ($p<0.05$), but ADIM showed no difference in both positions. ADIM and AB can produce contraction of the lateral abdominal muscles (LAM) in both sitting and standing positions.

Conclusion: ADIM and AB can produce contraction of the LAM in both sitting and standing positions. However, ADIM can activate TrA and IO muscles better than AB. Therefore, clinicians may initially consider using ADIM to teach individuals who need core stability exercises such as low back pain.

Introduction

Core stability exercises improve the stability of the lumbopelvic region. Spinal stability is achieved through a synchronized system of active spinal muscles, passive spinal-column architecture, and neurological components.[1,2] The active spinal muscles are otherwise called lateral abdominal muscles (LAM), which include transversus abdominis (TrA), internal oblique (IO), and external oblique (EO). The contraction of the LAM is primarily responsible for core stability during movements and weight-bearing tasks.[3,4] The contraction of LAM is achieved through movements such as abdominal drawing-in maneuver (ADIM) and abdominal bracing (AB).[5]

The most common technique for strengthening deep muscles is ADIM. It has been mentioned that ADIM

∗ **Corresponding contributor.**
Author's Address: *Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: aatit.p@cmu.ac.th
doi: 10.12982/JAMS.2024.059
E-ISSN: 2539-6056

potentially activates TrA while minimally contracting the IO and EO muscles.[6] ADIM also raises abdominal pressure by drawing the abdominal walls inward so that the TrA and oblique muscles are contracted, through which trunk stability is accomplished effectively. Similarly, AB is an action that improves trunk stability by tightening the abdominal muscles simultaneously (i.e., inflating the abdominal cavity). The AB has also been shown to increase intra-abdominal pressure, which is essential for trunk stability.[7] Thus, both the ADIM and AB are used to teach core stability exercises for patients with low back pain and are sometimes used as an exercise regime in some sports training such as weightlifting.[8,9]

Sitting and standing are common functional positions involving routine activities of daily living.[10] Prolonged sitting and improper standing positions are significant risk factors for LBP. Maintaining a neutral lordotic posture in the lumbar spine while sitting and standing is frequently recommended for reducing the load on the disc and spine.[10] Most individuals spend approximately 55% of their working hours in these sedentary postures, which can cause spine instability. In addition, in clinical practice, exercises to contract LAM are first taught in basic functional positions, such as sitting and standing, and eventually progressed to other positions, such as bridge position, when the patient has achieved sufficient spinal stability. Past studies show various positions such as supine lying, hook lying, supine with 90° flexed knee and hip, supine with stretched knees and 90° flexed hips, side plank, plank, bridge position, and bridge with one stretched knee position used to investigate the contraction of the LAM during the ADIM and AB.[11-13] Evidence shows that the bridge position has the highest activation of the TrA.[13]. Therefore, it is imperative that the contraction of LAM during the ADIM and AB needs to be studied in the primary functional positions such as sitting and standing.

Currently, there is a paucity of studies that have investigated the contraction of LAM during the ADIM and AB between the sitting and standing positions.

Therefore, the study's main aim was to investigate the contraction of LAM in sitting and standing positions during ADIM and AB among healthy adults to avoid pain interference. The study findings provide knowledge for clinicians to choose the appropriate position and movement (ADIM versus AB) for teaching LAM exercises to improve core stability.

Materials and methods

Study characteristics

This is a cross-sectional quantitative study conducted in a public university's radiological laboratory in a teaching hospital between December 2022 and 2023. The study procedure was explained to the participants using the participation information sheet. The study participants gave informed written consent before they participated in the study. Ethical approval was obtained from an institutional ethics committee before the study (Reference number - AMSEC-64EX-073).

Participant Characteristics:

Twenty-four healthy participants (12 males and 12 females) aged 18-45 years with a body mass index (BMI) of 18.5-24.9 kg/m^2 were recruited for the study. The sample size was calculated based on the previous study in healthy subjects while performing ADIM and AB in various positions. The G-power program, version 3.1.9.7, was used for sample size calculation, with an effect size for LAM thickness of 0.25, a power level set at 0.80, and a p-value of 0.05.[14] The study details were advertised in the communities around the university and inside the university campus using a recruitment poster. The participants were selected using a pre-determined study selection criterion. The participants were included in the study if they were healthy, 18-45 years old, with BMI within normal (18.5-24.9 kg/m^2), and had no form of deformity. The participants were excluded if they had a history of low back pain (acute or chronic), any types of orthopedic problems, any history of medication for pain, and those who did not attend the familiarization session for ADIM and AB.

Study Instrument

Real-time ultrasonography was used to investigate the contraction of the LAM by measuring the muscle thickness of the TrA, IO, and EO. A diagnostic ultrasound device (Canon, model Xario – 100, linear head transducer, multifrequency 5-12 MHz) with B-mode was used to measure the lateral abdominal muscle thickness. The ultrasonographic images were captured and analyzed using the Image J program (version 1.36b, http://rsb.info.nih. gov/ij/). A reliability assessment (intra-rater and inter-rater) of the technique and procedure was conducted before the main study. A total of ten participants (5 males and 5 females) who matched the inclusion criteria participated in the reliability procedure. Ultrasonographic images were captured, and the reliability value was calculated before the main data collection.

Procedure

The procedure for the ultrasound imaging to measure the thickness of the TrA, IO, and EO was followed as per the previously established protocols.[15] All measurements were taken on the same day. All the participants completed a familiarization session the same day before testing the thickness of TrA, IO, and EO. The main reason for familiarization was to ensure the participants correctly conversed with the exercises. For ADIM, participants were asked to pull their abdomen tight and navel towards the spine as hard as they could, holding for 10 seconds while controlling the neutral position of the lumbopelvic region.[14] To perform the AB, the participants were asked to tighten their abdominal muscles, exerting as much force as they could (i.e., inflating the abdominal cavity), holding for 10 seconds while also controlling the neutral position of the lumbopelvic region.[14] A pressure biofeedback unit was used to provide biofeedback (a belt was wrapped around the participant's abdomen just above the anterior superior iliac spine (ASIS), and

a pressure biofeedback unit was placed by the side of the belt and inflated to 40 mmHg. During ADIM, it was deflated to 30 mmHg, while for AB, it was inflated to 50 mmHg). This is to ensure that the participants understand and correctly contract the LAM.[16] The procedures were carried out in both the sitting and standing positions. To standardize the exercise, participants were in neutral positions while standing and sitting on a chair without back support; one bar was placed at the xiphoid process, and another was placed at the posterior superior iliac spine (PSIS) to avoid bending/stooping. The ultrasound transducer head was placed halfway between the lower end of the rib cage and the upper border of the iliac crest in the anterior mid-auxiliary line, where the thickness part of the TrA, IO, and EO muscles were measured.[14,16] All measurements were taken on the right side of the body in sitting (a seat with no backrest) and standing positions; there was a resting phase of 30 seconds between each repetition and 60 seconds between each position. So, each position was entirely scanned within 5 minutes. The contraction of the TrA, IO, and EO was measured three times at the thickest portion of the muscle as visualized in the image (an average value of 3 measures was used for further analysis). The order of tests and measurements was randomized. The measurements were obtained at rest and while performing ADIM and AB exercises with a resting phase of 5 seconds between each repetition and 60 seconds between each position during the procedure. All ultrasound imaging was taken, and muscle thickness was measured by the same person (a qualified physiotherapist trained in real-time ultrasonography).

Statistical analysis

The sample size was calculated using G-power, version 3.1.9.7, with an effect size of 0.25, a power level of 0.80, and a p-value of 0.05.[14] The result showed that 24 participants were required. SPSS version 26.0 was used to analyze the data. The intra-rater and inter-rater reliability of muscle thickness measuring was evaluated using the Intra-class Correlation Coefficient (ICC). ICC value greater than 0.75 was considered high reliability (Table 1). The normality of the data was examined using the Shapiro-Wilk test. Since the data was not a normal distribution, a non-parametric statistic (Wilcoxon Signed-Ranks test) was used to compare the differences in muscle thickness at rest and during the exercises (i.e., ADIM and AB) in both sitting and standing positions. The level of statistical significance was set at $p<0.05$.

Results

A total of 24 healthy participants (12 males and 12 females) with a mean age of 27.58±6.70 years; height, 163.92±8.30 cm; weight, 55.18±7.73 kg; and BMI, 20.44±1.28 kg/m^2 participated in the study. The results showed that the real-time ultrasonography of the contraction of the LAM had high reliability. Table 1 shows the results of the intra-rater and inter-rater reliability of the real-time ultrasound imaging of the lateral abdominal thickness during ADIM and AB in both standing and sitting positions.

Table 1. Intra and inter-rater reliability for lateral abdominal muscle thickness during rest, ADIM, and AB.

Intra rater reliability		**Intraclass correlation coefficients ($ICC_{3,3}$)**		
Position		**TrA**	**IO**	**EO**
Standing	Rest	0.84	0.95	0.80
	ADIM	0.79	0.92	0.88
	AB	0.85	0.95	0.85
Sitting	Rest	0.95	0.85	0.94
	ADIM	0.85	0.91	0.92
	AB	0.84	0.86	0.93
Inter rater reliability		**Intraclass correlation coefficients ($ICC_{2,3}$)**		
Position		**TrA**	**IO**	**EO**
Standing	Rest	0.83	0.95	0.79
	ADIM	0.81	0.98	0.84
	AB	0.85	0.95	0.82
Sitting	Rest	0.93	0.91	0.91
	ADIM	0.80	0.98	0.82
	AB	0.83	0.89	0.78

Note: *TrA: transverse abdominis, IO: internal abdominal oblique, EO: external abdominal oblique, ADIM: abdominal drawing-in maneuver, AB: abdominal bracing.*

The thickness of the TrA and IO were significantly higher during ADIM than those of the rest and AB in both positions. In the sitting position, the thickness of the TrA and IO was significantly higher during the AB when compared to the thickness at rest, and the EO was found to be significantly higher at rest than ADIM in the sitting position (Table 2). Additionally, the thickness of the IO muscle was markedly higher during AB than at rest in a standing position (Table 3). However, when compared between the sitting and standing positions, the TrA muscle thickness was significantly higher during AB in the sitting position than the standing position (Table 4).

Table 2. Comparison of lateral abdominal muscle thickness during rest, ADIM, and AB in sitting position.

Muscle thickness (mm)	Sitting		
	Rest	ADIM	AB
TrA	34.24±13.60	53.52±17.93*†	44.77±15.66#
IO	78.04±26.50	102.33±36.50*†	96.41±41.74#
EO	70.09±16.98	65.08±15.03*	65.63±17.88

Note: *TrA: transverse abdominis, IO: internal abdominal oblique, EO: external abdominal oblique, ADIM: abdominal drawing-in maneuver, AB: abdominal bracing, *muscle thickness between ADIM and rest (p<0.05), #: muscle thickness between AB and rest (p<0.05), †comparison for muscle thickness between ADIM and AB using Wilcoxon signed rank test (p<0.05).*

Table 3. Comparison of lateral abdominal muscle thickness during rest, ADIM, and AB in standing position.

Muscle thickness (mm)	Standing		
	Rest	ADIM	AB
TrA	30.60±9.03	45.77±11.26*†	34.41±11.12
IO	83.89±25.08	102.95±34.04*†	94.36±26.60#
EO	74.08±19.85	69.59±16.54	69.12±16.28

Note: *TrA: transverse abdominis, IO: internal abdominal oblique, EO: external abdominal oblique, ADIM: abdominal drawing-in maneuver, AB: abdominal bracing, *muscle thickness between ADIM and rest (p<0.05), #muscle thickness between AB and rest (p<0.05), †comparison for muscle thickness between ADIM and AB using Wilcoxon signed rank test (p<0.05).*

Table 4. Lateral abdominal muscle thickness during ADIM and AB between standing and sitting positions.

Exercise		Sitting (mm)		Standing (mm)		*p* value
		Mean SD	Mean diff.	Mean±SD	Mean diff.	
TrA	Rest	34.23±13.60		30.60±9.03		
	ADIM	53.52±17.93	19.28±15.65	45.77±11.26	15.18±11.22	0.414
	AB	44.78±15.66	10.54±18.72	34.41±11.12	3.82±10.78	0.004*
IO	Rest	78.00±26.50		83.89±25.08		
	ADIM	102.33±36.50	24.33±24.03	102.95±34.04	19.06±19.05	0.102
	AB	96.41±41.74	18.41±30.37	94.36±26.60	10.48±17.96	0.683
EO	Rest	70.09±16.98		74.08±19.85		
	ADIM	65.08±15.03	5.01±10.71	69.59±16.54	4.49±14.32	0.221
	AB	65.63±17.88	4.46±10.83	69.12±16.28	4.96±12.62	0.683

*TrA: transverse abdominis, IO: internal abdominal oblique, EO: external abdominal oblique, ADIM: abdominal drawing-in maneuver, AB: abdominal bracing, Mean diff: mean difference between exercises and rest, *comparison between sitting and standing positions using Wilcoxon signed rank test (p<0.05).*

Discussion

This study aimed to investigate the lateral abdominal muscle thickness during ADIM and AB in sitting and standing positions. TrA and IO muscle thickness increased statistically during ADIM than AB and at rest in both positions. It was found that TrA and IO thickness increased statistically, and the thickness of the muscles was still less during AB than ADIM exercise, which may be due to the nature of the muscles during AB because of the co-contraction of the muscles, it did not directly focus on the contraction of any one muscle. TrA muscle thickness during ADIM exercise was statistically greater than AB, similar to a study by Madokoro *et al.,*[14] which reported that TrA and IO thickness increased significantly during ADIM. The result was consistent with other studies.[17,18] Based on the main anatomical function of the TrA muscle, the TrA muscle is responsible for pulling the navel towards the spine[19,20] and in addition, during ADIM, the TrA and IO muscles (i.e., deep core muscles) work together, which is associated with increased spinal stability.[21] Computed tomography (CT) was used to evaluate ADIM and AB exercises, and it was discovered that ADIM increased TrA activity more while AB enhanced rectus abdominis, IO, and EO.[22] In the sitting position, the thickness of the TrA and IO was significantly higher during the AB when compared to the thickness at rest, and the thickness of the IO muscle was significantly higher during AB than at rest in the standing position. It has been noted in earlier research that IO can be activated during AB. Nonetheless, it has been noted that the IO muscle is crucial for performing abdominal bracing exercises; it was found to have higher activation than the rectus abdominis, EO, and erector spine.[23] The EO muscle thickness was significantly increased at rest than during ADIM in the sitting position, it could be due to the state of the body when resting or not when exercising muscular contractions. All muscles will be tense (muscle tone) at a certain level to perform control posture to resist gravity and help respond to forceful movements.[24] Furthermore, increasing IO muscle activity may cause muscular thickness to suppress or overlap the EO muscle, making it appear smaller.[17] Therefore, the EO muscle thickness may decrease during ADIM and AB compared to rest. Comparison of TrA, IO, and EO between sitting and standing during ADIM and AB shows that TrA muscle thickness significantly increased during AB in the sitting position than in the standing position. On the contrary, Madokoro *et al*.[14] discovered the EO was significantly thicker during AB in the sitting position using real-time ultrasound imaging. Variance from the results may be due to the sitting posture of the participants during the measurements; in this study, the participants sat on a seat without a backrest. Further studies are recommended to confirm the assumptions.

Real-time ultrasonography is one of the techniques used to measure lateral abdominal muscle activity.[25-27] The radiological procedure used in the current study to measure the thickness of the LAM using real-time ultrasonography would be helpful in clinical practice to identify core muscle contraction among low back patients in future studies. The procedure for ultrasonography in the current study was tested for reliability before the data collection of the LAM thickness. The procedure's reliability showed that the technique was highly reliable (Table 1). The study protocol was followed as per a previously established protocol, which could have contributed to its high reliability. Also, the researcher in the current study received additional training on using real-time ultrasonography from an expert. Eventually, intra- and inter-rater reliability was established to measure the accuracy of the LAM thickness. While the procedure might require some prior training, once an individual has trained, applying the procedure in clinical practice would be effective and more accessible. Thus, real-time ultrasonography imaging might be used to evaluate core muscle activity compared to other procedures, such as electromyography, as the procedure is reliable and useful to clinical practice.[28]

Practice Implications

The current study has several practice implications for clinical practice. The study findings provide additional knowledge and understanding of the LAM muscle activity during the ADIM and AB. ADIM produced a noticeable contraction of the IO and TrA, the key muscles that contribute to the stability of the lumbopelvic region. Secondly, clinicians may consider teaching core stability exercises to the patients using ADIM as increased muscle thickness of LAM was observed during the ADIM. Thirdly, the core stability exercises can be best taught in a sitting position instead of a less-stable standing position. Also, it is clinically possible to monitor and assess the muscle thickness of LAM using real-time ultrasonography; hence, the procedure and the methods described in this study apply to diagnostic imaging and rehabilitation of low back pain. Again, the body mass index should be kept in the normal range to facilitate the visibility of LAM muscles. Finally, the EO muscle was also found to change the sitting position significantly. This is a new finding apart from previous studies. This may result from the alignment of the EO muscle, which was pulled by TrA contraction, as well as the compression by IO and gastric contents. Further studies are recommended to confirm the assumption.

Limitation

Nevertheless, the above recommendations might be taken to practice with caution as there are few study limitations. For example, the study population involved healthy participants; hence, the transferability of the study findings to the low-back pain population needs to be interpreted with caution. Future studies are warranted to investigate the activation of LAM during ADIM and AB between sitting and standing positions among individuals with low back pain. The procedure of real-time ultrasonography to measure the muscle thickness of LAM needs some practice and time. However, the procedure is highly reliable in clinical practice, as demonstrated by the reliable findings of this study. Another point of limitation that should be considered for clinical application was that both types of exercise (ADIM, AB) were tested in this study

for a few repetitions in a short period (not a prolonged training exercise program), so muscle fatigue was not evidenced. It was limited information to suggest of the appropriate intensity, frequency, and duration of training. Further research is recommended on this notion.

Conclusion

The study concludes that both ADIM and AB produced contraction of the lateral abdominal muscles in sitting and standing positions. However, ADIM can activate TrA and IO muscles better than AB. Therefore, in clinical practice, ADIM has the potential to be used to teach core stability exercises.

Conflict of interest

The author(s) declared no potential conflicts of interest in this research study.

Acknowledgements

The authors express their gratitude to all study participants and the Department of Physical Therapy and the Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Thailand, for generously providing the essential infrastructure for conducting this research. We also sincerely appreciate the CMU Presidential Scholarship for the Graduate Master's Degree Program in Movement and Exercise Sciences (AMS-CMU).

References

[1] Panjabi MM. Clinical spinal instability and low back pain. J Electromyogr Kinesiol. 2003; 13(4): 371-9. doi: 10.1016/S1050-6411(03)00044-0.

[2] Panjabi MM. The stabilizing system of the spine. Part I. Function, dysfunction, adaptation, and enhancement. J Spinal Disord. 1992; 5(4): 383-9. doi: 10.1097/00002517-199212000-00001.

[3] Ishida H, Watanabe S. Maximum expiration activates the abdominal muscles during side bridge exercise. J Back Musculoskelet Rehabil. 2014; 27(4): 481-4. doi: 10.3233/BMR-140469.

[4] Ehsani F, Arab AM, Assadi H, Karimi N, Shanbehzadeh S. Evaluation of pelvic floor muscles activity with and without abdominal maneuvers in subjects with and without low back pain. J Back Musculoskelet Rehabil. 2016; 29(2): 241-7. doi: 10.3233/BMR-150620.

[5] Akuthota V, Ferreiro A, Moore T, Fredericson M. Core stability exercise principles. Curr Sports Med Rep. 2008; 7(1): 39-44. doi: 10.1097/01.CSMR.0000308663.13278.69.

[6] Urquhart DM, Hodges PW, Allen TJ, Story IH. Abdominal muscle recruitment during a range of voluntary exercises. Man Ther. 2005; 10(2): 144-53. doi: 10.1016/J.MATH.2004.08.011.

[7] Tayashiki K, Takai Y, Maeo S, Kanehisa H. Intra-abdominal pressure and trunk muscular activities during abdominal bracing and hollowing. Int J Sports Med. 2016; 37(2): 134-43. doi: 10.1055/S-0035-1559771.

[8] Paungmali A, Joseph LH, Sitilertpisan P, Pirunsan U, Uthaikhup S. Lumbopelvic core stabilization exercise and pain modulation among individuals with chronic nonspecific low back pain. Pain Practice. 2017; 17(8): 1008-14. doi:10.1111/papr.12552.

[9] Joseph LH, Hancharoenkul B, Sitilertpisan P, Pirunsan U, Paungmali A. Comparison of effects between core stability training and sports massage therapy among elite weightlifters with chronic non-specific low back pain: a randomized cross-over study. Asian J Sports Med. 2018; 9(1): e58644. doi: 10.5812/asjsm.58644.

[10] Lis AM, Black KM, Korn H, Nordin M. Association between sitting and occupational LBP. Eur Spine J. 2007; 16(2): 283-98. doi: 10.1007/s00586-006-0143-7.

[11] Tsartsapakis I, Pantazi GA, Konstantinidou A, Zafeiroudi A, Kellis E. Spinal muscle thickness and activation during abdominal hollowing and bracing in CrossFit® athletes. Sports. 2023; 11(8): 159. doi: 10.3390/sports11080159.

[12] Escamilla RF, Lewis C, Pecson A, Imamura R, Andrews JR. Muscle activation among supine, prone, and side position exercises with and without a Swiss ball. Sports Health. 2016; 8(4): 372-9. doi: 10.1177/1941738116653931.

[13] Moghadam N, Ghaffari MS, Noormohammadpour P, Rostami M, Zarei M, Moosavi M, *et al*. Comparison of the recruitment of transverse abdominis through drawing-in and bracing in different core stability training positions. J Exerc Rehabil. 2019; 15(6): 819-25. doi: 10.12965/jer.1939064.352.

[14] Madokoro S, Yokogawa M, Miaki H. Effect of the abdominal draw-in maneuver and bracing on abdominal muscle thickness and the associated subjective difficulty in healthy individuals. Healthcare. 2020; 8(4): 496. doi: 10.3390/healthcare8040496.

[15] Joseph LH, Hussain RI, Naicker AS, Htwe O, Pirunsan U, Paungmali A. Pattern of changes in local and global muscle thickness among individuals with sacroiliac joint dysfunction. Hong Kong Physiother J. 2015; 33(1): 28-33. doi: 10.1016/j.hkpj.2014.12.003.

[16] Lee SH, Kim TH, Lee BH. The Effect of abdominal bracing in combination with low extremity movements on changes in thickness of abdominal muscles and lumbar strength for low back pain. J Phys Ther Sci. 2014; 26(1): 157-60. doi: 10.1589/jpts.26.157.

[17] Teyhen DS, Miltenberger CE, Deiters HM, *et al*. The use of ultrasound imaging of the abdominal drawing-in maneuver in subjects with low back pain. J Orthop Sports Phys Ther. 2005; 35(6): 346-55. doi: 10.2519/jospt.2005.35.6.346.

[18] Hides J, Wilson S, Stanton W, McMahon S, Keto H, McMahon K, *et al*. An MRI investigation into the function of the transversus abdominis muscle during "drawing-in" of the abdominal wall. Spine (Phila Pa 1976). 2006; 31(6): E175-8. doi: 10.1097/01.brs.0000202740.86338.df.

[19] Teyhen DS, Bluemle LN, Dolbeer JA, *et al*. Changes in lateral abdominal muscle thickness during the abdominal drawing-in maneuver in those with lumbopelvic pain. J Orthop Sports Phys Ther. 2009;

39(11), 791–8. doi.org/10.2519/JOSPT.2009.3128.

[20] Cresswell AG, Oddsson L, Thorstensson A. The influence of sudden perturbations on trunk muscle activity and intra-abdominal pressure while standing. Exp Brain Res. 1994; 98(2): 336-41. doi: 10.1007/BF00228421.

[21] Richardson CA, Jull GA. Muscle control–pain control. What exercises would you prescribe? Man Ther. 1995; 1(1): 2-10. doi: 10.1054/math.1995.0243.

[22] Koh HW, Cho SH, Kim CY. Comparison of the effects of hollowing and bracing exercises on cross-sectional areas of abdominal muscles in middle-aged women. J Phys Ther Sci. 2014; 26(2): 295-9. doi: 10.1589/JPTS.26.295.

[23] Maeo S, Takahashi T, Takai Y, Kanehisa H. Trunk muscle activities during abdominal bracing: comparison among muscles and exercises. J Sports Sci Med. 2013; 12(3): 467-74.

[24] Chantry J, Crombie S. Functional posture. In: Handbook of electronic assistive technology. Elsevier; 2019:53-80. doi: 10.1016/B978-0-12-812487-1.00003-X.

[25] Bunce SM, Hough AD, Moore AP. Measurement of abdominal muscle thickness using M-mode ultrasound imaging during functional activities. Man Ther. 2004; 9(1): 41-4. doi: 10.1016/S1356-689X(03)00069-9.

[26] Kermode F. Benefits of utilising real-time ultrasound imaging in the rehabilitation of the lumbar spine stabilising muscles following low back injury in the elite athlete - A single case study. Phys Ther Sport. 2004; 5(1): 13-6. doi: 10.1016/J.PTSP.2003.08.001.

[27] Mew R. Comparison of changes in abdominal muscle thickness between standing and crook lying during active abdominal hollowing using ultrasound imaging. Man Ther. 2009; 14(6): 690-5. doi: 10.1016/j.math.2009.05.003.

[28] Whittaker JL, Stokes M. Ultrasound imaging and muscle function. J Orthop Sports Phys Ther. 2011; 41(8): 572-80. doi: 10.2519/jospt.2011.3682.

Journal of Associated Medical Sciences 2024; 57(3): 184-191

JAMS
Journal of Associated Medical Sciences

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

Journal of Associated Medical Sciences

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

Review article: Telerehabilitation interventions for rehabilitation and management of post-COVID-19 patients: A comprehensive review

Simaporn Promsarn[1*] Wilawan Ketpan[2] Rattanapond Pankratuk[1] Pairoj Suraprapapich[3]

[1]Pulmonary Function Test Unit, Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.
[2]Nursing Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.
[3]Division of Physical Therapy, Faculty of Physical Therapy, Mahidol University, Nakhon Pathom, Thailand.

ARTICLE INFO

Article history:
Received 23 September 2023
Accepted as revised 1 August 2024
Available online 2 August 2024

Keywords:
COVID-19, telerehabilitation, post-acute, rehabilitation, functional capacity, quality of life.

ABSTRACT

The COVID-19 pandemic has posed significant challenges to healthcare systems worldwide, impacting patients during the disease's acute and post-acute phases. Telerehabilitation has emerged as a promising approach to address the rehabilitation needs of individuals recovering from COVID-19. This comprehensive review examines the effectiveness and feasibility of telerehabilitation strategies for managing post-acute COVID-19 symptoms and promoting recovery. Through a systematic review and meta-analysis, various studies have demonstrated the potential benefits of telerehabilitation interventions in improving functional capacity, exercise perception, and quality of life among post-COVID-19 patients. Telerehabilitation programs have been successfully implemented across primary care settings, offering supervised home-based exercise training that targets physical and respiratory functions.

Moreover, innovative tools such as mobile applications have been employed to enhance patient engagement and track progress. While the evidence indicates positive outcomes, there is a need for further investigation to explore the effects of telerehabilitation on additional variables such as cardiopulmonary function, anxiety, depression, and long-term outcomes. Despite limitations, telerehabilitation has shown promise in addressing the persistent symptoms and complications associated with post-acute COVID-19 syndrome. The integration of telerehabilitation into clinical care frameworks has the potential to improve patient access, engagement, and overall recovery during the ongoing pandemic and beyond. Continued research and development are essential to optimize telerehabilitation strategies for long-term COVID-19 management.

Introduction

The COVID-19 pandemic, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has had profound global implications since its emergence in late 2019.[1] This unprecedented public health crisis has resulted in significant morbidity and mortality, overwhelming healthcare systems, and precipitating a myriad of physical and psychological challenges for individuals and societies worldwide. One of the understudied yet crucial aspects of the pandemic's impact is its long-term consequences on individuals who have recovered from the acute phase of the disease, commonly referred to as post-COVID-19 patients.[2]

The COVID-19 pandemic has led to a surge in the number of post-COVID-19 patients, individuals who have recovered from the acute phase of the disease but continue to experience a range of persistent physical and psychological symptoms.[2,3] These symptoms can encompass a wide spectrum, including but not limited to dyspnea,

∗ **Corresponding contributor.**
Author's Address: *Pulmonary Function Test Unit, Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.*
E-mail address: simaporn.but@mahidol.edu
doi: 10.12982/JAMS.2024.060
E-ISSN: 2539-6056

fatigue, musculoskeletal pain, cognitive impairment, and mood disturbances.[3,4] For many post-COVID-19 patients, these symptoms pose significant challenges to their daily lives, hindering their ability to return to work, engage in physical activities, and enjoy a good quality of life.[3,5] The rationale for adopting telerehabilitation as a solution to address the diverse physical and psychological challenges faced by post-COVID-19 patients stems from several key considerations. First, the COVID-19 pandemic has strained healthcare systems, limiting many individuals' access to in-person rehabilitation services. Second, the persistent nature of post-COVID-19 symptoms necessitates ongoing rehabilitation and support. Third, minimizing the risk of virus transmission has underscored the importance of remote healthcare delivery.[1,6] Telerehabilitation, which involves using technology to provide rehabilitation services at a distance, offers a promising solution to these challenges.

The importance of addressing the diverse physical and psychological challenges faced by post-COVID-19 patients cannot be overstated. Post-COVID-19 syndrome, often referred to as Long COVID, has emerged as a complex and multifaceted condition that requires a multidisciplinary approach to rehabilitation and care.[3,4] Many post-COVID-19 patients experience limitations in their physical functioning, including reduced exercise capacity and muscle weakness.[7,8] Psychological challenges, such as anxiety and depression, are also prevalent among this population.[7,9] Therefore, a comprehensive and patient-centered approach to rehabilitation is essential to address the unique needs of post-COVID-19 patients and promote their recovery and well-being. In summary, the COVID-19 pandemic has had a profound impact on individuals who have recovered from the acute phase of the disease, resulting in a growing population of post-COVID-19 patients with diverse physical and psychological challenges. Telerehabilitation has emerged as a promising solution to provide ongoing care and support to these individuals. This essay will explore the role of telerehabilitation in addressing motor symptoms in COVID-19 patients. To identify relevant articles, a systematic search was conducted across multiple databases, including PubMed, Scopus, and Google Scholar, using keywords such as "telerehabilitation," "COVID-19," and "post-COVID syndrome." A total of 33 articles were initially identified. The selection criteria included articles published between 2021 and 2023, focusing on telerehabilitation interventions for COVID-19 patients with motor symptoms. After screening titles and abstracts, duplicates were removed, resulting in a final selection of 24 articles for inclusion in this review. These articles encompassed various study designs, including systematic reviews, randomized controlled trials, narrative reviews, and pilot studies, providing a comprehensive overview of the effectiveness and feasibility of telerehabilitation in this population. Additionally, the references of selected articles were manually searched to identify any additional relevant studies for inclusion.

Telerehabilitation modalities and approaches

Telerehabilitation has emerged as a pivotal component in the continuum of care, especially in the context of the COVID-19 pandemic. Various studies have highlighted the efficacy and feasibility of telerehabilitation modalities in addressing the diverse needs of patients post-COVID-19. For instance, a telerehabilitation program focusing on respiratory exercises improved outcomes for COVID-19 patients in the acute phase, demonstrating significant improvements in dyspnea, physical performance, and quality of life.[10] Similarly, tele-exercise interventions have been shown to enhance physical activity, functional capacity, and quality of life in patients recovering from COVID-19 and those with chronic diseases such as breast cancer.[9,11] Implementation frameworks like the A3E framework have facilitated the widespread adoption of telerehabilitation practices, allowing for increased accessibility and engagement in rehabilitation services during and beyond the pandemic.[12] Furthermore, telerehabilitation has been identified as a viable option for addressing persistent post-COVID-19 symptoms, with evidence suggesting its effectiveness in improving physical function and reducing dyspnea in affected individuals.[13] These examples underscore the versatility and effectiveness of telerehabilitation in catering to the rehabilitation needs of diverse patient populations affected by COVID-19.

Moreover, telerehabilitation has been recognized as a valuable tool for addressing the long-term sequelae of COVID-19, commonly referred to as Long-term COVID-19. Studies have demonstrated the effectiveness of telerehabilitation interventions in improving physical performance, dyspnea, and overall quality of life in Long COVID patients.[4,14] Additionally, telerehabilitation programs utilizing innovative technologies such as virtual reality games have shown promise in enhancing the physical and mental aspects of rehabilitation for post-COVID-19 syndrome.[3] These findings underscore the potential of telerehabilitation to alleviate the symptoms of Long COVID and improve the overall well-being of affected individuals.

Furthermore, integrating telecommunication technologies and remote monitoring systems has facilitated the delivery of comprehensive telerehabilitation programs tailored to individual patient needs.[7] Meta-analyses have indicated the effectiveness and safety of telerehabilitation in improving outcomes such as dyspnea, muscle strength, ambulation capacity, and depression in patients with COVID-19[7]. The scalability and accessibility of telerehabilitation make it a valuable adjunct to traditional rehabilitation services, especially in the context of global health crises like the COVID-19 pandemic. As telerehabilitation continues to evolve, ongoing research and technological advancements are poised to enhance its effectiveness further and reach, ultimately improving the delivery of rehabilitation services to patients worldwide.[15]

Clinical outcomes and benefits

The COVID-19 pandemic has brought unprecedented challenges to healthcare systems worldwide, with many individuals experiencing persistent pulmonary deficits after recovering from the acute phase of the disease. Pulmonary Rehabilitation (PR) has emerged as a vital component of post-acute COVID-19 management in response to these challenges. Given the need for physical therapy while minimizing direct contact between healthcare providers and patients, telemedicine-driven PR, also known as telerehabilitation, has become an effective alternative. This section delves into telerehabilitation's clinical outcomes and benefits, focusing on various parameters, including the 6-minute walk distance (6MWD), muscle strength, dyspnea, and quality of life. It also explores statistically significant improvements achieved through telerehabilitation and investigates differences in outcomes based on the type and intensity of telerehabilitation interventions.

1. Clinical outcomes of tTelerehabilitation

- ***6-mMinute walk distance (6MWD):*** The 6MWD is a crucial measure of physical capacity and functional ability in post-acute COVID-19 patients. Several studies have reported significant improvements in 6MWD following telerehabilitation[8.,16,17] One study showed an increase from 17.1 to 30.2 meters in 6MWD.[18] This improvement indicates enhanced exercise tolerance and functional capacity in telerehabilitation patients.
- ***Muscle strength:*** Muscle strength is another critical parameter in post-acute COVID-19 patients, as muscle weakness is a common symptom. Studies have consistently demonstrated statistically significant improvements in muscle strength through telerehabilitation interventions, particularly in programs involving strength exercises.[8,16]
- ***Dyspnea:*** Dyspnea, or shortness of breath, is a debilitating symptom experienced by many post-acute COVID-19 patients. Telerehabilitation has proven effective in reducing dyspnea levels, as evidenced by improved parameters like the Borg scale and Multidimensional Dyspnoea-12 questionnaire scores.[9,16]
- ***Quality of life:*** Quality of life is a multifaceted aspect of patient well-being. Telerehabilitation has consistently enhanced the quality of life in post-acute COVID-19 patients. Improvements have been reported in scores from the Short Form Health Survey (SF-12), St. George's Respiratory Questionnaire (SGRQ), and other relevant assessments.[9,18,19]

2. Statistical significance in telerehabilitation

Statistical significance is critical to assessing the efficacy of telerehabilitation in post-acute COVID-19 patients. Several studies have confirmed statistically significant improvements in the outcomes above.[16,17,19] These findings underscore the clinical relevance of telerehabilitation to achieve meaningful improvements in patients' health.

3. Type and intensity of telerehabilitation interventions

Telerehabilitation programs have shown remarkable diversity in their approaches. Some involve aerobic and resistance training, breathing exercises, functional activities, and muscle strengthening.[18] Others have focused on specific exercises, such as strength exercises or breathing exercises.[8,16] The variation in intervention types demonstrates flexibility in tailoring rehabilitation programs to individual patient needs. Furthermore, telerehabilitation programs have shown varying durations, ranging from 4 to 10 weeks. The choice of program duration and intensity may depend on the severity of the patient's condition and their response to treatment.

Telerehabilitation has emerged as a practical and viable approach for improving clinical outcomes in post-acute COVID-19 patients. It consistently leads to statistically significant improvements in parameters such as 6MWD, muscle strength, dyspnea, and quality of life. The type and intensity of telerehabilitation interventions vary, offering flexibility in tailoring treatment to individual patient needs. Overall, telerehabilitation is promising to enhance the recovery and quality of life of post-acute COVID-19 patients, calling for further exploration of its long-term effects, cost-effectiveness, and best practices.

Comparison with conventional rehabilitation

In the wake of the COVID-19 pandemic, healthcare systems worldwide faced unprecedented challenges in delivering rehabilitation services to patients with post-acute COVID-19 symptoms. Traditional in-person rehabilitation programs were strained due to healthcare system stress, leading to the exploration of alternative approaches such as telerehabilitation. This section aims to compare telerehabilitation outcomes with conventional in-person rehabilitation, examine adherence rates and patient preferences for remote rehabilitation, and consider telerehabilitation programs' potential cost-effectiveness and scalability.

1. Comparison of telerehabilitation outcomes with traditional in-person rehabilitation

Several studies have compared its outcomes with traditional in-person rehabilitation to evaluate the effectiveness of telerehabilitation. One systematic review found that telemedicine-driven PR significantly improved physical health, as measured by step test scores, 6-minute walk distance (6MWD), and quality of life.[18] Additionally, some parameters of pulmonary function, such as the Modified Medical Research Council (mMRC) score, Short Physical Performance Battery (SPPB) score, and Maximum Voluntary Ventilation (MVV), also showed improvement in post-acute COVID-19 patients.[16]

Furthermore, telerehabilitation interventions have improved functional capacity, exercise perception, and various health parameters among

COVID-19 patients. The improvements include 6MWD, 30-second sit-to-stand test (30STS), Borg Scale scores, and dyspnea.[17] Similar findings were observed in other studies, indicating that telerehabilitation can enhance physical function, reduce dyspnea, and improve overall quality of life.[21]

2. Examination of adherence rates and patient preferences for remote rehabilitation

Adherence rates in telerehabilitation programs have been a subject of investigation. One study reported a high completion rate of telerehabilitation interventions, with the most common reason for withdrawal being lost to follow-up or uncooperativeness. The study demonstrated that telerehabilitation can be implemented with minimal adverse events and high patient engagement.[17] Similarly, another trial found that telerehabilitation had a 90% adherence rate, indicating a strong patient commitment to remote rehabilitation programs.[8]

Patient preferences for telerehabilitation have also been explored. A qualitative study revealed that patients isolated due to COVID-19 infection highly appreciated telerehabilitation programs. Patients highlighted the importance of applying these programs in public health systems, indicating a positive reception of telerehabilitation as a viable alternative to traditional in-person care.[22]

3. Consideration of potential cost-effectiveness and scalability of telerehabilitation programs

To assess the cost-effectiveness and scalability of telerehabilitation programs, comparing unit costs between telerehabilitation and traditional in-person rehabilitation services is crucial. Telerehabilitation has shown effectiveness in improving outcomes for post-acute COVID-19 patients, with high completion rates and minimal adverse events.[17] Integrating telerehabilitation into existing healthcare systems can optimize resource utilization.[1] Studies have indicated that telemedicine-driven pulmonary rehabilitation interventions can reduce direct costs compared to in-person rehabilitation. However, a detailed analysis of unit costs is essential to understand the economic implications fully. Further research should focus on conducting comprehensive economic evaluations to inform decision-making regarding telerehabilitation implementation.[1,17]

The comparison of telerehabilitation with traditional in-person rehabilitation reveals promising outcomes for telerehabilitation in the management of post-acute COVID-19 patients. Telerehabilitation has shown effectiveness in improving physical and mental health, quality of life, and pulmonary function. Moreover, high adherence rates and positive patient preferences indicate its acceptability and potential for scalability. Considering the cost-effectiveness and increasing demand for remote healthcare solutions, telerehabilitation emerges as a valuable approach to rehabilitation, especially in managing post-acute COVID-19 symptoms. Further research and exploration are warranted to optimize and effectively integrate telerehabilitation into healthcare systems.

Challenges and considerations

The acute phase of COVID-19 often leaves patients with persistent pulmonary deficits, making PR a recommended part of post-acute management. Telerehabilitation, encompassing a range of remote interventions, has gained prominence in delivering PR to these patients. However, the effective implementation of telerehabilitation for diverse patient populations is fraught with challenges.

1. Challenges in implementing telerehabilitation

- ***Technology access:*** One of the foremost challenges is ensuring that all patients, irrespective of their socioeconomic status or geographic location, have access to the necessary technology.[17] This includes access to smartphones, computers, and a stable internet connection. Disparities in access may result in unequal healthcare delivery.
- ***Patient engagement:*** Sustaining patient engagement in a remote setting is another significant hurdle. Patients often require motivation to adhere to telerehabilitation programs, which may lack the direct supervision and social support of traditional in-person rehabilitation.[21] The potential for patient fatigue and disinterest exacerbates this challenge.
- ***Healthcare provider training:*** Healthcare providers must be proficient in delivering telerehabilitation services. This necessitates training in using telehealth platforms, understanding the nuances of remote patient assessment, and adapting to the virtual environment.[12] Ensuring that providers are adequately prepared is crucial for the success of telerehabilitation programs.
- ***Quality of care:*** Concerns linger about the quality of care provided through telerehabilitation. Clinicians may face limitations in their ability to perform physical assessments remotely, potentially leading to inaccurate diagnoses or treatment plans.[7] Ensuring that telerehabilitation maintains a high standard of care is paramount.

2. Considerations in addressing challenges

- ***Equitable access:*** To address technology access disparities, healthcare systems should invest in providing necessary devices and internet connectivity to underserved populations.[15] Additionally, strategies like mobile clinics or community centers equipped for telehealth can bridge the accessibility gap.[16]
- ***Patient-centered approach:*** Patient engagement can be enhanced through a patient-centered approach that incorporates patient preferences and goals into telerehabilitation programs. Gamification, tele-supervision, and remote monitoring tools can maintain patient interest and motivation.[20]

- ***Provider training:*** Robust training programs for healthcare providers should be developed, covering the technical and interpersonal skills required for telerehabilitation. Continuous education and proficiency assessments can ensure providers are well-equipped.[12]
- ***Quality assurance:*** To maintain the quality of care, telerehabilitation programs should incorporate standardized assessment tools and continuous quality monitoring.[17] This can help identify and rectify potential inaccuracies in remote assessments.

Implementing telerehabilitation for diverse patient populations recovering from COVID-19 presents a promising avenue for improving post-acute care. However, it is not without its challenges. Ensuring equitable access, patient engagement, healthcare provider training, and high-quality care are vital considerations for successfully integrating telerehabilitation into healthcare systems. Tailored strategies that consider the unique needs of different patient populations are crucial for realizing the full potential of telerehabilitation in post-acute COVID-19 management. These strategies must be informed by empirical evidence and a commitment to equitable healthcare delivery.

Safety and adverse events

Telerehabilitation, a burgeoning facet of modern healthcare delivery, has garnered significant attention, particularly during the COVID-19 pandemic. As traditional healthcare systems faced unprecedented challenges in delivering rehabilitation services, telerehabilitation emerged as a promising alternative to bridge the gap between patients and essential care.[22] Despite its growing popularity, the safety considerations associated with telerehabilitation remain paramount. Adverse events, though rare, present a critical aspect of patient care and must be thoroughly evaluated to ensure the efficacy and safety of telerehabilitation interventions.[19] An adverse event, such as a musculoskeletal injury or worsening of existing symptoms, can occur during telerehabilitation sessions, highlighting the importance of meticulous risk assessment and mitigation strategies.[16] Therefore, a comprehensive understanding of adverse event rates and safety protocols is imperative for healthcare practitioners and policymakers to optimize patient outcomes and minimize potential risks associated with telerehabilitation.[17]

1. Evaluation of adverse events

To assess the safety of telerehabilitation, it is essential to analyze adverse events associated with this mode of care delivery. The available evidence from systematic reviews and clinical studies offers valuable insights. For example, a systematic review that evaluated the effectiveness of telemedicine-driven PR in post-acute COVID-19 patients reported no adverse events during the telerehabilitation programs.[23] Similarly, another study examining the short-term effects of a telerehabilitation program in confined COVID-19 patients in the acute phase found no adverse events.[8] These findings provide initial reassurance regarding the safety of telerehabilitation in the context of COVID-19 rehabilitation.

2. Comparison of adverse event rates

A critical aspect of assessing telerehabilitation safety is comparing adverse event rates between telerehabilitation and conventional rehabilitation. To date, the available evidence points towards the safety of telerehabilitation. In a randomized controlled trial meta-analysis, telerehabilitation was effective and safe for patients with COVID-19, with no severe adverse events reported.[7] Furthermore, compared to traditional in-person care, patients treated using telerehabilitation had similar satisfaction levels with their treatment results, suggesting that telerehabilitation does not compromise patient safety or satisfaction.[15]

3. Strategies to ensure patient safety

While the evidence is promising, strategies to mitigate further risks associated with telerehabilitation are essential. These strategies should encompass various aspects of telerehabilitation, including technology, patient assessment, and care delivery.

- ***Robust technology infrastructure:*** It is paramount to ensure a stable and secure technology platform for telerehabilitation. This includes reliable internet connectivity, secure data transmission, and user-friendly interfaces to minimize technical issues that could lead to adverse events.
- ***Comprehensive patient assessment:*** Telerehabilitation should start with a thorough assessment of the patient's clinical condition and needs. This assessment should include the patient's physical health, technological literacy, and access to necessary devices.
- ***Patient education:*** Patients should receive comprehensive education about the telerehabilitation process, including clear instructions on performing exercises or using rehabilitation tools safely. This education can help prevent accidents or injuries during remote sessions.
- ***Continuous monitoring:*** Implementing mechanisms for real-time monitoring of patients during telerehabilitation sessions can provide an added layer of safety. Monitoring can be achieved through video calls, wearable devices, or other remote monitoring tools.
- ***Emergency protocols:*** Telerehabilitation providers should have clear protocols, including procedures for handling adverse events or medical emergencies during remote sessions.

Telerehabilitation has emerged as a valuable and safe mode of delivering rehabilitation services, particularly in the context of COVID-19 recovery. The evidence suggests that adverse events associated with

telerehabilitation are minimal, and patient satisfaction is comparable to conventional rehabilitation. However, to ensure patient safety, it is essential to maintain a robust technology infrastructure, conduct comprehensive patient assessments, provide patient education, enable continuous monitoring, and establish emergency protocols. As telerehabilitation continues to evolve, further research and ongoing evaluation of safety considerations will be crucial to optimize patient outcomes in the post-COVID-19 era and beyond.

Future directions and implications

The COVID-19 pandemic has brought significant challenges to healthcare systems worldwide, not only in the acute management of the disease but also in addressing the long-term consequences experienced by survivors. A growing body of evidence suggests that a substantial proportion of post-COVID-19 patients, including those with long-term COVID-19, continue to suffer from various physical, psychological, and functional impairments even after the acute phase of the disease. In response to these challenges, telerehabilitation has emerged as a promising approach to providing ongoing care and support for post-COVID-19 patients. This essay discusses the potential long-term effects of telerehabilitation on post-COVID-19 patients, the role of telerehabilitation in managing long COVID, and provides recommendations for future research in this critical area.

1. Long-term effects of telerehabilitation

- ***Physical health:*** Several studies have shown that telerehabilitation can significantly improve physical health among post-COVID-19 patients. For instance, telerehabilitation programs have been associated with enhanced exercise capacity, as demonstrated by improvements in the 6-minute walk distance (6MWD).[16,17,21] Telerehabilitation has improved muscle strength, as evidenced by the 30-second sit-to-stand test (30STS) results.[16,17] These findings are particularly important, as many post-COVID-19 patients report persistent fatigue and exercise intolerance.[20] Furthermore, telerehabilitation has demonstrated its effectiveness in improving lung function parameters such as the Modified Medical Research Council (mMRC) scale.[16,17] This is particularly relevant given that many post-COVID-19 patients experience respiratory symptoms.[23] Therefore, telerehabilitation may be crucial in mitigating the long-term pulmonary effects of COVID-19.
- ***Mental health:*** Mental health is another crucial aspect of post-COVID-19 care, as many patients experience anxiety, depression, and psychological distress.[20,24] Telerehabilitation has shown promise in addressing these mental health challenges. Studies have reported significant improvements in mental health outcomes, including reductions in anxiety and depression scores. [16,17,24] Enhanced mental well-being is vital for the overall quality of life of post-COVID-19 patients and may also contribute to better adherence to rehabilitation programs.
- ***Quality of life:*** Improvements in the quality of life have been consistently observed in post-COVID-19 patients who undergo telerehabilitation.[16,17,24] These improvements are reflected in various quality-of-life assessment tools, such as the Short Form-12 (SF-12) and the St. George's Respiratory Questionnaire (SGRQ). A better quality of life is not only a fundamental patient-centered outcome but also an indicator of the holistic impact of telerehabilitation on post-COVID-19 recovery.

2. Role of telerehabilitation in managing long-term COVID-19

Long COVID-19, characterized by persistent symptoms and functional limitations, presents a substantial challenge for healthcare providers.[3,20] Telerehabilitation holds promise in addressing the complex and diverse symptoms associated with long-term COVID-19. By providing tailored exercise programs, breathing exercises, and psychological support through remote platforms, telerehabilitation can offer continuous care for individuals experiencing lingering symptoms. The adaptability of telerehabilitation programs allows for personalized interventions based on the evolving needs of long COVID patients.

3. Recommendations for future research

While the existing evidence highlights the benefits of telerehabilitation for post-COVID-19 patients, further research is essential to enhance our understanding and optimize its use. To this end, several recommendations for future research are proposed.

- ***Long-term follow-up studies:*** Conduct longitudinal studies to assess the sustained effects of telerehabilitation on post-COVID-19 patients, including those with long-term COVID-19. Such studies should track changes in physical health, mental health, and quality of life over an extended period.
- ***Diversity and inclusivity:*** Ensure that future research includes diverse populations to account for variations in patient characteristics, such as age, comorbidities, and disease severity. This will help tailor telerehabilitation programs to the specific needs of different patient groups.
- ***Cost-effectiveness analysis:*** Evaluate the cost-effectiveness of telerehabilitation compared to traditional in-person rehabilitation. This analysis is crucial for healthcare systems seeking efficient ways to provide care to a growing population of post-COVID-19 patients.
- ***Optimal program design:*** Investigate telerehabilitation programs' optimal design and components for post-COVID-19 care. This includes determining the ideal duration, frequency, and content of remote rehabilitation sessions.
- ***Telemonitoring:*** Explore integrating telemonitoring technologies to track patients' progress and adjust interventions in real time. This can enhance the

effectiveness of telerehabilitation by enabling timely modifications to treatment plans.

- ***Patient-centered outcomes:*** Emphasize patient-centered outcomes in research, including patient-reported symptoms, functional improvements, and satisfaction with telerehabilitation programs.

Telerehabilitation has emerged as a valuable tool in the comprehensive care of post-COVID-19 patients, offering improvements in physical, mental, and overall quality of life. Its potential in managing long COVID is particularly promising. However, continued research is necessary to refine telerehabilitation programs, ensure cost-effectiveness, and address the diverse needs of post-COVID-19 patients. As the healthcare landscape evolves, telerehabilitation stands as a vital component in the ongoing battle against the long-term effects of the COVID-19 pandemic.

Conclusion

In the wake of the global COVID-19 pandemic, telerehabilitation has emerged as a crucial tool in the management of post-acute COVID-19 patients. A systematic review of studies investigating telerehabilitation's effectiveness has shed light on its remarkable potential. Telerehabilitation encompasses various interventions, including tele-supervised home-based exercise training, breathing exercises, and digital physiotherapy practice. These interventions have demonstrated significant improvements in various health parameters, including physical health, mental health, quality of life, and pulmonary function.

One key finding is that telerehabilitation significantly enhances physical health, as evidenced by improved parameters like the step test score, 6-minute walking distance, and muscle strength. Additionally, mental health benefits are notable, with reduced fatigue, dyspnea, and improved psychological well-being. Quality of life measurements, including the SF-12, SGRQ, and CAT scores, showed favorable outcomes after telerehabilitation. Moreover, pulmonary function, as assessed by parameters like mMRC, STST, and MVV, displayed positive changes.

The implications of these findings are significant. Telerehabilitation is a feasible and practical approach to enhance the recovery of post-acute COVID-19 patients. Its integration into routine post-COVID-19 care is strongly recommended. This recommendation is supported by the high completion rates and minimal adverse events associated with telerehabilitation interventions. Nonetheless, further research is needed to investigate its effects on cardiopulmonary function, cost-effectiveness, and its potential impact on variables like anxiety and depression.

In conclusion, telerehabilitation has demonstrated its potential to play a pivotal role in post-acute COVID-19 care. The evidence suggests that it can improve physical and mental health, enhance quality of life, and contribute to patients' recovery. As the world continues to grapple with the long-term effects of the COVID-19 pandemic, telerehabilitation stands out as a promising approach that merits further exploration and integration into standard care protocols.

References

[1] Cerfoglio S, Capodaglio P, Rossi P, *et al*. Telerehabilitation interventions for motor symptoms in COVID-19 Patients: A narrative review. Bioengineering (Basel) 2023; 10(6): 650. doi:10.3390/bioengineering10060650

[2] Nalbandian A, Sehgal K, Gupta A, *et al*. Post-acute COVID-19 syndrome. Nat Med. 2021; 27: 601-15. doi:10.1038/s41591-021-01283-z

[3] Valverde-Martínez MÁ, López-Liria R, Martínez-Cal J, Benzo-Iglesias MJ, Torres-Álamo L, Rocamora-Pérez P. Telerehabilitation, A viable option in patients with persistent post-COVID syndrome: A systematic review. Healthcare (Basel). 2023; 11(2): 187. doi:10.3390/healthcare11020187

[4] Bernal-Utrera C, Montero-Almagro G, Anarte-Lazo E, Gonzalez-Gerez JJ, Rodriguez-Blanco C, Saavedra-Hernandez M. Therapeutic exercise interventions through telerehabilitation in patients with post COVID-19 symptoms: A systematic review. J Clin Med. 2022; 11(24): 7521. doi:10.3390/jcm11247521

[5] Pehlivan E, Palalı İ, Atan SG, Turan D, Çınarka H, Çetinkaya E. The effectiveness of POST-DISCHARGE telerehabilitation practices in COVID-19 patients: Tele-COVID study-randomized controlled trial. Ann Thorac Med. 2022; 17(2): 110-7. doi:10.4103/atm.atm_543_21

[6] Li J, Xia W, Zhan C, *et al*. A telerehabilitation programme in post-discharge COVID-19 patients (TERECO): a randomised controlled trial. Thorax. 2022; 77(7): 697-706. doi:10.1136/thoraxjnl-2021-217382

[7] Huang J, Fan Y, Zhao K, *et al*. Do patients with and survivors of COVID-19 benefit from telerehabilitation? A meta-analysis of randomized controlled trials. Front Public Health .2022; 10: 954754. doi:10.3389/fpubh.2022.954754

[8] Rodriguez-Blanco C, Gonzalez-Gerez JJ, Bernal-Utrera C, Anarte-Lazo E, Perez-Ale M, Saavedra-Hernandez M. Short-Term Effects of a Conditioning Telerehabilitation Program in Confined Patients Affected by COVID-19 in the Acute Phase. A Pilot Randomized Controlled Trial. Medicina (Kaunas). 2021; 57(7): 684. doi:10.3390/medicina57070684

[9] Vieira AGDS, Pinto ACPN, Garcia BMSP, Eid RAC, Mól CG, Nawa RK. Telerehabilitation improves physical function and reduces dyspnoea in people with COVID-19 and post-COVID-19 conditions: a systematic review. J Physiother. 2022; 68(2): 90-8. doi:10.1016/j.jphys.2022.03.011

[10] Gonzalez-Gerez JJ, Saavedra-Hernandez M, Anarte-Lazo E, Bernal-Utrera C, Perez-Ale M, Rodriguez-Blanco C. Short-term effects of a respiratory telerehabilitation program in confined COVID-19 patients in the acute phase: A pilot study. Int J Environ Res Public Health. 2021; 18(14): 7511. doi:10.3390/ijerph18147511

[11] Peng Y, Zhang K, Wang L, *et al*. Effect of a telehealth-based exercise intervention on the physical activity of patients with breast cancer: A systematic review and meta-analysis. Asia Pac J Oncol Nurs. 2022; 9(12): 100117. doi:10.1016/j.apjon.2022.100117

[12] Kim SY, Daley K, Pruski AD, *et al*. Implementation of a framework for telerehabilitation in clinical care across the continuum during COVID-19 and beyond. Am J Phys Med Rehabil. 2022; 101(1): 53-60. doi:10.1097/PHM.0000000000001904

[13] Samper-Pardo M, León-Herrera S, Oliván-Blázquez B, Méndez-López F, Domínguez-García M, Sánchez-Recio R. Effectiveness of a telerehabilitation intervention using ReCOVery APP of long COVID patients: a randomized, 3-month follow-up clinical trial. Sci Rep. 2023; 13(1): 7943. doi:10.1038/s41598-023-35058-y

[14] Zasadzka E, Trzmiel T, Pieczyńska A, Hojan K. Modern technologies in the rehabilitation of patients with multiple sclerosis and their potential application in times of COVID-19. Medicina (Kaunas). 2021; 57(6): 549. doi:10.3390/medicina57060549

[15] Werneke MW, Deutscher D, Grigsby D, Tucker CA, Mioduski JE, Hayes D. Telerehabilitation during the COVID-19 pandemic in oOutpatient rehabilitation settings: A descriptive study. Phys Ther. 2021; 101(7): pzab110. doi:10.1093/ptj/pzab110

[16] Rodríguez-Blanco C, Bernal-Utrera C, Anarte-Lazo E, *et al*. Breathing exercises versus strength exercises through telerehabilitation in coronavirus disease 2019 patients in the acute phase: A randomized controlled trial. Clin Rehabil .2022; 36(4): 486-97. doi:10.1177/02692155211061221

[17] Seid AA, Aychiluhm SB, Mohammed AA. Effectiveness and feasibility of telerehabilitation in patients with COVID-19: a systematic review and meta-analysis. BMJ Open. 2022; 12(10): e063961. doi:10.1136/bmjopen-2022-063961

[18] Pescaru CC, Crisan AF, Marc M, *et al*. A systematic review of telemedicine-driven pulmonary rehabilitation after the acute phase of COVID-19. J Clin Med. 2023; 12(14): 4854. doi:10.3390/jcm12144854

[19] Chen JJ, Cooper DM, Haddad F, Sladkey A, Nussbaum E, Radom-Aizik S. Tele-exercise as a promising tool to promote exercise in children with cystic fibrosis. Front Public Health. 2018; 6: 269. doi:10.3389/fpubh.2018.00269

[20] Dalbosco-Salas M, Torres-Castro R, Rojas Leyton A, *et al*. Effectiveness of a primary care telerehabilitation program for post-COVID-19 patients: A feasibility study. J Clin Med. 2021; 10(19): 4428. doi:10.3390/jcm10194428

[21] Hajibashi A, Sarrafzadeh J, Amiri A, Salehi R, Vasaghi-Gharamaleki B. Effect of progressive muscle relaxation as an add-on to Pulmonary telerehabilitation in discharged patients with COVID-19: A randomised controlled trial. Complement Ther Clin Pract. 2023; 51: 101730. doi:10.1016/j.ctcp.2023.101730

[22] Bernal-Utrera C, Anarte-Lazo E, De-La-Barrera-Aranda E, *et al*. Perspectives and attitudes of patients with COVID-19 toward a telerehabilitation programme: A qualitative study. Int J Environ Res Public Health. 2021; 18(15): 7845. doi:10.3390/ijerph18157845

[23] Martin I, Braem F, Baudet L, *et al*. Follow-up of functional exercise capacity in patients with COVID-19: It is improved by telerehabilitation. Respir Med. 2021; 183: 106438. doi:10.1016/j.rmed.2021.106438

[24] Estebanez-Pérez MJ, Martín-Valero R, Vinolo-Gil MJ, Pastora-Bernal JM. Effectiveness of digital physiotherapy practice compared to usual care in long COVID patients: A systematic review. Healthcare (Basel). 2023; 11(13): 1970. doi:10.3390/healthcare11131970

Journal of Associated Medical Sciences 2024; 57(3): 192-203

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

A-Speak: Augmentative and alternative communication application for Thai individuals with complex communication needs

Nittaya Kasemkosin[1] Worawan Wattanawongsawang[2] Saowaluck Kaewkamnerd[3] Rachaporn Keinprasit[3] Alisa Suwannarat[3] Wansiya Kamonsitichai[1*]

[1]Department of Communication Sciences and Disorders, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.
[2]Thonburi Hospital 1, Bangkok, Thailand.
[3]National Science and Technology Development Agency, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 1 April 2024
Accepted as revised 1 August 2024
Available online 5 August 2024

Keywords:
A-Speak, augmentative and alternative communication, application, High-tech AAC.

ABSTRACT

Background: Augmentative and alternative communication (AAC) is the approach that enhances communication competence in individuals with complex communication needs. With the advancement of technology, there are varieties of AAC applications with colored-graphic symbols and speech output, improving communication's intelligibility compared to low-tech AAC systems. However, those AAC applications had some features that were not entirely suitable for Thai users, such as symbol appearance, speech intonation, etc.

Objective: This study aimed to develop the first version of the Thai AAC application, A-Speak, based on Thai culture, lexicon, and intonation and remove other constraints that other AAC applications had, such as variation in voice-output age and gender. The proficiency of A-Speak regarding communication functions was also examined.

Materials and methods: The participants comprised 15 individuals with cerebral palsy and complex communication needs. The participants were trained to use the A-Speak application, installed on a tablet, to communicate. The training procedures consisted of 3 phases: Phase 1: Train to select icons; Phase 2: Shift to different categories; and Phase 3: Use A-Speak to communicate. The researchers trained the participants to achieve adequate operational skills (i.e., Phases 1 and 2) before beginning Phase 3. In Phases 1 and 2, switches were employed to facilitate participants with limited mobility to operate A-Speak by finger. The researchers also taught the participants' caregivers to continue training them at home. The researchers collected the participants' communication abilities regarding communication functions in the recorded form. The data was reported into code numbers according to communication proficiency.

Results: After receiving A-Speak training, all participants showed improvement in their communication abilities across a variety of communication functions. Participants showed significant progress in 10 out of 12 communication functions. The communication function in which participants exhibited the most improvement was explaining skills, whereas the communication function that showed the least development was storytelling skills.

Conclusion: A-Speak AAC application reduced the constraints that possibly influenced communication intelligibility in the Thai language. Nevertheless, A-Speak still had a few drawbacks that required to be corrected to increase the productivity of this program. The findings indicated that participants gained communication skills through A-Speak as a means of communication.

* **Corresponding contributor.**
Author's Address: *Department of Communication Sciences and Disorders, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.*
E-mail address: wansiya.kam@mahidol.ac.th
doi: 10.12982/JAMS.2024.061
E-ISSN: 2539-6056

Introduction

A communication disorder is an impairment that affects individuals' speech and/or language comprehension and expression abilities.[1] Without adequate communication skills, individuals cannot clearly understand others' speech, respond appropriately in social situations, and/or adequately express their thoughts and feelings.[2] Moreover, communication skill deficits affect their living and social skills, preventing them from succeeding.[3,4] Nevertheless, the degree of these difficulties depends on the severity level of communication impairments. Individuals with profound communication impairments may have limited use and comprehension of speech due to concomitant impairments in intellectual, sensory, motor, and other areas, also known as complex communication needs. The severity might vary from being completely unable to speak to being able to speak with limited intelligibility. These might be the result of congenital disabilities (e.g., cerebral palsy, apraxia of speech, intellectual disabilities, etc.), acquired disabilities (e.g., acquired brain injuries, cerebrovascular accidents, laryngectomy, etc.), or neurological differences (e.g., autism, etc.).[1] There are several approaches to help individuals improve their communication skills, including augmentative and alternative communication (AAC).

The primary purpose of the AAC system is to help individuals with communication disorders promote their communicative competence and develop their language skills, including literacy skills.[5-7] No research currently exists that reports AAC systems impede speech development; instead, multiple studies cite that these systems may even promote speech development.[8-10] Furthermore, several studies reported that AAC reduces challenge behaviors due to the limitation of verbal communication[6, 11,12] AAC are classified into two main systems: (1) Unaided systems that communicate primarily through physical movement and do not require additional tools (e.g., manual signs, face expression, gestures, etc.); and (2) Aided systems that communicate through additional tools. Aided systems include low-tech (e.g., writing, communication boards, letter board, etc.) and high-tech systems (e.g., speech-generating devices (SGDs), iPad-based speech output technologies, Android-based speech output technologies, etc.).[1,13,14] High-tech systems offer features such as adjustable vocabulary, dynamic storage, and voice output which enhance the individuals' communication intelligibility.[15]

Many studies reported that AAC systems with speech-output technologies could improve the communication skills of people with communication difficulties.[16-19] Wansiya and Goldstein reported that most Thai speech-language pathologists in the study preferred administering speech-output technologies to individuals with severe speech impairment or complex communication needs.[20] However, the available applications, such as SymboTalk, Leeloo, Funjai, and Cboard, do not entirely support the Thai language. Thai voice output in those applications presented inaccurately pronounced intonation, which would result in unclear messages. Besides, the voice synthesis offering for the application was limited to adult-female voices, which might not be suitable for male and young users. Wickenden indicated that speech-output options (e.g., age, gender appropriate, accent, etc.) play a role in individuals' identity related to selfhood and the personhood of AAC users.[21] Pullin and Hennig also demonstrated that voice should be age- and gender-appropriate to the user to reduce confusion among communication partners due to the voice's mismatch in terms of age and gender.[22]

Therefore, this study aimed to develop an AAC application specifically for Thai individuals with complex and limited verbal communication needs. A-Speak or All-Speak has been designed as an Android-based speech output technology compatible with external input devices such as switches, keyboards, eye-tracking, speakers, etc. Therefore, A-Speak would facilitate communication competence for individuals with multiple disabilities, especially those who have a limited range of physical movement.

Materials and methods

Research design

This study employed quantitative descriptive research. The intervention consisted of three main phases. The first two phases involved operational skill training (e.g., selecting symbols and changing pages). The third phase focused on teaching communication skills through the A-Speak application. In the third phase, data collection regarding the changes in communication abilities was conducted three times on separate days. The data were gathered in recorded form.

Materials

Recorded form

The recorded form contained 2 main parts. The first part involved participants' demographic information**,** including gender, age, education level, and diagnosis. The second part involved participants' communication abilities using A-Speak application according to 12 different communication intentions. The data in the second part was collected in three separate days. The data was collected in the form of code (i.e., 0 = No communication; 1 = One occasion; 2 = More than one occasion; and 3 = Communication on all occasions) and written in the recorded form developed explicitly for this study. Three speech-language pathologists who were not involved in this study were assigned to check the content validity of the recorded form. They were asked to rate the importance of each communication function that was required to be evaluated to determine whether the A-Speak application fulfilled general communication needs for the participants (i.e., -1 = completely irrelevant; 0 = fair; and 1 = very relevant). Only 12 out of 15 communication functions that the three speech-language pathologists had a consensus that the study should consider. The recorded form is depicted in Appendix A.

A-Speak AAC application

The Thai National Electronics and Computer Technology Center (NECTEC) collaborated with researchers who were speech-language pathologists, disability specialists, and AAC professionals to create the AAC application, A-Speak. The application is currently available on the Android mobile operating system. A-Speak had two main modes: (a) Graphic-symbol board and (b) keyboard.

Graphic-symbol board

This system displayed colored-graphic symbols with text identifying the meaning of the symbols in the form of grid displays. Those symbol icons were arranged in the form of a board in both schematic and taxonomic grid displays, as shown in Figure 1. The number of grids per page could be adjusted to 3x5, 5x7, 6x9, or 7x11, as shown in Figure 2. The screen display was divided into three main parts: (a) main symbolic icons (i.e., taxonomic grid displays), (b) vocabulary categories (semantic-syntactic grid displays), and (c) visual message feedback bar.

The new user might feel overwhelmed by initially learning to use A-Speak. To reduce this problem, A-Speak is programmed with the "vocabulary mask" feature. This feature allows speech-language pathologists to fade out the symbol icons, which the user does not presently require, from the main board.

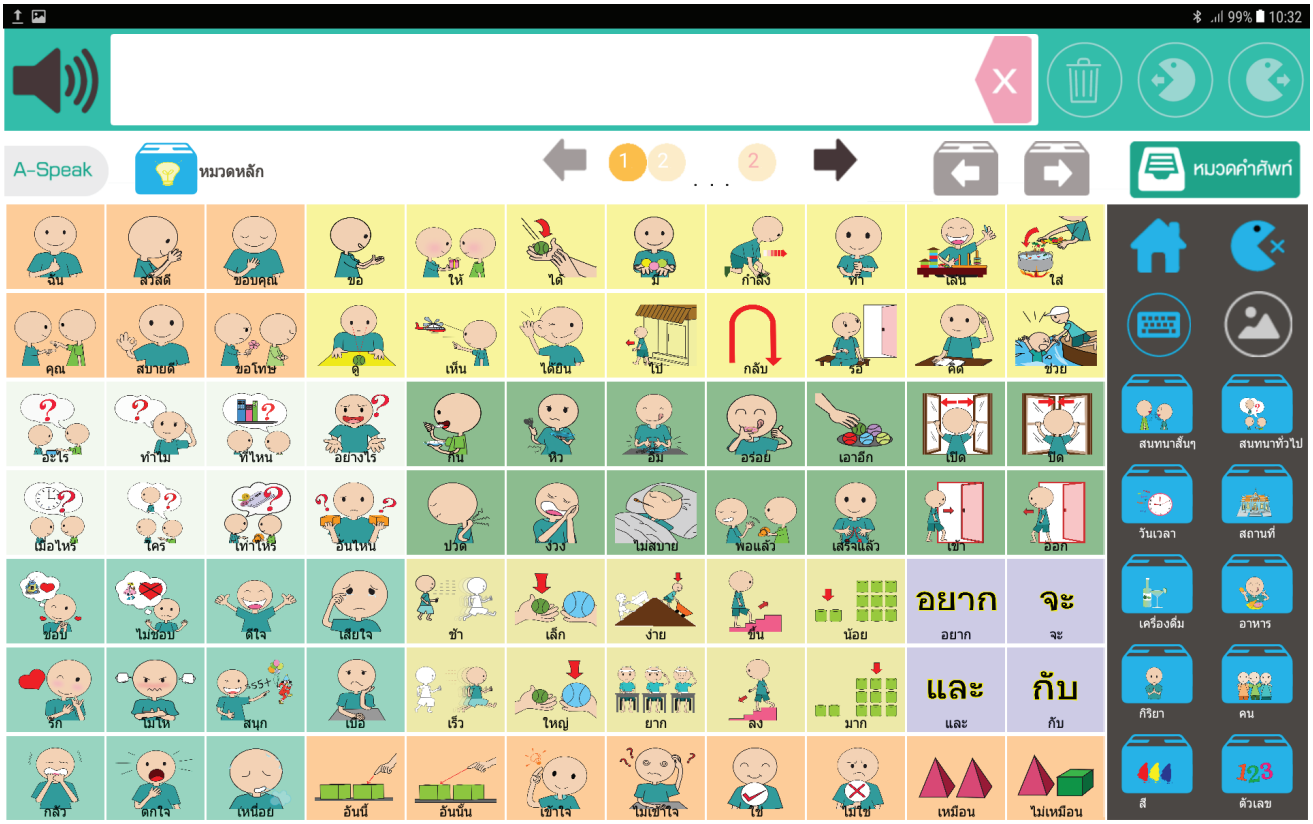


Figure 1. *The taxonomic grid displays.*

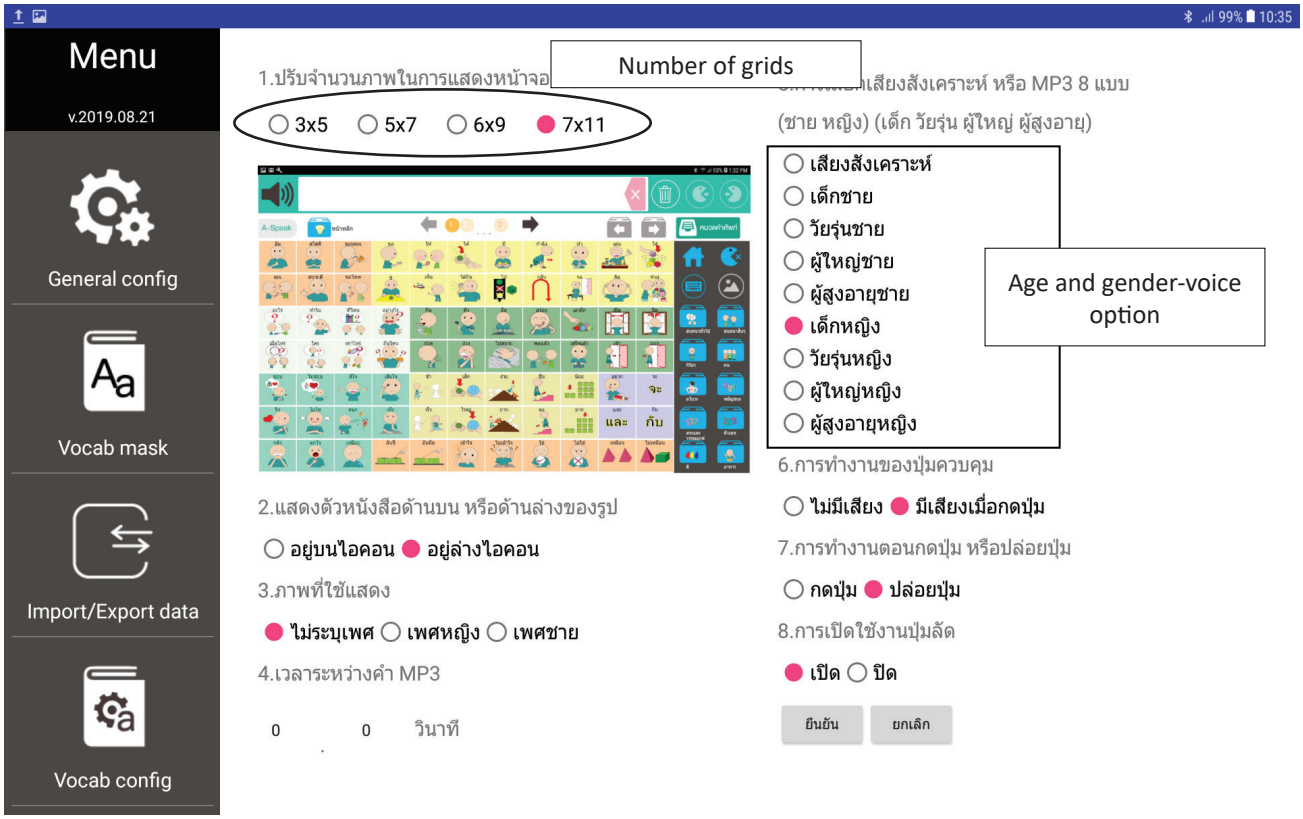


Figure 2. *The setting of grid display and voice output.*

There were more than 1,422 symbolic icons and 23 categories provided in A-Speak. Those symbol icons served as the vocabulary for basic communication functions (e.g., calling attention, protesting, requesting objects, etc.).[23]

The vocabulary was adapted from the Thai for Beginners book and Thai culture.[24] The vocabulary words and phrases in this book are arranged according to several communication functions (e.g., asking, answering, introducing oneself, etc.) in many circumstances that fulfill young individuals' and adults' basic communication needs. Therefore, this book's vocabulary words and phrases were selected to serve as the foundational concepts in A-Speak. According to the guidelines, the main page was recommended to consist of core vocabulary (i.e., words adaptable to a range of settings and can be utilized for several communication purposes) and fringe vocabulary (i.e., words unique to specific situations and settings).[25,26] 77 symbol icons were displayed on the main page. These icons comprised nouns, pronouns, verbs, adverbs, adjectives, questions, and expressions. The icons associated with the same categories had the same background color and were positioned adjacent to one another; for instance, WH-question terms (i.e., what, where, who, when, why, how much) were arranged together with the same background color (i.e., light green). The number of icons depended on the display grid size, which was evaluated and determined by speech-language pathologists.

Female prisoners drew all symbol icons in A-Speak from the Information Technology Foundation under the Initiative of Her Royal Highness Princess Maha Chakri Sirindhorn. Every icon has a written text that represents its meaning attached. The written text could be set to appear either above or below the icon. Furthermore, the text and icons' background color were adjustable to support the users with vision impairments (e.g., low visual acuities, narrow visual field, etc.).

According to Beukelman and Fager, visual attention was influenced by user relevance to the age and gender of the graphic symbols.[27] Therefore, A-Speak provided graphics symbols representing both male and female characters. The user can choose the gender of the symbol to symbolize his or her gender, as shown in Figure 2. The users were also able to add new categories and pictures.

Keyboard mode

There were two alphabets: Thai and English. The Thai and English keyboards were arranged in a QWERTY platform, as shown in Figure 3. This mode also offered the writing mode, which the user might prefer.

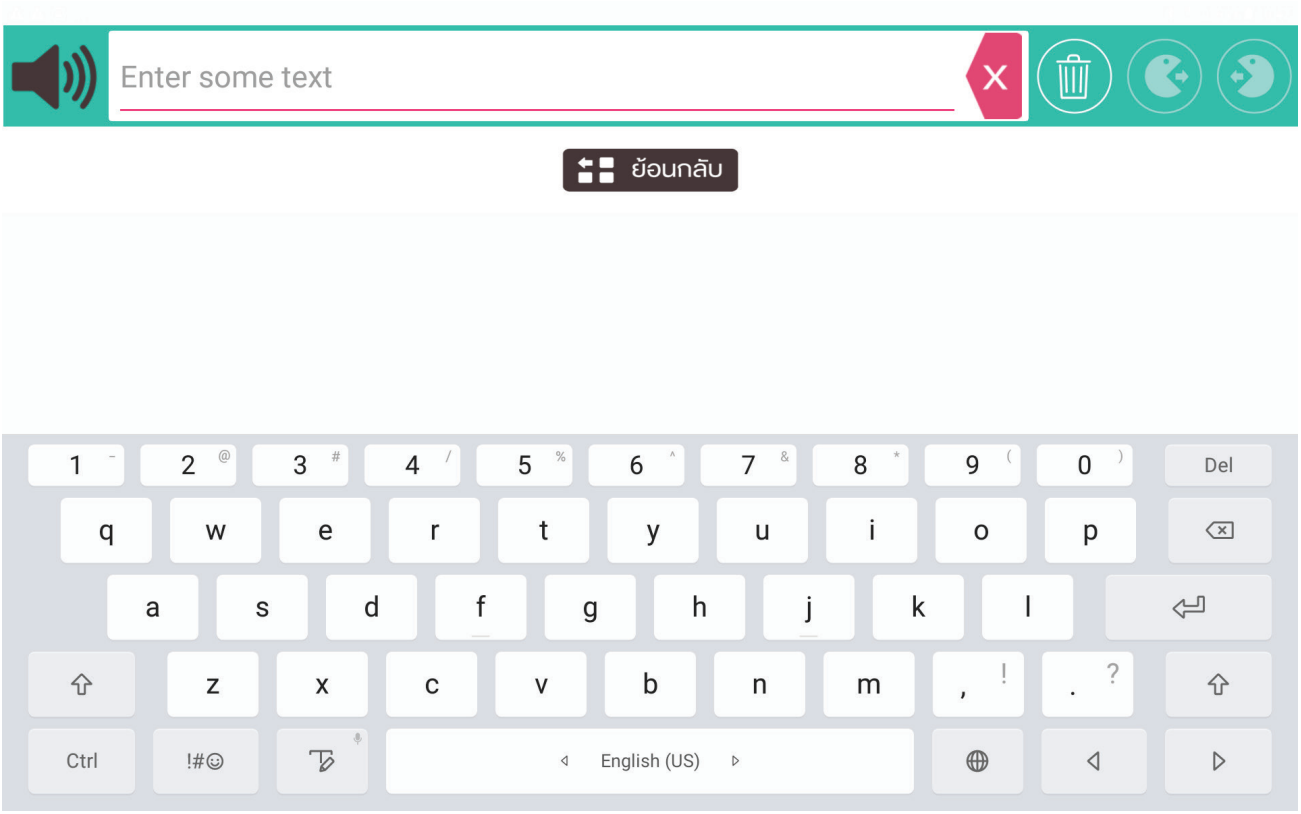


Figure 3. *Keyboard mode*

Feedback

Activation feedback

In graphic-symbol board mode, the click sound would be produced as auditory activation feedback since the icon was selected. Likewise, the click sound would be created in keyboard mode when the letter was tapped. This auditory activation feedback could be muted as needed.

Message feedback

Two types of message feedback were included in A-Speak (i.e., visual message feedback and auditory message feedback). The visual feedback messages would show the sequences of symbols as each symbol was selected in the screen display, as shown in Figure 4. In addition, the auditory message feedback (i.e., synthesized speech) would be produced word by word in the form of word echo. The user might select the voice gender of the synthetic speech output (i.e., male or female) based on his/her gender. The synthetic speech was recorded by the developer word by word. Consequently, there was no distortion of Thai voice tones in graphic-symbol board mode.

A-Speak operated with the external input devices (e.g., single switch, eye-tracking, etc.). In this study, a single switch was used in scanning mode with participants who lacked motor control and could not directly select the icons on the screen. The type of scanning pattern used in this study was group-item scanning with automatic scanning as a selection control technique. The cursor automatically ran from top to bottom rows (i.e., group selection) and left to right sides (i.e., icon selection). The first pressing the switch would choose the group, and then pressing the switch again would select the icon within that group. To create a phrase or sentence, the users repeated the selection process. For further information, including instruction video, about A-Speak, go to https://aspeak.kid-bright.org

Figure 4. *Visual feedback messages.*

Mechanical switches

Two switch models (Big Buddy and Specs) were used in this study. The Big Buddy Button from AbleNet (i.e., IT Outsourcing Company) had a 4.5-inch diameter and was wired to the apparatus (i.e., tablet). The Specs switch from Ablenet had a 1.375-inch diameter and was wired to the tablet. Both models provided activate feedback, including auditory click and tactile.

Tablet

For this project, A-Speak application was installed in Samsung Galaxy Tab A (2016) (SM-P585Y).

Participants

Out of all individuals in Pakkret Home for Children with Disabilities, only 15 individuals between 11 and 24 met the inclusion criteria and participated in this study. The inclusion criteria included: (a) Their primary language was Thai; (b) Their vision and hearing were normal; (c) They were considered to have complex communication needs; (d) They had never used hi-tech AAC; (e) Their receptive language development was equal to or above 4 years of age. Thai Speech and Language Assessment: Pediatric Standardized Test (0-4 years) was employed to evaluate participants 'receptive language development, and (f) They could remain in a sitting position.[28] The exclusion criteria included their speech and language development, which were within normal limits. Individuals who failed to attain the Phase 1 mastery level within two months would be withdrawn from the study. All participants had intellectual disabilities and cerebral palsy at different levels. Three of the fifteen participants had limited upper-body movement. Therefore, those three participants required training to use a scanning mode as an operating technique. The participants' demographic information is displayed in Table 1.

Table 1. The demographic data of the participants. (N=15)

Characteristics	N	%
Gender		
Male	5	33.33
Female	10	66.67
Age, Mean (SD)	17.4 (4.34)	
Type of cerebral palsy		
Spastic	13	86.67
Ataxic	2	13.33
Highest education level		
Grade 10	2	13.33
Grade 11	3	20
Grade 12	10	66.67

Training Procedures

The training procedures consisted of three main phases: Phase 1: Train to select icons; Phase 2: Shift to different categories; and Phase 3: Use A-Speak to communicate. Two researchers were responsible for training throughout all phases, while one engineer was in charge of setting up the switch with the tablets for participants. The researchers visited the Pakkret Home for Children with Disabilities once a month to provide the intervention for participants and train their caregivers to operate A-Speak to be able to assist participants (e.g., configuring the display, adding new icons, modeling participants to communicate through A-Speak, interacting with participants by using A-Speak, etc.). After the training session, the caregivers were assigned to encourage participants to use A-Speak to communicate with caregivers in their daily routines.

Phase 1: 12 individuals with precise finger control were taught to tab the tablet's screen with their fingers to select the desired icons. On the other hand, three participants who had trouble using their fingers effectively were instructed to utilize the scanning mode via the single switch. The single switch was used to activate A-Speak, and the participants were trained to have sufficient capacity (i.e., the switch was activated accurately when the cursor stopped at the target item). The type of scanning pattern used was group-item scanning with automatic scanning as a selection control technique. The cursor automatically ran from top to bottom rows (i.e., group selection) and left to right sides (i.e., icon selection). The first press of the switch would choose the group, and then pressing the switch again would select the icon within that group. To create a phrase or sentence, the users repeated the selection process. The participants were assessed motor skills by a physiotherapist, who was working in Pakkret Home for Children with Disabilities, to select individual discrete-motor control. They each pressed the switch with different parts of their, including the right temple, right shoulder blade (i.e., scapula), and left elbow. Two used the Big Buddy Button, and another used the Specs Switch. The participants were required to master the skill before proceeding to Phase 2. Those who utilized a switch needed two training sessions and spent between 30 and 40 minutes per session to be proficient at selecting icons. The participants who could use a finger to select the icons required one training session and spent between 15 and 20 minutes each.

Phase 2 involved training the participants to shift between different categories. The participants needed to switch between categories to generate a desired phrase or sentence. All participants were encouraged to answer questions or make requests in a word or sentence. Initially, the process was trained by navigating from the main page to a different category. Switching between multiple categories was taught afterward. The participants were supported in acquiring this capacity before moving to Phase 3. Like Phase 1, the participants acquire adequate proficiency in this skill before continuing to the next phase. The participants who used a switch required one training session and spent between 30 and 40 minutes per session to acquire adequate skills to switch between pages. Those who could use a finger to operate required one training session and spent between 10 and 15 minutes each.

Phase 3 was about training participants to operate A-Speak independently and motivating them to use A-Speak as a communication tool in their daily. To encourage individuals to use A-Speak for communication, the researchers verbally asked different kinds of questions (e.g., open-ended questions, yes/no questions, etc.) and developed circumstances for participants to communicate through various communication functions (e.g., requesting objects, telling stories, imitating conversation, denying, expressing feeling, etc.) while they were engaged in activities. The complexity of questions varies depending on participants' language proficiency, interests, and circumstances in a particular instant. As a result, the question lists that each participant received were unique. Participants who used switches and those who did not use a switch needed to attend three training sessions; each session lasted 60 minutes for participants and 60 minutes for caregivers. (a total of two hours per session).

Data collection and analysis

According to Phase 3, the researchers created the circumstances to encourage participants to communicate 12 communication functions, including requesting objects,

requesting actions, denying, answering yes/no questions, answering WH-questions, naming, asking questions, storytelling, explaining, telling emotions, greeting, and initiating conversation.

Two researchers with more than 18 years of experience as speech-language pathologists administered the data collection process. To reduce bias, the researchers rotated the participants throughout three sessions. In two sessions (sessions 1 and 3), each researcher was assigned to conduct an intervention with the same participants as the primary therapist, and in just one session (session 2), with a different participant. Both researchers individually collected the participants' performance in their record form during the intervention process without discussion. After every intervention session, the data in the researchers' record form were discussed to verify the agreement of each participant's outcomes.

The data was collected based on observation and recorded in the form of code (e.g., 0, 1, 2, and 3) according to communication performance in three sessions. This study used descriptive statistical analysis reporting frequency (e.g., number, and percentage). The significance of mean differences between the three intervention sessions was examined using the Friedman and Wilcoxon tests. The mean differences of 13 different communication functions were analyzed independently.

Results

The Friedman test indicated that the mean differences between the three sessions were significant for 10 out of 12 communication functions, including requesting objects (X^2=10.129, p=0.006), requesting actions (X^2=12.057, p=0.002), denying (X^2=14.205, p=0.001), answering yes/no questions (X^2=9.579, p=0.008), answering WH-questions, naming (X^2=18.865, p<0.001), asking (X^2=8.4, p=0.015), explaining (X^2=20.6, p < .001), telling emotions (X^2=12.067, p=0.002), and greeting (X^2= 13.231, p=0.001). The communication function of answering WH-questions was analyzed separately in two domains: (a) The questions did not require descriptive answers (i.e., what, where, when, and who questions; and (b) The questions required descriptive answers (i.e., why and how questions). Only WH-questions showed statistical significance, including what, where, when, and who questions (X^2=8.588, p=0.014). Some communication functions consisted of 14 participants since some did not wholly participate in three intervention sessions. Table 2 presents the overall data of statistical differences as determined by the Friedman test.

Table 2. The Friedman test analyzed mean differences between the three intervention sessions and statistical significance.

Measurement	Mean			N	Chi-Square	df	*p* value
	S1	S2	S3				
Requesting objects	1.60	1.97	2.43	15	10.129	2	0.006
Requesting actions	1.46	2.04	2.50	14	12.057	2	0.002
Denying	1.32	2.25	2.43	14	14.205	2	0.001
Answering yes/no questions	1.68	1.96	2.36	14	9.579	2	0.008
Answering what, where, when, and who questions	1.71	1.96	2.32	14	8.588	2	0.014
Answering why and how questions	0.64	2.04	2.32	14	5.871	2	0.053
Naming	1.43	1.90	2.67	15	18.865	2	0.000
Asking	1.57	2.07	2.37	15	8.400	2	0.015
Storytelling	1.87	1.97	2.17	15	2.800	2	0.247
Explaining	1.40	1.87	2.73	15	20.600	2	0.000
Telling emotions	1.63	1.87	2.5	15	12.067	2	0.002
Greeting	1.6	1.93	2.47	15	13.231	2	0.001
Initiating conversation	1.68	2.07	2.25	14	5.360	2	0.069

***Note**: S1: session 1, S2: session 2, S3: session 3, *p<0.05.*

Wilcoxon test was used as a post-hoc test. **Requesting objects**. The mean difference between Sessions 1 (M=1.6) and 3 (M=2.43) was significant, Z=2.549, p=0.011. **Requesting actions**. The mean differences demonstrated in both Session 1 (M=1.46) compared to Session 2 (M=2.04), Z=2.333, p=0.02, and Session 1 compared to Session 3 (M=2.5), Z=2.724, p=0.006, were significant. **Denying.** There was statistical significance in the mean differences between Sessions 1 (M=1.32) and 2 (M=2.25), as well as in Sessions 1 and 3 (M=2.43), the statistical significance illustrated, Z=2.887, p=0.004, and Z=2.804, p=0.005, respectively. **Answering yes/no questions**. The mean differences between Sessions 2 (M=1.96) and 3 (M=2.36), Z=2.251, p=0.024, as well as between Sessions 1 (M=1.68) and 3, Z=2.07, p=0.038, were statistically significant. **Answering what, where, when, and who questions.** The outcomes of Session 1 (M=1.71) compared to Session 3 (M=2.32) showed a statistical difference, Z=2.06, p=0.039. **Naming.** The means among all comparisons were significantly different: Session 1 (M=1.43) compared to Session 2 (M=1.9), Z=2.236, p=0.025; Session 1 compared to Session 3 (M=2.67), Z=3.134, p=0.002; and Session 2 compared to Session 3, Z=2.558, p=0.011. **Asking.** The significant differences were presented between

Sessions 1 (*M*=1.57) and 2 (*M*=2.07), *Z*=2.121, *p*=0.034, as well as between Sessions 1 and 3 (*M*=2.37), *Z*=2.565, *p*=0.01. **Explaining.** The means among all comparisons were significantly different: Session 1 (*M*=1.4) compared to Session 2 (*M*=1.87), *Z*=2.449, *p*=0.014; Session 1 compared to Session 3 (*M*=2.73), *Z*=3.095, *p*=0.002; and Session 2 compared to Session 3, *Z*=2.842, *p*=0.004. **Telling emotions**. There were significant differences between Sessions 1 (*M*=1.63) and 3 (*M*=1.87), *Z*=2.762, *p*=0.006, as well as Sessions 2 (*M*=2.5) and 3, *Z*=2.165, *p*=0.03. **Greeting.** The means among all comparisons were significantly different: Session 1 (*M*=1.6) compared to Session 2 (*M*=1.93), *Z*=2, *p*=0.046; Session 1 compared to Session 3 (*M*=2.47), *Z*=2.598, *p*=0.009; and Session 2 compared to Session 3, *Z*=2.333, *p*=0.02. The total data of statistical differences as determined by the Wilcoxon test is illustrated in Table 3.

The satisfaction of using A-Speak was not directly surveyed. Two participants used keyboard mode to express their gratitude. At the same time, A-Speak helped them explain the brief stories they had experienced and let their caregivers understand how they truly felt.

Table 3. Statistical significance of mean differences between the three comparisons of the three intervention sessions as analyzed by the Wilcoxon test.

Measurement	S1 compared to S2		S1 compared to S3		S2 compared to S3	
	z	*p* value	z	*p* value	z	*p* value
Requesting objects	1.890	0.059	2.549	0.011	1.681	0.093
Requesting actions	2.333	0.020	2.724	0.006	1.628	0.103
Denying	2.887	0.004	2.804	0.005	0.750	0.453
Answering yes/no questions	1.732	0.083	2.251	0.024	2.070	0.038
Answering what, where, when, and who questions	1.732	0.083	2.060	0.039	1.841	0.066
Naming	2.236	0.025	3.134	0.002	2.558	0.011
Asking	2.121	0.034	2.565	0.010	1.557	0.120
Explaining	2.449	0.014	3.095	0.002	2.842	0.004
Telling emotions	1.134	0.257	2.762	0.006	2.165	0.030
Greeting	2.000	0.046	2.598	0.009	2.333	0.020

***Note**: S1: session 1, S2: session 2, S3: session 3, *p<0.05.*

Discussion

Several AAC applications are attributed to facilitating the communication competencies of individuals with complex communication needs. Unfortunately, those applications have considerable limitations on speech-output features. Therefore, this study aims to develop AAC applications specifically for Thai individuals. The application was A-Speak or All-Speak. Two main communication options were available in A-Speak, including a graphic-symbol board and keyboard.

A-Speak was mainly similar to Funjai; Funjai is also a Thai AAC application that consists of colored-graphic symbols as a primary operating system with voice output. Both applications could add extra pictures saved on the device to serve as symbols. However, there were significant differences between these two AAC applications. Funjai could be installed on various operating systems, including Huawei, iOS, and Android; however, only Android was supported for A-Speak. There were additional differences in the voice output systems employed in these applications. Funjai's text-to-speech technique was utilized in its voice output, which was limited to female voices. This resulted in intonation distortion on transliterated words such as hamburger, French fried, and carrot. Another distinction was that Funjai did not offer keyboard mode as a backup communication option.

Both Symbotalk and A-Speak included colored-graphic symbols with voice output. They were also able to adjust the number of grids according to the user's level of proficiency. However, SymboTalk showed intonation distortion in Thai voice output. This could be the outcome of SymboTalk's use of voice synthesis through text-to-speech generation. On the other hand, the voice synthesis of A-Speak was created by word-by-word recording, which eliminated intonation distortion in Thai voice output. To add more graphic symbols to SymboTalk, the application automatically connects to graphic-symbol resources with an internet connection. Nevertheless, this feature was not available in A-Speak.

This limitation was similar to Leeloo. Graphic symbols were specifically created and attached to Leeloo, so the number of symbols was limited. However, if users wanted to add more symbols, they could download them by registering for the premium version. Likewise, graphic symbols were mainly developed and contributed to A-Speak. However, if users want to add more symbols, they can download them online or take pictures and set them to A-Speak.

Cboard and SymboTalk were comparable. These applications instantly linked to online symbol resources and allowed users to install those symbols on their boards. Nevertheless, the Thai voice output from these two applications was limited regarding intonation distortion and the lack of diversity in voice age and gender. A-Speak, on the other hand, offered a more incredible selection of voice choices for age and gender that matched the characters of users.

In addition to those features, A-Speak provided a list of folders organized vocabulary into different categories on the display's right side. This feature facilitated users to shift to another category quicker. A-Speak also offered literate individuals the option to use the keyboard mode. A-Speak's symbols have been deliberately designed to correspond with Thai culture. Nevertheless, A-Speak was currently compatible only with Android.

The results of A-Speak training showed that participants' communication efficacy had improved using graphic-symbol mode, even though there were few communication functions that the participants did not attain within the training period. The communication functions that the participants showed improvement during the A-Speak intervention included requesting objects, requesting actions, denying, answering yes/no questions, answering WH-questions, naming, asking, explaining, telling emotions, and greeting. However, the participants did not significantly attain communication functions of telling stories and initiating conversation within the three intervention sessions. The findings also showed that explanation abilities were the communication function that improved the most. On the other hand, storytelling abilities were the communication function that exhibited a minor development. Requesting activities, refusing, naming, asking, and greeting were likely the communication functions the participants found the simplest to learn, as the participants showed a noticeable improvement. The results agreed with the language development milestones in children, such as denying, which starts at 16 months of age, while responding to WH-questions, which begins at 2; 5 years of age.

Conversely, narrative storytelling begins to show at the age of four.[29,30] The explanation for these would be related to children's cognitive development, as they start to think intuitively at a later age.[31] However, this study measured communication efficiency in terms of communication functions only and recorded the outcomes in code numbers without collecting the details of language components (i.e., form of language, content of language, and use of language). Besides this limitation, since this project focused on graphic-symbol mode, participants' communication efficacy in keyboard mode was not investigated. Finally, the researchers had training sessions with the participants once a month, and there were only three training sessions, which were too short to explore the result.

Nevertheless, the main objective of this study was to develop the Thai AAC application and tentatively investigate the effectiveness of the application on individuals with complex communication needs. The future research would focus on further developing the A-Speak application to minimize its existing constraints, such as a limitation about inserting additional symbols, and evaluate the proficiency of the application with individuals with complex communication needs. The evaluation would extend the number of training sessions and use multiple baseline designs as a research design. Moreover, the satisfaction of A-Speak users and communication partners with the quality of voice output was also considered to be investigated in future research.

Conclusion

A-Speak is an AAC application developed especially for Thai individuals with complex communication needs. Although several AAC applications that included Thai-speech output were available, those applications had some drawbacks that possibly reduced the intelligibility and competence of communication. With the training, all participants improved their abilities to communicate in various communication functions while using A-Speak.

Acknowledgements

We are sincerely grateful to Nitha Ungsuprasert, an AAC specialist, for her guidance in developing the A-Speak application. We are also grateful that Pakkret Home for Children with Disabilities allowed us to collect the research data.

Funding

This research is funded by the Educational Promotion and Development Fund for Handicapped Group, Special Education Bureau, Ministry of Education.

Conflict of interest

The authors declare that there is no conflict of interest.

Ethics approval

The Ethical Committee of the Ramathibodi Hospital approved the participants in this study. Since all participants had disabilities and could not provide informed consent themselves, informed consent was obtained from their caregivers.

References

[1] American Speech-Language-Hearing Association [Internet]. Augmentative and Alternative Communication (AAC) [cited 2024 March 18]. Available from: https://www.asha.org/practice-portal/professional-issues/augmentative-and-alternative-communication/#collapse_1.

[2] Law J, Parkinson A, Tamhne R. Communication difficulties in childhood: A practical guide. 1st Ed. Oxon: CRC Press; 2000.

[3] Park CJ, Yelland GW, Taffe JR, Kylie MG. Brief report: The relationship between language skills, adaptive behavior, and emotional and behavior problems in pre-schoolers with autism. J Autism Dev Disord. 2012; 42: 2761-6. doi:10.1007/s10803-012-1534-8.

[4] Walker VL, Snell ME. Effects of augmentative and alternative communication on challenging behavior: a meta-analysis. Augment Altern Commun. 2013; 29(2): 117-31. doi:10.3109/07434618.2013.785020.

[5] Prizant BM, Wetherby AM, Rubin E, Laurent AC. The SCERTS model: A transactional, family-centered approach to enhancing communication and socioemotional abilities of children with autism spectrum disorder. Inf Young Child. 2003; 16(4): 296-316.

[6] Romski M, Sevcik RA, Barton-Hulsey A, Whitmore

AS. Early Intervention and AAC: What a Difference 30 Years Makes. Augment Altern Commun. 2015; 31(3): 181-202. doi:10.3109/07434618.2015.1064163.

[7] Sevcik RA, Barton-Hulsey A, Romski M. Early intervention, AAC, and transition to school for young children with significant spoken communication disorders and their families. Semin Speech Lang. 2008; 29(2): 92-100. doi:10.1055/s-2008-1079123.

[8] Charlop-Christy MH, Carpenter M, Le L, LeBlanc LA, Kellet K. Using the picture exchange communication system (PECS) with children with autism: Assessment of PECS acquisition, speech, social-communicative behavior, and problem behavior. J Appl Behav Anal. 2002; 35(3): 213-31. doi:10.1901/jaba.2002.35-213.

[9] Leech ER, Cress CJ. Indirect facilitation of speech in a late talking child by prompted production of picture symbols or signs. Augment Altern Commun. 2011; 27(1): 40-52. doi: org/10.3109/07434618.2010.550062.

[10] Stahmer A, Ingersoll B. Inclusive programming for toddlers with autism spectrum disorders: Outcomes from the children's toddler school. J Posit Behav Interv. 2004; 6: 67-82. doi:10.1177/109830070400600202.

[11] Branson D, Demchak M. The use of augmentative and alternative communication methods with infants and toddlers with disabilities: a research review. Augment Altern Commun. 2009; 25(4): 274-86. doi: 10.3109/07434610903384529.

[12] Ganz JB, Earles-Vollrath TL, Heath AK, Parker RI, Rispoli MJ, Duran JB. A meta-analysis of single case research studies on aided augmentative and alternative communication systems with individuals with autism spectrum disorders. J Autism Dev Disord. 2012; 42(1): 60-74. doi:10.1007/s10803-011-1212-2

[13] Caron J, Light J, Davidoff BE, Drager KDR. Comparison of the effects of mobile technology AAC apps on programming visual scene displays. Augment Altern Commun. 2017; 33(4):2 39-48. doi:10.1080/07434618.2017.1388836.

[14] Mirenda P. Toward Functional Augmentative and alternative communication for students with autism manual signs, graphic symbols, and voice output communication aids. Lang, Speech, and Hear Serv School. 2003; 34: 203-16. doi:10.1044/0161-1461(2003/017).

[15] Drager KD, Light JC, Speltz JC, Fallon KA, Jeffries LZ. The performance of typically developing 2 1/2-yearolds on dynamic display AAC technologies with different system layouts and language organizations. J Speech Lang Hear Res. 2003; 46(2): 298-312. doi:10.1044/1092-4388(2003/024)

[16] Boesch MC, Wendt O, Subramanian A, Hsu N. Comparative efficacy of the picture exchange communication system (PECS) versus a speech-generating device: effects on social-communicative skills and speech development. Augment Altern Commun. 2013; 29(3): 197-209. doi: 10.3109/07434618.2013.818059.

[17] Choi H, O'Reilly M, Sigafoos J, Lancioni G. Teaching requesting and rejecting sequences to four children with developmental disabilities using augmentative and alternative communication. Res Dev Disabil. 2010; 31(2): 560-7. doi: 10.1016/j.ridd.2009.12.006.

[18] Schäfer MCM, Sutherland D, McLay L, Achmadi D, van der Meer L, Sigafoos J, et al. Research note: attitudes of teachers and undergraduate students regarding three augmentative and alternative communication modalities. Augment Altern Commun. 2016; 32(4): 312-9. doi:10.1080/07434618.2016.1244561.

[19] van der Meer L, Rispoli M. Communication interventions involving speech-generating devices for children with autism: a review of the literature. Dev Neurorehabil. 2010; 13(4): 294-306. doi: 10.3109/17518421003671494.

[20] Kamonsitichai W, Goldstein H. Speech-language pathologists' perceptions of augmentative and alternative communication in Thailand. Augment Altern Commun. 2023; 39(4): 230-40. doi:10.1080/07434618.2023.2208222.

[21] Wickenden M. Whose voice is that?: Issues of identity, voice and representation arising in an ethnographic study of the lives of disabled teenagers who use augmentative and alternative communication (AAC). Disab Stud Quart. 2011; 31(4). doi:10.18061/dsq.v31i4.1724.

[22] Pullin G, Hennig S. 17 Ways to say yes: toward nuanced tone of voice in AAC and speech technology. Augment Altern Commun. 2015; 31(2): 170-80. doi: 10.3109/07434618.2015.1037930.

[23] Burkhart JL. Total augmentative communication in the early childhood classroom. Illinois, Don Johnston, Inc., 1993.

[24] de Groot JY. Thai for Beginners. Phuket: Prince of Songkhla University; 2010.

[25] Dodd JL. Augmentative and alternative communication intervention: an intensive, immersive, socially based service delivery model. 1st Ed. San Diego: Plural Publishing; 2017.

[26] Beukelman D, Light J. Augmentative and alternative communication: Supporting children and adults with complex communication needs. 5th Ed. Pennsylvania: Paul H. Brookes Publishing; 2020.

[27] Beukelman DR, Fager S, Eds. Selecting visual scene displays: Personal relevance for age and gender. The State of the Science Conference of the Rehabilitation Engineering Research Center on Augmentative and Alternative Communication; 2018 July 11; Arlington, USA.

[28] Angsupakorn N. Comparison of the reliability of parental reporting and the direct test of the Thai speech and language test. J Med Assoc Thai. 2012; 95(11): S67-S72. PMID: 23961623.

[29] Chapman RS. Children's language learning: an interactionist perspective. J Child Psychol Psychiatry. 2000; 41(1): 33-54. PMID: 10763675.

[30] Shipley KG, McAfee JG. Assessment in speech-language pathology: A resource manual. 6th Ed. San Diego: Plural Publishing; 2023.

[31] Piaget J. Part I: Cognitive development in children: Piaget development and learning. J Res Sci Teach. 1964; 2(3): 176-86. doi:10.1002/tea.3660020306.

Appendix A

Part 1: Demographic information

Participant code_____ Age_____ Gender_____
Education level_____ Type of cerebral palsy_____
Caregiver name_____

Part 2: Communication functions

No.	Communication functions	1st session	2nd session	3rd session
1	Requesting objects			
2	Requesting actions			
3	Denying			
4	Answering yes/no questions			
5	Answering WH-questions			
6	Naming			
7	Asking			
8	Storytelling			
9	Explaining			
10	Telling emotions			
11	Greeting			
12	Initiating conversation			

Scoring rubric

0 = No communication
1 = One occasion
2 = More than one occasion
3 = Communication on all occasions

Journal of Associated Medical Sciences 2024; 57(3): 204-211

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

Educational media utilization for rehabilitation among community-dwelling stroke survivors and their caregivers: a pilot study

Issaree Kongsri[1] Navin Chuaynoo[2] Pisak Chinchai[3] Pornpen Sirisatayawong[3] Sopida Apichai[3] Waranya Chingchit[3*]

[1]BRYC BLOSSOM KIDS, Udon Thani Province, Thailand.
[2]Meaning-kids clinic, Kanchanaburi Province, Thailand.
[3]Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 29 May 2024
Accepted as revised 1 August 2024
Available online 5 August 2024

Keywords:
Stroke survivors, caregivers, community, educational media, active information.

ABSTRACT

Background: Recently, there has been an increasing number of stroke patients. Those who survive still need continuous rehabilitation after being discharged from the hospital. The information and knowledge on stroke rehabilitation at home is crucial for these patients and their caregivers. However, there has never been a study of educational media on rehabilitation for those patients and their caregivers in Thailand.

Objective: To explore the educational media used in stroke rehabilitation, list their benefits, and the most satisfying media type for stroke participants and their caregivers after hospital discharge.

Materials and methods: Twenty-eight stroke patients and twenty-one primary caregivers from two Subdistrict Municipalities and two Districts in Chaing Mai province were recruited using purposive sampling methods. All participants were aged 18 years and over and could communicate in Thai. All the stroke participants had no cognitive impairment screening by the Mental Status Examination-10 (MSET-10) and had self-rehabilitation therapy at home. A Questionnaire on the use of educational media in stroke rehabilitation with the index of item objective congruence (IOC) range between 0.67-1.00 was used for data collection. The statistics used were descriptive.

Results: Stroke patients used information obtained primarily from medical personnel in the community (78.57%), while caregivers used advice from medical personnel the most in the hospitals (57.14%). Regarding the highest satisfaction with educational media, 59% of stroke patients and 42% of caregivers were satisfied with advice from medical personnel in the community. Stroke patients revealed that using a combination of advice from medical professionals, brochure-based media, and only document media was the most beneficial in providing knowledge about rehabilitation. For caregivers, the use of a combination of advice from medical professionals along with paper and video media, and the use of knowledge gained from online channels were the most beneficial.

Conclusion: Rehabilitation education materials used by stroke patients and their caregivers at home offer the highest level of satisfaction for the patients and caregivers if the medium of people who can give advice and communicate on issues meets their needs. In addition, patients and caregivers agreed that rehabilitation education should include more than one form of media for maximum benefit. Therefore, health professionals should consider the format of educational media appropriate to the needs of service recipients to be able to use them most effectively.

∗ **Corresponding contributor.**
Author's Address: *Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.*
E-mail address: waranya.chingchit@cmu.ac.th
doi: 10.12982/JAMS.2024.062
E-ISSN: 2539-6056

Introduction

In Thailand, the rate of stroke in the population aged 15 years and over was 278.49 per 100,000 people in 2017 and increased to 330.72 per 100,000 people in 2021.[1]

Stroke is the second leading cause of death among chronic non-communicable diseases, and those who survive are left with physical disabilities that cause difficulty for them. Many have to be cared for by immediate family members or relatives.[2, 3] A previous study has shown that only 18% of Thai stroke patients receive rehabilitation while in hospitals.[4] This means that there are still many stroke patients living at home who do not receive direct care and rehabilitation from medical personnel.

Additionally, these people do not receive knowledge and rehabilitation skills from rehabilitation institutions. Stroke patients and their family members should receive information and knowledge about many aspects of stroke. For example, issues about the causes of stroke, symptoms, risk factors, prevention of complications, consequences after a stroke, treatment, and rehabilitation when returning home should be better understood by the patients and their caregivers.[5] Therefore, health professionals should provide rehabilitation education to stroke patients and their caregivers before discharging them from hospitals.

The clinical practice guideline for stroke rehabilitation specifies that health professionals should provide information and knowledge about stroke, its complications, processes, and goals for rehabilitation, both through interactive conversation and in a written format.[6] This is consistent with a study on the rehabilitation information needs of stroke patients and their caregivers after being discharged from the hospital, which found that these individuals wanted to receive knowledge about stroke and care through mixed media methods that included both verbal communication such as interactive or face-to-face conversation and in written formats such as documents or pamphlets that can be stored and reread when needed.[7, 8]

Thailand still encounters problems in treating stroke patients when they must return to their homes since the caregivers do not receive knowledge and training in basic rehabilitation skills for practice with the patients. Some stroke survivors and their caregivers may receive some advice from therapists, but it is not sufficient for self-practice. Another issue is that the person caring for the patient in the hospital is not the primary caregiver. Therefore, when the patient returns home, the primary caregiver lacks knowledge and confidence in providing rehabilitation to the patient and cannot perform therapy correctly.[9]

Rehabilitation professionals, which include occupational therapists (OTs), play essential roles in encouraging physical and mental conditions for stroke patients. These therapists also promote the patient's ability to perform daily activities by directly providing knowledge and basic rehabilitation skill training to the patients and their caregivers.[10-12] Health professionals should use appropriate media to provide knowledge for self-rehabilitation at home that meets the needs of stroke patients and caregivers. Notably, each form of media has different advantages and limitations.

However, to our knowledge, no studies have been conducted on the use of educational media for rehabilitation in stroke patients living at home in Thailand, including determining the level of satisfaction with the media received.

Therefore, the researchers conducted a pilot study on media used for rehabilitation, examined the benefits of using the media, and determined the level of satisfaction with current media used in stroke rehabilitation from the perspective of stroke patients and their caregivers.

Materials and methods

This research is a pilot study of the use of educational media regarding the rehabilitation of people with stroke in the community. Forty-nine subjects participated in the study, consisting of 28 stroke patients and 21 caregivers who lived in Nong Pa Khrang Subdistrict Municipality, Mueang District, and Ban Waen Subdistrict Municipality, Hang Dong District, Chiang Mai Province. The inclusion criteria were as follows:

Inclusion criteria for stroke patients

1. be a stroke patient who has self-rehabilitation therapy at home.
2. aged 18 years and over.
3. do not have impaired cognitive function (using the Thai version of the Mental State Examination 10 or MSET10).
4. able to communicate in Thai

Inclusion criteria for caregivers of stroke patients

1. be the primary caregiver for a stroke patient at home.
2. aged 18 years and over.
3. able to communicate in Thai

Instruments

1. MSET10 is used to screen stroke patients for impaired thinking and understanding. It has good psychometric properties, with a sensitivity and specificity of 100.0 and 98.4-99.4, respectively.[13]
2. A Questionnaire on the use of educational media in stroke rehabilitation. It is an instrument developed by the research team by asking the patient to choose the type of educational media they use after being discharged from the hospital. The questionnaire consists of three parts, including 1) the type of media used, 2) the benefits of the media, and 3) the most satisfying media type to utilize for rehabilitation. Regarding the benefits of using media, each question has a minimum score of 1, meaning it is least helpful, and the highest score of 5, meaning it is most beneficial. The final section is an open-ended question about the highest satisfaction with an educational media type of educational media used for stroke rehabilitation. This questionnaire has been checked for content validity by 3 experts. Content validity was determined by using the index of item objective congruence (IOC) technique. The analysis found that every question in the questionnaire had an acceptable score, ranging from 0.67 to 1.00.

Procedure

1. After the research project received approval to proceed from the Human Research Ethics Committee, Faculty of Associated Medical Sciences, Chiang Mai University, the principal researcher coordinated

with officials of Nong Pa Khrang Subdistrict Municipality, Mueang District, and Ban Waen Subdistrict Municipality, Hang Dong District, Chiang Mai Province, to request permission to conduct research in the area.

2. After receiving permission, the researcher publicized the project by placing posters on the public relations boards in the target areas. The posters specified details of the research project, including methods for contacting researchers, so those interested can apply to participate in the study.
3. The main researcher performed preliminary cognition screening of stroke patients using MSET10 before data collection to prevent errors in the study's results. This is because most stroke patients are elderly and often experience impaired cognitive function.

Data collection

The study was conducted between November and December 2022 using face-to-face interviews. Before collecting data, subjects received information about the research project and all signed consent forms before participating.

Data analysis

This study analyzed data using descriptive statistics, reporting frequency, percentage, mean, and standard deviation (SD).

Results

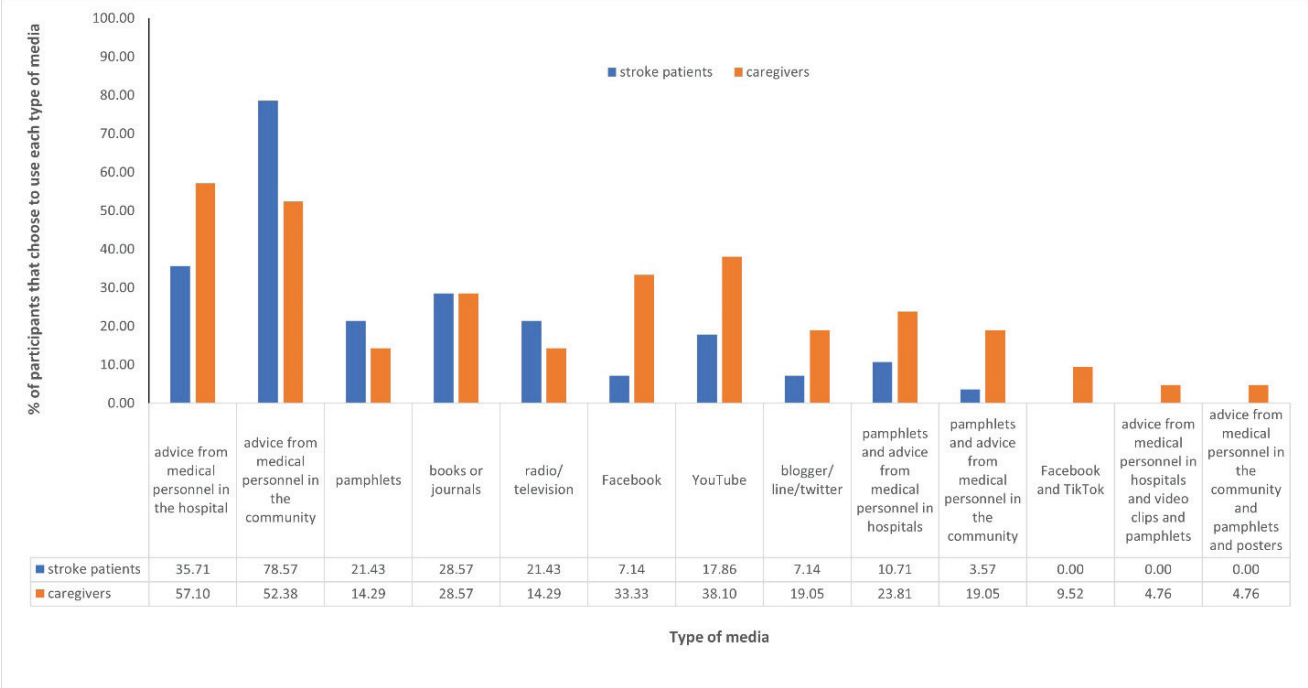
Data from Table 1 demonstrated that most stroke patients have had symptoms for more than 24 months. It determines whether they are unemployed, have insufficient income, have finished primary school education, or are not ready for communication equipment and internet connection.

Table 1. Sociodemographic data of stroke participants and caregivers.

Variables	Stroke participants (N=28) Frequency (percentage)	Caregivers (N=21) Frequency (percentage)
Age (years)		
Maximum	86	71
Minimum	40	18
Mean (SD)	65.21 (11.96)	46.95 (12.56)
Times since stroke onset/Times since care for stroke survivors (months)		
1-12	6 (21.43)	4 (19.05)
13-24	3 (10.71)	3 (14.28)
>24	19 (67.86)	14 (66.67)
Sex		
Male	18 (64.29)	5 (23.81)
Female	10 (35.71)	16 (76.19)
Occupation		
Paid employment	9 (32.15)	19 (90.48)
No employment	19 (67.85)	1 (4.76)
Students	0 (0.00)	1 (4.76)
Sufficiency of incomes		
Sufficient	8 (28.57)	13 (61.90)
Insufficient	20 (71.43)	8 (38.10)
Education levels		
No education	1 (3.57)	1 (4.76)
Elementary	16 (57.14)	3 (14.29)
Secondary	5 (17.86)	7 (33.33)
College/University	6 (21.43)	10 (47.62)
Own devices that can connect to the internet (smartphone/iPad)		
Yes	11 (39.29)	21 (100.00)
No	17 (60.71)	0 (0.00)
Availability of the home internet (monthly WIFI/internet)		
Available	9 (32.14)	19 (90.48)
Not available	19 (67.86)	2 (9.52)
Relationship with stroke survivors under care		
Son/daughter	-	14 (66.67)
Spouses	-	3 (14.29)
Mother/Father	-	1 (4.75)
Relatives	-	3 (14.29)
Responsibility of care		
Primary caregiver	-	9 (42.86)
Co-caregiver	-	12 (57.14)

Most stroke caregivers were female and had provided care to patients for more than 24 months. Most were the patients' sons or daughters who could pursue a career and earn enough income. Most caregivers graduated from college and universities and have their communication equipment with an internet connection.

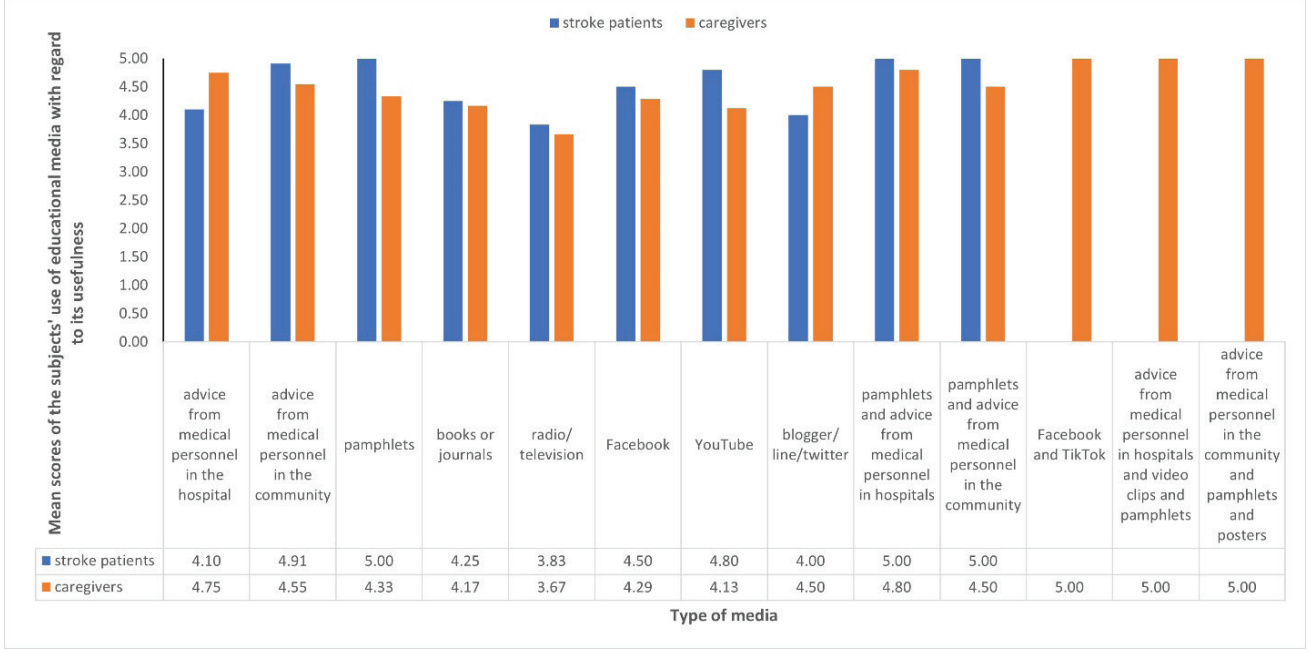
Data from Figure 1 shows that stroke patients choose to use educational media through verbal communication from medical personnel in the community the most (78.57%), followed by information from medical personnel in hospitals (35.71%) and media such as books or journals (28.51%). Caregivers chose to use media from medical personnel in hospitals the most (57.14%), followed by media from medical personnel in the community (52.38%) and YouTube internet media (38.10%).



	advice from medical personnel in the hospital	advice from medical personnel in the community	pamphlets	books or journals	radio/ television	Facebook	YouTube	blogger/ line/twitter	pamphlets and advice from medical personnel in hospitals	pamphlets and advice from medical personnel in the community	Facebook and TikTok	advice from medical personnel in hospitals and video clips and pamphlets	advice from medical personnel in the community and pamphlets and posters
stroke patients	35.71	78.57	21.43	28.57	21.43	7.14	17.86	7.14	10.71	3.57	0.00	0.00	0.00
caregivers	57.10	52.38	14.29	28.57	14.29	33.33	38.10	19.05	23.81	19.05	9.52	4.76	4.76

Figure 1. *Types of educational media and the number of samples that choose to use each type of media.*

The usefulness of educational media used by the participants in each category was calculated as an average score (highest score of 5 and lowest score of 1), as shown in Figure 2.



	advice from medical personnel in the hospital	advice from medical personnel in the community	pamphlets	books or journals	radio/ television	Facebook	YouTube	blogger/ line/twitter	pamphlets and advice from medical personnel in hospitals	pamphlets and advice from medical personnel in the community	Facebook and TikTok	advice from medical personnel in hospitals and video clips and pamphlets	advice from medical personnel in the community and pamphlets and posters
stroke patients	4.10	4.91	5.00	4.25	3.83	4.50	4.80	4.00	5.00	5.00			
caregivers	4.75	4.55	4.33	4.17	3.67	4.29	4.13	4.50	4.80	4.50	5.00	5.00	5.00

Figure 2. *Mean scores of the subjects' use of educational media concerning its usefulness.*

Figure 2 demonstrates that from the patient's point of view, the use of pamphlets, receiving advice from medical personnel in the hospital using pamphlets, and receiving advice from medical personnel in the community along with using pamphlets were the most useful, with an average score of 5 points. The second was receiving advice from medical personnel in the community (score of 4.91), followed by YouTube (score of 4.80). The caregiver's view was that using Facebook in combination with TikTok, receiving advice from medical personnel in the hospital in combination with using pamphlets and videos, and receiving advice from medical personnel in the community in combination with using pamphlets and educational posters were the most valuable (score 5), followed by receiving advice from medical personnel in the hospital along with using pamphlets (score 4.80).

Figure 3 displays the percentages of educational media forms that the patients were most satisfied with using during their recovery. Figure 4 displays the percentage of media types caregivers were pleased with employing in the patient's rehabilitation.

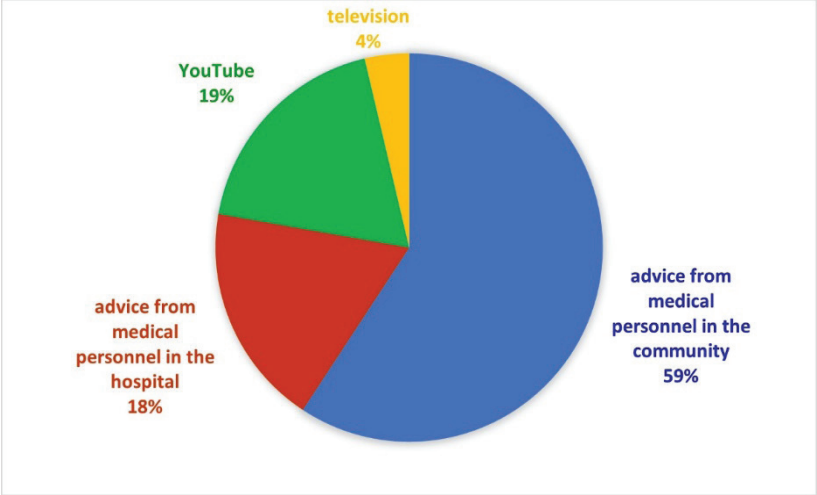


Figure 3. *The percentage of media types that patients feel most satisfying media type to utilize for rehabilitation.*

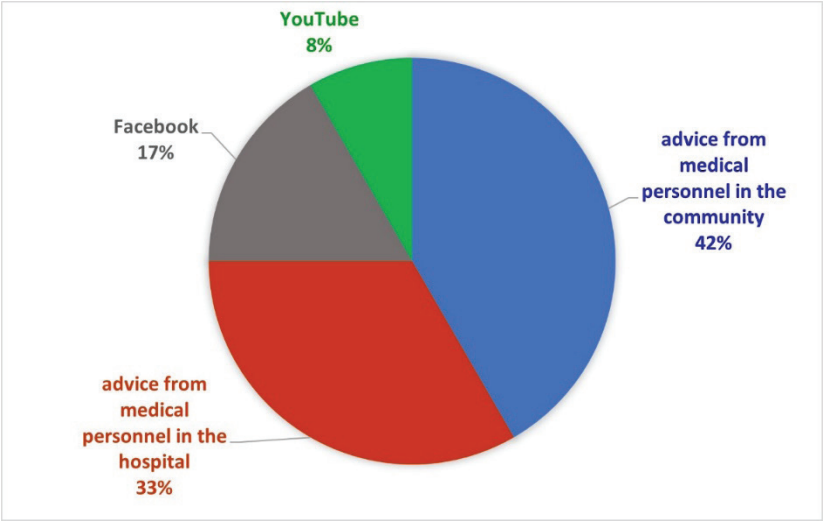


Figure 4. *The percentage of media types that caregivers feel most satisfying media type to utilize for rehabilitation.*

According to data shown in Figure 3, well over half of the patients liked the media that offered advice from community health experts and hospital medical staff.

According to the data shown in Figure 4, most caregivers expressed the most satisfaction with media featuring recommendations from community-based medical professionals, followed by recommendations from hospital-based medical professionals.

Discussion

The present study explores the types and benefits of educational media used in rehabilitation, including the degree of satisfaction with using each type of media among stroke patients living at home and their family caregivers in Chiang Mai Province, Thailand.

The study's results demonstrated that stroke patients and their caregivers used educational media in verbal, active

information communication the most (Figure 1). Patients chose to follow advice from medical personnel in the community the most, while caregivers chose to use advice from medical personnel in the hospital the most.

This may be because many stroke patients received services at rehabilitation centers in their communities. Therefore, they received advice from medical personnel working in that rehabilitation center, which includes occupational therapists, physical therapists, village health volunteers, etc. In contrast, the caregivers did not receive services at a rehabilitation center in the community because community rehabilitation centers have vehicles to pick up persons with disabilities from their homes directly to the rehabilitation center, yet do not provide transportation for the caregivers. Therefore, most caregivers did not come to receive counseling from medical personnel in their community rehabilitation centers. However, caregivers often receive advice on health and rehabilitation care only from medical personnel in hospitals such as the subdistrict health promotion hospitals, the district community hospitals, and the provincial general hospitals. This is because caregivers can meet these people when accompanying stroke patients under care to see doctors, nurses, physical therapists, occupational therapists, etc.

This study's results also showed that stroke patients and their caregivers are most satisfied with the type of educational media and two-way verbal communication between themselves and medical personnel, as shown in Figures 3 and 4. According to the study's findings, most stroke patients and their caregivers choose to use verbal communication media based on their level of satisfaction with it.

These findings are consistent with the studies of Eames *et al.* and Eames *et al*. on providing educational information for stroke patients in community health services in Australia, which found that most rehabilitation knowledge was provided through interactive verbal communication and verbal communication combined with paper-written forms.[14,15] The type of education stroke patients and their caregivers were most satisfied with was a combination of clinical information and practical management strategies. As for receiving knowledge about healthy lifestyles after hospital treatment, patients are most confident with face-to-face communication with medical personnel. Another form of health education that stroke patients were satisfied with, rated second after the strategies mentioned above, was using passive information in written format. At the same time, caregivers were confident using active information media such as a telephone conversation.

Verbal communication, in which patients and caregivers can talk and interact directly with health professionals, helps promote personal participation in receiving information better than receiving knowledge communicated one way (passive information), such as information in pamphlets.[16] Furthermore, stroke patients have distinct symptoms, and both patients and caregivers have different needs.[8] Therefore, it is the type of media that people can communicate verbally and interact with each other, which can better respond to the specific needs of each individual. As a result, it is the media that patients and caregivers prefer and are most satisfied with.

Previous studies demonstrated the benefits of rehabilitation education programs recommended by occupational therapists or trained health village volunteers for improving ADL performances in community-dwelling stroke survivors.[17,18] Although the most common media used by the patients and caregivers were advice from community health professionals and advice from hospital health professionals, stroke patients and their caregivers agreed that mixed media provides the most significant benefits for rehabilitation. The patients with cerebrovascular disease noted that two-way verbal communication and written media (pamphlets) give the most essential benefits for rehabilitation. The caregivers maintained that using personal communication media and media in written letters (pamphlets or posters) and a combination of online media (Facebook, TikTok) are the most beneficial, as shown in Figure 2.

This is consistent with the study results regarding satisfaction with media used in rehabilitation shown in Figures 3 and 4, where both stroke patients and their caregivers choose to use personal communication media the most. In addition, some stroke patients expressed their opinions that media in written form (pamphlets) was most helpful, and some caregivers expressed the idea that using online media through two applications (Facebook and TikTok) was the most useful.

This opinion difference may be due to patients' and their caregivers' age and educational level. Most stroke patients are elderly, while most caregivers are middle-aged and are still working. Furthermore, most stroke patients need more available equipment and an internet connection. In contrast, almost all caregivers own online equipment using a smartphone and computer and have an internet connection, as shown in Table 1. This allows caregivers to use online media, and they are more satisfied with it than stroke patients, who are primarily elderly and are less tech-savvy.

This is in line with the study of Teuschl and Brainin, which found that sociodemographic factors, including being female, middle-aged, and possessing a high level of education, correlate with more excellent knowledge of stroke. [19] However, no study has concluded which media or methods are most effective in educating people about knowledge of stroke, care, and methods of rehabilitation after stroke. Practical knowledge should be culturally adapted and transmitted appropriately according to the social context.

This study's findings are consistent with an earlier study done in Thailand on providing knowledge about stroke and care after leaving the hospital. This study found that caregivers wanted to receive information about the symptoms and methods of rehabilitation; the information obtained must be explained verbally, and relevant documents should be kept for further usage.

Limitation

This study had some limitations, including 1) the

small sample size of participants. Therefore, this study's results may be used as evidence for further studies but may need to be more generalizable with a broad sample of the exact nature. 2) most stroke patients have had symptoms for more than 24 months, which is considered a group with chronic symptoms; therefore, the use of rehabilitation education methods may be different from those who have recently experienced stroke. Recruitment that ensures greater homogeneity of the samples' characteristics should be considered in future studies. In addition, further research should investigate the use of educational media for rehabilitation in patients with similar sociodemographic characteristics and use the data to produce specific conclusions for each group. 3) The types and formats of media used by patients and caregivers for rehabilitation in this study do not cover all types of educational media. This depends on the media source; for example, some patients or caregivers may not have received advice from healthcare professionals in hospitals or community rehabilitation centers or have not received knowledge pamphlets. Therefore, the study's results regarding the benefits and satisfaction with these media types are missing.

Conclusion

A study of the use of educational media for rehabilitation in 28 stroke patients and 21 of their caregivers living at home revealed that verbal communication offers the most excellent satisfaction to stroke patients and caregivers. Results also showed that the educational media that provided the most benefit from the perspective of stroke patients and their caregivers was a combination of advice from medical personnel and media in the form of a written format. The results of the present study indicated that providing education to stroke patients and caregivers in the matter of rehabilitation while at home should use a combination of educational media, namely giving advice or communicating verbally together with providing knowledge in the form of documents such as pamphlets, where people can read them at any time after that. Additionally, the patient's physical symptoms and needs, including their sociodemographic characteristics, should be considered when providing appropriate educational media to these individuals.

Conflict of interest

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

Ethical approval

Ethical approval was obtained from the Human Research Ethics Committee, Faculty of Associated Medical Sciences, Chiang Mai University, Thailand. Project number AMSEC-65EX-087, and ethics clearance number 579/2565.

Funding

This study did not receive funding.

Acknowledgements

Researchers would like to thank all participants for their valuable time in responding to the questionnaire and thank all Nong Pa Khrang Subdistrict Municipality staff and Ban Waen Subdistrict Municipality for their cooperation.

References

[1] Tiamkao S. Incidence of stroke in Thailand. Thai J Neurol 2022; 39: 39-46.
[2] Murphy SJ, Werring DJ. Stroke: causes and clinical features. Medicine (Abingdon). 2020; 48(9): 561-6. doi:10.1016/j.mpmed.2020.06.002.
[3] Suwanwela NC. Stroke epidemiology in Thailand. J Stroke. 2014; 16(1): 1-7. doi:10.5853/jos.2014.16.1.1.
[4] Suksathien R. Accessibility to Medical Rehabilitation Service for Acute stroke at Maharat Nakhon Ratchasima Hospital: Related Factors and Outcomes. J Thai Rehabil Med. 2014; 24(2): 37-43.
[5] Pindus DM, Mullis R, Lim L, Wellwood I, Rundell AV, Abd Aziz NA, *et al*. Stroke survivors' and informal caregivers' experiences of primary care and community healthcare services - A systematic review and meta-ethnography. PLoS One. 2018; 13(2): e0192533. doi:10.1371/journal.pone.0192533.
[6] Duncan PW, Zorowitz R, Bates B, Choi JY, Glasberg JJ, Graham GD, *et al*. Management of Adult Stroke Rehabilitation Care: a clinical practice guideline. Stroke. 2005; 36(9): e100-43. doi:10.1161/01.STR.0000180861.54180.FF.
[7] Changmai S. Supportive care needs of family caregiver. CUTJ. 2016; 22(3): 424-35.
[8] Finch E, Minchell E, Cameron A, Jaques K, Lethlean J, Shah D, *et al*. What do stroke survivors want in stroke education and information provision in Australia? Health Soc Care Community. 2022; 30(6): e4864-e72. doi:10.1111/hsc.13896.
[9] Pitthayapong S. Situations, problems, and barriers of post-stroke care in the transitional period from hospital to home. TRC Nurs J. 2018; 11: 26-39.
[10] Gillen G, Nilsen DM, Attridge J, Banakos E, Morgan M, Winterbottom L, *et al*. Effectiveness of interventions to improve occupational performance of people with cognitive impairments after stroke: an evidence-based review. Am J Occup Ther. 2015; 69(1): 6901180040p1-9. doi:10.5014/ajot.2015.012138.
[11] 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.
[12] Jolliffe L, Lannin NA, Cadilhac DA, Hoffmann T. Systematic review of clinical practice guidelines to identify recommendations for rehabilitation after stroke and other acquired brain injuries. BMJ Open. 2018; 8(2): e018791. doi:10.1136/bmjopen-2017-018791.
[13] Boongird P. Mental State Examination T10. Dementia Association of Thailand Newsletter. 2018; (10).
[14] Eames S, Hoffmann T, Worrall L, Read S. Stroke patients' and carers' perception of barriers to accessing stroke

information. Top Stroke Rehabil. 2010; 17(2): 69-78. doi:10.1310/tsr1702-69.

[15] Eames S, Hoffmann T, Worrall L, Read S. Delivery styles and formats for different stroke information topics: patient and carer preferences. Patient Educ Couns. 2011; 84(2): e18-23. doi:10.1016/j.pec.2010.07.007.

[16] Crocker TF, Brown L, Lam N, Wray F, Knapp P, Forster A. Information provision for stroke survivors and their carers. Cochrane Database Syst Rev. 2021; 11(11): CD001919. doi:10.1002/14651858.CD001919.pub4.

[17] Pakdee P, Chinchai P. The influence of home visit program on functional abilities and quality of life in persons with disabilities resulting from stroke. Bull Chiang Mai Assoc Med Sci. 2016; 49(2): 276-85. doi:10.14456/jams.2016.24.

[18] Chinchai P, Jindakham N, Punyanon T. Effect of rehabilitation education to village volunteers toward activities of daily living performance of people with disabilities resulting from stroke. Bull Chiang Mai Assoc Med Sci. 2015; 48(3): 241-50. doi:10.14456/jams.2015.15.

[19] Teuschi Y BM. Stroke Education: Discrepancies among Factors Influencing Prehospital Delay and Stroke Knowledge. Int J Stroke. 2010; 5(3): 187-208. doi:10.1111/j.1747-4949.2010.00428.x.

Journal of Associated Medical Sciences 2024; 57(3): 212-220

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

A systematic review of melodic intonation therapy used by speech therapists on speech recovery for patients with non-fluent aphasia

Natwipa Wanicharoen[1] Vich Boonrod[2*] Thanasak Kalaysak[1] Sirapit Samueanjai[3]

[1]Communication Sciences and Disorders Division, Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai Province, Thailand.
[2]Department of Music, Faculty of Humanities, Naresuan University, Phitsanulok Province, Thailand.
[3]Maeai Hospital, Chiang Mai Province, Thailand.

ARTICLE INFO

Article history:
Received 9 May 2024
Accepted as revised 14 August 2024
Available online 16 August 2024

Keywords:
Melodic intonation therapy; aphasia; speech recover; systematic review.

ABSTRACT

Background: Many studies have reported positive results regarding the benefits of melodic intonation therapy (MIT) in patients with non-fluent aphasia. Currently, there is no specific inclusion of speech therapists (STs) in MIT research. Investigating effective speech therapy (ST) techniques to address the language functions hindered by non-fluent aphasia could yield evidence for aphasia rehabilitation research.

Objective: This systematic review (SR) examines the effectiveness of the traditional MIT protocol used by STs on speech recovery for patients with non-fluent aphasia after stroke. It also discusses other characteristics of the traditional MIT, such as the participants, the MIT protocol applied, the therapy intensity, and the role of STs.

Materials and methods: This SR followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) 2020 statement. The four computerized databases (PubMed, Embase, ICTRP, and Google Scholar) were searched in February 2024 to review all empirical findings. We also conducted a hand search in relevant journals. The search yielded 538 studies, of which 2 met the criteria and underwent review. The methodological quality of the included studies was evaluated using the Cochrane Collaboration's tool for assessing the risk of bias. Furthermore, the protocol was registered in PROSPERO under the reference CRD42024508733.

Results: This review included 2 randomized controlled trials (RCTs) involving 44 patients. We found evidence that MIT significantly improved speech recovery, precise language repetition, and functional communication in patients with non-fluent aphasia. STs were interventionists in MIT research and used MIT following the American manual, and they had previously received MIT training.

Conclusion: Our review provides some evidence of the effectiveness of MIT on speech recovery in patients with non-fluent aphasia after stroke. MIT may be a practical alternative to standard ST. There is some indication that MIT requires music therapy (MT) skills and training; therefore, STs must also have these abilities.

Introduction

Aphasia is an acquired loss or impairment of the language system caused by a symptom of brain damage.[1] It has a variety of causes. Cerebrovascular accidents, also known as stroke, are the most common cause. The left hemisphere of the brain typically experiences the pathology, and approximately one-third of stroke patients experience aphasia.[2,3] In the United States, the prevalence of aphasia is 1 in 250.[4]

∗ **Corresponding contributor.**
Author's Address: *Department of Music, Faculty of Humanities, N aresuan University, Phitsanulok Province, Thailand.*
E-mail address: vichb@nu.ac.th
doi: 10.12982/JAMS.2024.063
E-ISSN: 2539-6056

Non-fluent aphasia, also known as expressive aphasia or Broca's aphasia, is one type. This occurs in the damaged area of the inferior frontal lobe in the Broca area, located in Brodmann areas 44 and 45, which are centers for motor control, speech, and communication. Generally, about 96 to 99 percent of right-handed people and about 60 percent of left-handed have their language abilities located in the brain's left hemisphere.[5]

People with non-fluent aphasia or Broca's aphasia do not have or have fewer problems with language comprehension.[5] They can understand speech or conversation but have difficulty expressing words in speech and writing. They have difficulty putting their thoughts into words. They can only produce short utterances when speaking or writing.[6] MIT is a specific treatment for patients with aphasia. Studies have reported positive results regarding the benefits of traditional MIT for aphasic patients.[7-12] MIT is a combination of ST and MT.[13] STs utilize the MIT program as a therapeutic intervention to assist patients with aphasia.[7-12]

MIT involves speaking rhythmically and tapping the left hand to enhance speech prosody and fluency.[14] Traditional MIT is a hierarchically structured therapy that utilizes three key elements: (1) melodic intonation (singing), (2) rhythmic speech, and (3) the use of common phrases (formulaic language). The melodic and rhythmic structures only allow for two notes (high and low) and two durations (long and short). MIT consists of three linguistic tiers. The first two tiers consist of multisyllabic words and short phrases, while the third tier comprises more phonologically complex phrases. The interventionist will teach the patient to speak slowly in high- and low-note sequences that mimic everyday speech.[14]

MIT's use is based on stimulating music processing in the brain's right hemisphere to increase language ability. Research has revealed a correlation between MIT's success and the integrity of the right arcuate fasciculus, which compensates for damage in the brain's left hemisphere, which is associated with language and speech.[15-18]

A recent SR from Mata determined the number of studies involving music therapists (MTs) and their involvement and contributions to the field.[19] Currently, extensive literature focuses on the effectiveness of MIT for patients with non-fluent aphasia or Broca's aphasia. However, STs are not specifically included in MIT research. Investigating effective ST techniques to address the language functions hindered by non-fluent aphasia could yield evidence for aphasia rehabilitation research.

Therefore, the primary aim of this SR was to examine the emergence in the scientific literature of the effectiveness of the traditional MIT protocol used by STs on the speech recovery of patients with non-fluent aphasia after stroke. Furthermore, the secondary aim was to examine other characteristics of the traditional MIT, such as the participants, the MIT protocol applied, the therapy intensity, and the role of STs. The authors intended to answer the following research questions:

(1) What is the effectiveness of the traditional MIT protocol used by speech therapists on the speech recovery of patients with non-fluent aphasia after stroke?

(2) What are other characteristics of traditional MIT, such as the participants, the MIT protocol applied, the therapy intensity, and the role of speech therapists?

Materials and methods

This SR was reported according to the PRISMA 2020 statement to guide the methodology of this research.[20] Additionally, the protocol was registered in PROSPERO under the reference CRD42024508733 (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=508733).

Inclusion and exclusion criteria

Table 1 illustrates a selection criterion for this SR. The inclusion and exclusion criteria based on the population, intervention, comparator, outcome, and study design (PICOS) principles to capture well-prepared studies of interventions are addressed.

Table 1. Inclusion and exclusion criteria.

PICOS	Inclusion	Exclusion
P	Patients with non-fluent aphasia (Broca's aphasia) with a history of stroke (e.g., ischemic stroke, hemorrhagic stroke)	Patients with other types of aphasia (e.g., Wernicke, Global, Isolation, Transcortical motor, Transcortical sensory, Conduction, and Anomic) and Apraxia of speech
I	Melodic intonation therapy	Modified melodic intonation therapy and other types of music therapy
C	Speech therapy, treatment as usual, waiting list, no control group, no intervention group.	N/A
O	Language skills (e.g., information content, fluency, auditory comprehension, repetition and naming)	Primary outcomes were not focus on improvement in language skills
S	RCTs	Other types of study design (e.g., quasi-experimental studies, single-group studies, single case studies, case series, editorials, opinions and commentaries, and qualitative studies)

Search strategy

The four computerized databases were conducted in February 2024 using the following databases: PubMed (75), Embase (77), ICTRP (75), and Google Scholar (295). The researchers also conducted a hand search in relevant music therapy journals to find other potentially eligible studies (16). Furthermore, we used backward citations to check reference lists of all relevant articles. The search was limited to articles written in the English language. Articles published up-to-date information on the existing literature from 2013 onward.

Table 2 demonstrates comprehensive search syntax. The search terms were combined using the Boolean operator OR, and each PICO was combined using the Boolean operator AND.

Table 2. Search syntax.

PICOS	Search terms
P	"Stroke" OR "Aphasia" OR "non-fluent aphasia" OR "Broca aphasia"
I	"Melodic intonation therapy"
C	"Speech therapy" OR "Waiting list" OR "No control group"
O	"Language" OR "Iinguistic skills" OR "Information content" OR "Auditory comprehension" OR "Fluency" OR "Repetition" OR "Naming" OR "Communication"
S	"Randomized controlled trials"

Study selection and Data extraction

Two researchers (N.W. and T.K.) independently screened the titles and abstracts of all identified articles. For studies meeting the eligibility criteria, full-text articles were reviewed. Disagreements about whether a study should be included were discussed until a consensus was reached, involving a third researcher (V.B.) where necessary.

Data were then extracted by the two researchers (N.W. and T.K.) using a data extraction sheet for study design, participants, MIT protocol applied, intervention, duration, language measurements, role of speech-language therapist, and results. The two researchers (N.W. and T.K.) independently extracted each data extraction throughout the entire data extraction process. Disagreements about data extraction were discussed until consensus was reached, involving a third researcher (V.B.) where necessary.

Quality assessment

The methodological quality of the included studies was evaluated using the Cochrane Collaboration tool to assess the risk of bias for RCTs.[21] The two researchers (N.W. and T.K.) independently reviewed and scored the studies. After applying the methodological quality framework to all research, the ratings were translated into percentages to allow quality comparisons across papers. In the disagreement, a third researcher (V.B.) was consulted. However, there was no disagreement to resolve. Table 3 provides a summary of the risk of bias for RCTs.

Table3. Risk of bias for RCTs.

	Van der Meulen *et al*.[7]	Van der Meulen *et al*.[8]
Random sequence generation (selection bias)	+	+
Allocation concealment (selection bias)	+	+
Blinding of participants and personnel (performance bias)	-	-
Blinding of outcome assessment (detection bias)	+	+
Incomplete outcome data (attrition bias)	+	+
Selective reporting (reporting bias)	+	?
Other bias	?	?

Results

Figure 1 shows a PRISMA statement chart that details the selection process. A total of 538 records were screened through the database (N=522) and hand searches (N=16). After removing duplicates, 160 articles remained. The researchers analyzed the abstracts of these articles, resulting in 26 potentially eligible full-text articles; we retained 2 studies for inclusion in the review.

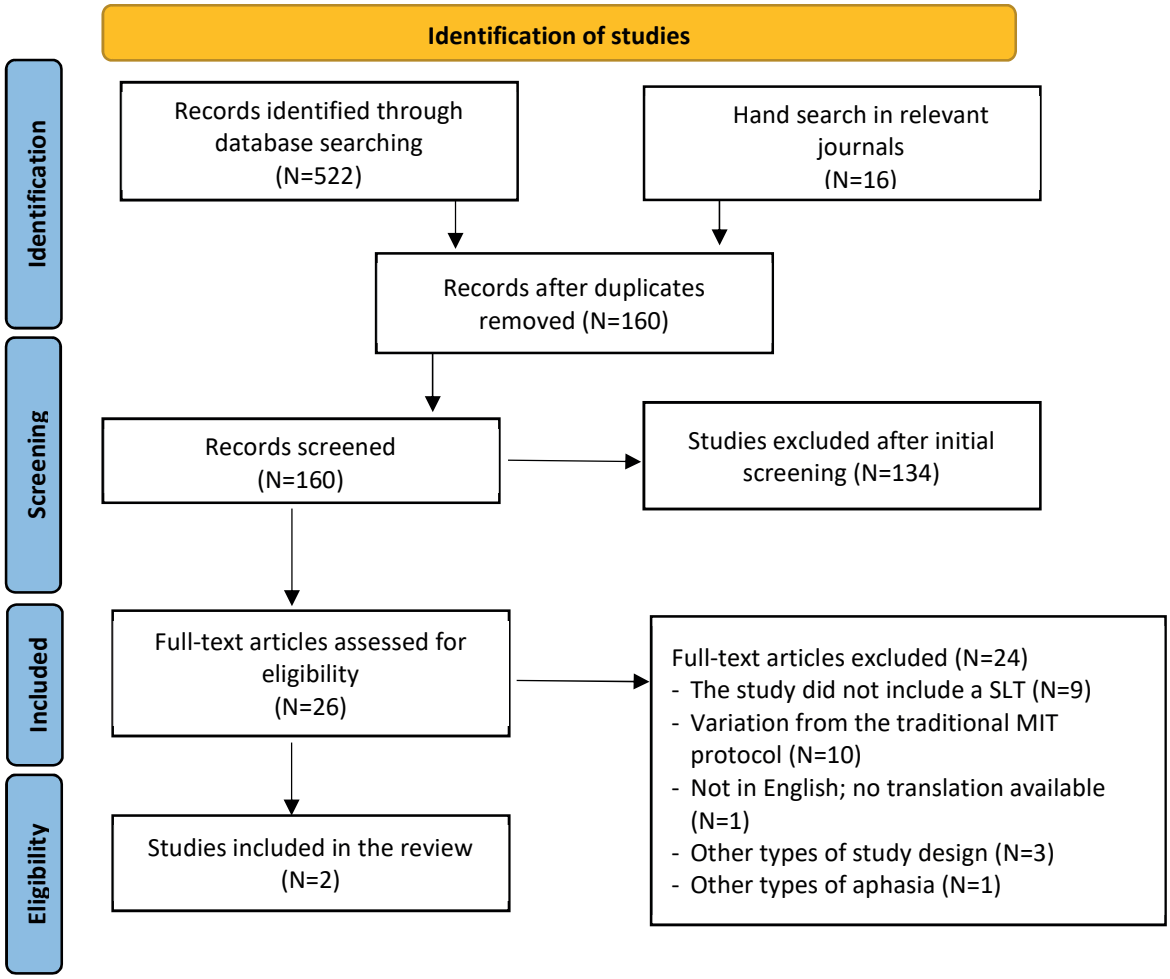


Figure 1. *Methodology flowchart.*

Table 4 presents the included studies. Two studies were conducted in the Netherlands.[7,8] However, it should be noted that the same group of researchers conducted these studies.[7,8] Furthermore, it is essential to note that the selection criteria for the methodological design were RCTs. This review included a total of 2 randomized comparisons.[7,8] All studies were generally highly scientific (level I evidence). Van der Meulen *et al.*[7] conducted a multicenter, waiting-list RCT with a crossover design.

Table 4. Included studies.

Study/ Year/ Country	Study design	Participant	Intervention	Duration	Language measurements	Role of speech therapist	Language benefits
Van der Meulen *et al*.[7] (Netherlands)	Randomized, waiting list, single-blind, multicenter clinical trial	Total: 27 participants (N=16 in the experimental group and N=11 in the control group) Inclusion criteria: - Aphasic after left hemisphere stroke - 2 to 3 months after stroke - Premorbid right-handed - Age between 18 and 80 years - Native language Dutch - MIT candidacy (non-fluent aphasia, articulation deficits, repetition severely affected, and moderate to good auditory language comprehension)	Experimental group: the traditional MIT protocol Control group: delayed MIT	5 hours per week (a minimum of 3 hours per week plus iPod-base homework) for 6 weeks	- Sabadel story retelling task - Amsterdam-Nijmegen Everyday Language Test (ANELT) - Aachen Aphasia Test (AAT: subtests repetition and naming) - MIT repetition task	Interventionist	Improvement in functional communication and language repetition
Van der Meulen *et al*.[8] (Netherlands)	Randomized, waiting list, single-blind, multicenter clinical trial	Total: 17 participants (n = 10 in the experimental group and n = 7 in the control group) Inclusion criteria: - Aphasic after left hemisphere stroke - Up to 12 months after stroke - Premorbid right-handed - Age between 18 and 80 years - Native language Dutch - MIT candidacy (non-fluent aphasia, poorly articulated speech, poor - language repetition, and - moderate to good auditory - language comprehension)	Experimental group: the traditional MIT protocol Control group: delayed MIT	5 hours per week (a minimum of 3 hours per week plus iPod-based homework) for 6 weeks	- Sabadel story retelling task - Amsterdam-Nijmegen Everyday Language Test (ANELT) - Aachen Aphasia Test (subtests repetition and naming) - MIT repetition task	Interventionist	Improvements in language repetition of trained material.

Van der Meulen *et al*. also used a multicenter waiting-list RCT design.[8] A total of 44 patients were included in this systematic review.[7,8] Van der Meulen *et al*.[7] included 27 participants: 16 in the experimental group and 11 in the control group. Van der Meulen *et al*.[8] included 17 participants: 10 in the experimental group and 7 in the control group. All participants included only people with non-fluent aphasia because of stroke. At this point, there was a wide age range, from 18 years old to 80. All authors described their participants' etiologies: strokes in the left hemisphere. However, the stroke took between 2-3 months and 12 months.[7,8]

The traditional MIT intervention was used in two studies.[7,8] Each participant adhered to the traditional protocol. No minor variations were permitted in delivery; the Dutch language was used for intervention. The two studies provided the exact intervention dosage. They did, however, use the various types of homework equipment provided. In Van der Meulen *et al*., the MIT sessions were 5 hours per week (a minimum of 3 hours per week plus homework for 6 weeks).[7] Conversely, in Van der Meulen *et al.,* the MIT sessions lasted 5 hours per week (a minimum of 3 hours per week, plus iPod-based homework for 6 weeks).[8] In these studies, STs previously receiving training at MIT provided the traditional intervention. Crucially, MIT was used following the American manual.[7,8]

All studies reported standardized language measures.[7,8] Van der Meulen *et al*.[7] and Van der Meulen *et al*.[8] used the Sabadel story retelling task and the Amsterdam-Nijmegen Everyday Language Test (ANELT) to evaluate verbal communication. Also, the authors also used the Aachen Aphasia Test (subtests repetition and naming) and the MIT repetition task to assess the ability to repeat and name.[7,8]

The study of Van der Meulen *et al*. demonstrated that MIT benefits language production in severe nonfluent aphasia in the subacute phase poststroke.[7] The experiment group improved verbal communication, suggesting a generalization of verbal communication capabilities in daily life. Also, the experimental group improved language repetition in the trained (MIT test) and the untrained (AAT subtest repetition), suggesting a generalization to untrained material. The result found that the control group received delayed MIT, resulting in less improvement in training material repetition.

The study of Van der Meulen *et al*.[8] demonstrated improvements in both trained and untrained items. The authors also compared the experimental group's language improvement post-MIT to the control group. They discovered that MIT only improved training material repetition. It did not affect word retrieval, everyday verbal communication, or generalization to untrained material. After finishing MIT, patients were unable to maintain their MIT-related language gains.

Discussion

The primary aim of this systematic review was to examine the emergence in the scientific literature of the effectiveness of the traditional MIT protocol used by STs on the speech recovery of patients with non-fluent aphasia.

The main findings showed that MIT significantly improved speech recovery, specifically language repetition[7,8] and functional communication in patients with non-fluent aphasia.[7] The outcome evaluations used in the different clinical trials reported to date assess the efficacy of MIT in post-stroke aphasia. These two studies used the AAT to assess repetition and the ANELT test to assess functional communication.[7,8] Similar to van de Sandt-Koenderman *et al*., the findings indicated that the participants demonstrated significant improvement in repetition using the AAT and everyday verbal communication, as measured by the ANELT.[10]

Some studies have used the communicative activity log (CAL) questionnaire to assess communication. The findings indicated that stroke survivors with nonfluent aphasia have improved their communication skills.[9,11] One study has used the Japanese version of the Western Aphasia Battery (WAB) to assess repetition. The result showed that MIT improved language output (as indicated by the spontaneous speech, repetition, and naming subscores of the WAB).[12] In contrast to Plukwongchuen *et al*., the study demonstrated the positive effects of MIT on spontaneous speech and naming, but not repetition, in Thai stroke patients with aphasia.[22] This study used the Thai version of the WAB test to assess linguistic function. Note that this study used a different measurement tool and did not define the type of aphasia or stroke duration.

Remarkably, Van der Meulen *et al*.,[7] found that patients with subacute severe nonfluent aphasia in the control group experienced less improvement in the repetition of trained material when they received 6-week delayed MIT. Meanwhile, there was an overall more significant improvement in the early MIT group. Therefore, timing influences therapy outcomes. Earlier treatment led to more significant improvement. This is consistent with studies demonstrating that spontaneous recovery occurs within the first three months after a stroke.[23,24]

The secondary aim of this systematic review also included examining other aspects of the traditional MIT, such as participants, MIT protocol, therapy intensity, and the role of the STs.

Participants

In the study by Van der Meulen *et al*., patients with severe nonfluent aphasia persist until 2 to 3 months poststroke.[7] In the study by Van der Meulen *et al*., patients with severe nonfluent aphasia persist for more than 1-year poststroke.[8] The effect of MIT in chronic aphasia is more restricted than its effect in earlier stages post-stroke. When the etiology of aphasia is a stroke, recovery of language function peaks in the first three months.[23,24] Also, Hojo *et al*. found an association between the initial severity of aphasia and recovery rates.[25] People with severe aphasia recovered less than people with mild aphasia, and this trend was obvious in Wernicke and Broca aphasia. Furthermore, in terms of the aphasia type, conduction aphasia had the highest recovery rates, followed by Anomic, Wernicke, and Broca aphasia. Global aphasia

showed much lower recovery rates. Future studies should explore a correlation between severity, aphasia type, and recovery rates.

Additionally, the two included studies provided a lack of detailed documentation of the size and location of the lesion, which had significant neurological characteristics.[7,8] van de Sandt-Koenderman *et al*.[10] found no consistent shift in language activation between the left and right hemispheres. However, subacute patients exhibited symmetrical or right-lateralized language activation before therapy, which tended to become more right-lateralized after treatment. Some chronic patients exhibited left-lateralized language activation, which became stronger following therapy. Furthermore, Tabei *et al*. found that after MIT-J training, the right hemisphere showed decreased activation in correct naming trials, but incorrect trials remained stimulated similarly.[12] Patients with severe chronic non-fluent aphasia have improved neural processing efficiency and decreased cognitive workload.

Besides, there was a wide age range, from 18 to 80. The studies did not provide solid outcome information for each specific age group.[7,8] Hojo *et al*. found that age and recovery rates showed a significant negative correlation.[25] Younger patients recovered more rapidly, and this trend was particularly noticeable in Wernicke aphasics but not in Broca aphasics. The future study should explore a correlation between age and recovery rates.

In addition, the sample of participants met the criteria for good candidacy for MIT therapy, and they benefited from the therapy.[7,8] Similar to previous studies, a positive response to MIT therapy was associated with left hemisphere lesions.[15-18] MIT is based on three main components: singing, rhythmic speech, and common phrases. Singing activates the right hemisphere, which takes over the function of the left-brain speech areas when severe damage occurs. However, the two included studies did not explore language lateralization and neuroplastic reorganization after or during MIT.[7,8] Therefore, the participants in the MIT trial and the clinical benefits resulting from functional and structural changes in the right hemisphere remain unknown.

MIT protocol applied

Speech therapists used MIT following the American manual and the protocol in Dutch target utterances.[7,8] Remarkably, all studies employed the traditional MIT methodology.[7,8] MIT involved both formulaic statements, such as "How are you?" and nonformulaic statements, such as "The ministers are talking nonsense" and "A thunderstorm is coming our way." As a result, they could not fully understand the influence of the MIT components and could not conclude the role of formulaic and nonformulaic language. Evidence supports the idea that the right hemisphere facilitates formulaic language processing.[26] On the other hand, the left hemisphere facilitates non-formulaic language.[27] Consequently, left-hemisphere stroke patients often retain the ability to produce formulaic expressions.[28]

In addition, the researchers provided homework assignments in both studies to ensure therapy intensity. They captured the target utterances in a short video on an iPod application so patients could sing along with the video or repeat the utterances later.[7,8]

Therapy intensity

In the study by Van der Meulen *et al*. and Van der Meulen *et al*., the face-to-face therapy time was 5 hours per week (minimum 3 hours per week plus homework) for 6 weeks.[7,8] This range of therapy intensity showed improvement in functional communication[7] and repetition tasks.[7,8] However, the question of the intensity of therapy remains open. A systematic review by Brady *et al*.[29] found that studies with high-intensity treatment had higher drop-out rates than those with less frequent treatment. Some aphasic patients may not be able to commit to high-intensity therapy. Additionally, the range of optimal timing for aphasia therapy, precisely the post-stroke period, may have led to generalizations about verbal communication. MIT may be more beneficial in the subacute stage than the chronic stage after a stroke.

Role of speech therapist

Speech therapists were interventionists in MIT research.[7,8] They had previously received MIT training and delivered the traditional MIT intervention. According to many studies, STs were licensed MTs who administered the MIT.[9-12] Contrary to Mata, the role of MTs in the MIT study was that of practitioners.[19] They frequently use their professional music composition and singing expertise to modify MIT. ST and MT were separated because MIT necessitates music therapy skills. However, the two disciplines can collaborate in a clinical context to perform MIT or work separately, but the combined expertise of both experts would result in best research practices. Our study reveals that MIT necessitates MT skills and training, which STs must also fulfill.

Limitations and future research

Our study has several limitations. First, the low number of included studies (N=2) and the total number of participants (N=44) used in this review may not represent the general population. To corroborate the findings, a larger sample size was required, as the limited sample size was a limitation and, therefore, insufficient to draw definite conclusions. Second, this study did not provide information about MIT's role in neurobiological mechanisms that promote language recovery and neuroplastic reorganization. Future research should investigate language lateralization and neuroplastic reorganization during or after MIT to determine how language lateralization and neuronal reorganization change in patients with non-fluent aphasia during or after MIT.

Despite the several limitations of this review, the research provides a solid basis for further investigations into MIT's noteworthy—and possibly unexpected—

contribution to speech recovery. It is also crucial to consider the clinical implications of the results in speech therapy.

This study found a significant correlation between treatment intensity and therapy outcomes after MIT. Earlier treatment may also lead to more substantial improvement. Importantly, this finding suggests STs must also possess music therapy skills and training when using MIT.

Conclusions

This systematic review demonstrates that MIT may benefit language repetition and functional communication in patients with non-fluent aphasia. Additionally, MIT may be a practical alternative to standard speech therapy in encouraging speech recovery for these patients. However, MIT requires music therapy skills and training; speech therapists must also have these abilities.

Conflict of interest

The authors declare no conflict of interest, and the review did not require full board ethics approval because it was a systematic review. The Ethics Committee, Faculty of Associated Medical Sciences, Chiang Mai University (CMU), Thailand (AMSEC-66EM-024) approved the study for exemption review. The data that support this study's findings are available on request from the corresponding author.

Funding

The authors declared that no financial support from any public, commercial, or non-profit organization was given for the writing of this article.

References

[1] Benson DF, Ardila A. Aphasia: A clinical perspective. Oxford University Press, USA; 1996.

[2] Engelter ST, Gostynski M, Papa S, Frei M, Born C, Ajdacic-Gross V, Gutzwiller F, Lyrer PA. Epidemiology of aphasia attributable to first ischemic stroke: incidence, severity, fluency, etiology, and thrombolysis. Stroke. 2006; 37(6): 1379-84. doi:10.1161/01.STR.0000221815.64093.8c

[3] Laska AC, Hellblom A, Murray V, Kahan T, Von Arbin M. Aphasia in acute stroke and relation to outcome. J Intern Med. 2001; 249(5): 413-22. doi:10.1046/j.1365-2796.2001.00812.x

[4] National Institute on Deafness and Other Communication Disorders. NIDCD fact sheet: Aphasia. 2015 [cited 2024 Apr 26]. Available from: https://www.nidcd.nih.gov/sites/default/files/Documents/health/voice/Aphasia.pdf.

[5] Acharya AB, Wroten M. Broca Aphasia. 2023 [Updated 2023 Feb 13; cited 2024 Apr 26]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK436010/

[6] Kernich CA. MSN, RN. Aphasia. The Neurologist. 2004; 10(3): 169-70. doi: 10.1097/01.nrl.0000126591.23625.6e

[7] Van der Meulen I, van de Sandt-Koenderman WM, Heijenbrok-Kal MH, Visch-Brink EG, Ribbers GM. The efficacy and timing of melodic intonation therapy in subacute aphasia. Neurorehabil Neural Repair. 2014; 28(6): 536-44. doi:10.1177/1545968313517753

[8] Van Der Meulen I, De Sandt-Koenderman V, Mieke WM, Heijenbrok MH, Visch-Brink E, Ribbers GM. Melodic intonation therapy in chronic aphasia: Evidence from a pilot randomized controlled trial. Front Hum Neurosci. 2016; 10: 221011. doi:10.3389/fnhum.2016.00533

[9] Haro-Martínez AM, Lubrini G, Madero-Jarabo R, Díez-Tejedor E, Fuentes B. Melodic intonation therapy in post-stroke nonfluent aphasia: a randomized pilot trial. Clin Rehabil. 2019; 33(1): 44-53. doi:10.1177/0269215518791004

[10] van de Sandt-Koenderman MW, Mendez Orellana CP, van der Meulen I, Smits M, Ribbers GM. Language lateralisation after melodic intonation therapy: an fMRI study in subacute and chronic aphasia. Aphasiology. 2018; 32(7): 765-83. doi:10.1080/02687038.2016.1240353

[11] Haro-Martínez AM, García-Concejero VE, López-Ramos A, Maté-Arribas E, López-Táppero J, Lubrini G, Díez-Tejedor E, Fuentes B. Adaptation of melodic intonation therapy to Spanish: a feasibility pilot study. Aphasiology. 2017; 31(11): 1333-43. doi:10.1080/02687038.2017.1279731

[12] Tabei KI, Satoh M, Nakano C, Ito A, Shimoji Y, Kida H, Sakuma H, Tomimoto H. Improved neural processing efficiency in a chronic aphasia patient following melodic intonation therapy: A neuropsychological and functional MRI study. Front Neurol. 2016; 7: 148. doi:10.3389/fneur.2016.00148

[13] Albert ML, Sparks RW, Helm NA. Melodic intonation therapy for aphasia. Arch Neurol. 1973; 29(2): 130-1. doi:10.1001/archneur.1973.00490260074018

[14] Zumbansen A, Peretz I, Hébert S. Melodic intonation therapy: back to basics for future research. Front Neurol. 2014; 5: 7. doi:10.3389/fneur.2014.00007

[15] Helm-Estabrooks N. Exploiting the right hemisphere for language rehabilitation: Melodic Intonation Therapy. In: Perecman E, ed. Cognitive Processing in the Right Hemisphere. New York, NY: Academic Press; 1983: 229-240.

[16] Schlaug G, Marchina S, Norton A. Evidence for plasticity in white-matter tracts of patients with chronic Broca's aphasia undergoing intense intonation-based speech therapy. Ann N Y Acad Sci. 2009; 1169: 385-94. doi:10.1111/j.1749-6632.2009.04587.x

[17] Schlaug G, Norton A, Marchina S, Zipse L, Wan CY. From singing to speaking: facilitating recovery from nonfluent aphasia. Future Neurol. 2010; 5(5): 657-65. doi: 10.2217/fnl.10.44

[18] Zipse L, Norton A, Marchina S, Schlaug G. Singing versus speaking in nonfluent aphasia. NeuroImage. 2009; 47: S119. doi:10.1016/S1053-8119(09)71121-8.

[19] Mata HL. A systematic review of melodic intonation

therapy that involved music therapists (Doctoral dissertation).

[20] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021; 372: n71. doi:10.1136/bmj.n71

[21] Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savović J, Schulz KF, Weeks L, Sterne JA. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011; 343: d5928. doi:10.1136/bmj.d5928

[22] Plukwongchuen T, Klubchai P, Punchasuwan P, Uma M, Kanlaya S. Effect of Melodic Intonation Therapy on Scoring of Spontaneous Speech, Repetition and Naming in Thai Non-fluent Aphasic Patient. Poster session presented at Annual academic conference 32nd "Healthcare in a Changing World", 2016 April 3-5; Songkla, Thailand.

[23] El Hachioui H, Lingsma H, Van de Sandt-Koenderman WME, Dippel D, Koudstaal P, Visch-Brink EG. Recovery from aphasia after stroke: a one year follow-up study. J Neurol. 2013; 260: 166-71. doi:10.1007/s00415-012-6607-2

[24] Bakheit A, Shaw S, Carrington S, Griffiths S. The rate and extent of improvement with therapy from different types of aphasia in the first year after stroke. Clin Rehabil. 2007; 21: 941-9. doi:10.1177/0269215507078452

[25] Hojo K, Watanabe S, Tasaki H, Sato T, Metoki H, Saito M. [Recovery in aphasia (Part 1)]. No To Shinkei. 1985; 37(8): 791-7. Japanese. PMID: 4074584.

[26] Van Lancker Sidtis D, Postman WA. Formulaic expressions in spontaneous speech of left-and right-hemisphere-damaged subjects. Aphasiology. 2006 May 1; 20(5): 411-26. doi:10.1080/02687030500538148

[27] Meinzer M, Flaisch T, Breitenstein C, Wienbruch C, Elbert T, Rockstroh B. Functional re-recruitment of dysfunctional brain areas predicts language recovery in chronic aphasia. Neuroimage. 2008; 39(4): 2038-46. doi:10.1016/j.neuroimage.2007.10.008

[28] Lum CC, Ellis AW. Is "nonpropositional" speech preserved in aphasia?. Brain and Lang. 1994; 46(3): 368-91. doi:10.1006/brln.1994.1020

[29] Brady MC, Kelly H, Godwin J, Enderby P, Campbell P. Speech and language therapy for aphasia following stroke. Cochrane Database Syst Rev; 2016; 6: CD000425. doi:10.1002/14651858.CD000425.pub4

Journal of Associated Medical Sciences 2024; 57(3): 221-229

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

Exploration of divalent metal transporter 1 (*DMT1*) gene intronic IVS4+44C/A polymorphisms in population exposed to cadmium

Sittiporn parnmen[1*] Nattakarn Nooron[1] Pornpanna Chonnakijkul[1] Sujitra Sikaphan[1] Dutsadee Polputpisatkul[1] Chutimon Uttawichai[1] Rungsaeng Chankunasuka[1] Sriprapa Phatsarapongkul[1] Chidkamon Thunkhamrak[1] Unchalee Nitma[1] Nisakorn Palakul[1] Archawin Rojanawiwat[2]

[1]Toxicology Center, National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Nonthaburi Province, Thailand.
[2]National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Nonthaburi Province, Thailand.

ARTICLE INFO

Article history:
Received 22 December 2023
Accepted as revised 17 August 2024
Available online 22 August 2024

Keywords:
Divalent metal transporter 1, genetic polymorphism, human biomonitoring, urinary cadmium.

ABSTRACT

Background: Cadmium exposure affects the expression of the *DMT1* gene and the function of its transporter protein, impacting the transport and accumulation. This study investigates genetic polymorphisms to understand better the pivotal role of genetic factors in cadmium-related diseases within environmental health research.

Objective: The study sought to examine the intronic IVS4+44C/A polymorphism in the divalent metal transporter 1 (*DMT1*) gene among individuals aged 35-60 residing in regions contaminated with cadmium.

Materials and methods: Blood samples were collected from 306 genetically unrelated individuals (158 females and 148 males). The *DMT1* IVS4+44C/A polymorphism was determined using restriction fragment length polymorphism (RFLP) and Sanger sequencing methods. Urinary cadmium levels were measured with graphite furnace atomic absorption spectrometry (GFAAS). Statistical analyses included Hardy-Weinberg equilibrium testing, analysis of variance (ANOVA), and student's *t*-tests.

Results: The geometric mean of urinary cadmium levels were significantly higher in females (4.03±4.15 µg/gm creatinine) than in males (2.62±2.73 µg/gm creatinine). Remarkably, 85% of females and 66% of males exceeded the reference values for urinary cadmium concentration set by the German Human Biomonitoring (HBM) Commission (HBM I and II). Genotype frequencies were 65.4% homozygote typical (CC), 31.0% heterozygote (CA), and 3.6% homozygote atypical (AA). The C allele frequency was 80.9%, while the A allele frequency was 19.1%. Notably, the *DMT1* IVS4+44C/A polymorphism significantly influenced urinary cadmium levels, with the CA genotype showing higher levels than CC and AA genotypes. Urinary cadmium levels were also statistically increased with the presence of the A allele (A^+= CA+AA) compared to its absence (A^-= CC). Furthermore, our analysis revealed that individuals with the CC genotype more frequently surpass the reference values for urinary cadmium in HBM I and II across all age groups despite their overall urinary cadmium levels not being high.

Conclusion: This study indicates that the CA genotype may signify susceptibility to prolonged cadmium exposure, given its association with elevated urinary cadmium levels. Additional research is essential for a thorough grasp of the implications of *DMT1* gene polymorphisms on health outcomes and for establishing monitoring measures for populations residing in cadmium-contaminated areas.

∗ **Corresponding contributor.**
Author's Address: *Toxicology Center, National Institute o f Health, Department o f Medical Sciences, Ministry of Public Health, Nonthaburi Province, Thailand.*
E-mail address: sittiporn.p@dmsc.mail.go.th
doi: 10.12982/JAMS.2024.064
E-ISSN: 2539-6056

Introduction

The largest zinc mining site in the country, formerly located in Mae Sot District, Tak Province, operated for over 32 years until its closure in 2017. During 2001-2004,

cadmium (Cd) accumulation as a byproduct of zinc smelting was observed in the environment and the food chain.[1–5] Natural cadmium occurring contains eight stables and two radioactive isotopes.[6] Cadmium primarily exists in the divalent cation (+2) oxidation state and exhibits slow oxidation.[6] Despite the cessation of mining activities, cadmium continues to pose environmental and health concerns in the region. Cadmium has an exceptionally long biological half-life, with significant accumulations in the kidneys (6-38 years) and liver (4-19 years) following initial distribution in the blood (75 to 128 days).[7,8] The remaining cadmium disperses throughout other tissues (9-47 years), with renal accumulations gradually excreted in the urine, making urinary cadmium levels an indicator of long-term exposure.[7–9] Concerning the homeostatic mechanisms governing cadmium regulation, metallothionein (*MT*) and the divalent metal transporter 1 (*DMT1*) genes are key players in the process. Together, they form an intricate system dedicated to maintaining cadmium homeostasis. These components collaborate to mitigate cadmium toxicity by sequestering it within cells, regulating its uptake, and ensuring controlled elimination, thereby safeguarding the body against the adverse consequences of excessive cadmium exposure.[10–15] Examining single nucleotide polymorphisms (SNPs) in genes associated with homeostatic mechanisms is crucial, as these genetic variations may be linked to the development of various diseases.[11,14–17]

In this study, the intronic region of *DMT1* gene was used to investigate the IVS4 +44C/A polymorphism in the general population residing in cadmium-contaminated areas. This position is in the intron 2 with C/A SNP in the intervening sequence 4+44 (IVS4+44) region.[10,11] Our study population consisted of 306 subjects (158 females and 148 males) aged 35-60 years from Mae Sot District, Tak Province in western Thailand. The area was reported to have high levels of cadmium in the environment and its accumulation in food chains.[1–3] The accumulation of cadmium in the body can be measured through blood or urinary cadmium levels, which are effective biomarkers for exposure. The study employed urinary cadmium levels in the population to indicate long-term exposure. The population was divided according to the German Commission on Human Biomonitoring (HBM) reference values for urinary cadmium exposure.[18] Investigating patterns of *DMT1* gene polymorphisms concerning cadmium exposure helps elucidate individual differences in susceptibility and improves risk assessments. Data on genetic polymorphisms in populations living in cadmium-contaminated areas in Thailand is inadequate. Such data could significantly enhance our understanding of the interactions between genetics and environmental exposures, contributing valuable knowledge for future research. Thus, this study aimed to explore the association between the *DMT1* IVS4+44C/A polymorphism and urinary cadmium levels in a population exposed to cadmium contamination

Materials and methods

Study subjects

In this study, 306 blood samples were collected from genetically unrelated and healthy individuals, including 158 females (mean age 48.84±6.65 years; range 35-60 years) and 148 males (mean age 50.57±6.45 years; range 35-60 years). The study was conducted in the western Thai population living in cadmium-contaminated areas. The exclusion criteria for the subjects included a medical history of renal failure, kidney stones, osteoporosis, and carcinoma. Both genders were divided into 3 different age groups: 35 to 44, 45 to 54, and 55 to 60.

DNA extraction and restriction fragment length polymorphism (RFLP) method

According to the manufacturer's instructions, blood genomic DNA was extracted using the FavorPrep™ Blood Genomic DNA Extraction Mini Kit (Favorgen, Taiwan). DNA was stored at -20 °C until required for polymerase chain reaction (PCR) analysis. PCR amplification was performed using the following forward primer: 5'-GAC ACA TGC AAT ATC TGA CAT TG-3' and reverse primer: 5'-AGG CTA CTA TCC AAC ATG CAG-3'.[10,11] Each PCR reaction of 25 µL contained 9.5 µL of OnePCR™ master mix with fluorescence dye (GeneDireX®, Taiwan), 2.5 µL of 10 µM of each primer, 1 µL of DNA template, and 9.5 µL of nuclease-free H_2O. The amplification was conducted using SimpliAmp™ Thermal Cycler (Thermo Fisher Scientific, USA) under the following conditions: initial denaturation at 94 °C for 5 min, followed by 35 cycles of denaturation at 94°C for 1 min, annealing at 60 °C for 1 min, extension at 72 °C for 1 min, and final extension at 72 °C for 5 min. For detection of the *DMT1* IVS4+44 C/A polymorphism (rs224589; GeneID: 4891; accession Number: NC_000012), PCR products were digested with *Mnl*I restriction enzyme (Thermo Fisher Scientific, USA) by incubating at 37 °C for 4 hr. Each reaction of 20 µL contained 10 µL of PCR product, 2 µL of 10X buffer, 7.75 µL of nuclease-free H_2O and 0.25 µL of *Mnl*I. Digested PCR products were examined using 3% (w/v) agarose gel electrophoresis. The PCR products of each variant were randomly selected and cleaned using the FavorPrep™ GEL/PCR Purification Mini Kit (Favorgen, Taiwan) for Sanger sequencing. Ten microliters cycle sequencing reactions consisted of 1 µL of BigDye™ Terminator v3.1 (Applied Biosystems, USA), 1 µL of BigDye™ 5X running buffer, 1 µL of 10 pM primer, 5 µL of nuclease-free H_2O and 2 µL of purified PCR product. The cycle sequencing conditions were as follows: 25 cycles of 96 °C for 10 sec, 50 °C for 5 sec, and 60 °C for 4 min. Samples were purified using BigDye™ XTerminator™ purification (Applied Biosystems, USA) and sequenced using the Applied Biosystem 3500xL Genetic Analyzer (Applied Biosystems, USA).

Urinary cadmium analysis

Cadmium levels and creatinine urine were measured by graphite furnace atomic absorption spectrometry (GFAAS) (Perkin Elmer, Analyst 600) and the automated

analyzer (Beckman coulter, AU 480 chemistry System) according to our previous protocols.[19] The limit of quantification was 0.5 µg/L. Cadmium level (µg/L) was divided by the determined creatinine amount (gm/L) to obtain creatinine-adjusted urinary cadmium levels (µg/gm creatinine).

Statistical analysis and reference values

The study design was cross-sectional. The frequencies of alleles and genotypes were directly observed, and the departure from the Hardy-Weinberg equilibrium was evaluated by the Chi-Square (χ^2) test. The analysis of variance (ANOVA) or student's *t*-tests was used to identify differences in means±SD of cadmium measures across categories of urinary cadmium level, age, and genotype. Urinary cadmium values used in this study were published by the German Human Biomonitoring (HBM) Commission.[18] The German Commission on HBM recommended two reference values, HBM I, and HBM II, for urinary cadmium based on toxicology and epidemiology studies. In adults (>25 years old), urinary cadmium levels below the lower HBM I alert level (2 µg/gm creatinine) are not considered to be a risk of advanced health effects. In contrast, action levels above HBM II (5 µg/gm creatinine) indicate an increased risk of adverse health effects in susceptible individuals in the general population.

Results

Of the 306 subjects enrolled in the study, the inclusion criteria required that participants have no history of kidney failure, kidney stones, osteoporosis, or carcinoma. The total geometric mean of urinary cadmium level was 4.03±4.15 µg/gm creatinine in females (ranging from 0.43 to 23.42 µg/gm creatinine) and 2.62±2.73 µg/gm creatinine in males (ranging from 0.25 to 14.21 µg/gm creatinine). Considering smoking history among individuals, non-smokers had lower urinary cadmium levels than smokers in both genders (Table 1). The female smokers had higher urinary cadmium levels than male smokers. However, there was no statistically significant difference between non-smokers and smokers among the female (p=0.406) and male (p=0.183) population living in the cadmium-contaminated areas.

Table 1. Baseline characteristics of the study population according to self-reported smoking exposure.

Gender	Variable	N	Age±SD (min-max) year	Geometric mean±SD of urinary Cd levels (min-max), µg/gm creatinine	*p* value
Female	Non-smoker	147	49±6.68 (35-60)	4.02±4.13 (0.43-23.42)	0.406
	Smoker	11	50±4.70 (43-58)	4.11±4.59 (0.95-14.75)	
Male	Non-smoker	73	51±7.45 (35-60)	2.43±2.78 (0.24-13.49)	0.183
	Smoker	75	51±5.35 (37-60)	2.81±2.67 (0.37-14.21)	

After total genomic DNA was extracted from the blood samples, the *DMT1* intron was amplified, and amplicons of 351 bp were obtained (Figure 1 and Figure 2). The IVS4+44C/A SNP was determined based on PCR-RFLP technique using *Mnl*I restriction enzyme with nonpalindromic sequence recognition (Figure 1 and Figure 2).

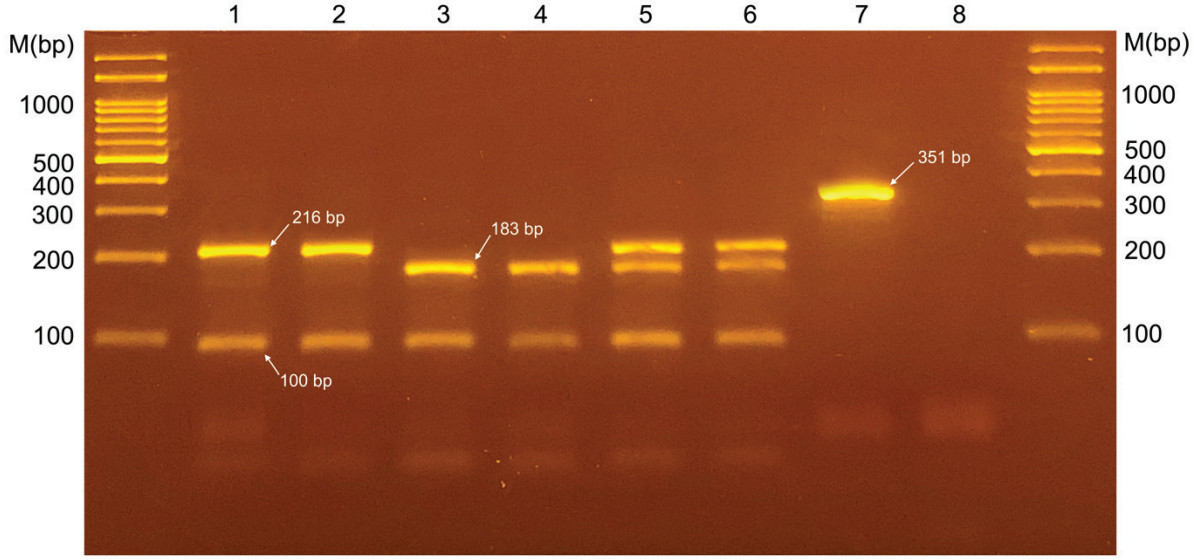


Figure 1. *Representative gel image of digested with MnlI restriction enzyme and undigested PCR products. (M 100 bp ladder), lanes 1, 2: AA genotype (216, 100 bp), lanes 3, 4: CC genotype (183, 100 bp), lanes 5, 6: CA genotype (216, 183 and100 bp), lane 7: undigested PCR product (351 bp), lane 8: negative control.*

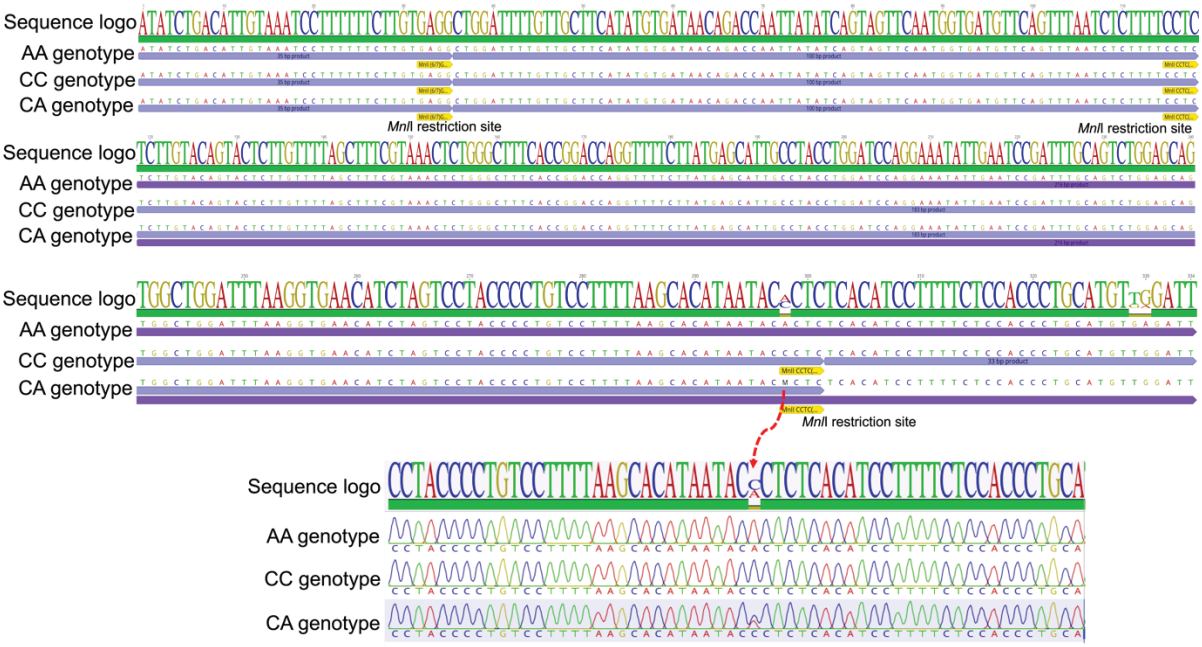


Figure 2. *DNA sequence and cutting of the 351 bp amplicon with MnlI for the presence of A and C alleles.*

Genotype frequencies among 306 subjects were 65.4% homozygote typical for CC (N=200), 31.0% heterozygote for CA (N=95), and 3.6% homozygote atypical for AA (N=11) (Table 2). The frequency of the C allele was 80.9%, while that of the A allele was 19.1%. The genotype and allele frequencies in both female and male individuals were consistent with the Hardy-Weinberg equilibrium (p=0.936 for females and p=0.938 for males) (Table 2). The frequency of the A allele in the female individuals was higher than in the male individuals; however, the results were not statistically significant (p=0.6812) (Figure 3).

Table 2. Genotype and allele frequencies of *DMT1* IVS4+44C/A polymorphism in female and male individuals.

Gender	Polymorphism	Observed frequency of genotypes		Predicted frequency by Hardy-Weinberg equilibrium		*p* value
		N	%	N	%	
Female	IVS4+44C/C	93	58.9	94.98	60.1	0.936
	IVS4+44C/A	59	37.3	55.05	34.8	
	IVS4+44A/A	6	3.8	7.98	5.1	
	IVS4+44C allele	245	77.5	-	-	
	IVS4+44A allele	71	22.5	-	-	
Male	IVS4+44C/C	107	72.3	105.6	71.3	0.938
	IVS4+44C/A	36	24.3	38.9	26.3	
	IVS4+44A/A	5	3.4	3.6	2.4	
	IVS4+44C allele	250	84.5	-	-	
	IVS4+44A allele	46	15.5	-	-	
Total	IVS4+44C/C	200	65.4	200.2	65.4	0.946
	IVS4+44C/A	95	31.0	94.6	30.9	
	IVS4+44A/A	11	3.6	11.2	3.7	
	IVS4+44C allele	495	80.9	-	-	
	IVS4+44A allele	117	19.1	-	-	

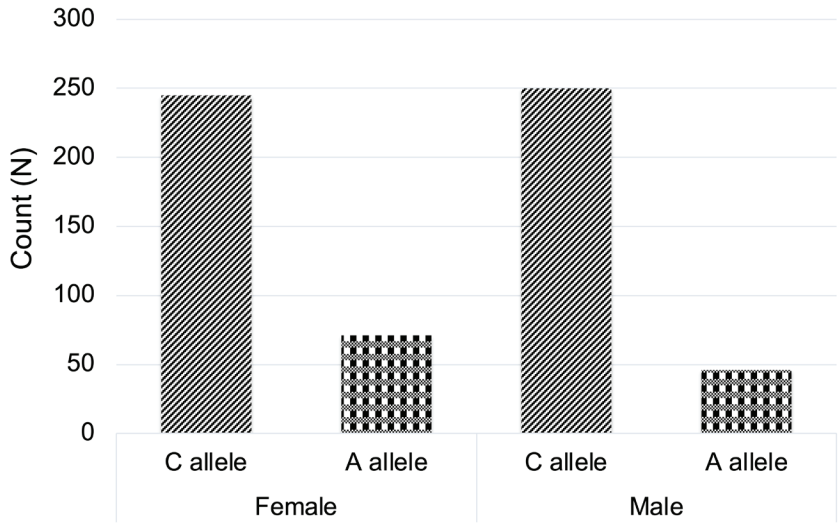


Figure 3. *Distribution of C and A allele frequencies in the female and male individuals.*

Individuals with the CA genotype had higher geometric mean urinary cadmium levels than the CC and AA genotypes. This study revealed that the cadmium levels were significantly higher in the combined male and female population with the presence of the A allele (A^+ = CA+AA) compared to those without the A allele (A^- = CC) (p=0.0007). Regarding gender, it was shown that female subjects with the CA genotype and male subjects with the AA genotype had higher urinary cadmium levels than those with the CC genotype. Females exhibited a statistically significant difference in cadmium levels (p=0.0001), while males did not show a statistically significant difference (p=0.0789). The values of urinary cadmium concentration based on HBM I (2 µg/gm creatinine) and HBM II (5 µg/gm creatinine) for adults older than 25 years old were exceeded by 73% for CC, 80% for CA, and 82% for AA genotypes (Table 3).

Table 3. *DMT1* IVS4+44C/A polymorphism and urinary cadmium (U-Cd) level.

Gender	Polymorphism	N	Geometric mean±SD of urinary Cd levels (µg/gm creatinine)	Human Biomonitoring (HBM) value Frequency, N (%)			*p* value
				<HBM I	HBM I (2 µg/gm creatinine)	HBM II (5 µg/gm creatinine)	
Female	CC	93	3.82±363	14 (15)	46 (50)	33 (35)	0.0001 (p<0.01)
	CA	59	4.45±4.79	9 (15)	23 (39)	27 (46)	
	AA	6	2.89±4.77	1 (17)	3 (50)	2 (33)	
	A(-)(CC)	93	3.82±363	14 (15)	46 (50)	33 (35)	
	A(+)(CA+AA)	65	4.43±4.78	10 (15)	26 (40)	29 (45)	
Male	CC	107	2.57±2.66	40 (37)	49 (46)	18 (17)	0.0789 (p>0.05)
	CA	36	2.67±2.64	10 (28)	13 (36)	13 (36)	
	AA	5	3.34±4.54	1 (20)	2 (40)	2 (40)	
	A(-)(CC)	107	2.57±2.66	40 (37)	49 (46)	18 (17)	
	A(+)(CA+AA)	41	2.75±2.88	11 (27)	15 (36.5)	15 (36.5)	
Total	CC	200	3.09±3.24	54 (27)	95 (48)	51 (25)	0.0007 (p<0.01)
	CA	95	3.72±4.24	19 (20)	36 (38)	40 (42)	
	AA	11	3.04±4.46	2 (18)	5 (46)	4 (36)	
	A(-)(CC)	200	3.09±3.24	54 (27)	95 (48)	51 (25)	
	A(+)(CA+AA)	106	3.64±4.26	21(20)	41 (38)	44 (42)	

A comparison between age ranges, genotype patterns, and levels of exposure revealed that both females and males showed high levels of cadmium exposure according to age (Figure 4). The increase in cadmium level showed a statistically significant difference with age and genotypes in both females (ANOVA, p=0.0020) and males (ANOVA p=0.0180). Focusing on the number of exposures to cadmium that exceeded the urinary cadmium reference values for HBM I and II across all ages, the results showed that individuals with the homozygote typical CC genotype more frequently exceeded these reference values (Figure 5).

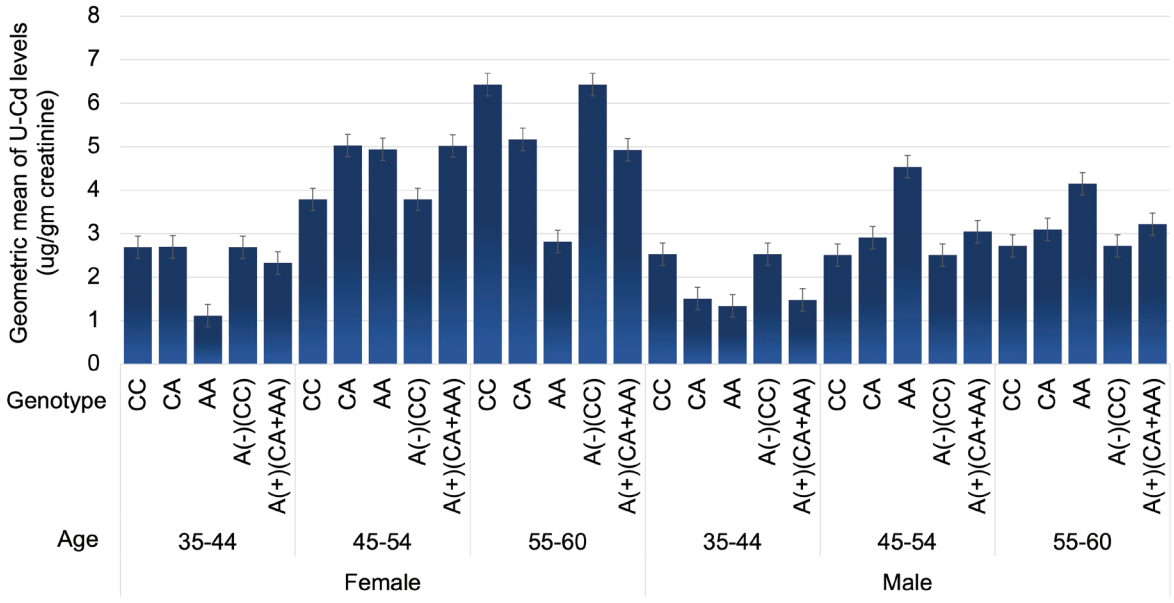


Figure 4. *Distribution of genotype and urinary cadmium levels in female and male subjects.*

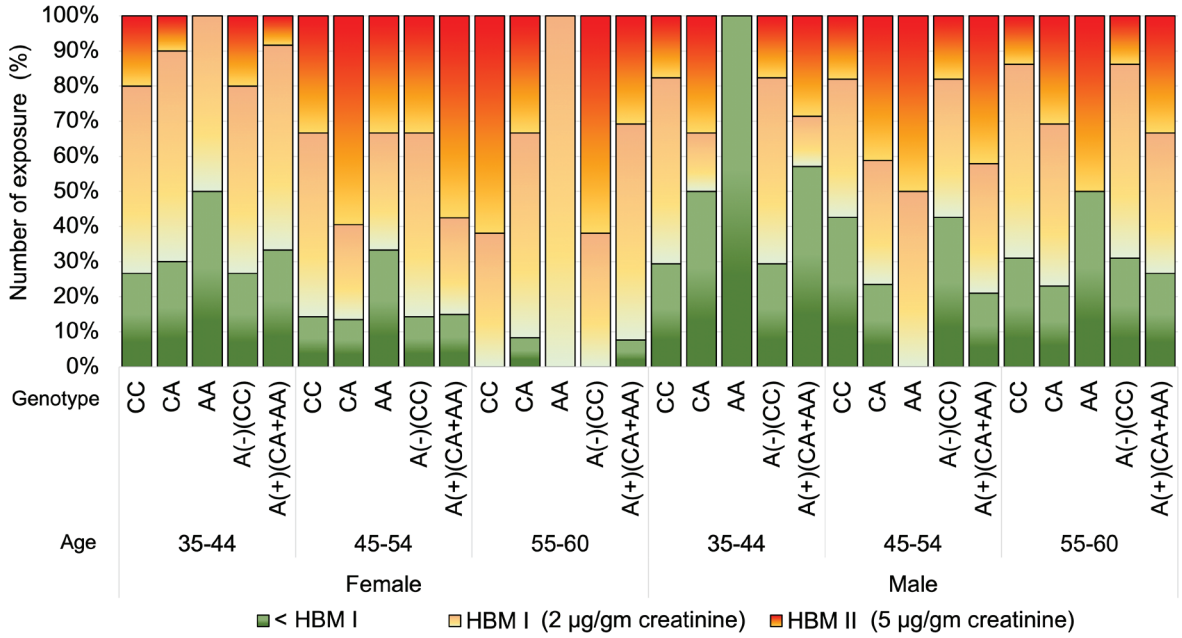


Figure 5. *Distribution of genotype and guideline values for urinary cadmium levels in female and male subjects.*

Discussion

Divalent metal transporter 1 (*DMT1*) gene is known as the natural resistance-associated macrophage protein 2 (*NRAMP2*) gene or the solute carrier family 1 member 2 (*SLC1A2*) gene on human chromosome 12q13.[10,11] This gene encodes a member of a family of solute transporter proteins which are ubiquitously expressed in human tissues, and responsible for iron uptake in duodenal enterocytes and kidneys,[11] iron-regulatory in the placenta,[20] and iron homeostasis in brain.[21] The *DMT1* is not only responsible for the uptake and translocation of iron from endosomes during the transferrin cycle but also absorbs and transports the other divalent cation metals, including manganese, copper, zinc, and some toxic metals such as lead and cadmium.[11,22,23] In humans, metal excretion relies on homeostasis mechanisms rather than specific pathways. Cadmium absorption starts in the gastrointestinal tract, where intestinal metallothionein (MT) is protective to mitigate cadmium toxicity.[12,13] The cadmium-metallothionein (Cd-MT) complex from the liver enters the bloodstream and is filtered by the kidneys into urine.[12,13,24] However, if the amount of cadmium absorbed by the body exceeds the metallothionein ability to form complexes, cadmium accumulates.[13,24] Additionally, the divalent metal transporter 1 (*1*) protein, responsible for iron absorption and divalent metal transport, including cadmium, can facilitate the transport of toxic heavy metals into cells.[10,11,14,15] This process increases oxidative stress, inflammation, and DNA damage.[14,15] Cadmium disrupts cellular function by interfering with essential metal transport processes, increasing oxidative stress, and promoting inflammation, ultimately contributing to cellular dysfunction and disease.[14-15,22,25-28] The *DMT1* gene polymorphisms have been linked to various diseases, including Parkinson's disease,[22] Wilson's disease,[25] hypertension,[26] iron deficiency anemia,[27] and age-related macular degeneration (AMD).[28]

In this study, we observed the *DMT1* IVS4+44 C/A polymorphism in a population historically contaminated by cadmium. Analysis of cadmium levels in urine, adjusted by urine creatinine, revealed a significantly higher geometric mean of urinary cadmium levels in females than males. Furthermore, urinary cadmium levels of non-smokers and smokers were not significantly different between females and males living in areas with a history of cadmium contamination. The genotype frequencies in 306 subjects were highest for the CC genotype (65.4%), followed by the CA genotype (31.0%) and the AA genotype (3.6%). The frequency of the C allele was higher than the A allele. This distribution was similar to studies conducted on the Turkish population.[10,11]

Limited studies have explored the impact of *DMT1* IVS4+44 C/A polymorphism on heavy metal levels. A survey by Kayaalti *et al*. suggested that individuals with the CC genotype might be more susceptible to increased blood iron, lead, and cadmium levels than those with the AA and CA genotypes.[11] Blood cadmium levels are reliable indicators of recent exposure,[7] implying that the CC genotype may heighten susceptibility to short-term cadmium exposure, particularly in the general population. Our results indicate that individuals with the CA genotype exhibited higher geometric mean urinary cadmium levels than those with the CC and AA genotypes. Furthermore, although the overall urinary cadmium levels were not high, the highest frequency of the CC genotype among 200 subjects exceeded urinary cadmium reference values for HBM I and II across all age groups. Thus, our findings suggest that *DMT1* IVS4+44 C/A polymorphisms contribute to inter-individual variations in urinary cadmium levels. The CC genotype may signify susceptibility to recent

cadmium exposure in the general population. In contrast, the CA genotype may indicate susceptibility to long-term cadmium exposure in populations residing in cadmium-contaminated areas.

Conclusions

This marks the initial investigation delving into the connection between the *DMT1* IVS4+44 C/A polymorphism and urinary cadmium levels among genetically unrelated individuals residing in cadmium-contaminated regions. These findings could have important implications for identifying susceptible individuals to cadmium toxicity and developing effective prevention strategies. However, this study has limitations. First, there is limited population size, which must be replicated in other populations. Second, a lack of blood cadmium detection indicates recent exposure, allowing a comparison with previous studies.

Ethical approval

This study was performed in line with the principles of the Declaration of Helsinki, which was approved by the Research Ethics Committee for Research in Human Subjects at the Department of Medical Sciences, Ministry of Public Health, Thailand (Study code. 4/2561, date of approval: 17 February 2023).

Conflict of interest

The authors declare that there is no conflict of interest.

References

[1] Swaddiwudhipong W, Limpatanachote P, Mahasakpan P, *et al*. Cadmium-exposed population in Mae Sot District, Tak Province: 1. Prevalence of high urinary cadmium levels in the adults. J Med Assoc Thai. 2007; 90(1): 143-8.

[2] Boonprasert K, Kongjam P, Limpatanachote P, *et al*. Urinary and blood cadmium levels in relation to types of food and water intake and smoking status in a Thai population residing in cadmium-contaminated areas in Mae Sot. Southeast Asian J Trop Med Public Health. 2011; 42(6): 1521-30.

[3] Pollution Control Department. Cadmium contamination in Mae Tao Creek, Mae Sot District, Tak Province. Bangkok, Thailand: Ministry of Natural Resources and Environment. 2004.

[4] Sriprachote A, Kanyawongha P, Ochiai K, Matoh T. Current situation of cadmium-polluted paddy soil, rice and soybean in the Mae Sot District, Tak Province, Thailand. Soil Sci Plant Nutr. 2012; 58(3): 349-59. doi.org/10.1080/00380768.2012.686435.

[5] La-Up A, Wiwatanadate P, Pruenglampoo S, Uthaikhup S. Recommended rice intake levels based on average daily dose and urinary excretion of cadmium in a cadmium-contaminated area of northwestern Thailand. Toxicol Res. 2017; 33(4): 291-7. doi: 10.5487/TR.2017.33.4.291.

[6] Huff J, Lunn RM, Waalkes MP, Tomatis L, Infante PF. Cadmium-induced cancers in animals and in humans. Int J Occup Environ Health. 2007; 13(2): 202-12. doi: 10.1179/oeh.2007.13.2.202.

[7] Agency for Toxic Substances and Disease Registry. Toxicological profile for Cadmium. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service. 2012. Available from: https://wwwn.cdc.gov/TSP/ToxProfiles/ToxProfiles.aspx

[8] Järup L, Rogenfelt A, Elinder CG, *et al*. Biological half-time of cadmium in the blood of workers after cessation of exposure. Scand J Work Environ Health. 1983; 9(4): 327-31. doi: 10.5271/sjweh.2404.

[9] Vacchi-Suzzi C, Kruse D, Harrington J, Levine K, Meliker JR. Is urinary cadmium a biomarker of long-term exposure in humans? A Review. Curr Environ Heal reports. 2016; 3(4): 450-8. doi: 10.1007/s40572-016-0107-y.

[10] Kayaalti Z, Odabaşi M, Söylemezoğlu T. Genotype and allele frequencies of divalent metal transporter 1 polymorphism in Turkish population. Mol Biol Rep. 2011; 38(4): 2679-84. doi: 10.1007/s11033-010-0410-x.

[11] Kayaaltı Z, Akyüzlü DK, Söylemezoğlu T. Evaluation of the effect of divalent metal transporter 1 gene polymorphism on blood iron, lead and cadmium levels. Environ Res. 2015; 137: 8-13. doi: 10.1016/j.envres.2014.11.008.

[12] Klaassen CD, Liu J, Diwan BA. Metallothionein protection of cadmium toxicity. Toxicol Appl Pharmacol. 2009; 238(3): 215–20. doi: 10.1016/j.taap.2009.03.026.

[13] Nordberg M, Nordberg GF. Metallothionein and cadmium toxicology-historical review and commentary. Biomolecules. 2022; 12(3). doi: 10.3390/biom1203 0360.

[14] Lane DJR, Merlot AM, Huang ML-H, *et al*. Cellular iron uptake, trafficking and metabolism: Key molecules and mechanisms and their roles in disease. Biochim Biophys Acta. 2015; 1853(5): 1130-44. doi: 10.1016/j.bbamcr.2015.01.021.

[15] Thévenod F, Lee W-K, Garrick MD. Iron and cadmium entry into renal mitochondria: physiological and toxicological implications. Front cell Dev Biol. 2020; 8: 848. doi: 10.3389/fcell.2020.00848.

[16] Menon AV, Chang J, Kim J. Mechanisms of divalent metal toxicity in affective disorders. Toxicology. 2016; 339: 58-72. doi: 10.1016/j.tox.2015.11.001.

[17] Tallkvist J, Bowlus CL, Lönnerdal B. *DMT1* gene expression and cadmium absorption in human absorptive enterocytes. Toxicol Lett. 2001; 122(2): 171-7. doi: 10.1016/s0378-4274(01)00363-0.

[18] Jakubowski M, Trzcinka-Ochocka M. Biological monitoring of exposure: trends and key developments. J Occup Health. 2005; 47(1): 22-48. doi: 10.1539/joh.47.22.

[19] Sikaphana S, Boonthum R, Leudang S, Parnmen S. Associations between urinary excretion of cadmium with alpha-1 microglobulin and microalbuminuria: a cross-sectional study in northwestern Thai population. J Assoc Med Sci. 2021; 54(2): 35-41.

Retrieved from https://he01.tci-thaijo.org/index.php/bulletinAMS/article/view/247385

[20] Chong WS, Kwan PC, Chan LY, *et al*. Expression of divalent metal transporter 1 (*DMT1*) isoforms in first trimester human placenta and embryonic tissues. Hum Reprod. 2005; 20(12): 3532-8. doi: 10.1093/humrep/dei246.

[21] Ingrassia R, Garavaglia B, Memo M. *DMT1* expression and iron levels at the crossroads between aging and neurodegeneration. Front Neurosci. 2019; 13:575. doi: 10.3389/fnins.2019.00575.

[22] He Q, Du T, Yu X, *et al*. *DMT1* polymorphism and risk of Parkinson's disease. Neurosci Lett. 2011; 501(3): 128-31. doi: 10.1016/j.neulet.2011.07.001.

[23] Garrick MD, Singleton ST, Vargas F, *et al*. DMT1: which metals does it transport? Biol Res. 2006; 39(1): 79-85. doi: 10.4067/s0716-97602006000100009.

[24] Peana M, Pelucelli A, Chasapis CT, *et al*. Biological effects of human exposure to environmental cadmium. Biomolecules. 2022; 13(1): 36. doi: 10.3390/biom13010036.

[25] Przybyłkowski A, Gromadzka G, Członkowska A. Polymorphisms of metal transporter genes *DMT1* and *ATP7A* in Wilson's disease. J Trace Elem Med Biol. 2014; 28(1): 8-12. doi: 10.1016/j.jtemb.2013.08.002.

[26] Kim H-K, Lee H, Kim H-J. A polymorphism in *DMT1* is associated with lead-related hypertensive status. Mol Cell Toxicol. 2013; 9(4): 415-20. doi: 10.1007/s13273-013-0051-y.

[27] Tolone C, Bellini G, Punzo F, *et al*. The *DMT1* IVS4+44C>A polymorphism and the risk of iron deficiency anemia in children with celiac disease. PLoS One. 2017; 12(10): e0185822.

[28] Wysokinski D, Zaras M, Dorecka M, *et al*. An association between environmental factors and the IVS4+44C>A polymorphism of the *DMT1* gene in age-related macular degeneration. Graefe's Arch Clin Exp Ophthalmol. 2012; 250(7): 1057-65. doi: 10.1007/s00417-012-1966-z.

Journal of Associated Medical Sciences 2024; 57(3): 230-236

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

The study on verbal fluency in older adults in Nonthaburi Province

Isara Suttichujit Nicha Kripanan* **Somjit Ruamsuk**

Sirindhorn National Medical Rehabilitation Institute, Nonthaburi Province, Thailand.

ARTICLE INFO

Article history:
Received 5 June 2024
Accepted as revised 20 August 2024
Available online 23 August 2024

Keywords:
Aphasia, verbal fluency, Thai older, Nonthaburi, Thai language.

ABSTRACT

Background: Verbal fluency assesses cognitive function in dementia and word retrieval in aphasia. However, the lack of data on verbal fluency among healthy older Thai individuals hinders comparisons with patient results. This is particularly concerning because older individuals are at higher risk for stroke and dementia. As Thailand has transitioned into an aging society, addressing this gap in data is essential.

Objective: This study aims to examine verbal fluency in Thai individuals aged 60 and above, focusing on factors such as age, gender, and education level affecting word count. The goal is to update Thai data, provide more assessment options, and enhance understanding of related factors. This will improve the accuracy of result interpretation and inform treatment planning.

Materials and methods: The study involved 147 healthy Thai individuals aged 60-89 living in Nonthaburi, Thailand. Participants must pass the Thai Mental State Examination (TMSE), visual screening by naming a picture and reading text, hearing screening through finger rubbing, oral reading (Noo Jaew Passage), and oral motor examinations to include only healthy participants without speech impairment. The speech-language pathologist, as an examiner, asks participants to generate words within a minute for each category (randomly assigned: animal, object, and food). Each intelligible and correct word in their category was scored. Researchers transcribed the recordings and counted the words produced.

Results: The average age of the total participants was 70.59 (SD=7.25) years, with 110 women (75%). The average words are: 19.35 (SD=5.25) animals, 20.18 (SD=6.70) objects, and 15.02 (SD=4.56) foods. Participants aged 60-69 exhibited the highest verbal fluency for animal, object, and food categories at 20.63 (SD=5.02), 21.86 (SD=6.19), and 16.35 (SD=4.71), respectively.

Conclusion: The study investigates verbal fluency in older Thai individuals in Nonthaburi, focusing on animal, object, and food categories. Results show that fluency declines with age. While higher education enhanced performance in the animal and object categories, it did not affect the food category. Gender significantly impacted the food category, with females performing better, possibly due to cultural roles. The data can be helpful in clinical assessments and future research on cognitive aging in Thai populations.

Introduction

Thailand has become an "Aging Society," with over 10% of its population aged 60 years and older. Of the 66 million people, 13 million (19%) are in this age group.[1] In Nonthaburi, 20% of the population is 60 and above. Older individuals are at higher risk for stroke and dementia.[2] The prevalence of stroke in Thai adults aged 45 and above is 1.88%, with a mean onset age of 65.[3] The prevalence of dementia in Thailand varied from 2.35%

* **Corresponding contributor.**
Author's Address: *Sirindhorn National Medical Rehabilitation Institute, Nonthaburi Province, Thailand.*
E-mail address: slp.nicha@gmail.com
doi: 10.12982JAMS.2024.065
E-ISSN: 2539-6056

among adults aged 45 and above to 3.4-9.88% in those aged 60 and above.[4-6] Women had a greater dementia rate than men, and it increased with age, doubling about every five years until 85 years of age.[7]

Verbal fluency tests effectively assess cognitive impairments related to both stroke and dementia. These tests evaluate word retrieval ability in the context of aphasia and cognitive function within one minute and categorize them into semantic fluency (listing words within a category) and phonemic fluency (naming words starting with a given letter).[8] Examples of tests include verbal fluency; Western Aphasia Battery (WAB),[9] Boston Diagnostic Aphasia Examination (BDAE),[10] and Montreal Cognitive Assessment (MoCA).[11]

Phonemic fluency requires literacy skills, while semantic fluency is generally more effortless but can be influenced by educational levels.[12] Previous studies on Thai adults' verbal fluency in the animal category lack comparisons between early and late older populations.[12, 13] Rather than relying solely on the animal category, broadening categories will improve assessment options. This study compares verbal fluency data among Nonthaburi older people aged 60-69, 70-79, and 80-89, focusing on animal, object, and food categories. Researchers selected these categories to provide more options for repeated assessments and prevent memorization. The study updates Thai verbal fluency data for Nonthaburi, identifies factors affecting word counts, such as age, gender, and education, and guides the selection of assessment categories. These insights will help explain results and plan appropriate training.

Materials and methods

The study population included healthy, older Thai individuals living in Nonthaburi, Thailand. The sample size calculation was performed using the finite sample proportion in the n4Studies application,[14] utilizing the 2019 data from the Department of Older Persons, which indicated 239,410 older individuals in Nonthaburi.[15] The sample was stratified into age cohorts 60-69, 70-79, and 80-89, mirroring the demographic distribution.[16] The study included 147 people, with 84 participants aged 60-69, 45 participants aged 70-79, and 18 participants aged 80-89. The participants were chosen via purposive sampling, selecting healthy clients or caregivers receiving services at Sirindhorn National Medical Rehabilitation Institute and members from the Center for Older People's Quality of Life in Nonthaburi. The study was conducted through in-person contacts at these places, with participant recruitment starting in October 2020 and ending in July 2021. The Sirindhorn National Medical Rehabilitation Institute Human Ethics Committee, Nonthaburi Province, Thailand, accepted this study (63017).

Inclusion criteria: healthy Thai individuals aged 60-89 years, primarily using the Central Thai dialect, with no history of cerebral or neurological diseases, no severe visual or auditory impairments, and the ability to perform daily tasks independently. Exclusion criteria: TMSE score ≤23,[17] inability to read or repeat the passage intelligibly, and failure in the oral motor examination.

Participants had to pass dementia screening using the Thai Mental State Examination (TMSE).[17] Vision and hearing were initially screened by questioning participants about any existing vision or hearing problems. Participants were observed to ensure they could see images and read the text correctly for visual screening during the TMSE and Noo Jaew Passage assessments,[18] and hearing was evaluated using a finger-rubbing test.[19] Oral reading of the Noo Jaew Passage assessed intelligibility, and illiterate participants repeated the text after an examiner. Oral motor examinations evaluated the functioning of the speech organs. These screenings included only healthy participants with clear and intelligible speech. All participants provided written informed consent. Researchers and speech-language pathologists (SLPs) conducted the screenings and assessments.

Participants had one 1 minute per category (animal, object, and food) to generate as many words as possible without cues. Categories were randomly allocated, and no examples were given to prevent bias. The researchers instructed participants that "food" is anything edible, typically complete dishes. Each intelligible and correct word in its category received one score. Repeated words, incorrect target language, intrusions (words outside the category), and non-specific words (e.g., "cooked food," "fried food," or "fish" without specifying boiled/fried/grilled) are not credited. Variations (starting or ending with the same word) are credited up to two scores. For example, if "noodles" and "fish noodles" are mentioned, only "fish noodles" will be credited. Similarly, "fried chicken," "fried fish," and "fried meat" will receive only two scores. Synonyms like "dog" and "canine" (or "สุนัข" and "หมา" in Thai) or "TV" and "television" (or "ทีวี" and "โทรทัศน์" in Thai) will be credited once.

Researchers transcribed the recordings and counted the words produced. Percentage, mean, SD, and 95% confidence interval were used to examine demographic data. One-way ANOVA compared means across the three age groups; independent t-tests compared means between genders; and multiple regression analysis examined how age, TMSE score, gender, and education affected word count in each category. SPSS 29.0 was used for all analyses.

Results

The study included 147 healthy older people in Nonthaburi, 110 women and 37 men, 70 from the Center for Older People's Quality of Life, and 77 from Sirindhorn National Medical Rehabilitation Institute. The mean TMSE score for the entire sample was 28.27 (SD=1.52). One-way ANOVA revealed significant differences in TMSE scores among the three age groups (p=0.017[a]), as presented in Table 1. An independent t-test indicated that the 60-69 age group exhibited significantly higher TMSE scores than the 80-89 age group (p=0.024[b]).

Table 1. Demographic data and average words within 1 minute of each age group.

	60-69 (N=84)	70-79 (N=45)	80-89 (N=18)	Total (N=147)	*p* value[a]
Age (years)	65.07±2.67 (64.49, 65.65)	75.49±1.51 (75.03, 75.94)	84.11±1.93 (83.22, 85.00)	70.59±7.25 (69.42, 71.77)	<0.001**
Female (%)	67 (80%)	33 (73%)	56 (50%)	110 (75%)	0.096
TMSE	28.54±1.41 (28.23, 28.84)	28.07±1.50 (27.62, 28.52)	27.50±1.79 (26.67, 28.33)	28.27±1.52 (28.02, 28.51)	0.017*
Education (years)	13.52±4.45 (12.56, 14.49)	13.56±4.43 (12.22, 14.89)	12.33±5.72 (9.69, 14.98)	13.39±4.60 (12.64, 14.13)	0.586
Animal	20.63±5.02 (19.54, 21.72)	18.09±5.08 (16.56, 19.62)	16.56±5.15 (14.18, 18.93)	19.35±5.25 (18.50, 20.20)	0.001**
Object	21.86±6.19 (20.51, 23.20)	19.42±6.85 (17.37, 21.48)	14.28±4.91 (12.01, 16.55)	20.18±6.70 (19.10, 21.27)	<0.001**
Food	16.35±4.71 (15.32, 17.37)	13.71±3.21 (12.75, 14.68)	12.11±4.60 (9.99, 14.24)	15.02±4.56 (14.28, 15.76)	<0.001**

*Note: values are reported as mean±SD (95% CI lower, 95% CI upper), [a]tested via One-way ANOVA, *p<0.05, **p<0.01, considered statistically significant.*

The average years of education for the total sample was 13.39 (SD=4.60), with no significant differences observed among the three age groups (p=0.586[a]). The educational distribution of the participants is as follows: one participant is illiterate with non-formal education; 17 participants have completed 4-6 years of primary education; 38 participants have completed 7-12 years of secondary education; and 91 participants have received 13-21 years of tertiary education. Within the tertiary education group, 12 hold diplomas, 50 hold bachelor's degrees, 27 hold master's degrees, and 2 hold doctorate degrees.

Verbal fluency averaged 19.35 (SD=5.25) words for animals, 20.18 (SD=6.70) for objects, and 15.02 (SD=4.56) for food. There were significant differences across all three categories among age groups (p<0.001[a]). In Table 2, an independent t-test showed that the 60-69 age group produced more animal and food words than the 70-79 and 80-89 age groups, whereas the 80-89 age group created fewer object words.

Table 2. Comparative analysis of age, TMSE, education, and verbal fluency between different age groups.

	60-69 vs 70-79[b]	60-69 vs 80-89[b]	70-79 vs 80-89[b]
Age (years)	-10.41 (-11.44, -9.39) p<0.001**	-19.04 (-20.48, -17.6) p<0.001**	-8.63 (-10.17, -7.08) p<0.001**
TMSE	0.47 (-0.2, 1.13) p=0.269	1.04 (0.1, 1.97) p=0.024*	0.57 (-0.44, 1.57) p=0.521
Education (years)	-0.03 (-2.09, 2.03) p=1.000	1.19 (-1.71, 4.09) p=0.965	1.22 (-1.89, 4.34) p=1.000
Animal	2.54 (0.28, 4.8) p=0.022*	4.08 (0.9, 7.26) p=0.007**	1.53 (-1.88, 4.95) p=0.836
Object	2.43 (-0.37, 5.24) p=0.112	7.58 (3.64, 11.52) p<0.001**	5.14 (0.91, 9.38) p=0.011*
Food	2.63 (0.71, 4.56) p=0.003**	4.23 (1.53, 6.94) p<0.001**	1.6 (-1.3, 4.5) p=0.551

*Note: values are reported as mean difference (95% CI lower, 95% CI upper),[b]tested via independent T-test, *p< 0.05, **p<0.01, considered statistically significant*

A multiple linear regression analysis examining the factors influencing verbal fluency determined that increased age was associated with fewer words across all three categories. Conversely, higher TMSE scores were correlated with an increased number of words in each category. Additionally, more years of education were linked to an increased number of words in the animal and object categories; however, education level did not significantly affect the number of words in the food category (coefficient=0.11, p=0.158), as detailed in Table 3.

Table 3. Regression coefficients and significance levels of each verbal fluency.

	Animal	Object	Food
Age	-0.17 (*p*=0.003**)	-0.25 (*p*<0.001**)	-0.18 (*p*<0.001**)
TMSE	0.69 (*p*=0.020*)	1.17 (*p*=0.001**)	0.73 (*p*=0.004**)
Education	0.23 (*p*=0.017*)	0.47 (*p*<0.001**)	0.11 (*p*=0.158)
Gender	1.13 (*p*=0.224)	1.67 (*p*=0.113)	1.89 (*p*=0.016*)

*Note: values are reported as regression coefficients with p values, *p<0.05, **p<0.01, considered statistically significant.*

Gender did not significantly impact the number of words in the animal and object categories but did influence the food category (coefficient=1.89, *p*=0.016). When comparing genders, only the average age and the number of words in the food category demonstrated statistically significant differences (*p*=0.043 and *p*=0.008, respectively), as illustrated in Table 4.

Table 4. Comparison of characteristics by gender.

	Female (N=110)	Male (N=37)	*p* value[b]
Age	69.89±6.76 (68.63, 71.16)	72.67±8.31 (70.00, 75.35)	0.043*
TMSE	28.27±1.52 (27.99, 28.56)	28.24±1.53 (27.75, 28.74)	0.919
Education	13.08±4.78 (12.19, 13.97)	14.30±3.93 (13.03, 15.56)	0.165
Animal	19.69±5.39 (18.68, 20.70)	18.35±4.75 (16.82, 19.88)	0.181
Object	20.65±6.95 (19.35, 21.94)	18.81±5.75 (16.96, 20.66)	0.150
Food	15.59±4.60 (14.73, 16.45)	13.32±4.03 (12.03, 14.62)	0.008**

*Note: [b]tested via independent T-test, *p<0.05, **p<0.01, considered statistically significant.*

Discussion

This study is the first to investigate verbal fluency in older Thai individuals with the highest average educational attainment of 13.39 (SD=4.60) years, focusing on vocabulary in the Thai language's animal, object, and food categories. A literature review revealed no prior studies on word counts in the categories of objects or food, although related studies in categories such as furniture, clothing, vegetable, and fruit exist.[32,33] However, these are not directly comparable due to the broader definitions of object and food. Therefore, the comparison here is limited to the animal category, as shown in Table 5.

Table 5. Comparison of study data across different languages.

Study	Language	N	Age	Education	Animal
Teerapong[9]	Thai	30	52.5±6.8	10.6±5.7	17.7±3.4
Muangpaisan *et al.*[13]		30	63.7±7.3	6.7±3.2	17.3±6.4
Charernboon[12]		61	64.7±6.7	10.4±5.0	19.4±5.0
This study		147	70.6±7.3	13.4±4.6	19.4±5.3
Tombaugh *et al.*[20]	English	259	60-79	9-21	17.1±4.3
Brickman *et al.*[21]		55	61-82	13.0±3.2	17.4
Knight *et al.*[22]		272	73.7±5.8	0-16	18.8±4.7
Kempler *et al.*[23]	Vietnamese	60	71.6±5.8	8.6±4.2	17.3±5.2
Chan & Poon[24]	Chinese	156	71.4±5.4	7.5±6.1	14.5±4.2
Ostrosky-Solis *et al.*[25]	Spanish	181	60-90	7-12	17.9±4.7
Mathuranath *et al.*[26]	Malayalam	153	66.9±5.6	7.2±6.1	7.8±3.6
Kavé[27]	Hebrew	166	51-85	8-24	19.0±5.0
Pekkala *et al.*[28]	Finnish	30	66.7±5.5	9.7±3.3	18.9±4.7
Ryu *et al.*[29]	Korean	3025	71.7±6.7	7.4±5.2	12.9±4.2
Cavaco *et al.*[30]	Portuguese	487	60-89	0-20	14.8±5.1
Vogel *et al.*[31]	Danish	100	70.9±6.4	11.9±2.6	21.3±4.8

Note: values are reported as mean±SD.

Verbal fluency scores vary across languages and are influenced by age, education, and cultural context.[32, 33] In Thai studies, Teerapong reported 17.7 (SD=3.4) words,[9] Muangpaisan *et al*. found 17.3 (SD=6.4) words,[13] and Charernboon reported 19.4 (SD=5.0) words.[12] This study, with participants averaging 70.6 (SD=7.3) years of age and 13.4 (SD=4.6) years of education, found an average of 19.4 (SD=5.3) words, consistent with Charernboon's findings. This study's average TMSE score was 28.3 (SD=1.5), similar to the previous study by Muangpaisan *et al*., which reported a score of 28.1 (SD=1.8).[13]

Comparing these results with studies in other languages reveals diverse outcomes. English studies by Tombaugh et al. reported 17.1 (SD=4.3) words,[20] Brickman *et al.* found 17.4 words,[21] and Knight et al. reported 18.8 (SD=4.7) words.[22] Chinese studies by Chan and Poon reported 14.5 (SD=4.2) words,[24] and Spanish studies by Ostrosky-Solis *et al*. found 17.9 (SD=4.7) words.[25] Danish studies by Vogel *et al*. reported the highest average of 21.3 (SD=4.8) words.[31] The study conducted by Mathuranath *et al*. in the Malayalam language reported the lowest average of 7.8 (SD=3.6) words.[26]

These differences emphasize that verbal fluency scores are language-specific and cannot be directly compared across languages. This highlights the need for culturally and linguistically appropriate normative data to assess verbal fluency accurately.[32,33] Factors such as word length and cultural differences in language use affect performance.[32] For example, Kempler *et al.* found that Spanish speakers had lower verbal fluency scores than Vietnamese speakers, attributing this to the longer word length in Spanish.[23] Pekkala et al. noted that although Finnish words are generally longer than English words, there was no significant difference in verbal fluency scores between Finnish and English speakers.[28] In Thai, the nature of the language adds another layer of complexity. For example, the use of word-guides presents unique challenges and opportunities in verbal fluency tasks: words starting with "ปลา" (pla/fish): "ปลาทอด" (pla-tod/fried fish), "ปลานึ่ง" (pla-nueng/steamed fish), "ปลาต้ม" (pla-tom/boiled fish); words starting with "นก" (nok/bird): "นกยูง" (nok-yung/peacock), "นกพิราบ" (nok-pirab/pigeon), "นกอินทรี" (nok-insee/eagle)**;** Words ending with "ย่าง" (yang/grill): "ไก่ย่าง" (gai-yang/grilled chicken), "หมูย่าง" (moo-yang/grilled pork), "เนื้อย่าง" (nuea-yang/grilled beef). In this study, researchers defined precise criteria for variations of a term that share the same first or last word. These variations were considered valid only if they did not exceed two variations. Other common rhyme word groups include "หมู หมา กา ไก่" (moo-ma-ga-gai/pig-dog-crow-chicken) and "ช้าง ม้า วัว ควาย" (chang-ma-wua-khwai/elephant-horse-cow-buffalo). If no variations or intrusions exist, each word is counted as one point.

Impact of Education and Age

In agreement with most normative studies, performance on the semantic fluency test declined with age[34-36] and increased with education.[20,34,35] Education affects verbal fluency (animal) in Chinese, English, Greek, Hebrew, Malayalam, Portuguese, and Spanish.[20,23-27, 30,33,36] When using these normative data clinically, educational data should be considered. These findings confirm literacy's role in semantic fluency.[37] However, the association pattern is specific to the "animal" category. Literacy may affect other categories more or less, depending on their ecological or cultural relevance.[37] Our study found no significant correlation between education level and the food category.

Gender Differences

The gender effect is inconsistent across studies. No gender effect in the animal category was found in Dutch, English, Hebrew, and Malayalam.[20,26,27,33,38] Most studies in the literature are consistent with these negative findings.[20,34,35] However, a gender impact was observed in Chinese and Greek.[24,33,36,] presumably because women are less educated.[24] Some studies report significant gender effects in other categories similar to our findings.[36,38,39] For example, while men may perform better in naming tools, women name fruits better.[39] Our study found significant gender-related differences only in the food category, not animal or object categories. This may be because Thai women typically prepare food.[40] Cultural factors, including gender-specific responsibilities and educational opportunities, affect language fluency.

Conclusion

This study is one of the first to investigate verbal fluency in literate older Thai individuals in Nonthaburi and focuses on vocabulary in animal, object, and food categories. The results provide critical normative data for semantic fluency in the Thai language, highlighting the influence of age, education, and gender on verbal fluency performance.

Verbal fluency declined with age across all categories. The 60-69 age group exhibited significantly higher verbal fluency scores than the 70-79 and 80-89 age groups. Higher levels of education were associated with better performance in the animal and object categories. However, education did not significantly impact performance in the food category. Gender had a significant impact only on the food category, where females performed better, likely due to cultural roles in food preparation. These findings in the animal category align with previous research across various languages, emphasizing the importance of considering demographic factors when evaluating cognitive health in the elderly.

The normative data provided by this study can serve as a valuable reference for clinical assessments and future research on cognitive aging in Thai populations, particularly those with complete secondary education. However, the study sample included fewer participants aged 80 and above. Most participants were from Nonthaburi province, similar to Bangkok but may only partially represent Thailand's older population. This shortcoming may limit the findings' applicability to other sociocultural locations. Age, education, and culture should be considered when assessing verbal fluency.

Conflict of interest

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

Acknowledgements

This research was supported by the Pay for performance (P4P) of Sirindhorn National Medical Rehabilitation Institute in 2022. 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] Senanarong V, Harnphadungkit K, Poungvarin N, Vannasaeng S, Chongwisal S, Chakorn T, *et al*. The Dementia and Disability Project in Thai Elderly: rational design methodology and early results. BMC Neurol. 2013; 13: 3. doi: 10.1186/1471-2377-13-3.

[2] Foundation of Thai Gerontology Research and Development Institute (TGRI). Situation of the Thai Older Persons 2021. Nakhon Pathom: Institute for Population and Social Research, Mahidol University; 2021.

[3] Suwanwela NC. Stroke epidemiology in Thailand. J Stroke. 2014; 16(1): 1-7. doi: 10.5853/jos.2014.16.1.1.

[4] 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.

[5] 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.

[6] 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.

[7] Nichols E, Steinmetz JD, Vollset SE, Fukutaki K, Chalek J, Abd-Allah F, et al. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health. 2022; 7(2): e105-e125.doi: 10.1016/S2468-2667(21)00249-8.

[8] 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.

[9] 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.

[10] Gandour J, Dardarananda R, Buckingham H, Viriyavejakul A. A Thai adaptation of the Boston Diagnostic Aphasia Examination. Crossroads. 1986; 2(3): 1-39.

[11] Julayanont P, Tangwongchai S, Hemrungrojn S, Tunvirachaisakul C, Phanthumchinda K, Hongsawat J, *et al*. The Montreal cognitive assessment—basic: A screening tool for mild cognitive impairment in illiterate and low-educated elderly adults. J Am Geriatr Soc. 2015; 63(12): 2550-4. doi: 10.1111/jgs.13820.

[12] 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.

[13] 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. PMID: 26591404.

[14] Ngamjarus C. n4Studies: Sample Size Calculation for an Epidemiological Study on a Smart Device. Siriraj Med J. 2016; 68(3): 160-70.

[15] Statistics of Elderly Person in Thailand 2019 [Internet]. 2019. Available from: www.dop.go.th/download/knowledge/th1580099938-275_1.pdf.

[16] 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.

[17] Train the Brain Forum Committee. Thai mental state examination (TMSE). Siriraj Hosp Gaz. 1993; 45: 661-74.

[18] Sindermsuk D. The survey of speech defects among prathom 4 students in Mitsampan School Group [Thesis]. Nakhon Pathom: Mahidol University; 1986.

[19] Torres-Russotto D, Landau WM, Harding GW, Bohne BA, Sun K, Sinatra PM. Calibrated finger rub auditory screening test (CALFRAST). Neurology. 2009; 72(18): 1595-600. doi: 10.1212/WNL.0b013e3181a41280.

[20] 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. doi: 10.1093/arclin/14.2.167.

[21] Brickman AM, Paul RH, Cohen RA, Williams LM, MacGregor KL, Jefferson AL, *et al*. Category and letter verbal fluency across the adult lifespan: relationship to EEG theta power. Arch Clin Neuropsychol. 2005; 20(5): 561-73. doi: 10.1016/j.acn.2004.12.006.

[22] Knight RG, McMahon J, Green TJ, Skeaff CM. Regression equations for predicting scores of persons over 65 on the rey auditory verbal learning test, the mini-mental state examination, the trail making test, and semantic fluency measures. Br J Clin Psychol. 2006; 45(Pt 3): 393-402. doi: 10.1348/014466505x68032.

[23] Kempler D, Teng EL, Dick M, Taussig IM, Davis DS. The effects of age, education, and ethnicity on verbal fluency. J Int Neuropsychol Soc. 1998; 4(6): 531-8. doi: 10.1017/s1355617798466013.

[24] Chan AS, Poon MW. Performance of 7-to 95-year-old individuals in a Chinese version of the category

fluency test. J Int Neuropsychol Soc. 1999; 5(6): 525-33. doi: 10.1017/s135561779956606x.

[25] Ostrosky-Solis F, Gutierrez AL, Flores MR, Ardila A. Same or different? Semantic verbal fluency across Spanish-speakers from different countries. Arch Clin Neuropsychol. 2007; 22(3): 367-77. doi: 10.1016/j.acn.2007.01.011.

[26] Mathuranath PS, George A, Cherian PJ, Alexander A, Sarma SG, Sarma PS. Effects of age, education and gender on verbal fluency. J Clin Exp Neuropsychol. 2003; 25(8): 1057-64. doi: 10.1076/jcen.25.8.1057.16736.

[27] Kavé G. Phonemic fluency, semantic fluency, and difference scores: normative data for adult Hebrew speakers. J Clin Exp Neuropsychol. 2005; 27(6): 690-9. doi: 10.1080/13803390490918499.

[28] Pekkala S, Goral M, Hyun J, Obler LK, Erkinjuntti T, Albert ML. Semantic verbal fluency in two contrasting languages. Clin Linguist Phon. 2009; 23(6): 431-45. doi: 10.1080/02699200902839800.

[29] Ryu SH, Kim KW, Kim S, Park JH, Kim TH, Jeong HG, Kim JL, Moon SW, Bae JN, Yoon JC, Choo IH, Lee DW, Chang SM, Jhoo JH, Kim SK, Cho MJ. Normative study of the category fluency test (CFT) from nationwide data on community-dwelling elderly in Korea. Arch Gerontol Geriatr. 2012; 54(2): 305-9. doi: 10.1016/j.archger.2011.05.010.

[30] Cavaco S, Gonçalves A, Pinto C, Almeida E, Gomes F, Moreira I, Fernandes J, Teixeira-Pinto A. Semantic fluency and phonemic fluency: regression-based norms for the Portuguese population. Arch Clin Neuropsychol. 2013; 28(3): 262-71. doi: 10.1093/arclin/act001.

[31] Vogel A, Stokholm J, Jørgensen K. Normative data for eight verbal fluency measures in older Danish adults. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn. 2020; 27(1): 114-124. doi: 10.1080/13825585.2019.1593935.

[32] 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.

[33] 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.

[34] Benito-Cuadrado MM, Esteba-Castillo S, Böhm P, Cejudo-Bolívar J, Peña-Casanova J. Semantic Verbal Fluency of Animals: A Normative and Predictive Study in a Spanish Population. J Clin Exp Neuropsychol. 2002; 24(8): 1117-22. doi: 10.1076/jcen.24.8.1117.8376.

[35] Gladsjo JA, Schuman CC, Evans JD, Peavy GM, Miller SW, Heaton RK. Norms for Letter and Category Fluency: Demographic Corrections for Age, Education, and Ethnicity. Assessment. 1999; 6(2): 147-78. doi: 10.1177/107319119900600204.

[36] Kosmidis MH, Vlahou CH, Panagiotaki P, Kiosseoglou G. The verbal fluency task in the Greek population: Normative data, and clustering and switching strategies. J Int Neuropsychol Soc. 2004; 10(2): 164-72. doi: 10.1017/S1355617704102014.

[37] da Silva CG, Petersson KM, Faísca L, Ingvar M, Reis A. The Effects of Literacy and Education on the Quantitative and Qualitative Aspects of Semantic Verbal Fluency. J Clin Exp Neuropsychol. 2004; 26(2): 266-77. doi: 10.1076/jcen.26.2.266.28089.

[38] Van Der Elst WIM, Van Boxtel MPJ, Van Breukelen GJP, Jolles J. Normative data for the Animal, Profession, and Letter M Naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex. J Int Neuropsychol Soc. 2006; 12(1): 80-9. doi: 10.1017/S1355617706060115.

[39] Capitani E, Laiacona M, Barbarotto R. Gender Affects Word Retrieval of Certain Categories in Semantic Fluency Tasks. Cortex. 1999;35(2):273-8. doi: 10.1016/s0010-9452(08)70800-1.

[40] Chandrangam K. The relationship between gender and women's self-value in Bangkok metropolis [Thesis]. Bangkok: Chulalongkorn University; 2005.

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.

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