E-ISSN 2228-8082
Volume 77, Number 4, April 2025

PMAC
PRINCE MAHIDOL AWARD CONFERENCE 2025

SMJ

Siriraj Medical Journal

MONTHLY

Otorhinolaryngology

The world-leading biomedical science of Thailand

ORIGINAL ARTICLE REVIEW ARTICLE

Indexed by

Scopus DOAJ EBSCO

https://he02.tci-thaijo.org/index.php/sirirajmedj/index
E-mail: sijournal92@gmail.com

SMJ
Siriraj Medical Journal
The world-leading biomedical science of Thailand

Volume 77 Number 4
April 2025

ORIGINAL ARTICLE

SIRIRAJ MEDICAL JOURNAL

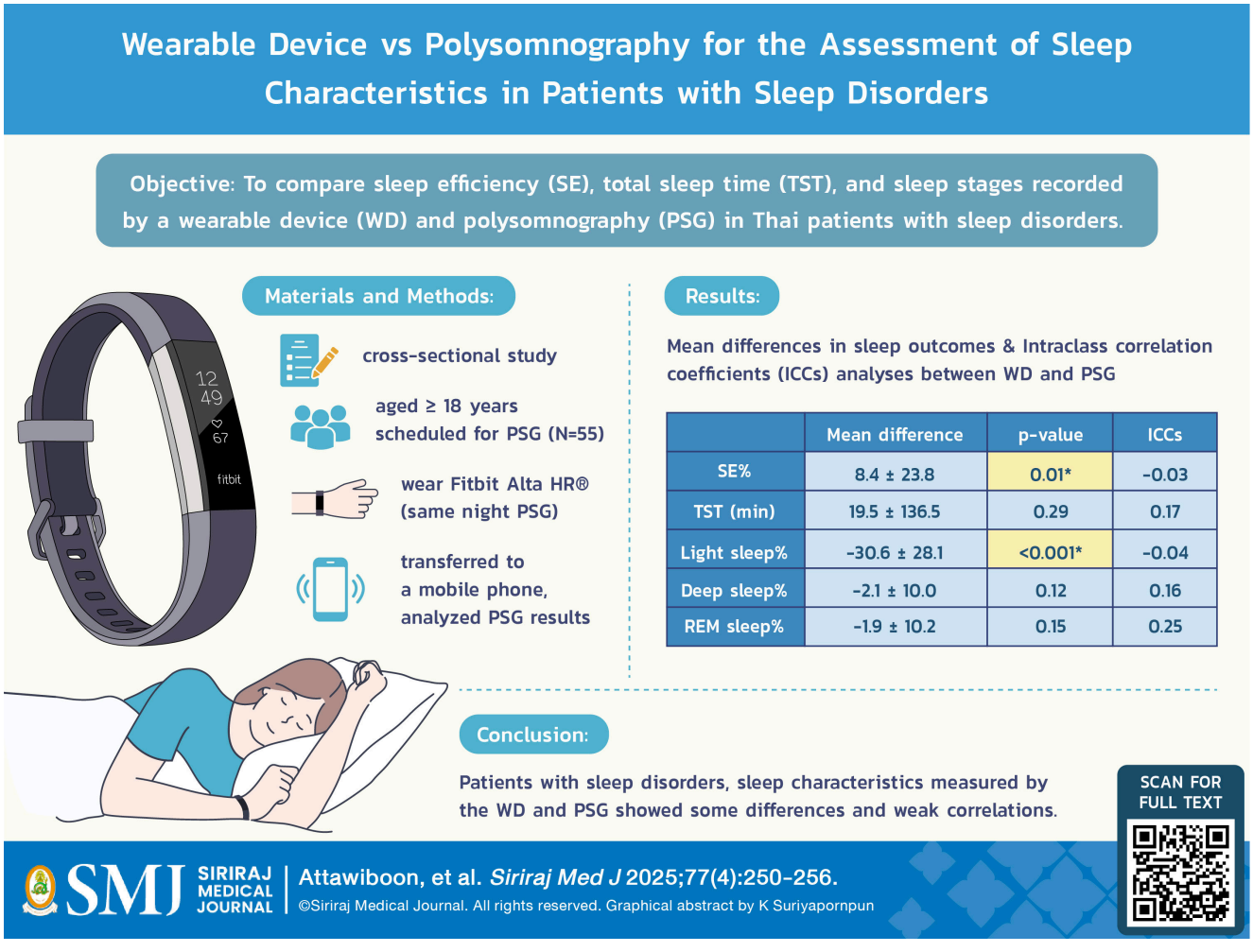
https://he02.tci-thaijo.org/index.php/sirirajmedj/index

Office: His Majesty the King's 80th Birthday Anniversary 5th December 2007 Building (SIMR), 2nd Fl., Room No.207
Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand
E-mail: sijournal92@gmail.com

Wearable Device versus Polysomnography for the Assessment of Sleep Characteristics in Patients with Sleep Disorders

Kankanok Attawiboon, M.D.[1], Wish Banhiran, M.D.[1,*], Phawin Keskool, M.D.[1], Wattanachai Chotinaiwattarakul, M.D.[2], Sarin Rungmanee, M.D.[1]

[1]*Department of Otorhinolaryngology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, 10700 Thailand,* [2]*Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, 10700, Thailand.*



*Corresponding Author: Wish Banhiran
E-mail: wishbanh@gmail.com
Received 4 December 2024 Revised 23 January 2025 Accepted 25 January 2025
ORCID ID:http://orcid.org/0000-0002-4029-6657
https://doi.org/10.33192/smj.v77i4.272562

ABSTRACT

Objective: To compare sleep efficiency (SE), total sleep time (TST), and sleep stages recorded by a wearable device (WD) and polysomnography (PSG) in Thai patients with sleep disorders.

Materials and Methods: Patients aged ≥18 years scheduled for PSG were included in this cross-sectional study. All research subjects completed questionnaires and wore a WD (Fitbit Alta HR®) on the same night they underwent PSG study. The data from the WD were transferred to a mobile phone and analyzed independently of PSG results, which were scored by sleep technicians. Bland-Altman plots and intraclass correlation coefficients (ICCs) were used for the analyses.

Results: Data from 55 patients (33 males, 22 females) were analyzed, with four patients excluded due to data errors. The mean differences between WD and PSG for SE (%) and light sleep were 8.4±23.8 and 43.6±26.4, respectively, both statistically significantly ($p<0.05$). The ICCs for SE and light sleep were -0.03 and -0.04, indicating poor reliability. However, the mean differences for TST, deep sleep, and REM sleep between the two methods were not statistically significant ($p>0.05$), with ICC values of 0.17, 0.16, and 0.25, respectively, all considered poor correlations.

Conclusion: In patients with sleep disorders, sleep characteristics measured by the WD and PSG showed some differences and weak correlations. As technology advances, the accuracy of wearable devices may improve. Further studies are needed to evaluate different devices and populations.

Keywords: Wearable device; Fitbit Alta HR®; sleep disorder; Thai (Siriraj Med J 2025; 77: 250-256)

INTRODUCTION

Sleep is an essential part of life, and maintaining good sleep hygiene can help reduce the risk of serious physical and mental illnesses linked to several sleep disorders affecting people worldwide. While polysomnography (PSG) is currently considered the standard diagnostic method for several sleep disorders, it has some limitations, such as its high cost, labor intensity, and complexity, which makes it less accessible to a large portion of general population.[1-3] In recent years, electronic wearable devices —often in the form of wristbands or smartwatches— have been introduced to help individuals track and record their sleep patterns. These devices measure sleep activity using mechanisms like actigraphy, which syncs with mobile phones or personal computers.[3-6] Their convenience and compatibility with modern lifestyles make them widely accepted as inexpensive and accessible tools for improving quality of life.

By facilitating remote patient monitoring enabling assessment and management of patients' sleep health without necessitating in-person visits, wearable devices can extend the reach of healthcare providers, enabling assessment and management of patients' sleep health without necessitating in-person visits. This approach not only reduces the strain on sleep centers but also enhances access to care for individuals who might otherwise face obstacles due to geographical or socioeconomic factors. Furthermore, the continuous data collection afforded by wearables allows for longitudinal monitoring, offering insights into sleep patterns over time, which is advantageous for both diagnosis and ongoing management.

Previous studies have shown that wristband-like wearable devices may be used to evaluate sleep, yielding varied results in young, healthy adults and some children with sleep disorders.[7-13] Much of the literature suggests that these devices may serve as alternative tools for sleep evaluation, demonstrating high sensitivity, but low specificity for distinguishing different sleep–wake stages.[7] However, studies comparing the performance of wearable devices with standard PSG in patients with sleep disorders remain limited, and none have been conducted on Thai populations. Therefore, the purpose of this study was to evaluate the sleep characteristics measured by a wearable device, namely the Fitbit Alta HR®, a popular and inexpensive wristband, and compare them with PSG measurements in Thai patients with sleep disorders.

MATERIALS AND METHODS

Study design

This cross-sectional study was conducted on Thai patients with sleep disorders at the snoring clinic, Department of Otorhinolaryngology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, between June 2019 and June 2020. It was approved by the Siriraj Institutional Review Board (SiRB), COA no. Si 451/2019. Written consent was obtained from all participants before enrollment.

Subjects

The inclusion criteria for the study were patients aged ≥ 18 years who were being treated at the snoring clinic and scheduled for diagnostic PSG. The exclusion criteria included patients with severe or unstable medical problems, such as recent myocardial infarction, stroke, epilepsy, neuromuscular disorders, Parkinsonism, schizophrenia, and sleep-related movement disorders. The final analysis included data from 55 patients (33 males, 22 females), aged 25 to 78 years.

Interventions

All participants completed pre-treatment questionnaires and wore the designated wearable device while undergoing routine diagnostic PSG throughout the night. As one sleep technician scored the PSG results, the data simultaneously recorded by the wearable device were transferred to a mobile phone. The data were analyzed by a research assistant blinded to the PSG results and by an independent assessor not involved in the clinical trial.

Wearable device

The wearable device used in this study was the Fitbit Alta HR® (Fitbit Inc., San Francisco, CA, USA) designed to track various personal metrics such as heart rate, motion, and sleep activity (Fig 1). Recording began automatically once the device was worn and stopped upon removal. Examples of the extracted data are shown in Fig 2.

Polysomnography

Overnight technician-attended PSG (SOMNOmedics, SOMNO HD PSG, DOMINO 3.0.0.3; Randersacker, Germany) was routinely performed at the sleep center in Siriraj Hospital. The recording channels included electroencephalogram (EEG), electro-oculogram (EOG), electromyogram (EMG), electrocardiogram (ECG), nasal pressure transducer, airflow thermistor, respiratory effort measurement, body position sensor, pulse oximetry, and real-time video recordings. The recording began when patients indicated they were ready to sleep (lights off) and concluded when they woke up (lights on). All PSG parameters were manually scored by well-trained sleep technologists and reviewed by certified sleep specialists.

Sleep-related parameters

The outcomes of this study included several important sleep-related parameters defined as follows: Sleep efficiency (SE), expressed as a percentage and calculated by dividing the total sleep time (TST) by the total time in bed (TIB) or total recording time (TRT). TIB or TRT refers to the time-period from light-offs to lights-on (from the beginning to the end of the recording). Light sleep, deep sleep, and REM sleep are also expressed as percentages. Light sleep is calculated from the time spent in sleep stages N1 and N2, deep sleep from stage N3, and REM sleep from the time spent in REM, all divided by the TST. In general, good sleep quality in a young adult population is characterized by a TST of ≥7 hours, SE of 80%, light sleep constituting 50%–55%, deep sleep of 20%, and REM sleep making up 20%–25%.

Statistical analysis

Categorical data are presented as numbers and

Fig 1. Model of Fitbit Alta HR® used in this study.

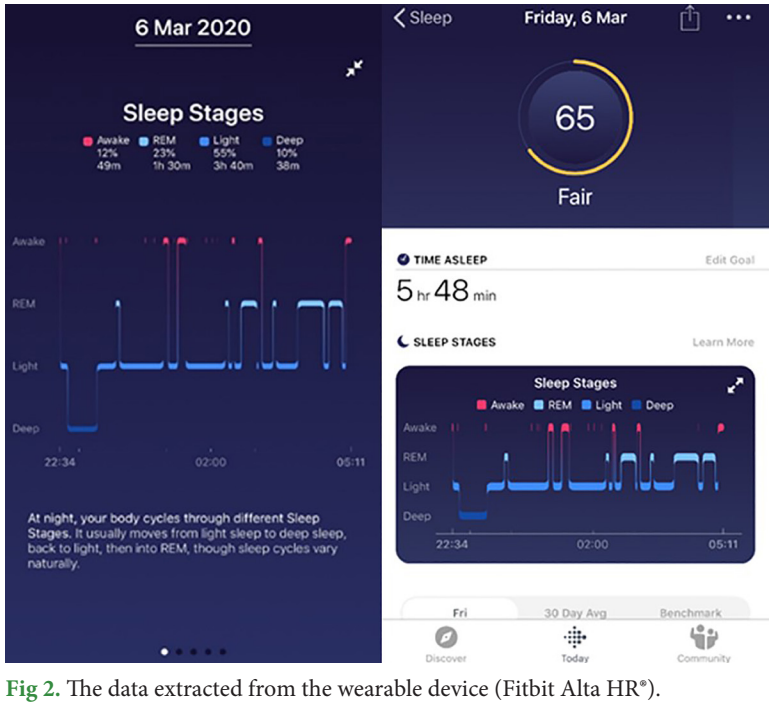


Fig 2. The data extracted from the wearable device (Fitbit Alta HR®).

percentages, while continuous data are expressed as means ± standard deviations (SD) or medians. Comparisons of SE, TST, and sleep stages between the Fitbit Alta HR® and PSG were analyzed using paired t-tests. Intraclass correlation coefficients (ICCs) with a two-way random-effects model with absolute agreement and single measures, and Bland-Altman plots to depict any systematic bias and identify outliers were used to evaluate the agreement between the two assessments. An ICC of <0.4 was considered indicative of poor reliability, while an ICC of 0.4–0.74 indicated a moderate level of reliability, and an ICC of ≥0.75 indicated excellent reliability.[14] All statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS)® software, version 18.0, with significance set at $p < 0.05$ for two-tailed tests.

RESULTS

Initially, 59 patients were enrolled in the study, but 4 patients were excluded due to technical issues (data recording errors). Further details about the patients' characteristics are provided in Table 1. The primary outcomes, mean differences, and agreement (ICC) of SE between the wearable device and PSG are shown in Table 2. The secondary outcomes, including mean differences, and agreement (ICC) of TST, light sleep, deep sleep, and REM sleep are presented in Tables 2 and Table 3. The Bland–Altman plots comparing SE between the wearable device and PSG are displayed in Fig 3, showing the mean differences ± standard deviations (SD) and lower and upper limits of agreement as 8.4 ± 23.8, -38.2, and 55, respectively. It reveals a mean bias, indicating that Fitbit's SE measurements systematically

TABLE 1. Baseline characteristics of study participants.

Characteristic	Data
Age, year	50.4 ± 13.2
BMI, kg/m²	24.9 ± 4.1
AHI, events/h	32.8 ± 23.6
Non-OSA	5 (9.1)
Mild OSA	10 (18.2)
Moderate OSA	13 (23.6)
Severe OSA	27 (49.1)
Nighttime symptoms	
Sleep apnea	16 (28.1)
Snoring	49 (86)
Light sleeper	27 (47.4)
Chocking	18 (31.6)
Daytime symptoms	
Headache	28 (49.1)
Daytime sleepiness	39 (68.4)
Loss of productivity	30 (52.6)
Depression	7 (12.3)
Underlying disease	
Coronary artery disease	5 (8.5)
Diabetes	3 (5.1)
Dyslipidemia	22 (37.3)
Hypertension	18 (30.5)

Continuous data are presented as mean ± standard deviation; categorical data are presented as numbers (percentages); AHI = apnea-hypopnea index (events/h), Non-OSA = AHI <5, Mild OSA = AHI 5 to <15, Moderate OSA = AHI 15 to <30, and Severe OSA = AHI >30 events/h

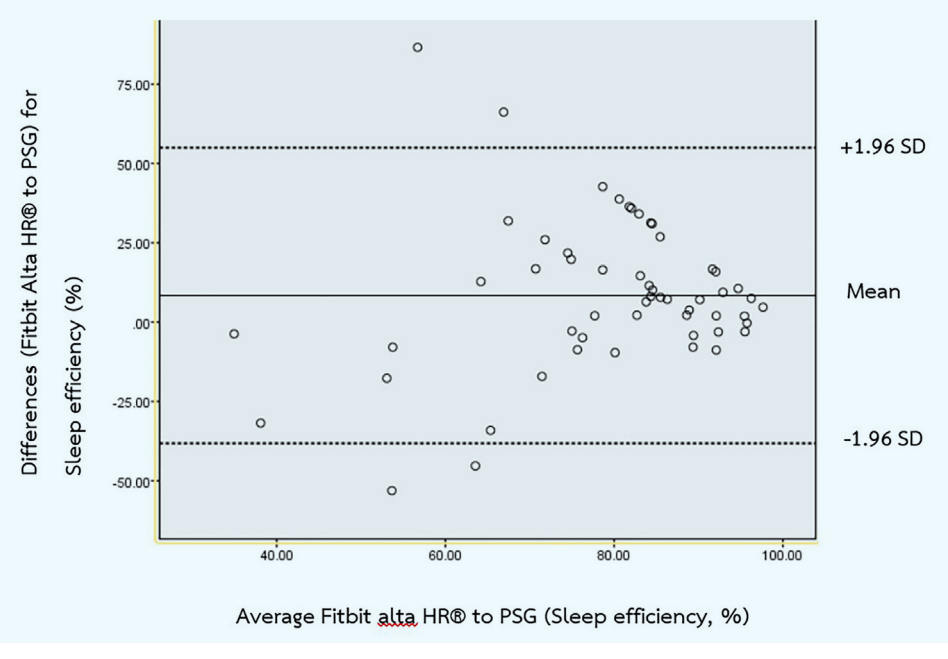


Fig 3. The Bland–Altman plots comparing sleep efficiency (SE) between the wearable device and polysomnography (PSG): Mean bias, the wide limits of agreement, the presence of outliers are shown.

TABLE 2. Mean differences in sleep outcomes between the wearable device (Fitbit Alta HR®) and polysomnography.

	PSG	Fitbit Alta HR®	Difference (Mean ± SD)	Lower limit	Upper limit	p-value
Sleep efficiency, %	74.9 ± 16.9	83.3 ± 19.9	8.4 ± 23.8	1.9	14.8	0.01*
Total sleep time, min	357.5 ± 82.1	377.1 ± 115.8	19.5 ± 136.5	-17.3	56.4	0.29
Light stage, %	74.3 ± 10.3	43.6 ± 26.4	-30.6 ± 28.1	-38.3	-23.0	<0.001*
Deep stage, %	10.2 ± 7.2	8.0 ± 6.3	-2.1 ± 10.0	-4.8	0.5	0.12
REM stage, %	10.2 ± 7.2	12.8 ± 9.0	-1.9 ± 10.2	-4.7	0.7	0.15

Abbreviations: PSG: Polysomnography, SD: Standard deviation.
**The p-value <0.05 indicates statistical significance.*

TABLE 3. Agreement of sleep outcomes between the wearable device (Fitbit Alta HR®) and polysomnography.

	ICC	95% CI
Sleep efficiency	-0.03	-0.22, 0.18
Total sleep time	0.17	-0.06, 0.39
Light stage	-0.04	-0.16, 0.10
Deep stage	0.16	-0.06, 0.38
REM stage	0.25	-0.06, 0.47

Abbreviations: ICC: intraclass correlation coefficient, CI: confidence interval

differ from PSG's. The wide limits of agreement suggest significant variability between the two methods, with discrepancies increasing at higher average SE levels, implying proportional bias. Additionally, the presence of outliers indicates instances where Fitbit's SE estimates substantially deviate from PSG measurements.

DISCUSSION

This is likely the first study to compare the performance of a wearable device (Fitbit Alta HR®) with PSG performed in adult Thai patients with sleep disorders, primarily obstructive sleep-disordered breathing (OSDB), addressing a knowledge gap in wearable device applicability. The study results revealed a statistically significant difference ($p < 0.05$) between the two devices in terms of mean sleep efficiency and light sleep, but not for mean TST, deep sleep, or REM sleep ($p > 0.05$). On average, the Fitbit Alta HR® overestimated SE and TST compared to PSG, with differences of 8.4 ± 23.8% and 19.5 ± 136.5 minutes, respectively. However, all sleep parameters measured by both methods showed poor agreement overall.

The findings of this study differ from those of *de Zambotti et al.*[10] who compared another wearable fitness-tracker device, the Jawbone UP®, with PSG in 18 healthy adults. Their study demonstrated good agreement between the wearable device and PSG for sleep estimation, including for TST and sleep onset latency, but less accuracy in detecting wake time after sleep onset (i.e., a poor ability to detect being awake). Similarly, *de Zambotti et al.*[8] conducted another study in 65 healthy adolescents using the same device and found good agreement with PSG. A systemic review by *Evenson et al.*,[7] reported a pilot study comparing the Fitbit Flex® with PSG in OSDB children, which showed high sensitivity for detecting sleep but low specificity for detecting wakefulness. The differences in results between our study and others are likely due to variations in subject age groups, health conditions, device models, and methods for measuring sleep parameters. In addition, factors such as dietary habits (high consumption of caffeinated beverages, particularly energy drinks), environmental conditions (co-sleeping), and cultural sleep patterns (varying sleep schedules) of Thai population might also influence findings.[16] However, these were not well recorded in our study.

Wearable devices like the Fitbit utilize proprietary algorithms to monitor sleep patterns, primarily through motion-based actigraphy and heart rate variability (HRV). These algorithms estimate sleep stages—light, deep, and REM sleep—by analyzing movement and physiological

signals, however, has certain limitations. It may inaccurately detect periods of wakefulness as sleep, leading to errors in tracking sleep duration and quality. Additionally, the device might struggle to accurately identify short naps or power naps, resulting in incomplete sleep data. Furthermore, heart rate can vary due to factors unrelated to sleep, such as stress or illness, which (may) potentially confoundsleep stage estimation.[12,15]

There are some potential limitations of this study. First, the results are based on a single model of a commercially available wearable device capable of detecting sleep characteristics in patients with sleep disorders. Therefore, the results may not apply to newer models or the latest technology from Fitbit or other smartwatch manufacturers, as devices are continuously evolving. Second, the study recorded sleep data from only a single night. Measuring over longer periods or across multiple nights might yield more practical outcomes. Third, the Bland–Altman plots are not true quantitative analyses. Bias and limits of agreement may be superimposed on its visual presentation of the data. Finally, the study did not include any healthy subjects as a control group. For future direction, we recommend that subsequent studies should focus on comparing multiple devices, integrating multi-sensor data, and employing machine learning models to enhance the accuracy of wearable devices. They should also include control groups and incorporate the latest technology to provide more comprehensive insights.

CONCLUSION

Wearable devices (Fitbit Alta HR®) offer a convenient means to monitor sleep patterns over extended periods in naturalistic settings. However, in patients with sleep disorders, it exhibited some discrepancies and weak correlations in sleep measurements compared to PSG. In real-world practice, integrating data from these devices can enhance such patients monitoring and engagement but should be interpreted with caution and in conjunction with comprehensive clinical evaluations. As technology continues to advance and future studies are coming, the outcomes of wearable devices for sleep monitoring may significantly improve.

Data Availability Statement

The data that support the findings of this study are available upon request from the corresponding author, [W.B.]. The data are not publicly available due to containing information that could compromise the privacy of research participants.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Miss Julaporn Pooliam, our consulting statistician, and all the patients who took part in this study. We also extend our thanks to Miss Jeerapa Kerdnoppakhun and Miss Ngamrat Treerassapanich for their assistance with manuscript preparation. Additionally, we thank the staff of the Department of Otorhinolaryngology as well as the sleep technicians at Siriraj Sleep Center, Faculty of Medicine Siriraj Hospital, Mahidol University.

DECLARATION

Grants and Funding Information

This study was funded by a grant from the Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand (Grant Number R016231067).

Conflict of Interest

The authors declare they have no conflicts of interest. All authors have seen and approved this manuscript.

Registration Number of Clinical Trials

TCTR20250204008

Author Contributions

K.A., W.B.; Conceptualization, Methodology. W.B.; Original draft preparation, K.A. K.A., P.W., W.C. and S.R.; Data Curation. W.B.; Writing - Review & Editing. All authors have accepted responsibility for the entire content of this manuscript and have approved its submission.

Use of Artificial Intelligence

None

REFERENCES

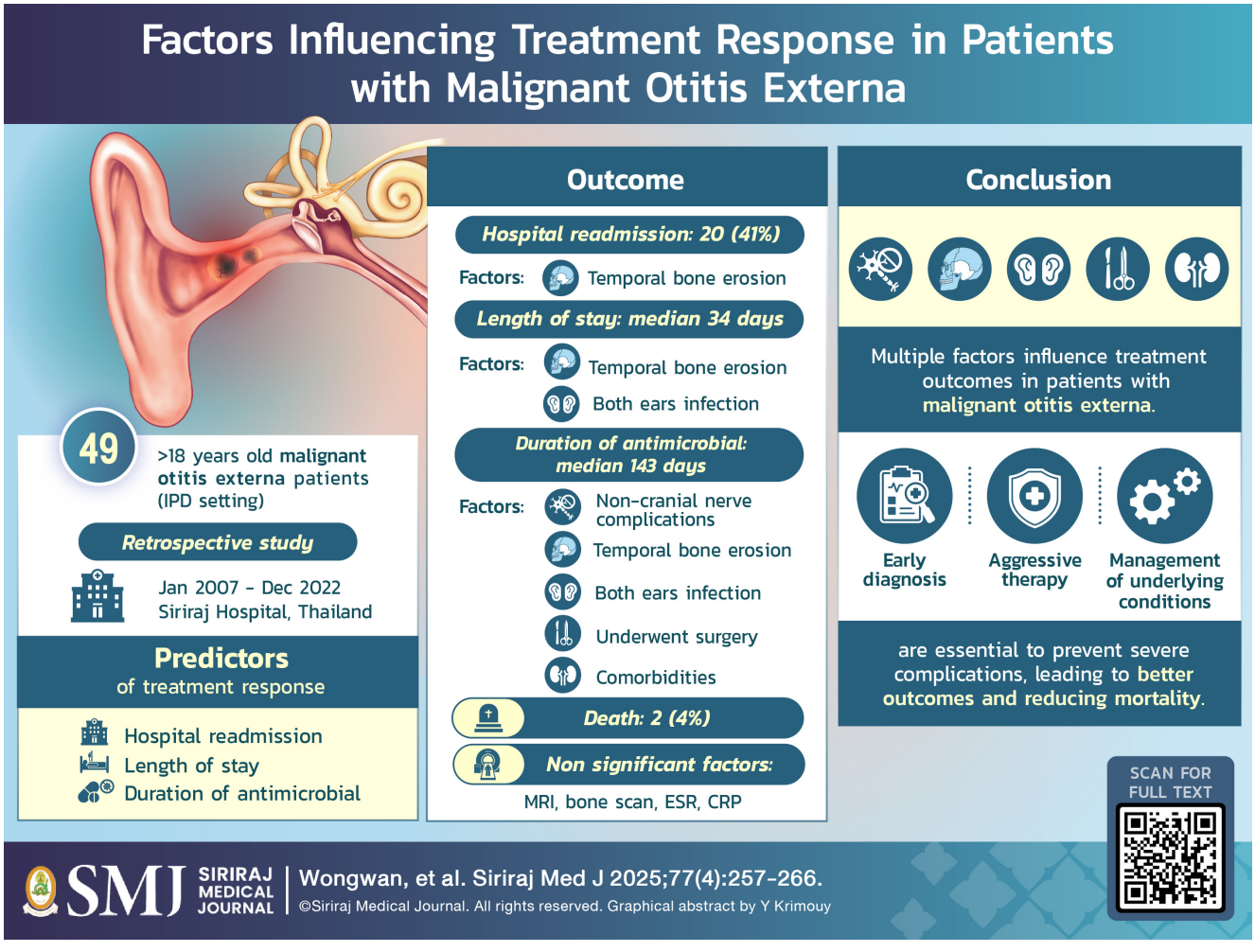
1. Practice parameters for the use of portable recording in the assessment of obstructive sleep apnea. Standards of Practice Committee of the American Sleep Disorders Association. Sleep. 1994;17:372-7.
2. Lee SA, Paek JH, Han SH. Sleep hygiene and its association with daytime sleepiness, depressive symptoms, and quality of life in patients with mild obstructive sleep apnea. J Neurol Sci. 2015;359:445-9.
3. Mehra R, Moul DE, Strohl KP. In: Kryger M, Roth T, Dement WC, editors. Principles and Practice of Sleep Medicine. 6th ed. Elsevier, Philadelphia; 2017.p.167-73.
4. Marino M, Li Y, Rueschman MN, Winkelman JW, Ellenbogen JM, Solet JM, et al. Measuring Sleep: Accuracy, Sensitivity, and Specificity of Wrist Actigraphy Compared to Polysomnography. Sleep. 2013;36:1747-55.
5. Martin JL, Hakim AD. Wrist actigraphy. Chest. 2011;139:1514-27.
6. Morgenthaler T, Alessi C, Friedman L, Owens J, Kapur V,

Boehlecke B, et al. Standards of Practice Committee; American Academy of Sleep Medicine. Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: an update for 2007. Sleep. 2007;30:519-29.
7. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act. 2015;12:159.
8. de Zambotti M, Baker FC, Colrain IM. Validation of Sleep-Tracking Technology Compared with Polysomnography in Adolescents. Sleep. 2015;38:1461-8.
9. Cook JD, Eftekari SC, Dallmann E, Sippy M, Plante DT. Ability of the Fitbit Alta HR to quantify and classify sleep in patients with suspected central disorders of hypersomnolence: A comparison against polysomnography. J Sleep Res. 2018;28(4): e12789.
10. de Zambotti M, Claudatos S, Inkelis S, Colrain IM, Baker FC. Evaluation of a consumer fitness-tracking device to assess sleep in adults. Chronobiol Int. 2015;32:1024-8.
11. Cook JD, Prairie ML, Plante DT. Ability of the Multisensory Jawbone UP3 to Quantify and Classify Sleep in Patients With Suspected Central Disorders of Hypersomnolence: A Comparison Against Polysomnography and Actigraphy. J Clin Sleep Med. 2018;14:841-8.
12. Lee HA, Lee HJ, Moon JH, Lee T, Kim MG, In H, et al. Comparison of Wearable Activity Tracker with Actigraphy for Sleep Evaluation and Circadian Rest-Activity Rhythm Measurement in Healthy Young Adults. Psychiatry Investig. 2017;14:179-85.
13. Osterbauer B, Koempel AJ, Davidson Ward SL, Fisher LM, Don DM, Ver Halen J. A Comparison Study of The Fitbit Activity Monitor And PSG For Assessing Sleep Patterns And Movement In Children. J Otolaryngol Adv. 2016;1:24-35.
14. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research [published correction appears in J Chiropr Med. 2017;16:346.
15. de Zambotti M, Cellini N, Goldstone A, Colrain IM, Baker FC. Wearable sleep technology in clinical and research settings. Med Sci Sports Exerc [Internet]. 2019;51(7):1538-57.
16. Banhiran W. Characteristics and clinical presentations of patients at the Siriraj snoring clinic. Siriraj Med J [Internet]. 2020 [cited 2025 Jan 22];72(3):202-8.

Factors Influencing Treatment Response in Patients with Malignant Otitis Externa

Pithiwat Wongwan, M.D.[1], Kanthong Thongyai, M.D.[2,*], Suvajana Atipas, M.D.[2], Sarun Prakairungthong, M.D.[2], Siriporn Limviriyakul, M.D.[2], Kanokrat Suvarnsit, M.D.[2], Herman Jenkins, M.D.[3]

[1]Department of Otorhinolaryngology, Chonburi Hospital, Chonburi, Thailand, [2]Department of Otorhinolaryngology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [3]Department of Otolaryngology-Head and Neck Surgery, University of Colorado, Anschutz Medical Campus, Aurora, Colorado, USA.



**Corresponding author: Kanthong Thongyai*
E-mail: kanthong@gmail.com
Received 6 December 2024 Revised 25 January 2025 Accepted 5 February 2025
ORCID ID:http://orcid.org/0000-0001-8987-6006
https://doi.org/10.33192/smj.v77i4.272587

ABSTRACT

Objective: To identify factors affecting treatment outcomes in patients with malignant otitis externa (MOE).

Materials and Methods: A retrospective review of MOE treatment was conducted in patients aged > 18 years admitted to Siriraj Hospital from January 2007 to December 2022. Predictors of treatment response chosen included duration of hospitalization, duration of antimicrobial treatment, disease-related re-admissions, and mortality.

Results: The study included 49 patients (33 males, 16 females) with a mean age of 65±12 years. Comorbidities were present in 90% of patients, with diabetes mellitus being the most common. *Pseudomonas aeruginosa* was identified in 37% of cases. Bilateral symptoms/infections were reported in 18%. Facial nerve palsy and non-cranial nerve complications were 51% and 12%, respectively. Computed tomography (CT) imaging revealed bony erosion in 75% of patient. Surgical management was performed in 73% of patients. The median duration of hospital stays and antimicrobial treatment were 34 days and 143 days, respectively. Readmissions due to disease progression occurred in 35%, and the mortality rate was 4%. Bone erosion on CT was associated with an increased likelihood of readmission. Prolonged hospital stays were associated with bilateral symptoms and positive CT findings. Extended antimicrobial treatment was linked to multiple comorbidities, bilateral symptoms, non-cranial nerve complications, positive CT findings, and surgical cases.

Conclusion: Various factors influence treatment outcomes in MOE patients. Early diagnosis, aggressive treatment, and management of prognostic factors are essential for preventing severe complications and improving survival outcomes.

Keywords: Otitis externa; necrotizing otitis externa; osteomyelitis; risk factors; patient outcome assessment; diabetes mellitus; malignant otitis externa (Siriraj Med J 2025; 77: 257-266)

INTRODUCTION

Malignant otitis externa (MOE), also known as necrotizing otitis externa (NOE), was first reported by Toulmouche in 1838 and later described by Chandler in 1968.[1,2] It is a rare, non-cancerous but potentially life-threatening infection of the external ear that can progressively spread via various anatomical pathways to adjacent structures, including the mastoid, skull base, and cranial nerves. This extension may result in serious complications, including cranial neuropathies, brain abscesses, meningitis, and dural venous sinus thrombosis.[3]

MOE primarily affects older adults, particularly males, and is more common among individuals with underlying diabetes mellitus (DM) or other immunocompromised conditions, including malignancies, use of immunosuppressive drugs, and Acquired Immune Deficiency Syndrome (AIDS). Immunosuppressed patients due to HIV or other non-diabetic factors may develop MOE at a younger age compared to those with diabetes.[4,5] The most common pathogen is *Pseudomonas aeruginosa*, but MOE may result from other pathogens, such as methicillin-resistant *Staphylococcus aureus* (MRSA), *Klebsiella*, and *Proteus mirabilis*. Fungal species, including Aspergillus and *Candida*, are also notable causative agents.[6,7]

Common presenting symptoms are persistent otorrhea and otalgia, particularly at night, that do not respond to topical antibiotic treatment. Physical examination often reveals a painful, erythematous, and edematous ear canal with discharge. The presence of granulation tissue at the bony-cartilaginous junction is suggestive of a *Pseudomonas* infection.[8] In addition to frequent involvement of the facial nerve, other cranial nerves, including the glossopharyngeal, vagus, spinal accessory, hypoglossal, abducens, and trigeminal nerves, may be affected. Laboratory and imaging studies are essential for diagnosis and management. In a recent systematic review identified 27 different diagnostic criteria for MOE.[9] Treatment generally requires prolonged systemic antimicrobial therapy, aural cleaning and blood sugar control, with some cases necessitating hyperbaric oxygen therapy. Surgery is generally reserved for severe cases that are unresponsive to medical management.[8]

Although rare, the incidence of MOE varies across studies, with recent systematic reviews estimating it at between 0.22 and 1.19 cases per 100,000 individuals.[10] No official incidence data is available for Thailand. This study aims to identify factors that influence MOE treatment outcomes in patients at Siriraj Hospital, thus providing an initial exploration of this condition in the Thai population.

MATERIALS AND METHODS

A retrospective review was conducted using the database of the Department of Otorhinolaryngology at Siriraj Hospital. The study included all patients over

18 years of age who were admitted for malignant otitis externa between January 1st, 2007, and December 31st, 2022. Diagnoses made by otologists in the department were based on four out of seven items from Levenson's criteria: refractory otitis externa, severe nocturnal otalgia, purulent otorrhea, granulation tissue in the external auditory canal, positive *Pseudomonas* culture, presence of an immunocompromised state or diabetes, and a positive technetium bone scan.[11]

Data collected from medical records included age, sex, comorbidities, physical examination findings, treatment details, biochemical results, culture and biopsy results, imaging studies, and follow-up information. Potential predictors of treatment response were defined as the duration of hospitalization, length of antimicrobial therapy, disease-related hospital readmission, and mortality. Cases with incomplete data or those managed exclusively on an outpatient basis were excluded from the study. Ethical approval for the study was obtained from the institutional Ethics Committee.

Statistical analyses

Descriptive statistics summarized demographic characteristics such as age and sex. Quantitative data are presented as mean ± standard deviation or median (interquartile range), while qualitative data are shown as frequencies or percentages.

To analyze hospital readmission rates, we used the independent t-test for normally distributed quantitative variables and the Mann-Whitney U test for skewed distributions. Pearson's chi-square or Fisher's exact test assessed associations between qualitative variables.

The Pearson's correlation coefficient was used for univariate analysis of hospitalization time and antimicrobial therapy. Variables with p-values less than 0.2 in univariate analysis were selected for multivariate analysis using logistic regression for readmission factors and linear regression for hospital and antimicrobial durations.

Survival data were analyzed using Kaplan-Meier survival curves.

All statistical analyses were performed using PASW Statistics 18 (SPSS Inc., Chicago, IL, USA), with significance set at $\mathbf{p < 0.05}$.

RESULTS

Patient characteristics

A total of 49 patients were included in the study, consisting of 33 men and 16 women, with ages ranging from 42 to 91 years, with a mean of 65±12 years. Of these, 44 patients (90%) had multiple comorbidities. The most common underlying condition was diabetes mellitus (98%), followed by hypertension (71%) and chronic kidney disease (39%). Two patients were diagnosed with malignancies (breast cancer and lymphoma).

Physical examination and investigations

The median duration of symptoms was approximately 60 days. Symptoms were predominantly observed in the right ear, with 23 patients (43%) presenting symptoms exclusively on the right side, while nine (18%) experiencing symptoms in both ears. On physical examination, purulent otorrhea was observed in 45 patients (92%), and otalgia was noted in 45 patients (92%). Granulation tissue was identified in 38 (78%) cases. Facial nerve paralysis, a common otologic complication, was documented in 25 patients (51%). Additionally, advanced MOE involved other cranial nerves: cranial nerve VI (two patients), X (one patient), IX (one patient), and XII (one patient). Six patients presented with other complications, including infratemporal abscess, sigmoid sinus thrombosis, and parotid abscess.

Inflammatory markers were useful for both diagnosis and follow-up in cases of MOE. The erythrocyte sedimentation rate (ESR) was elevated in 43 patients (88%), with values ranging from 17 to 119 mm/hour. C-reactive protein (CRP) levels were elevated in 27 patients (55%), ranging from 0.6 to 287 mg/L. Computed tomography (CT) scan revealed erosion of the temporal bone and soft tissue in the ear canal in 37 patients (75%). Technetium scans (Tc-99m) were performed on 29 patients, all of which showed positive findings. Additionally, 31 patients had absorption or positive gallium scans (Ga-67).

Pathogenic microorganisms and pathological diagnosis

The most frequently isolated pathogen was *Pseudomonas aeruginosa*, found in 18 patients (37%). Fungal pathogens were isolated in 16 patients, while cultures were negative in 15 cases. Granulation tissue and inflammation were present in almost all cases (86%; 42 out of 49).

Treatment

Medical therapy was the primary treatment approach, with a median treatment duration of 191 days. All patients received predominantly ceftazidime antibiotic therapy, with 92% requiring intravenous administration. Sixteen patients with fungal infections received antifungal treatment with voriconazole or itraconazole in addition to antibiotics. Surgical intervention was necessary for 36 patients (73%), mostly in the form of various types of mastoidectomy (27 cases). Other procedures included ear canal debridement, and local incision and drainage.

Prognosis and risk factors

Table 1 summarizes the overall response outcomes of the group. The median duration of hospital stay for the 49 patients was 34 days, with a median 143 days duration of antimicrobial treatment. The in-hospital mortality rate included two patients (4%), while rehospitalization was required in 20 patients (41%), including disease progression in 17 patients and Gallium scan preparation in 3 patients.

Disease-related hospital readmission

Twenty patients treated at Siriraj Hospital required readmission and the data for this group are presented in Table 2 and analysis of the different variables in Table 3. Notably, erosion of the temporal bone noted on CT scan (90%; 18 out of 20) was significantly associated with an increased likelihood of readmission ($p = 0.03$). Other factors, though not statistically significant, included symptoms, duration of more than 60 days, the non-cranial nerve involvement complications, positive technetium and gallium scans, and undergoing surgery.

Length of Hospital Stay (LOS)

The median length of hospitalization was 34 days (range: 18 to 58 days). Infections involving both ears and temporal bone erosion observed on CT scan were statistically significant factors contributing to a longer hospital stay (Table 4).

Duration of antimicrobial treatment

As shown in Table 5, the median duration of antimicrobial therapy was 143 days (range: 78 to 290 days). Patients with comorbidities, bilateral ear infections, non-cranial nerve complications, positive CT scan findings, and those who underwent surgery tended to have a significantly longer antimicrobial treatment.

In-Hospital mortality

The overall survival rate was 96%. Two patients died at 2 months and 14 months, respectively, due to progressive disease process (Fig 1). The first was a 75-year-old woman with diabetes mellitus, hypertension, and end-stage renal disease, dying from meningitis and sepsis. The second was a 90-year-old man with diabetes mellitus, hypertension, and a history of myocardial infarction, who succumbed to septic shock.

DISCUSSION

MOE is a rare but potentially life-threatening infection of the external ear, with a prevalence ranging between 0.22 and 1.19 cases per 100,000 patients.[10] Due to the condition's rarity, formal studies in Thailand are limited. This 16-year study, involving 49 patients, aimed to identify factors that influence treatment outcomes, providing valuable insights for future research.

Elderly patients (age ≥ 65 years) are consistently identified as being at higher risk for poor prognosis in numerous studies, with advanced age correlating with extended hospitalization and increased mortality rates.[3,6] However, in this study, age was not a predictor of poor outcomes. Diabetes mellitus is widely recognized as a significant risk factor for MOE, with reported prevalence between 65% and 85% in recent studies.[12,13] Some studies link diabetes mellitus with poor treatment outcomes, whereas others do not.[12,14,15] In this study, 98% of patients had diabetes mellitus, however, it did not significantly correlate to worse outcomes. Additionally, neither poorly nor well-controlled diabetes (as indicated by HbA1c and

TABLE 1. Treatment response outcomes.

Treatment response	Total (N=49) Numbers of cases (percentage) Days: Median (25%, 75%)
Hospital readmission (cases)	20 (41)*
Death (cases)	2 (4)
Hospitalization time (days)	34 (18, 58)
Duration of antimicrobial (days)	143 (78, 290)

*17(35) readmission due to disease progression, 3(6) readmission due to Gallium scan

TABLE 2. Baseline characteristics of patients with readmission outcomes.

	Total	Re-admission		p-value
		No	Yes	
	N = 49 (%)	N = 29 (%)	N = 20 (%)	
Sex				
Male	33 (67)	20 (69)	13 (65)	0.77
Female	16 (33)	9 (31)	7 (35)	
Age (years)				
<65 years	25 (51)	15 (52)	10 (50)	0.91
≥ 65 years	24 (49)	14 (48)	10 (50)	
Diabetes Mellitus	48 (98)	28 (97)	20 (100)	1.00
HbA1c				
Good control	14 (31)	8 (31)	6 (32)	0.95
Poor control (≥7 mg%)	31 (69)	18 (69)	13 (68)	
Fasting blood sugar:				
Good control	13 (27)	7 (25)	6 (30)	0.75
Poor control (≥130 mg/dl)	35 (73)	21 (75)	14 (70)	
Non-diabetic underlying disease	44 (90)	25 (86)	19 (95)	0.64
BMI (kg/m^2)	23 ± 4	22 ± 4	23 ± 5	0.33
Affected ears:				
Right	21 (43)	13 (45)	8 (40)	0.61
Left	19 (39)	12 (41)	7 (35)	
Bilateral	9 (18)	4 (14)	5 (25)	
Duration day (days) (First day of symptom to the first day of hospital visit)	60 (14, 60)	30 (7, 60)	60 (30, 105)	0.05
Duration of ≥ 60 days	25 (51)	12 (41)	13 (65)	0.10
Clinical presentation				
Otalgia	45 (92)	27 (93)	18 (90)	1.00
Purulent otorrhea	45 (92)	25 (86)	20 (100)	0.13
Granulation tissue	38 (78)	25 (86)	13 (65)	0.09
Cranial nerve involvement	27 (55)	15 (52)	12 (60)	0.57
Non-cranial nerve involvement complications	6 (12)	2 (7)	4 (20)	0.21
Inflammatory markers				
Erythrocyte sedimentation rate (ESR) (mm/hour)	81 (63, 98)	80 (51, 103)	83 (72, 98)	0.61
C-reactive protein (CRP) (mg/l)	23 (8, 47)	30 (6, 47)	22 (9, 57)	0.62
Pathogens				
Pseudomonas aeruginosa	18 (37)	10 (34)	8 (40)	0.69
Fungus	16 (33)	11 (38)	5 (25)	0.34
Pathology				
Granulation and inflammation	42 (86)	24 (83)	18 (90)	0.68
CT scan: positive result	37 (76)	19 (66)	18 (90)	0.09
Technetium scan: positive	29 (59)	14 (48)	15 (75)	0.06
Gallium scan: positive	31 (63)	16 (55)	15 (75)	0.16
Surgery	36 (73)	19 (66)	17 (85)	0.13

TABLE 3. Re-hospitalization based on different variables.

	Crude Odds ratio (95%CI)	P-value	Adjusted Odds ratio (95%CI)	P-value
Duration range ≥ 60 days	2.63 (0.81, 8.55)	0.11		
Granulation tissue	0.30 (0.07, 1.20)	0.09	0.26 (0.05, 1.31)	0.10
Non-cranial nerve involvement complications	3.38 (0.55, 20.55)	0.18	4.96 (0.63, 39.36)	0.13
CT scan: positive	4.73 (0.91, 24.65)	0.07	8.62 (1.16, 63.86)	**0.03**
Technetium scan: positive	3.21 (0.92, 11.18)	0.07		
Gallium scan: positive	2.43 (0.69, 8.49)	0.16	2.81 (0.69, 11.40)	0.15
Surgery	2.98 (0.70, 12.67)	0.14		

*Logistic regression was used for multivariate analyses for **all p-values less than 0.2** in univariate analysis with backward selection variable (p-value for removal equal to 0.17)

TABLE 4. Hospitalization duration in relation to different variables.

	Univariable analysis		Multivariable analysis	
	Coefficient (95%CI)	P-value	Coefficient (95%CI)	P-value
Female	-29.41 (-63.32, 4.49)	0.09		
Well-controlled HbA1c	-27.74 (-65.11, 9.63)	0.14		
Bilateral ears affected	47.34 (7.31, 87.37)	0.02	59.76 (19.67, 99.86)	**0.00**
Fungal infection	24.69 (-9.53, 58.92)	0.15		
CT scan: positive result	39.17 (2.79, 75.55)	0.03	47.08 (11.13, 83.03)	**0.01**
Technetium scan: positive	36.91 (5.33, 68.48)	0.02		
Gallium scan: positive	26.38 (-6.76, 59.52)	0.12		
Surgery	30.80 (-5.24, 66.85)	0.09	27.26 (-7.32, 61.85)	0.12

*Linear regression was used for multivariate analyses for **all p-values less than 0.2** in univariate analysis with backward selection variable (p-value for removal equal to 0.17)

TABLE 5. Duration of antimicrobials in relation to different variables.

	Univariable analysis		Multivariable analysis	
	Coefficient (95%CI)	P-value	Coefficient (95%CI)	P-value
Non-diabetic underlying diseases	111.64 (-54.94, 278.22)	0.18	146.01 (17.01, 275.01)	**0.03**
Otorrhea	126.25 (-57.74, 310.25)	0.17	101.63 (-43.36, 246.61)	0.16
Bilateral ears affected	113.51 (-14.96, 241.98)	0.08	142.32 (36.56, 248.08)	**0.01**
Non-cranial nerve involvement complications	146.68 (-4.07, 297.42)	0.06	121.06 (3.01, 239.11)	**0.04**
CT scan: positive result	110.62 (-4.39, 225.63)	0.06	126.97 (27.21, 226.72)	**0.01**
Technetium scan: positive	134.92 (38.14, 231.69)	0.00		
Gallium scan: positive	130.32 (30.82, 229.83)	0.01		
Surgery	154.62 (47.43, 261.82)	0.00	149.03 (52.85, 245.21)	**0.00**

*Linear regression was used for multivariate analyses for **all p-values less than 0.2** in univariate analysis with backward selection variable (p-value for removal equal to 0.17)

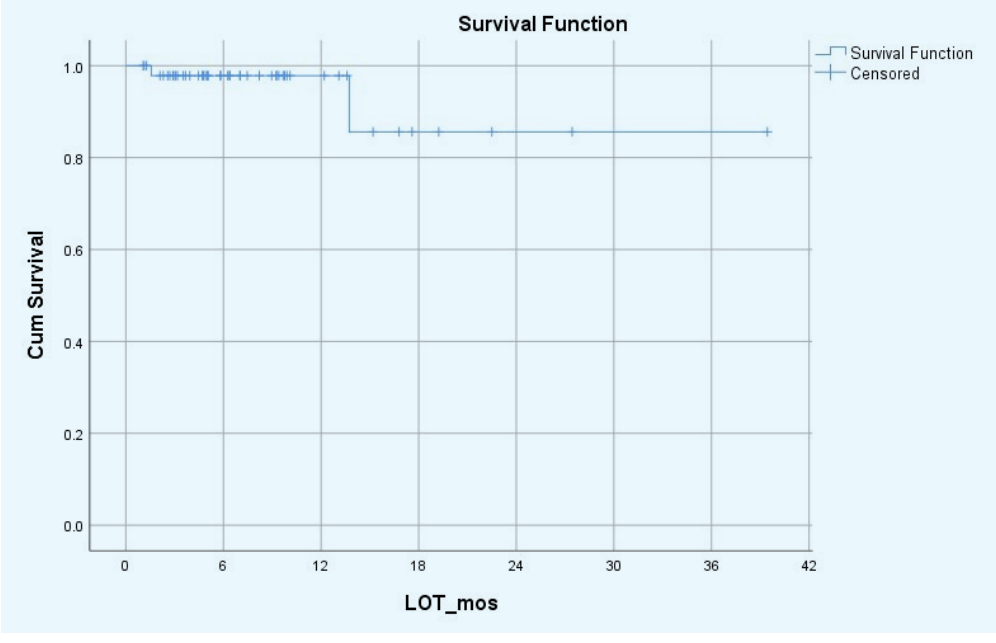


Fig 1. Kaplan-Meier curve showing association between cumulative survival rate and length of treatment (in months). * Cum Survival = cumulative survival rate, ** LOT mos = length of treatment in the month.

fasting blood sugar levels) were not determined to be a prognostic factor. Patients with multiple comorbidities, such as end-stage renal disease and hypertension, often have compromised immune systems and microvascular complications, potentially leading to poor treatment responses. In this study, multiple comorbidities were significantly associated with longer antimicrobial treatment durations, in aligning with findings from Arsovic et al.[16] These findings underscore the importance of controlling underlying comorbid conditions when treating MOE.

Common symptoms included otalgia, otorrhea, and granulation tissue. Bilateral ear involvement was statistically associated with longer antimicrobial treatment and extended hospital stays. Facial nerve involvement was the most frequent cranial nerve complication. The prognostic significance of facial nerve palsy is debated; *Stevens et al.* reported increased mortality with facial nerve palsy, while other studies found no difference in survival.[17-19] In this study, facial nerve palsy occurred in 25 patients (51%) but did not predict poorer prognosis. In contrast, non-cranial nerve complications, such as parotid abscess and transverse sigmoid sinus thrombosis, were associated with prolonged antimicrobial treatment.

Pseudomonas aeruginosa is reported as the most common pathogen in MOE, implicated in 50-90% of cases.[2] In this study, *Pseudomonas aeruginosa* was found in 37% of cases, fungal pathogens in 32%, and 31% of cultures were negative. This variation may be attributed to the study's tertiary care center setting, where cases are often complicated or partially treated.

Inflammatory markers, such as elevated ESR (88%) and CRP (55%), were frequently observed. Although these markers are useful in diagnosing and monitoring MOE, they were not significant prognostic factors for treatment outcomes in this study, which aligns with other reports.[20] Imaging plays a critical role in diagnosis and treatment planning. CT scans are valuable for assessing disease extent and identifying anatomical structures for surgical planning. Increased uptake of Technetium-99m (TC-99m) scans at the temporal bone area helps diagnosing MOE due to being more sensitive test than CT scan for detecting osteogenic activity especially in early or unclear cases. However, Tc99m uptake persists for an infinite period. Gallium-67 (Ga-67), which binds to white blood cells, can detect infection and inflammation. The Ga-67 scans are beneficial for follow-up, as they normalize after resolution of infection. Therefore, the number of patients who needed TC-99m scans for diagnosis are less than the patients who had Ga-67 scans for follow up if they did not need TC-99m for fulfilling Levenson's criteria.

However, this retrospective study collects data from the past 15 years, during which Ga-67 was not commonly available. As a result, only 30 patients received a follow-up Ga-67 scan. Among those who underwent follow-up with Ga-67, the readmission rate was 36% (11 patients). Most patients who received a Ga-67 scan were followed up in the outpatient department (OPD) setting, as they were transitioned to oral antibiotics once clinical improvement was observed. Therefore, the primary purpose of Ga-67 scans in our setting is to reassess the active inflammatory phase of the disease and decide whether to discontinue antibiotic treatment. The scheduling of Ga-67 scans at our center requires long appointment durations, which means that some patients treated as outpatients may develop progressive disease and require rehospitalization before receiving the scan. Of the 11 patients, two presented with new lesions in another ear after the resolution of the initial lesion. Two patients did not show clinical improvement and continued to exhibit active uptake after discharge. Two required intravenous antibiotics and readmission following pathogen identification. Additionally, two patients were re-hospitalized due to increased Ga-67 uptake, and three patients had progressive lesions before their Ga-67 scans. However, none of the patients identified by the negative Ga-67 scan required readmission due to disease recurrence.

In this study, bony erosion or soft tissue abnormalities on CT scans were significantly associated with poorer outcomes. including higher rates of hospital readmission, longer antimicrobial treatment durations, and extended hospital stays. This finding is consistent with studies linking temporal bone erosion to prolonged hospitalizations.[6,16] However, CT imaging may have limitations in detecting early-stage infections, probably explaining why changes were more frequently observed in advanced cases.[21,22]

The primary treatment for MOE remains systemic antimicrobial therapy, combined with strict control of underlying conditions. Although the optimal duration of antimicrobial treatment is not definitively established, most studies recommend 4 to 6 weeks. Several factors are considered when deciding to discontinue treatment for malignant otitis externa, including resolution of signs and symptoms, negative cultures, normalized inflammatory markers and radionuclide scanning improvement. Due to the aggressive nature of malignant otitis externa, disease progression can still occur despite treatment, highlighting the importance of close monitoring throughout the management process.

At our center, if clinical improvement is observed,

such as resolving painful swollen ear canal, returning to normal in inflammatory markers (ESR and C-reactive protein), and no uptake in a Ga-67 scan, this serves as evidence of antimicrobial discontinuation. Conversely, the presence of suspected active lesions clinically and confirmed by sustained or progressive uptake of Ga-67 inform the decision to pursue further aggressive treatment for MOE.

Surgical intervention is required in severe cases, particularly those involving complications or failure to respond to antimicrobial treatment, to obtain tissue for culture and debride necrotic tissue.[23] At our center, the surgical intervention rate is notably high at 73%, which was significantly associated with longer antimicrobial treatment. The most common procedure performed in our setting is simple mastoidectomy as it offers superior tissue samples for culture and pathological evaluation, that is, more helpful in diagnosis and treatment. Mostly our patients are complex cases that were referred from the other hospitals, for example, partially undergone medical treatment or no previous tissue culture. Consequently, when patients experience persistent pain and unresponsiveness to treatment, a definitive tissue diagnosis is necessary. Additionally, granulation tissue in the middle ear and mastoid cavity were primary operative findings. Some studies suggest that early debridement of granulation and necrotic tissue can reduce pain, shorten antibiotic therapy, and expedite healing.[24]

Previous studies have reported mortality rates of 17.8% to 21%.[25] In this study, the in-hospital mortality rate in this study was 4%, suggesting effective management of MOE patients at our institution. Numerous studies have explored factors influencing MOE treatment outcomes, and the findings from this study provide valuable insights for clinicians, supporting the development of effective treatment strategies that result in favorable outcomes. However, this study has limitations, including its retrospective design, relatively small sample size, and restriction to a single-center experience.

CONCLUSION

Multiple factors influence treatment outcomes in patients with malignant otitis externa. Early diagnosis, aggressive therapy, and management of underlying conditions are essential to prevent severe complications, leading to better outcomes and reducing mortality. This is the first formal study of malignant otitis externa in Thailand. Future multicenter studies with larger sample sizes are recommended to expand the understanding of this disease from various perspectives.

Data Availability Statement

The data that support the findings of this study are available upon request from the corresponding author, [K.T.]. The data are not publicly available due to containing information that could compromise the privacy of research participants.

ACKNOWLEDGEMENTS

We extend our heartfelt gratitude to Asst. Prof. Dr. Orawan Supapueng, statistician, and Ms. Jeerapa Kerdnoppakhun for their invaluable assistance in developing this manuscript. We are also deeply thankful to all the patients involved in this study and the staff of the Department of Otorhinolaryngology, Faculty of Medicine Siriraj Hospital, Mahidol University.

DECLARATION

Grants and Funding Information

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

Conflict of Interest

The authors declare no potential conflicts of interest regarding the research, authorship, or publication of this article.

Registration Number of Clinical Trial

Not registered in the clinical trial due to retrospective study.

Author Contributions

P.W.; Conceptualization, Methodology, Formal analysis, Data Curation, Writing - Review & Editing, Supervision S.A, S.P., S.L., K.S., H.J.; Resources, Methodology, Data Curation, Writing - Original Draft K.T.; Conceptualization, Data Curation, Writing - Review & Editing.

Use of Artificial Intelligence

None

Ethics Committee Approval

Ethics committee approval was obtained for this study from the Siriraj Institutional Review Board (COA no. Si 253/2023).

REFERENCES

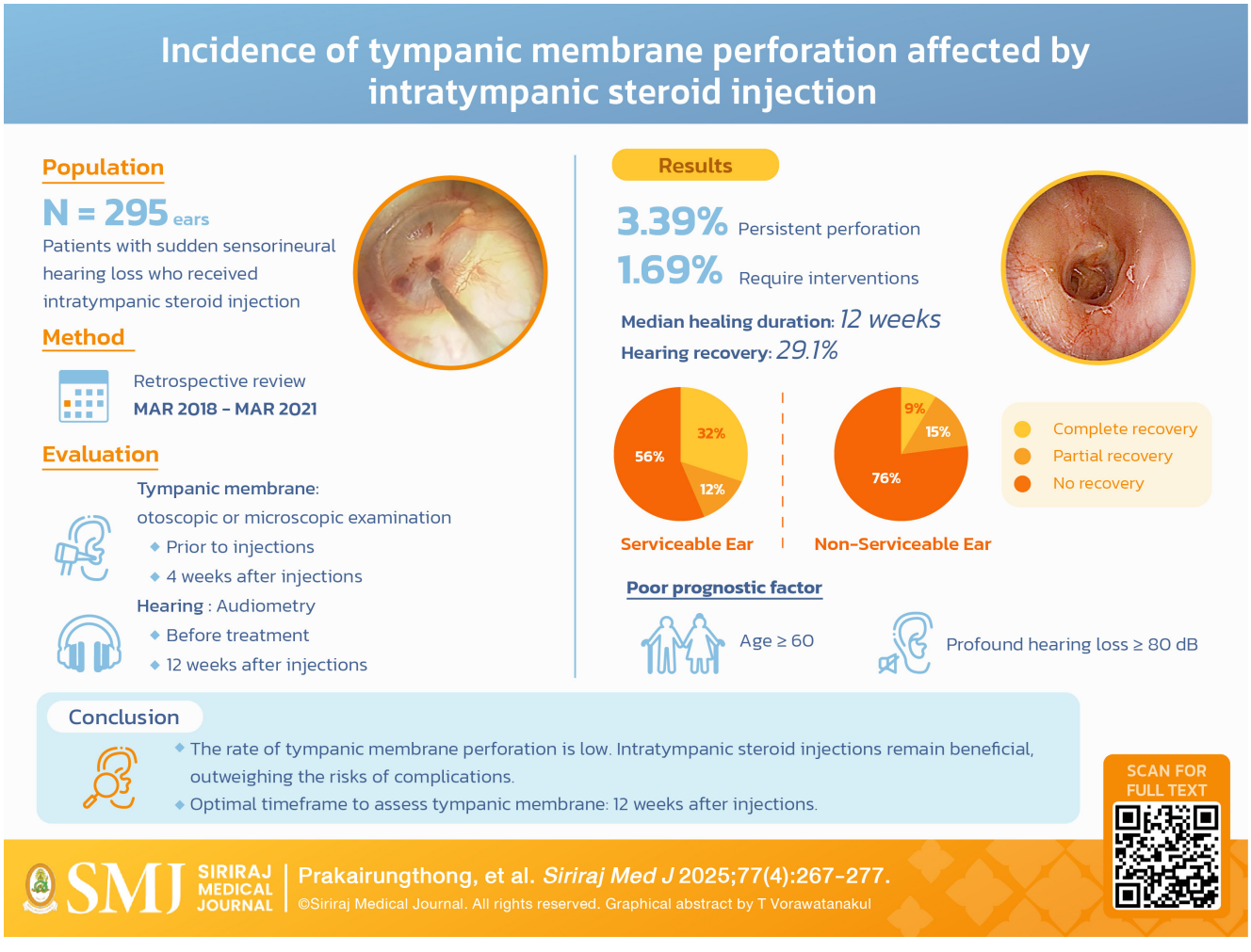
1. Dabholkar JP, Sheth A. Malignant otitis externa. Indian J Otolaryngol Head Neck Surg. 2001;53(1):55-6.
2. Treviño González JL, Reyes Suárez LL, Hernández de León JE. Malignant otitis externa: An updated review. Am J Otolaryngol. 2021;42(2):102894.

3. Sylvester MJ, Sanghvi S, Patel VM, Eloy JA, Ying YM. Malignant otitis externa hospitalizations: Analysis of patient characteristics. Laryngoscope. 2017;127(10):2328-36.
4. Slattery WH, 3rd, Brackmann DE. Skull base osteomyelitis. Malignant external otitis. Otolaryngol Clin North Am. 1996; 29(5):795-806.
5. Rubin J, Yu VL. Malignant external otitis: insights into pathogenesis, clinical manifestations, diagnosis, and therapy. Am J Med. 1988;85(3):391-8.
6. Stern Shavit S, Soudry E, Hamzany Y, Nageris B. Malignant external otitis: Factors predicting patient outcomes. Am J Otolaryngol. 2016;37(5):425-30.
7. Hamzany Y, Soudry E, Preis M, Hadar T, Hilly O, Bishara J, et al. Fungal malignant external otitis. J Infect. 2011;62(3): 226-31.
8. Al Aaraj MS, Kelley C. Necrotizing (Malignant) Otitis Externa. StatPearls. Treasure Island (FL) ineligible companies. Disclosure: Cecylia Kelley declares no relevant financial relationships with ineligible companies.: StatPearls Publishing Copyright© 2024, StatPearls Publishing LLC.; 2024.
9. Mahdyoun P, Pulcini C, Gahide I, Raffaelli C, Savoldelli C, Castillo L, et al. Necrotizing otitis externa: a systematic review. Otol Neurotol. 2013;34(4):620-9.
10. Takata J, Hopkins M, Alexander V, Bannister O, Dalton L, Harrison L, et al. Systematic review of the diagnosis and management of necrotising otitis externa: Highlighting the need for high-quality research. Clin Otolaryngol. 2023;48(3): 381-94.
11. Levenson MJ, Parisier SC, Dolitsky J, Bindra G. Ciprofloxacin: drug of choice in the treatment of malignant external otitis (MEO). Laryngoscope. 1991;101(8):821-4.
12. Franco-Vidal V, Blanchet H, Bebear C, Dutronc H, Darrouzet V. Necrotizing external otitis: a report of 46 cases. Otol Neurotol. 2007;28(6):771-3.
13. Glikson E, Sagiv D, Wolf M, Shapira Y. Necrotizing otitis externa: diagnosis, treatment, and outcome in a case series. Diagn Microbiol Infect Dis. 2017;87(1):74-8.
14. Joshua BZ, Sulkes J, Raveh E, Bishara J, Nageris BI. Predicting outcome of malignant external otitis. Otol Neurotol. 2008;29(3): 339-43.
15. Chen JC, Yeh CF, Shiao AS, Tu TY. Temporal bone osteomyelitis: the relationship with malignant otitis externa, the diagnostic dilemma, and changing trends. ScientificWorldJournal. 2014; 2014:591714.
16. Arsovic N, Radivojevic N, Jesic S, Babac S, Cvorovic L, Dudvarski Z. Malignant Otitis Externa: Causes for Various Treatment Responses. J Int Adv Otol. 2020;16(1):98-103.
17. Stevens SM, Lambert PR, Baker AB, Meyer TA. Malignant Otitis Externa: A Novel Stratification Protocol for Predicting Treatment Outcomes. Otol Neurotol. 2015;36(9):1492-8.
18. Soudry E, Joshua BZ, Sulkes J, Nageris BI. Characteristics and Prognosis of Malignant External Otitis With Facial Paralysis. Arch Otolaryngol Head Neck Surg. 2007;133(10):1002-4.
19. Hatch JL, Bauschard MJ, Nguyen SA, Lambert PR, Meyer TA, McRackan TR. Malignant Otitis Externa Outcomes: A Study of the University HealthSystem Consortium Database. Ann Otol Rhinol Laryngol. 2018;127(8):514-20.
20. Yigider AP, Ovunc O, Arslan E, Sunter AV, Cermik TF, Yigit O. Malignant Otitis Externa: How to Monitor the Disease in Outcome Estimation? Medeni Med J. 2021;36(1):23-9.
21. Ahmed AA, Rashid S, Gupta VK, Molony NC, Gupta KK. The diagnostic conundrum in necrotizing otitis externa. World J Otorhinolaryngol Head Neck Surg. 2024;10(1):59-65.
22. Kim DH, Kim SW, Hwang SH. Predictive value of radiologic studies for malignant otitis externa: a systematic review and meta-analysis. Braz J Otorhinolaryngol. 2023;89(1):66-72.
23. Singh J, Bhardwaj B. The Role of Surgical Debridement in Cases of Refractory Malignant Otitis Externa. Indian J Otolaryngol Head Neck Surg. 2018;70(4):549-54.
24. Raines JM, Schindler RA. The surgical management of recalcitrant malignant external otitis. Laryngoscope. 1980;90(3):369-78.
25. Chawdhary G, Pankhania M, Douglas S, Bottrill I. Current management of necrotising otitis externa in the UK: survey of 221 UK otolaryngologists. Acta Otolaryngol. 2017;137(8): 818-22.

Incidence of Tympanic Membrane Perforation Affected by Intratympanic Steroid Injection: A Retrospective Review

Vanthana Charoensawatsiri, M.D., Sarun Prakairungthong, M.D., Suvajana Atipas, M.D., Kanthong Thongyai, M.D., Siriporn Limviriyakul, M.D., Phoupong Phousamran, M.D., Kanokrat Suvarnsit, M.D.*

Department of Otorhinolaryngology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.



*Corresponding author: Kanokrat Suvarnsit
E-mail: kanokrat.suv@gmail.com
Received 6 December 2024 Revised 7 January 2025 Accepted 15 January 2025
ORCID ID:http://orcid.org/0000-0003-3112-3911
https://doi.org/10.33192/smj.v77i4.272595

ABSTRACT

Objective: This study aims to determine the incidence of persistent tympanic membrane perforation following intratympanic steroid injection and to identify potential factors associated with delayed healing. Additionally, it aims to estimate the time required for perforation closure in prolonged cases to avoid unnecessary interventions.

Materials and Methods: Data from patients who underwent intratympanic steroid injections were reviewed. The primary outcome was the incidence of tympanic membrane perforation lasting beyond four weeks post-injection. Secondary outcomes included identifying factors affecting healing duration and closure time in prolonged cases.

Results: Of 295 ears treated between March 2018 and March 2021, 3.39% (10/295) experienced persistent perforation at four weeks. Of these, 1.69% (5/295) required intervention, while the rest healed spontaneously. The median closure time was 12 weeks. All patients with persistent perforation were female. No statistically significant differences were found between groups. Hearing recovery was achieved in 29.1% (74/254) of patients with sudden sensorineural hearing loss. Younger age and non-profound hearing loss were favorable prognostic factors.

Conclusion: The incidence of perforation in this study is lower than that previously reported but consistent with other studies in the literature. Intratympanic steroid injections remain beneficial, outweighing the risks of complications.

Keywords: Tympanic membrane perforation; intratympanic steroid injection; sudden sensorineural hearing loss (Siriraj Med J 2025; 77: 267-277)

INTRODUCTION

Idiopathic sudden sensorineural hearing loss (ISSNHL) is a critical condition in otological practice, affecting individuals of all ages, with incidence rates increasing with age. Globally, ISSNHL affects 7 to 27 individuals per 100,000 population, though its incidence in Thailand remains undocumented.[1,2] Observations from our neuro-otology clinic between 2006 and 2008 revealed that 7.67% of patients were diagnosed with sudden sensorineural hearing loss (SSNHL).[3] While some cases were non-idiopathic, such as those associated with Meniere's disease, the management approach remains similar.

Systemic corticosteroid has long been the primary treatment for hearing loss, with oral corticosteroids being a common systemic approach.[4] However, systemic corticosteroid use is associated with adverse effects, including weight gain, sleep disturbances, gastrointestinal symptoms, and an increased risk of infection. Additionally, systemic steroids pose limitations for certain patient groups, such as those with uncontrolled diabetes mellitus or a history of peptic ulcer disease. As a result, local steroid application has emerged as an alternative approach.

The intratympanic administration of steroids was first introduced by *Itoh et al.* in 1991 for Meniere's disease, and later applied by *Silverstein et al.* in 1996 for ISSNHL.[5,6] This method allows higher concentrations of steroids to reach the inner ear, reducing systemic adverse effects. Numerous studies have since demonstrated its effectiveness in treating ISSNHL.[7-11] Intratympanic steroid injection, using steroids such as methylprednisolone and dexamethasone, has been used to treat various conditions, including Meniere's disease and labyrinthitis, serving as both first-line and salvage therapy.[12]

While generally safe, intratympanic injections can cause complications, including transient pain, dizziness, or tinnitus.[13] Persistent tympanic membrane perforation is the most concerning side effect, though studies report a low incidence of 1-2%.[14-16] However, a retrospective study in Thailand found a higher incidence rate of 7.18%, potentially due to differences in definition and criteria used.[17] This study aims to determine the incidence of tympanic membrane perforation at 4 weeks post-intratympanic steroid injection and to investigate potential risk factors, such as steroid type, needle gauge, and injection frequency, that may affect tympanic membrane perforation.

MATERIALS AND METHODS

This study was approved by the Institutional Review Board (IRB), with approval code COA no. Si 218/2022.

Sample size

The sample size was calculated using an estimation for categorical outcome with the following formula.

$$n = \frac{z^2_{1-\frac{\alpha}{2}}\, p(1-p)}{d^2}$$

p = expected proportion in population based on literature review[16] = 0.0114
z = standard normal variate P values < 0.05
d = absolute error = 0.0126

The calculated sample size was 273. After adjusting for a 10% allowance for missing data, the final expected N was 300.

Study population

A retrospective review was conducted on individuals who received intratympanic steroid injections at a tertiary care center from March 2018 to March 2021. Patients were identified using the International Classification of Diseases (ICD-10) code for "injection of tympanum" from the electronic database. Indications for steroid injections included ISSNHL, Meniere's disease, and labyrinthitis. Subjects with incomplete data, those lost to follow-up, tympanic membrane perforation at the injected or ventilation site before injection, and those who received gentamycin injections were excluded from the study.

Outcome assessment

The primary outcome was the percentage of persistent tympanic membrane perforations, defined as a perforation present at least 4 weeks after the last injection, confirmed through otoscopic or microscopic examination. The healing time in the protracted closure group was analyzed descriptively in weeks. Factors such as age, gender, diabetes mellitus, smoking, radiation exposure, type of steroid, number of injections, and prior systemic steroid use were considered as potential predictors. These factors were statistically analyzed to compare outcomes between the spontaneous tympanic membrane closure group and the perforation group.

Statistical analysis

The perforation rate was reported as a percentage, with comparisons between the normal and the perforation groups made using the Mann-Whitney U test and Fisher Exact test for continuous and categorical variables, respectively. A p-value <0.05 was considered statistically significant. Data collection and statistical analysis were conducted using SPSS version 22.

Procedure

After obtaining informed consent, the patient was positioned in the supine position with the head tilted toward the unaffected side. Local anesthesia with lidocaine was applied to the external ear canal. Under microscopic visualization, a ventilation hole was created at the anterosuperior portion of the tympanic membrane using a 25G needle tip. Subsequently, the steroid was then injected into the posteroinferior part of the drum, using the same 25G needle. Each injection administers 0.4-0.8 ml of steroid into the middle ear cavity. Patients were instructed to maintain in a lateral position for 30 minutes, without talking or swallowing. The injection was administered up to four times within a two-week period for one course. However, some subjects underwent fewer than four injections due to complications, but these patients were still included in the study. Hearing evaluations, including an audiogram, were conducted at initial diagnosis and repeated within 12 weeks after the last injection.

Operational definition

- Idiopathic sudden sensorineural hearing loss (ISSNHL) is defined as a hearing loss of ≥ 30 dB across at least three consecutive frequencies within a 72-hour period, following the clinical practice guidelines for sudden hearing loss published in 2019.[18]
- Recovery of SSNHL was assessed based on improvement in pure tone average (PTA; the mean hearing level at 500, 1,000, 2,000, and 4,000 Hz), and is categorized into three group according to the clinical practice guidelines[18]:
  - Complete recovery: PTA within 10 dB of the unaffected ear and WRS within 5-10% of the unaffected ear
  - Partial recovery: Defined based on hearing levels during the SSNHL episode.
    - Nonserviceable level: Return to serviceable hearing (PTA ≤ 50 dB and Word recognition score ≥ 50%)
    - Serviceable level: Improvement of PTA by ≥ 10-dB HL or WRS ≥ 10%
  - No Recovery: <10-dB HL improvement
- Persistent perforation is defined as a perforation present 4 weeks or more after the last injection, confirmed through otoscopic or microscopic examination.

RESULTS

A total of 320 subjects coded for the 'injection of tympanum' procedure, as per ICD-10 criteria, were initially recruited from the electronic database for this study. However, 42 subjects were excluded for various reasons: 27 had inadequate follow-up, 10 received injections with different medications, such as gentamicin, two were excluded due to incorrect ICD coding, and three had incomplete data. Ultimately, this left 278 subjects who met the study's inclusion criteria.

Among the 278 subjects, 15 received more than one course of injections. Of these, three subjects presented with bilateral SSNHL, resulting in steroid injections

administered in both ears. Additionally, one subject, with Meniere's disease, underwent three courses of steroid injections over the study period. Furthermore, 11 patients experienced fluctuating hearing loss, with 10 receiving two courses of injections, and 1 receiving three courses. In total, 295 ears from 278 subjects were included in the analysis.

The demographic characteristics of the study population are summarized in Table 1. Females made up 63.1% of the subjects, with a median age of 58 years (range: 9 to 86 years). The median time from onset of hearing loss to treatment was 7 days (range: 1 to 60 days). At the onset of hearing loss, the median pure tone average (PTA) was 65 dB (range: 10 to 120 dB). The primary indication for intratympanic steroid injections was SSNHL, accounting for 86% of cases, while the remaining cases included Meniere's disease (13%) and labyrinthitis (1%).

TABLE 1. Demographic data.

Characteristic	Number of ears (%)
Age (years; min, max)	58 (9, 86)
Gender	
Male	109 (36.9)
Female	186 (63.1)
Diabetes mellitus	
DM	79 (26.8)
Non-DM	216 (73.2)
Side of affected ear	
Right	160 (54.2)
Left	135 (45.8)
Objective	
Primary	104 (35.3)
Salvage	191 (64.7)
Type of steroid	
Methylprednisolone	257 (87.1)
Dexamethasone	33 (11.2)
Both	5 (1.7)
Onset of hearing loss (days; min, max)	7 (1, 60)
PTA before treatment	
< 80 dB	193 (65.4)
≥ 80 dB	102 (34.6)

Abbreviations: DM, Diabetes mellitus; PTA, Pure tone average

At 4 weeks after the final injection , 10 subjects (3.39%) exhibited persistent tympanic membrane perforation. Of these, five cases (50%) achieved spontaneous closure, though the healing period extended beyond the 4-week cutoff, with a median healing time of 12 weeks (range: 6 to 50 weeks). The remaining five subjects required procedural intervention to close the perforation. Three of these subjects underwent office-based paper patch myringoplasty, one underwent tympanoplasty, and one subject, despite having a persistent perforation, declined any intervention.

All subjects in the persistent perforation group were female. Other factors did not show statistically significant differences between the spontaneous closure group and those with persistent perforation, as shown in Table 2. Ex-smokers who had abstained for more than 4 weeks were classified as non-smokers, as previous literature suggests that smoking cessation for at least four weeks improves wound healing.[19-21] None of the subjects in the perforation group had a history of radiation exposure to the head and neck region. Methylprednisolone was the most commonly used steroid in both the perforation group (80%) and the normal-healing group (87.4%). Although a full course of treatment generally involves four injections, eight subjects (2.8%) in the spontaneous closure group received fewer than four injections. No significant differences were observed between subjects who required procedural intervention for tympanic membrane perforation and those who did not, including those without perforation and those with delayed spontaneous closure, as shown in Table 3. Detailed data on subjects with perforations persisting beyond 4 weeks are presented in Table 4.

All subjects in the group requiring procedural intervention for perforation received four injections, with 90% having a history of systemic steroid use. Among these 10 cases, 8 (80%) received methylprednisolone injections, while two (20%) received dexamethasone. Both subjects who received dexamethasone had large perforations, measuring 20% and 30%, respectively. One notable case involved a 59-year-old female with no underlying disease, who presented with SSNHL and underwent tympanoplasty after receiving four doses of dexamethasone as salvage treatment. Her tympanic membrane perforation measured 20% at the injection site and did not close after 27 weeks of follow-up.

Hearing recovery rates were assessed in 254 subjects for whom SSNHL was the indication for intratympanic steroid injection. The overall recovery rate was 29.1% (74/254), with 14.6% (37/254) achieving complete recovery and 14.6% (37/254) partial recovery. The remaining 70.9%

TABLE 2. Comparison of potential factors associated with persistent TM perforation and spontaneous TM closure at 4 weeks after last injection.

	Persistent TM perforation (N = 10) Number of ears (%)	Spontaneous TM closure (N = 285) Number of ears (%)	P value	OR (95%CI)
Age (years; min, max)	57.5 (42,77)	58 (9,86)	0.550	-
Gender				
Female	10 (100)	176 (61.8)	**0.015**	-
Male	0 (0)	109 (38.2)		
Diabetes mellitus				
DM	2 (20)	77 (27)	1.000	0.68
Non-DM	8 (80)	208 (73)		(0.14-3.25)
Smoking*				
Current smoker	0 (0)	4 (2.6)	1.000	-
Non-smoker or quit > 4 weeks	5 (100)	152 (97.4)		
Radiation (Head and neck)				
Yes	0 (0)	3 (1.1)	1.000	-
No	10 (100)	282 (98.9)		
Prior systemic steroid				
Yes	9 (90)	186 (65.3)	0.173	4.79
No	1 (10)	99 (34.7)		(0.60-38.36)
Type of steroid				
Dexamethasone	2 (20)	36 (12.6)	0.623	1.73
Methylprednisolone	8 (80)	249 (87.4)		(0.35-8.47)
Number of injections				
4 times	10 (100)	277 (97.2)	1.000	-
< 4 times	0 (0)	8 (2.8)		

*missing data 45.4%

Abbreviations: TM, Tympanic membrane; DM, Diabetes mellitus; OR, Odd ratio

(180/254) showed no recovery. Fig 1 illustrates a subgroup analysis of recovery rates, categorizing subjects into two groups based on initial hearing levels: serviceable and non-serviceable hearing. The serviceable hearing group had an overall recovery rate of 44% (29/66), with 32% (21/66) achieving complete recovery and 12% (8/66) achieving partial recovery. In contrast, the non-serviceable group had a complete recovery rate of 9% (16/188) and a partial recovery rate of 15% (29/188), with an overall recovery rate of 24% (45/188).

The mean improvement in pure tone average within 12 weeks after the last injection was 11.55 ± 20.57 dB. Factors potentially associated with hearing recovery are presented in Table 5. Age and initial hearing level were found to be significantly associated with hearing recovery. Subjects aged 60 years and older had a relative risk of non-recovery hearing of 1.30 (95% CI: 1.11 to 1.53) compared to younger subjects. Those with profound hearing loss at the onset had a higher risk of non-recovery, with a relative risk of 1.35 (95% CI: 1.17 to 1.57) compared to those with an initial PTA of less than 80 dB. Gender, underlying conditions such as diabetes mellitus, type of steroid used, the number of injections, and onset of hearing loss were not significantly associated with recovery.

TABLE 3. Comparison of potential factors associated with TM perforation in subjects undergoing the procedure and no procedure group.

	TM perforation underwent procedure (N = 5) Number of ears (%)	No procedure (N = 290) Number of ears (%)	P value	OR (95%CI)
Age (years; min, max)	59 (51,72)	58 (9,86)	0.621	-
Gender				
Female	5 (100)	181 (62.4)	0.162	-
Male	0 (0)	109 (37.6)		
Diabetes mellitus				
DM	1 (20)	78 (26.9)	1.000	0.68
Non-DM	4 (80)	212 (73.1)		(0.08-6.17)
Smoking*				
Current smoker	0 (0)	4 (2.5)	1.000	-
Non smoker or quit > 4 weeks	2 (100)	155 (97.5)		
Radiation (Head and neck)				
Yes	0 (0)	3 (1)	1.000	-
No	5 (100)	287 (99)		
Prior systemic steroid				
Yes	5 (100)	190 (65.5)	0.171	-
No	0 (0)	100 (34.5)		
Type of steroid				
Dexamethasone	2 (40)	36 (12.4)	0.126	4.70
Methylprednisolone	3 (60)	254 (87.6)		(0.76-29.12)
Number of injection				
4 times	5 (100)	282 (97.2)	1.000	-
< 4 times	0 (0)	8 (2.8)		

Abbreviations: TM, Tympanic membrane; DM, Diabetes mellitus

DISCUSSION

Steroid administration via intratympanic injection is an effective treatment for ISSNHL and other inner ear conditions such as Meniere's disease and labyrinthitis. However, this procedure can lead to complications, the most significant being tympanic membrane perforation. Previous studies including retrospective reviews, systematic reviews, and meta-analyses, have documented the rate of tympanic membrane perforation following intratympanic injections to range between 1.14% to 2.9%.[14-16]

In Thailand, a study by *Limviriyakul et al.* using a retrospective chart review at a tertiary care hospital and reported a perforation rate of 7.18%, significantly higher than rates observed in international studies.[17] This discrepancy may be attributed to differences in criteria for defining persistent perforation or the type of steroid used. In our current study, the incidence of persistent tympanic membrane perforation was 3.39% across 295 ears, which is lower than Limviriyakul's findings but more closely aligns with rates reported in international literature.[14-17]

Several factors could explain the lower perforation rate observed in this study. First, intratympanic steroid injections have been increasingly used in our clinical setting over the past few years, and the sample size for this study was calculated to adequately evaluate the rate of tympanic membrane perforation. Secondly, Limviriyakul's study did not specify the timeframe for evaluating tympanic

TABLE 4. Details of subjects with persistent perforation more than 4 weeks after last injection.

	Age (years)	Sex	Risk	Disease	Prior steroid	Injection Count (times)	Steroid type	Size	Site	Duration (weeks)	Procedure
1	54	F	None	SSNHL	Yes	4	Methylpred	pinpoint	Injected	23	Paper patch
2	51	F	None	SSNHL	Yes	4	Methylpred	pinpoint	Injected	12	No
3	77	F	DM (HbA1C = 6)	SSNHL	Yes	4	Methylpred	pinpoint	Ventilation	16	No
4	59	F	None	SSNHL	Yes	4	Dexa	20%	Injected	27	T-plasty
5	73	F	None	SSNHL	No	4	Methylpred	pinpoint	Injected	12	No
6	51	F	None	SSNHL	Yes	4	Dexa	30%	Injected	9	Paper patch
7	72	F	DM (HbA1C = 8.1)	SSNHL	Yes	4	Methylpred	20%	Injected	*	Follow up
8	56	F	None	Meniere	Yes	4	Methylpred	pinpoint	Ventilation	50	No
9	42	F	None	SSNHL	Yes	4	Methylpred	10%	Injected	6	No
10	67	F	None	Meniere	Yes	4	Methylpred	pinpoint	Injected	7	Paper patch

Abbreviations: F, Female; DM, Diabetes mellitus; SSNHL, Sudden sensorineural hearing loss; Methlypred, Methylprednisolone; Dexa, Dexamethasone

*Subject with persistent perforation refused intervention

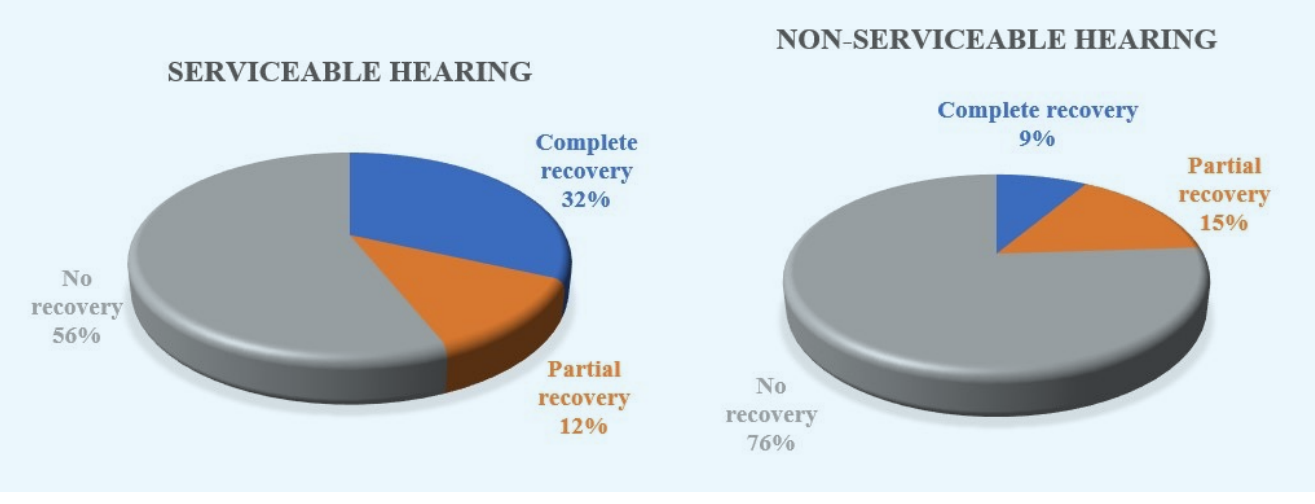


Fig 1. Recovery rates in initial serviceable hearing vs non-serviceable hearing group.

TABLE 5. Comparison of potential factors associated with recovery vs. non-recovery groups

	Non-recovery (N = 180) Number of ears (%)	Recovery (N = 74) Number of ears (%)	P value	Relative risk (95% CI)
Age				
≥ 60 years	99 (80.5)	24 (19.5)	**0.001**	**1.30**
< 60 years	81 (61.8)	50 (38.2)		**(1.11 to 1.53)**
Gender				
Female	111 (68.9)	50 (31.1)	0.375	0.93
Male	69 (74.2)	24 (25.8)		(0.79 to 1.09)
Underlying disease				
Yes	110 (75.3)	36 (24.7)	0.068	1.16
No	70 (64.8)	38 (35.2)		(0.98 to 1.37)
Diabetes mellitus				
DM	54 (74)	19 (26)	0.489	1.06
Non-DM	126 (69.6)	55 (30.4)		(0.90 to 1.26)
Type of steroid				
Dexamethasone	19 (70.4)	8 (29.6)	0.952	0.99
Methylprednisolone	161 (70.9)	66 (29.1)		(0.77 to 1.28)
Objective				
Salvage	119 (69.6)	52 (30.4)	0.521	0.95
Primary	61 (73.5)	22 (26.5)		(0.80 to 1.11)
Number of injections				
4 times	175 (70.9)	72 (29.1)	1.000	0.99
< 4 times	5 (71.4)	2 (28.6)		(0.62 to 1.60)
Onset of hearing loss*				
>14 days	42 (79.2)	11 (20.8)	0.112	1.17
≤ 14 days	134 (68)	63 (32)		(0.99 to 1.38)
PTA before injection				
≥ 80 dB	82 (84.5)	15 (15.5)	**<0.001**	**1.35**
< 80 dB	98 (62.4)	59 (37.6)		**(1.17 to 1.57)**

*missing data 1.6%

Abbreviations: DM, Diabetes mellitus; PTA, Pure tone average

membrane status post-injection, whereas this study assigned a clear 4-week post-injection evaluation period. In the literature, timeframes for evaluating persistent tympanic membrane perforations vary, ranging from 4 to 12 weeks.[14,16] In this study, the evaluation was set at 4 weeks or later after the last injection, based on findings from *Simani et al.'s* systemic review, which indicated that most perforations would heal within this period.[16] However, our observations indicated a median healing time of 12 weeks, with three subjects initially categorized as having persistent perforations eventually experiencing spontaneous healing within 12 weeks. This suggests that a 12-week evaluation period may be more appropriate for assessing persistent perforation in clinical practice.

Among the 10 subjects in the perforation group, five required procedural intervention to close the perforation, resulting in an incidence rate of 1.69% for persistent tympanic membrane perforations requiring a procedure.

However, two of these patients underwent interventions at 7- and 9-weeks post-injection, and it is possible that their perforations might have healed spontaneously with a longer observation period. Predictors of persistent perforation are divided into two categories: host-related factors affecting immune status and wound healing, and procedure-related factors. The first category included diabetes mellitus, smoking, radiation exposure, and HIV infection, while the second group involved variables such as prior steroid usage, number of injections, and type of steroid. These factors were analyzed by comparing the persistent perforation group with the spontaneous closure group, as well as the procedural and non-procedural groups.

In the first group of factors, gender was the only variable that showed a statistically significant difference between the persistent perforation group and the spontaneous closure group. This finding contradicts other studies, which found no association between gender and perforation healing.[22-24] This discrepancy is likely coincidental and not clinically relevant.[14,15] Other factors, including smoking, radiation, and HIV status, showed no statistical significance, consistent with previous literature. Due to incomplete medical records for 45.4% of subjects, smoking data could not be fully analyzed, as detailed smoking histories were rarely documented in outpatient visits. In the second group of procedural factors, there was no statistical difference between the persistent perforation and spontaneous closure groups, nor between the procedural and non-procedural groups. This may be attributed to the relatively low incidence of perforation in this cohort.

Most of the studies in the systemic literature reviews used dexamethasone for injections, whereas in our practice, methylprednisolone has been the preferred steroid since 2007 due to its higher concentration and longer duration of action.[25,26] Although we hypothesized that steroid type might influence the perforation rate, this study found no statistical association between steroid type and perforation. However, Table 5 shows that two subjects who received dexamethasone injections had larger perforations that did not heal spontaneously. Topf's study suggested that multiple injections might increase the risk of persistent perforation.[14] In this study, most subjects received four injections, and while eight subjects in the spontaneous closure group received fewer than four, all subjects in the persistent perforation group received the full four injections. This may lend some support to the hypothesis that multiple injections could increase perforation risk, though the small number of persistent perforations limits the strength of this conclusion.

Diabetes mellitus was expected to be associated with delayed perforation healing. Although this study showed no significant association between diabetes and perforation healing, subjects with poorly controlled diabetes appeared to have a higher risk of persistent perforation, which is consistent with previous studies.[14,22,27] Two subjects in the persistent perforation group had diabetes. One had well-controlled diabetes with an HbA1c level of 6.0 and experienced perforation healing within 16 weeks. The other subject, with poorly controlled diabetes and an HbA1c level of 8.1, did not experience spontaneous healing.

In addition to evaluating perforation rates, this study also assessed hearing recovery outcomes. According to the 2019 Clinical Practice Guideline, the prognosis for hearing recovery depends on factors such as age, the presence of vertigo at onset, degree of hearing loss, audiometric configuration, and the duration between onset and treatment.[18] In this study, factors including age, hearing level, and onset-to-treatment duration were evaluated, along with additional factors like gender, underlying disease, diabetes mellitus, type of steroid used, purpose of injection, and number of injections. Subjects aged 60 years or older, and those with an initial PTA of greater than 80 dB, had a higher risk of no hearing recovery after the injection.[17,28-31] These findings align with other studies that identified age and degree of hearing loss at onset as poor prognostic factors. However, factors such as onset of hearing loss and underlying disease, often considered relevant to hearing recovery, were not statistically significant in this study.[17,28] This may be due to the study's sample size, which was calculated to evaluate tympanic membrane perforation repair rates, potentially limiting its power to detect smaller differences in hearing recovery.

The total overall hearing recovery rate in subjects with SSNHL was 29.1%, a rate comparable to the spontaneous recovery rates reported in the literature.[29,32] A subgroup analysis showed a 44% recovery rate in subjects with serviceable hearing at onset, compared to a 24% recovery rate in those with non-serviceable hearing, further supporting the degree of hearing loss at onset as a critical prognostic factor for recovery.

Limitation

A previous study by *Simani et al* reported time to perforation healing in days after injection and compared healing duration between risk and non-risk groups.[16] In this study, healing time was recorded in weeks, as patients were scheduled for weekly follow-ups in our practice. Data on healing duration for the spontaneous perforation

closure group were unavailable, as most patients were seen for tympanic membrane examination 1-2 weeks post-injection, depending on the physician's discretion. At the first post-treatment visit, medical records noted either a normal tympanic membrane or a hematoma at the injection site without perforation, making It difficult to determine the exact timing of perforation closure in the spontaneous perforation closure group.

Previous ear conditions, including prior trauma, history of infection, and the status of the eustachian tube, may influence perforation healing time. However, due to the retrospective nature of this study, data on these factors are limited and therefore excluded from the analysis.

Expected Benefit and application

The findings from this study enable physicians to better counsel patients on the expected outcomes of intratympanic injection procedures. Comprehensive medical advice, including potential complications and anticipated recovery rates, can be provided and included in the consent form prior to the procedure. For future studies, a 12-week evaluation period after the final injection is recommended for assessing tympanic membrane status.

CONCLUSION

The rate of persistent tympanic membrane perforation following intratympanic steroid injection is low, indicating a favorable risk-benefit ratio. Based on the data gathered from this study, 12 weeks appears to be the optimal timeframe for assessing tympanic membrane status, serving as a cutoff to differentiate subjects with and without perforations. Regarding hearing recovery, younger patients and those with less severe hearing loss at onset were more likely to experience recovery following intratympanic injection.

Data Availability Statement

The data that support the findings of this study are available upon request from the corresponding author, [K.S.]. The data are not publicly available due to containing information that could compromise the privacy of research participants.

ACKNOWLEDGEMENTS

We extend our heartfelt appreciation to Asst. Prof. Dr. Chulaluk Komoltri, our consulting statistician, and Ms Jeerapa Kerdnoppakhun for their assistance in developing this manuscript, as well as all patients involved in this study.

DECLARATION

Grants and Funding Information

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

Conflict of Interest

The authors have no conflicts of interest to disclose.

Registration number of clinical trial

None

Author Contributions

S.P.; Conceptualization, Methodology, Investigation, Data Curation, Writing - Original Draft, Review & Editing, Supervision V.C.; Methodology, Data Curation, Formal analysis, Writing - Original Draft P.P.; Supervision,Writing - Review & Editing, Validation S.A.; Methodology, Writing - Review & Editing, Validation K.T.; Conceptualization, Resources, Formal analysis, Writing - Original Draft, Validation S.L.; literature review, Methodology, Formal analysis, Writing - Original Draft, Validation K.S.; Conceptualization, Data Curation, Writing - Review & Editing.

Use of Artificial Intelligence

None

REFERENCES

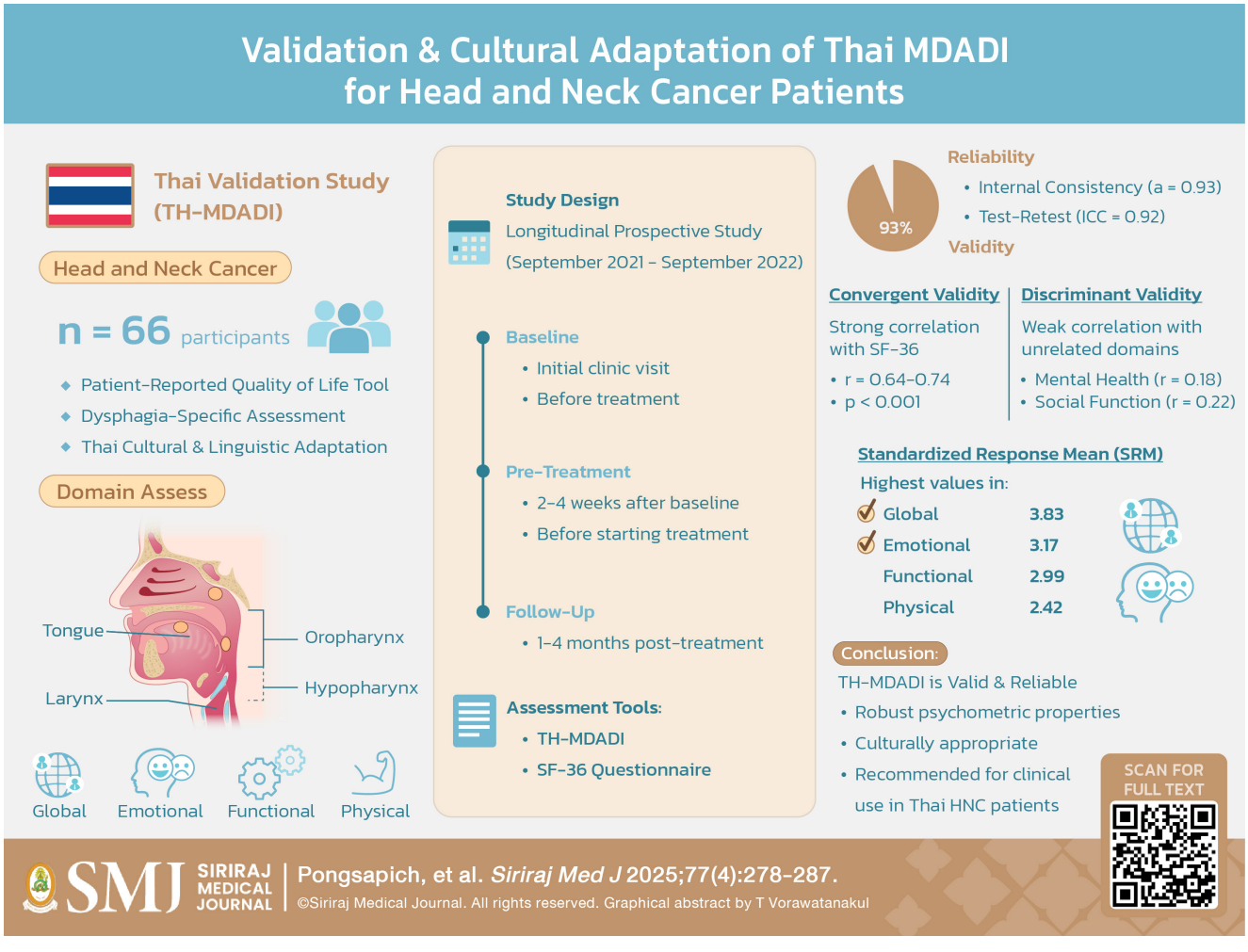
1. Alexander TH, Harris JP. Incidence of sudden sensorineural hearing loss. Otol Neurotol. 2013;34(9):1586-9.
2. Wu CS, Lin HC, Chao PZ. Sudden sensorineural hearing loss: evidence from Taiwan. Audiol Neurootol. 2006;11(3):151-6.
3. Swain SK, Das S, Lenka S. Sudden Sensorineural Hearing Loss among COVID-19 Patients-Our Experiences at an Indian Teaching Hospital. Siriraj Medical Journal. 2021;73(2):77-83.
4. Wilson WR, Byl FM, Laird N. The efficacy of steroids in the treatment of idiopathic sudden hearing loss. A double-blind clinical study. Arch Otolaryngol. 1980;106(12):772-6.
5. Itoh A, Sakata E. Treatment of vestibular disorders. Acta Otolaryngol Suppl. 1991;481:617-23.
6. Silverstein H, Choo D, Rosenberg SI, Kuhn J, Seidman M, Stein I. Intratympanic steroid treatment of inner ear disease and tinnitus (preliminary report). Ear Nose Throat J. 1996;75(8):468-71, 74, 76 passim.
7. Chandrasekhar SS. Intratympanic dexamethasone for sudden sensorineural hearing loss: clinical and laboratory evaluation. Otol Neurotol. 2001;22(1):18-23.
8. Ho HG, Lin HC, Shu MT, Yang CC, Tsai HT. Effectiveness of intratympanic dexamethasone injection in sudden-deafness patients as salvage treatment. Laryngoscope. 2004;114(7):1184-9.
9. Xenellis J, Papadimitriou N, Nikolopoulos T, Maragoudakis P, Segas J, Tzagaroulakis A, et al. Intratympanic steroid treatment in idiopathic sudden sensorineural hearing loss: a control study. Otolaryngol Head Neck Surg. 2006;134(6):940-5.

10. Kiliç R, Safak MA, Oğuz H, Kargin S, Demirci M, Samim E, et al. Intratympanic methylprednisolone for sudden sensorineural hearing loss. Otol Neurotol. 2007;28(3):312-6.
11. Filipo R, Covelli E, Balsamo G, Attanasio G. Intratympanic prednisolone therapy for sudden sensorineural hearing loss: A new protocol. Acta Otolaryngol. 2010;130(11):1209-13.
12. Lavigne P, Lavigne F, Saliba I. Intratympanic corticosteroids injections: a systematic review of literature. Eur Arch Otorhinolaryngol. 2016;273(9):2271-8.
13. Liu YC, Chi FH, Yang TH, Liu TC. Assessment of complications due to intratympanic injections. World J Otorhinolaryngol Head Neck Surg. 2016;2(1):13-6.
14. Topf MC, Hsu DW, Adams DR, Zhan T, Pelosi S, Willcox TO, et al. Rate of tympanic membrane perforation after intratympanic steroid injection. Am J Otolaryngol. 2017;38(1):21-5.
15. Kim YH, Lee DY, Lee DH, Oh S. Tympanic Membrane Perforation After Intratympanic Steroid Injection: A Systematic Review and Meta-analysis. Otolaryngol Head Neck Surg. 2022;166(2):249-59.
16. Simani L, Shilo S, Oron Y, Eta RA, Handzel O, Muhanna N, et al. Residual Perforation Risk Assessment of Intratympanic Steroids via Tympanostomy Tube Versus Transtympanic Injections. Laryngoscope. 2021;131(9):E2583-e91.
17. Limviriyakul S. TK, Prakairungthong S, Atipas S, Chongvisal S. Methylprednisolone intratympanic injection in idiopathic sudden sensorineural hearing loss: a retrospective review of 167 cases. J Med Assoc Thai. 2018(101(6)):707-12.
18. Chandrasekhar SS, Tsai Do BS, Schwartz SR, Bontempo LJ, Faucett EA, Finestone SA, et al. Clinical Practice Guideline: Sudden Hearing Loss (Update). Otolaryngol Head Neck Surg. 2019;161(1_suppl):S1-s45.
19. Kuri M, Nakagawa M, Tanaka H, Hasuo S, Kishi Y. Determination of the duration of preoperative smoking cessation to improve wound healing after head and neck surgery. Anesthesiology. 2005;102(5):892-6.
20. Chan LK, Withey S, Butler PE. Smoking and wound healing problems in reduction mammaplasty: is the introduction of urine nicotine testing justified? Ann Plast Surg. 2006;56(2):111-5.
21. Mills E, Eyawo O, Lockhart I, Kelly S, Wu P, Ebbert JO. Smoking cessation reduces postoperative complications: a systematic review and meta-analysis. Am J Med. 2011;124(2):144-54.e8.
22. Chen SL, Yang SW. Factors affecting the treatment outcomes of myringoplasty in patients with small tympanic membrane perforations. Eur Arch Otorhinolaryngol. 2019;276(11):3005-12.
23. Tachibana T, Kariya S, Orita Y, Makino T, Haruna T, Matsuyama Y, et al. Spontaneous closure of traumatic tympanic membrane perforation following long-term observation. Acta Otolaryngol. 2019;139(6):487-91.
24. Lou Z, Yang J, Tang Y, Xiao J. Risk factors affecting human traumatic tympanic membrane perforation regeneration therapy using fibroblast growth factor-2. Growth Factors. 2015;33(5-6):410-8.
25. Banerjee A, Parnes LS. The biology of intratympanic drug administration and pharmacodynamics of round window drug absorption. Otolaryngol Clin North Am. 2004;37(5):1035-51.
26. Parnes LS, Sun AH, Freeman DJ. Corticosteroid pharmacokinetics in the inner ear fluids: an animal study followed by clinical application. Laryngoscope. 1999;109(7 Pt 2):1-17.
27. Vrabec JT. Tympanic membrane perforations in the diabetic rat: a model of impaired wound healing. Otolaryngol Head Neck Surg. 1998;118(3 Pt 1):304-8.
28. Byl FM, Jr. Sudden hearing loss: eight years' experience and suggested prognostic table. Laryngoscope. 1984;94(5 Pt 1):647-61.
29. Fetterman BL, Saunders JE, Luxford WM. Prognosis and treatment of sudden sensorineural hearing loss. Am J Otol. 1996;17(4):529-36.
30. Kuhn M, Heman-Ackah SE, Shaikh JA, Roehm PC. Sudden sensorineural hearing loss: a review of diagnosis, treatment, and prognosis. Trends Amplif. 2011;15(3):91-105.
31. Wu H, Wan W, Jiang H, Xiong Y. Prognosis of Idiopathic Sudden Sensorineural Hearing Loss: The Nomogram Perspective. Ann Otol Rhinol Laryngol. 2023;132(1):5-12.
32. Mattox DE, Simmons FB. Natural history of sudden sensorineural hearing loss. Ann Otol Rhinol Laryngol. 1977;86(4 Pt 1):463-80.

Assessing Dysphagia in Head and Neck Cancer Patients: Validation and Cultural Adaptation of the Thai M.D. Anderson Dysphagia Inventory (TH-MDADI)

Warut Pongsapich, M.D., Narin Ratanaprasert, M.D.*, Prachya Maneeprasopchoke, M.D., Paveena Pithuksurachai, M.D., Cheerasook Chongkolwatana, M.D., Chanticha Chotigavanich, M.D., Paiboon Sureepong, M.D., Sunun Ongard, M.D., Phawin Keskool, M.D.

Department of Otorhinolaryngology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.



*Corresponding Author: Narin Ratanaprasert
E-mail: narinkup@gmail.com
Received 11 December 2024 Revised 18 January 2025 Accepted 18 January 2025
ORCID ID:http://orcid.org/0000-0002-4366-6414
https://doi.org/10.33192/smj.v77i4.272679

ABSTRACT

Objective: The validation study of the Thai version of the M.D. Anderson Dysphagia Inventory (TH-MDADI) addresses a critical gap in head and neck cancer (HNC) care in Thailand. At leading institutions like Siriraj Hospital, patient-reported outcomes have been historically neglected, with care primarily focusing on routine medical services. This study aimed to validate the TH-MDADI to enhance dysphagia assessment and to promote patient-centered care.

Materials and Methods: In this longitudinal prospective study, 66 HNC patients completed the TH-MDADI and SF-36 at baseline, pre-treatment, and post-treatment at Siriraj Hospital. Psychometric properties were evaluated, including internal consistency (Cronbach's α), test-retest reliability (intraclass correlation coefficient, ICC), convergent and discriminant validity (correlations with SF-36 domains), and responsiveness (standardized response mean, SRM).

Results: The TH-MDADI demonstrated excellent internal consistency (Cronbach's α = 0.93) and test-retest reliability (ICC = 0.92, 95% CI: 0.88–0.96). Strong correlations with related SF-36 domains (r = 0.64-0.74, $p < 0.001$) established convergent validity, while weak correlations with unrelated domains confirmed discriminant validity. High responsiveness to change post-treatment was observed, particularly in the Global (SRM = 3.83) and Emotional (SRM = 3.17) subscales.

Conclusion: The TH-MDADI demonstrates robust psychometric properties, establishing its value as a reliable tool for assessing dysphagia-related quality of life in Thai HNC patients. This validation represents a significant advancement for Thai cancer care, addressing the historical neglect of patient-reported outcomes. By providing clinicians with a validated assessment instrument, this study promotes a more systematic, patient-centered approach to HNC treatment in Thailand.

Keywords: Head and neck cancer; MDADI; SF-36; health-related quality of life (Siriraj Med J 2025; 77: 278-287)

INTRODUCTION

The evaluation of health-related quality of life (HRQOL) in individuals experiencing dysphagia, especially those affected by head and neck cancer (HNC), is essential for comprehending how the condition influences daily functioning and overall quality of life. The M.D. Anderson Dysphagia Inventory (MDADI) is a widely accepted assessment tool designed to measure quality of life in patients with dysphagia. It includes four subscales: global, emotional, functional, and physical.[1]

While the MDADI has been extensively validated and used in its original English version, there is an increasing need for validated translations in other languages to ensure its applicability across different cultural contexts. In Thailand, where the incidence of HNC is significant[2], having a properly validated Thai version of the MDADI is essential for accurate assessment and improved patient care.

Translating and adapting HRQOL questionnaires for different cultural settings is complex and requires careful consideration. *Herdman et al., 1997*[3] emphasized the importance of establishing equivalence between the original and translated versions of HRQOL assessments. They advocate for a standardized approach and terminology in cross-cultural adaptation, highlighting the need for conceptual equivalence, and suggesting that a universal perspective may be necessary to fully capture HRQOL across various cultures.[3] They advocated for a standardized terminology and approach in cross-cultural adaptations, highlighting the need for conceptual equivalence and suggesting that a universalist perspective may be necessary to fully capture HRQOL across different cultures.

This perspective is particularly relevant when validating the MDADI for use in Thailand. It underscores the importance of not only translating the questionnaire but also ensuring that the concepts being measured are relevant and understood similarly in the Thai cultural context. This process involves careful examination of linguistic nuances, cultural norms, and differences in the healthcare system to ensure that the Thai version of the MDADI maintains the conceptual equivalence of the original while being culturally appropriate and understandable to the Thai population.

To validate the Thai version of the MDADI, it is crucial to compare it with a well-established, general HRQOL measure. The Short Form-36 (SF-36)[4], which has already been translated and validated in Thai, serves as an ideal comparator. The SF-36 comprises eight domains: physical functioning, role limitations due to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems, and mental health.[5] The general objective of this study is to analyze the validity and reliability of the Thai MDADI by correlating it with relevant domains

in the SF-36 questionnaire. The additional objectives of this study were as follows:

a) Assess the convergent validity of the Thai MDADI by demonstrating its relationship with relevant SF-36 domains (e.g., physical functioning, role limitations, and overall health).
b) Evaluate the discriminant validity of the Thai MDADI by highlighting its distinction from unrelated SF-36 domains (e.g., mental health and social functioning).

By establishing these relationships and examining the Thai MDADI's performance over time, this study seeks to provide a validated Thai version of the MDADI, enhancing its utility in assessing Thai HNC patients with dysphagia. This validation will contribute to the growing body of evidence supporting the use of the Thai MDADI, ultimately leading to improved patient care and outcomes in Thailand and potentially informing similar efforts in other Southeast Asian countries.

MATERIALS AND METHODS

Study design

This study employed a longitudinal prospective design to validate the Thai version of the M.D. Anderson Dysphagia Inventory (TH-MDADI) in HNC patients. Data was collected from both the Out-Patient Department (OPD), and In-Patient Department (IPD) by two instructors from the Otorhinolaryngology Head and Neck Surgery Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand, over a one-year period from September 30, 2021, to September 29, 2022.

Ethical considerations

The study was approved by the Siriraj Institutional Review Board of the Faculty of Medicine Siriraj Hospital, Mahidol University (Protocol No. 629/2563(IRB4), COA No. Si 807/2020). Informed consent was obtained from all participants prior to their enrollment in the study. The study adhered to the Declaration of Helsinki and Good Clinical Practice guidelines.

Participants

Eligible participants were adult patients (≥18 years) diagnosed with HNC and scheduled to undergo either surgical intervention or primary concurrent chemoradiotherapy (CCRT) at our institution. Exclusion criteria included an inability to understand or read Thai, cognitive impairment preventing questionnaire completion, and presence of other conditions significantly impacting swallowing function unrelated to HNC.

Data collection

Clinical data were extracted from medical records at Siriraj Hospital, a tertiary referral center in Thailand. The data collected included patient demographics such as age, gender, clinical staging, and treatment modality. Participants were asked to complete the TH-MDADI along with the previously validated Thai version of the Short Form-36 (SF-36) at three time points:

1. Baseline: During the initial clinic visit, prior to the initiation of treatment.
2. Pre-Treatment Confirmation: 2-4 weeks after the baseline assessment, before the commencement of primary treatment (either surgery or concurrent chemoradiotherapy [CCRT].
3. Follow-Up: 1-4 months post-treatment completion.

The MDADI consists of 20 items across four domains: global, emotional, functional, and physical. Each item is scored on a 5-point Likert scale, with total scores ranging from 20 to 100. Higher scores indicate a better quality of life.

The SF-36 includes 36 items across eight domains: physical functioning, role limitations due to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems, and mental health.

Statistical analysis

All analyses were conducted using Python (version 3.12.6) with the following libraries:

- **Data Processing & Visualization:** *pandas, numpy, matplotlib, seaborn*
- **Statistical Analysis:** *scipy.stats, statsmodels, pingouin, fancyimpute*

Descriptive Statistics summarized demographic and clinical characteristics, presenting continuous variables as means ± SD or medians (IQR) and categorical variables as frequencies and percentages.

Reliability

Internal consistency was assessed using Cronbach's alpha (>0.70) calculated via the *pingouin. cronbach_alpha*. Intraclass correlation coefficients (ICC) were calculated for 30 participants using the *pingouin.intraclass_corr* package.

Construct validity

- **Convergent validity:** Pearson correlations were calculated between MDADI and related SF-36 domains (physical, role limitations, overall health) using *scipy.stats.pearsonr*.

- **Discriminant validity:** Pearson correlations between MDADI and unrelated SF-36 domains (mental health, social functioning).

Responsiveness

The Standardized Response Mean (SRM) was calculated as the mean of change in scores divided by the standard deviation (SD) scores.

Known-Groups Validity MDADI scores were compared across tumor stages and treatment modalities using one-way ANOVA (*statsmodels.stats.anova.anova_lm*) or independent t-tests (*scipy.stats.ttest_ind*).

The significance level of $p < 0.05$ was applied. For missing data exceeding 5% multiple imputation was performed using *fancyimpute.IterativeImputer*.

Translation process

The original English version of the M.D. Anderson Dysphagia Inventory (MDADI) was meticulously translated into Thai to ensure linguistic and conceptual equivalence. The translation process was conducted by The Center for Translation and Language Services (CTLS) at Mahidol University. The following steps were undertaken to ensure a culturally appropriate translation:

1. **Forward Translation:** Two independent bilingual translators, fluent in English and Thai, performed forward translations of the MDADI. The goal was to accurately convey the original content while adapting it to the Thai linguistic and cultural context.
2. **Reconciliation:** The two forward translations were reviewed and reconciled by a panel of experts, including linguists and specialists in swallowing disorders, to create a unified Thai version. Any discrepancies were discussed and resolved to maintain the integrity of the questionnaire.
3. **Backward Translation:** The reconciled Thai version was then translated back into English by two different bilingual translators who were not familiar with the original MDADI. This back-translation was compared with the original English version to identify and rectify any inconsistencies or deviations in meaning.
4. **Expert Committee Review:** An expert committee, consisting of clinicians, researchers, and language specialists, reviewed all versions of the MDADI to ensure that the conceptual meaning remained unchanged between the original and translated versions.
5. **Pre-Testing (Cognitive Interviewing):** The pre-final Thai version was pilot-tested on a sample of **30 Thai-speaking individuals** representative of the target population. Cognitive interviews were conducted to assess the clarity, relevance, and comprehensibility of each item. Participant feedback was incorporated to further refine the questionnaire.
6. **Finalization and Validation**
   Based on feedback from the pre-testing phase, the Thai MDADI was finalized. This version underwent further psychometric testing to confirm its reliability, validity, and responsiveness, ensuring its appropriateness for assessing dysphagia-related quality of life in Thai head and neck cancer patients.

RESULTS

Population characteristics

A total of 66 patients diagnosed with HNSCC were enrolled in the study. The mean age of participants was 65.4 ± 7.2 years, with 66.7% being male. Primary tumor sites were distributed as follows: oral cavity (53.0%), oropharynx (25.8%), larynx (12.1%), and hypopharynx (9.1%). Most presented with advanced disease, with 39.4% at Stage III and 33.3% at Stage IVA, as detailed in Table 1.

Questionnaire completion rates

Out of the initial 66 participants:

- **Baseline:** All 66 participants completed the questionnaires during the first clinic visit.
- **Pre-Treatment Confirmation:** 41 participants completed the questionnaires approximately 2–4 weeks after the baseline assessment, prior to the initiation of primary treatment.
- **Follow-Up:** 53 participants completed the questionnaires 1–4 months post-treatment.

Quality of life outcomes

Quality of life was assessed using both the MD Anderson Dysphagia Inventory (MDADI) and the Short Form-36 (SF-36) questionnaire.

MDADI scores:

Global: 76.8 ± 5.8
Emotional: 76.5 ± 5.7
Functional: 75.5 ± 6.2
Physical: 76.6 ± 5.3

SF-36 scores:

Physical Functioning: 76.1 ± 5.5
Role Limitations due to Physical Health: 75.2 ± 5.5
Overall Health: 76.1 ± 5.6
Mental Health: 79.2 ± 12.0
Social Functioning: 76.5 ± 11.9

TABLE 1. Baseline and clinical characteristics.

Characteristics	Value
Age (years)	Mean ± SD: 65.4 ± 7.2 Range: 27 - 81
Gender, n (%)	Male: 44 (66.7%) Female: 22 (33.3%)
Cancer Site, n (%)	Oral Cavity: 35 (53.0%) Oropharynx: 17 (25.8%) Larynx: 8 (12.1%) Hypopharynx: 6 (9.1%)
AJCC 8 Stage, n (%)	Stage I: 4 (6.1%) Stage II: 14 (21.2%) Stage III: 26 (39.4%) Stage IV: 22 (33.3%)
Primary Treatment Modality, n (%)	Definitive surgery: 4 (6.1%) Surgery with adjuvant CCRT: 37 (56.1%) Definitive radiotherapy: 3 (4.5%) Definitive chemoradiotherapy: 22 (33.3%)

Reliability analysis

Internal consistency

The TH-MDADI demonstrated excellent internal consistency, with an overall Cronbach's alpha of 0.93. Subscale-specific alpha values ranged from 0.87 in the Global domain to 0.91 in the Functional domain, confirming the tool's reliability across various quality-of-life dimensions, as detailed in Table 2. These findings underscore the TH-MDADI's robust consistency for assessing dysphagia-related quality of life in head and neck cancer (HNC) patients.

Test and retest reliability

Test-retest reliability was assessed using intraclass correlation coefficients (ICC) in a subset of 30 participants who completed the MDADI on two occasions within a two-week interval, with no treatment changes during this period. The ICC for the overall MDADI score was 0.92 (95% CI: 0.88–0.96), demonstrating excellent stability over time, with subscale ICCs from 0.88 to 0.94. These results, also in Table 2, confirm the instrument's reliability across domains.

In the context of the reliability analysis for the Thai MDADI, Fig 1 visually presents both the Cronbach's Alpha and ICC values for each subscale and the overall score. The figure highlights the high internal consistency of the MDADI, with Cronbach's alpha values ranging from 0.87 for the Global domain to 0.93 for the Overall score. Additionally, the ICC values, represented by the red line, confirm the strong test and retest reliability, with ICCs ranging from 0.90 to 0.94 across all domains.

Construct validity

Convergent validity

Convergent validity was assessed by examining Pearson correlation coefficients between MDADI domains and related SF-36 domains, including Physical Functioning, Role Limitations due to Physical Health, and Overall, Health. These correlations confirmed strong convergent validity.

Discriminant validity

Discriminant validity was demonstrated by examining MDADI correlations with unrelated SF-36 domains (Mental Health and Social Functioning), which were weak and statistically non-significant: Mental Health ($r = 0.18$, $p = 0.15$) and Social Function ($r = 0.22$, $p = 0.10$).

TABLE 2. Cronbach's Alpha and Intraclass Correlation Coefficients for the Thai MDADI.

Domain	Cronbach's Alpha	ICC (95% CI)
Global	0.87	0.90 (0.84–0.95)
Emotional	0.89	0.91 (0.86–0.96)
Functional	0.91	0.94 (0.90–0.97)
Physical	0.90	0.92 (0.88–0.96)
Overall	**0.93**	**0.92 (0.88–0.96)**

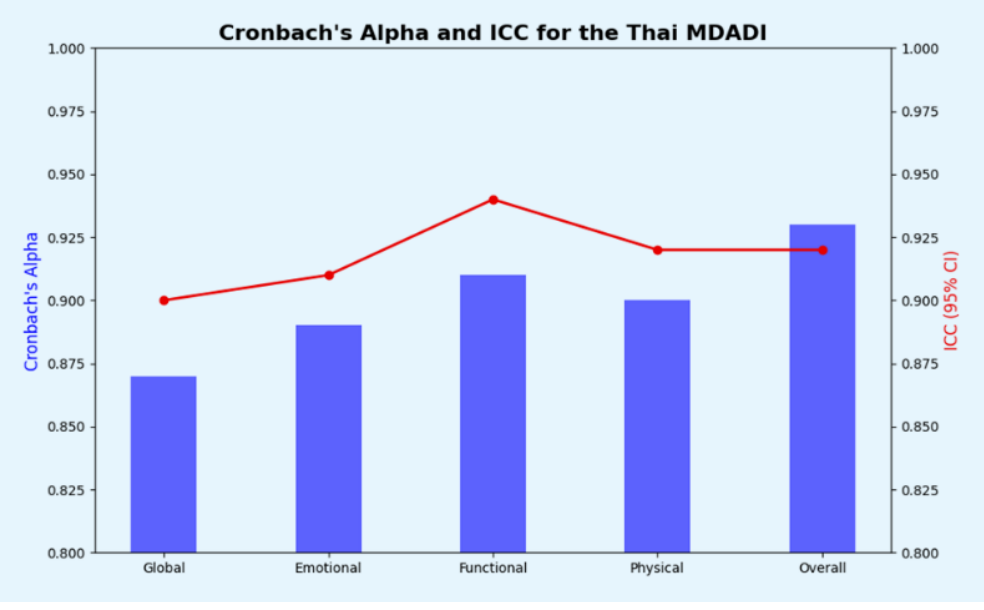


Fig 1. Cronbach's Alpha and Intraclass Correlation Coefficients (ICC) for the Thai MDADI Across Subscales and Overall Score.

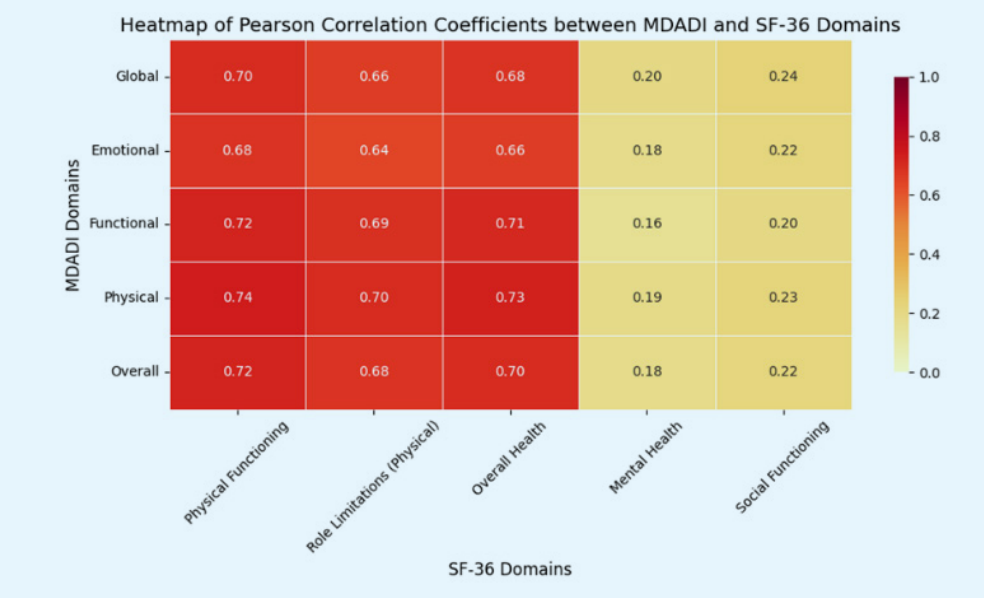


Fig 2. Heatmap of Pearson Correlation Coefficients Between MDADI and SF-36 Domains.

Responsiveness

Standardized Response Mean (SRM) values, assessing sensitivity to treatment effects, highlighted the TH-MDADI's responsiveness across domains. SRM values were highest in the Global domain (3.83), followed by the Emotional (3.17) and Functional (2.99) domains, indicating substantial responsiveness to treatment outcomes, especially in physical and emotional health. Table 3 summarizes these SRM values, alongside comparisons with SF-36 domains, demonstrating that both MDADI and SF-36 are effective in capturing post-treatment quality-of-life changes.

Standardized Response Mean (SRM) analysis

The Standardized Response Mean (SRM) was calculated for both the MDADI and SF-36 domains to assess sensitivity to treatment effects. Higher SRM values indicate greater responsiveness, with the MDADI showing very high SRM in the Global domain (3.83), high in Emotional (3.17) and Functional (2.99), and moderate in Physical (2.42). The SF-36 showed very high responsiveness in Physical Functioning (4.17), followed by Mental Health (3.65) and Role Limitations (3.13), with moderate responsiveness in Overall Health (2.50). Fig 3 further visualizes SRM values in a radar plot, comparing MDADI (green) and SF-36 (purple) responsiveness across domains. This figure highlights how each instrument captures changes in quality of life, especially in physical and emotional domains.

DISCUSSION

This study aimed to validate the TH-MDADI for assessing the quality of life in HNC patients with dysphagia. Our findings demonstrate that the Thai MDADI has strong psychometric properties, supporting its use in clinical and research settings in Thailand.

Reliability and validity

The Thai MDADI showed excellent internal consistency (Cronbach's alpha = 0.93) and test-retest reliability (ICC = 0.92), indicating a high degree of measurement stability. These results are comparable to those reported for the original English version and other validated translations, such as Dutch (Speyer et al., 2011)[6], Italian (Schindler, Borghi, Tiddia, Ginocchio, Felisati, & Ottaviani, 2008)[7], Swedish (Carlsson et al., 2012)[8], Korean (Kwon, Kim, Park, Oh, & Han, 2013)[9], and Brazilian Portuguese (Guedes, Angelis, Chen, Kowalski, & Vartanian, 2013).[10] The strong reliability suggests that the Thai MDADI consistently measures dysphagia-related quality of life across subscales over time.

The validity of TH-MDADI was supported by strong correlations between its domains and related domains of the SF-36, such as Physical Functioning, Role Limitations due to Physical Health, and Overall Health. The correlation coefficients (ranging from 0.64 to 0.74) indicate that the Thai MDADI effectively captures the physical and functional aspects of dysphagia-related quality of life. Conversely, weaker correlations with

TABLE 3. Standardized Response Mean (SRM) Values for MDADI and SF-36 Domains Post-Treatment.

Domain	Instrument	SRM	Interpretation
Global	MDADI	3.83	Very High
Emotional	MDADI	3.17	High
Functional	MDADI	2.99	High
Physical	MDADI	2.42	Moderate
Physical Functioning	SF-36	4.17	Very High
Role limitation	SF-36	3.13	High
Overall, Health	SF-36	2.5	Moderate
Mental Health	SF-36	3.65	High

The interpretations align correctly with the SRM values, where an SRM greater than 3.0 is considered very high, between 2.0 and 3.0 is high, and values between 1.5 and 2.0 are considered moderate.

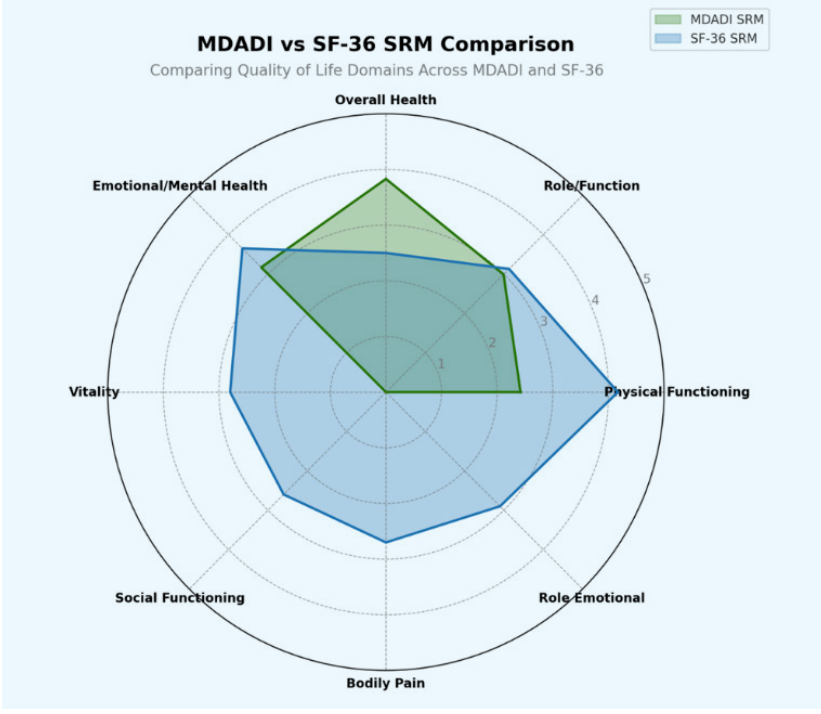


Fig 3. Comparison of Standardized Response Mean (SRM) Across Quality of Life Domains for MDADI and SF-36.

unrelated SF-36 domains (Mental and Social Function) demonstrate good discriminant validity, suggesting that the Thai MDADI can distinguish between physical and psychosocial aspects. Building on these findings, it is useful to consider prior studies that have evaluated dysphagia assessments in various populations. For instance, previous research examining the safety of specific food consistencies in stroke patients (Kientchockwiwat et al., 2022)[11] or the correlation between tongue strength and aspiration risk (Keskool et al., 2018)[12] highlights the importance of having robust, validated tools. Apart from investigations that have evaluated swallowing function and dysphagia outcomes in different populations, our study aims to fill a gap in the clinical assessment of dysphagia-related quality of life among Thai head and neck cancer patients. For instance, studies such as those examining the safety of specific food consistencies in stroke patients with dysphagia (Kientchockwiwat et al., 2022)[11] or the relationships between objective tongue strength measures and aspiration risk (Keskool et al., 2018)[12] underscore the importance of robust, validated instruments in both research and clinical practice. By validating the Thai MDADI, we provide clinicians and researchers a reliable tool to complement existing objective measures, enabling more comprehensive assessments and potentially guiding treatment strategies.

Responsiveness to change

High SRM values in MDADI's Global (3.83) and Emotional (3.17) subscales confirm sensitivity to dysphagia-related quality-of-life changes post-treatment, crucial for tracking treatment outcomes. Fig 3 illustrates these SRM comparisons, supporting the MDADI's role as a dysphagia-specific measure alongside SF-36 in assessing physical and emotional impacts.

Cultural adaptation and equivalence

The rigorous translation and cultural adaptation process used in this study, following the guidelines proposed by *Herdman et al.* (1997)[3], ensured that the Thai MDADI maintains conceptual and linguistic equivalence with the original English version. This approach addresses the challenges of cross-cultural adaptation highlighted in previous literature (*Guillemin et al.*, 1993; *Beaton et al.*, 2000)[13,14] and adds to the growing body of evidence supporting the use of the MDADI across different cultures and languages.

Clinical implications

The validation of the Thai MDADI provides clinicians and researchers in Thailand with a reliable and valid tool for assessing dysphagia-related quality of life in HNC patients. This instrument can facilitate patient care by:

1. Enabling systematic assessment of dysphagia impact on quality of life
2. Assisting in treatment planning and decision-making
3. Monitoring treatment outcomes and recovery progress
4. Facilitating communication between healthcare providers and patients about dysphagia-related concerns

Limitations and future directions

Despite demonstrating strong psychometric properties, our study has several limitations. First, while the sample size was sufficient for initial validation, future studies with larger, more diverse populations would enhance the generalizability of the findings. Second, although both Videofluoroscopic Swallow Study (VFSS) and Fiberoptic Endoscopic Evaluation of Swallowing (FEES) are utilized at Siriraj Hospital for assessing swallowing physiology, this study did not include data from either of these objective assessments. Integrating VFSS parameters (e.g., Penetration-Aspiration Scale, residue scores) and FEES findings (e.g., penetration, aspiration, pharyngeal residue) into future research could establish correlations with TH-MDADI scores, bridging subjective self-reports and objective swallowing outcomes. This multimodal approach would further validate the clinical utility of the TH-MDADI and enhance its application in tracking treatment outcomes.

CONCLUSION

The Thai MDADI demonstrates strong reliability, validity, and responsiveness, making it a valuable tool for evaluating dysphagia-related quality of life in Thai HNC patients with dysphagia. Its validation contributes to the global effort to standardize dysphagia assessments and improve patient care across diverse cultural contexts.

Data Availability Statement

The data that support the findings of this study are available upon request from the corresponding author, [N.R.]. The data are not publicly available due to containing information that could compromise the privacy of research participants.

ACKNOWLEDGEMENTS

The authors would like to express appreciation to Miss Jeerapa Kerdnopakun and Miss Sirakorn Pothipakdee, our research assistants, for their help with reference searches, manuscript formatting, and reference management.

DECLARATION

Grants and Funding Information

This study was funded by the Faculty of Medicine Siriraj Hospital, Mahidol University, approval no. (IO) R016432006.

Conflict of Interest

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

Registration Number of Clinical Trial

Not registered in the clinical trial.

Author Contributors

W.P.; Conceptualization, Methodology, Supervision. N.R.; Data Curation, Project administration. N.R.; Data Curation, Investigation. S.O., P.P., C.C., P.S., and P.K.; Validation, Writing - Review & Editing.

Use of Artificial Intelligence

None

REFERENCES

1. Chen AY, Frankowski R, Bishop-Leone J, Hebert T, Leyk S, Lewin J, et al. The development and validation of a dysphagia-specific quality-of-life questionnaire for patients with head and neck cancer: the M. D. Anderson dysphagia inventory. Arch Otolaryngol Head Neck Surg. 2001;127(7):870-6.
2. Tangjaturonrasme N, Vatanasapt P, Bychkov A. Epidemiology of head and neck cancer in Thailand. Asia Pac J Clin Oncol. 2018; 14(1):16-22.
3. Herdman M, Fox-Rushby J, Badia X. 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires. Qual Life Res. 1997;6(3):237-47.
4. Leurmarnkul W, Meetam P. Properties Testing of the Retranslated SF36 (Thai Version). The Thai Journal of Pharmaceutical Sciences. 2005;29.
5. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30(6):473-83.
6. Speyer R, Heijnen BJ, Baijens LW, Vrijenhoef FH, Otters EF, Roodenburg N, et al. Quality of life in oncological patients with oropharyngeal dysphagia: validity and reliability of the Dutch version of the MD Anderson Dysphagia Inventory and the Deglutition Handicap Index. Dysphagia. 2011;26(4): 407-14.
7. Speyer R, Heijnen BJ, Baijens LW, Vrijenhoef FH, Otters EF, Roodenburg N, et al. Quality of life in oncological patients with oropharyngeal dysphagia: validity and reliability of the Dutch version of the MD Anderson Dysphagia Inventory and the Deglutition Handicap Index. Dysphagia. 2011;26(4): 407-14.
8. Carlsson S, Rydén A, Rudberg I, Bove M, Bergquist H, Finizia C. Validation of the Swedish M. D. Anderson Dysphagia Inventory (MDADI) in patients with head and neck cancer and neurologic swallowing disturbances. Dysphagia. 2012;27(3):361-9.
9. Kwon CH, Kim YH, Park JH, Oh BM, Han TR. Validity and reliability of the Korean version of the MD Anderson Dysphagia inventory for head and neck cancer patients. Ann Rehabil Med. 2013;37(4):479-87.
10. Guedes RL, Angelis EC, Chen AY, Kowalski LP, Vartanian JG. Validation and application of the M.D. Anderson Dysphagia Inventory in patients treated for head and neck cancer in Brazil. Dysphagia. 2013;28(1):24-32.
11. Kientchockwiwat K, Thitisakulchai P, Tanvijit P. The Safety of Food and Drink Consistencies Based on a Fiberoptic Endoscopic Evaluation of Swallowing Study Results in Stroke Patients with

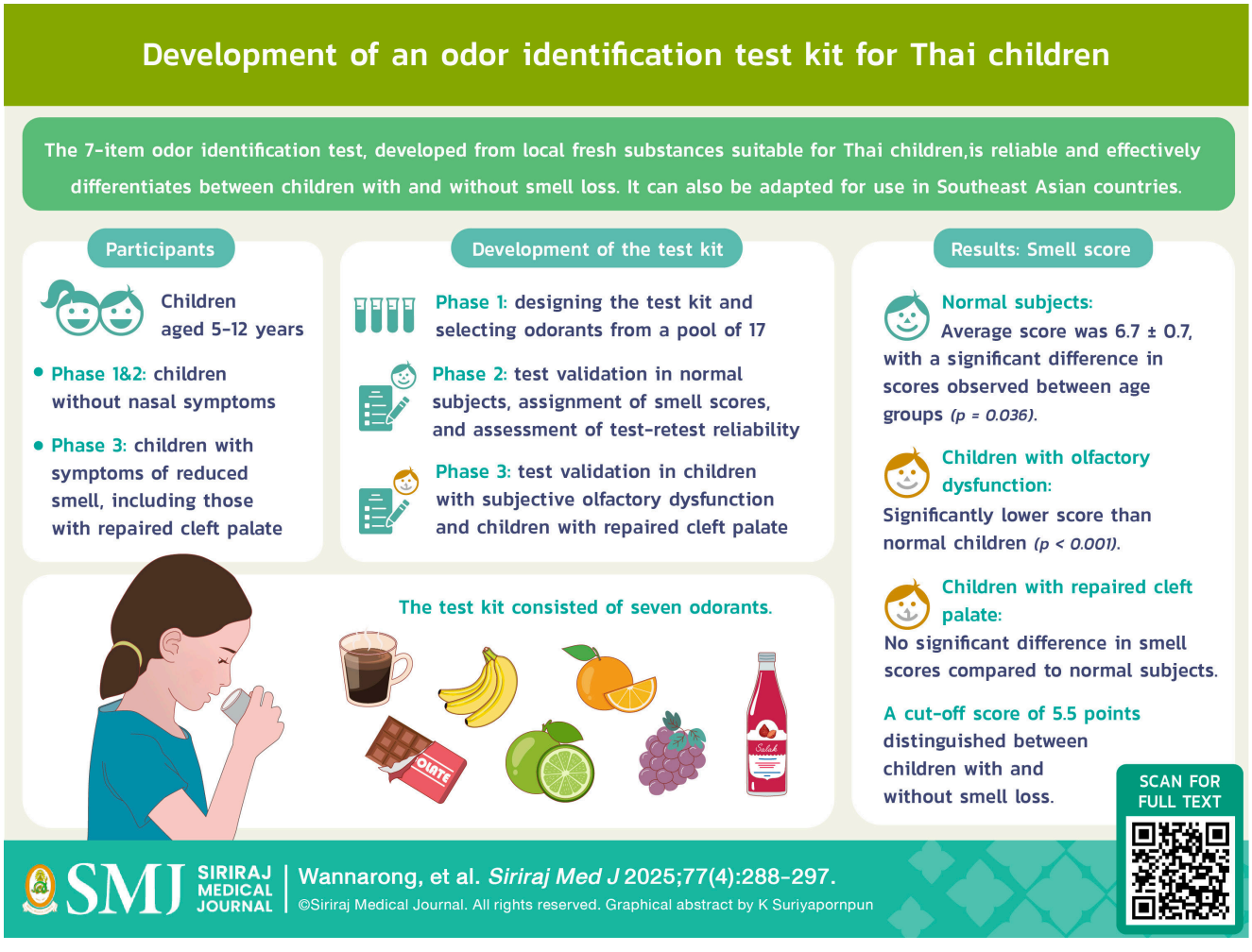
Dysphagia. Siriraj Med J. 2022;74(6):395-400.
12. Keskool P, Warnpeurch L, Ongard S, Pitaksurachai P, Nujchanart N, Kerdnoppakhun J. The relationships among objective measures of tongue strength and risk of aspiration. Siriraj Med J. 2018;70(4):302-9.
13. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol. 1993;46(12):1417-32.
14. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976). 2000;25(24):3186-91.

Development of an Odor Identification Test Kit for Thai Children

Odor Test for Children

Thanakrit Wannarong, M.D.[1], Sithatcha Wongkom, M.D.[2], Triphoom Suwanwech, M.D.[1], Archwin Tanphaichitr, M.D.[1], Vannipa Vathanophas, M.D.[1], Kitirat Ungkanont, M.D.[1,*]

[1]Department of Otorhinolaryngology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand, [2]Department of Otorhinolaryngology, Wattanapat Hospital, Aonang, Krabi, Thailand.



**Corresponding author: Kitirat Ungkanont*
E-mail: kitirat.ung@mahidol.ac.th
Received 3 December 2024 Revised 13 January 2025 Accepted 13 January 2025
ORCID ID:http://orcid.org/0000-0003-0923-1908
https://doi.org/10.33192/smj.v77i4.272527

ABSTRACT

Objective: To develop and validate an odor identification test kit for Thai children that can be adapted for use in Southeast Asian countries.

Materials and Methods: The test kit was developed in three phases, using local fresh substances. Phase 1 involved designing the test kit and selecting odorants from a pool of 17. Phase 2 focused on test validation in normal subjects, assignment of smell scores, and assessment of test-retest reliability. Phase 3 validated the test in children with subjective olfactory dysfunction and children with repaired cleft palate. Cut-off scores were determined using receiver operating curve analysis.

Results: The participants were children aged 5-12 years. Sample sizes in Phases 1, 2, and 3 were 53, 31, and 36, respectively. Seven odorants that met the selection criteria were chosen. The average score for normal subjects was 6.7 (SD 0.7), with a significant difference between age groups ($p = 0.036$). Children with olfactory dysfunction had an average score of 3.8 (SD 1.6), significantly lower than normal children ($p < 0.001$). Children with repaired cleft palate showed no significant difference in smell scores compared to normal subjects. A cut-off score of 5.5 points was used to distinguish between normal and abnormal olfactory function, with an area under the curve of 0.928.

Conclusion: Children aged 5-12 years were able to complete the 7-item odor identification test developed from local fresh substances. The test kit demonstrated good reliability and effectively distinguished between children with and without smell loss, using a cut-off score of 5.5.

Keywords: Smell loss; smell test; odor identification test; olfactory dysfunction (Siriraj Med J 2025; 77: 288-297)

INTRODUCTION

Olfaction plays a vital role in daily life. Without it, detecting specific dangers such as fire becomes challenging, and overall quality of life is impacted.[1] The perception of odors is essential for children's learning and development.[2] Common causes of olfactory dysfunction in children include sinonasal diseases, head trauma, and congenital conditions, although the prevalence differs from that in adult patients.[3] Congenital syndromic anomalies, such as Kallmann syndrome, can also cause anosmia in children.[4] Olfactory testing is crucial for evaluating olfactory function, with odor identification tests being commonly used in children due to their simple procedure, ease of understanding, and short duration.[3] Most tests for children incorporate pictures since some may not be able to read letters.[5] Descriptors with pictures and words are often included to assist children in completing the test.[5] The olfactory test for children should be engaging, culturally appropriate, and validated for the target age group.[3]

The Universal Sniff (U-sniff) test consists of 12 odors presented in pen-like Sniffin' Sticks for children aged 4-17 years.[6] The maximum score is 12 points. In children with normosmia, the mean score was 9.88 ±1.8, with different cut-off points for each age group.[6] The 10th percentile of normative data is used to distinguish between normosmia and hyposmia. The Pediatric Smell Wheel (PSW) consists of 11 odors delivered via micro-capsulated stickers for children aged 4-19 years.[7] The PSW was validated in Brazilian children aged 5-12, using the 10th percentile as the cut-off.[8] A score below 7 indicates hyposmia. Age-related cut-off points were set at 5 points for children under 8 years old, 6 points for those aged 8-10, and 7 points for those over 10 years old.[8]

Different cultures can impact odor familiarity.[6] Due to cultural differences and limited exposure to certain odors, existing tests using odorants are not suitable for Thai children. This makes it difficult to diagnose or assess odor dysfunction in this population. A cross-cultural study between Japanese and German women showed that culture-specific foods can significantly influence odor perception.[9] Some odor choices in current tests are unfamiliar to Thai children, such as peach in the U-sniff test, play-doh in the NIH Toolbox, mustard in the San Diego Odor Identification Test (SDOIT) and aniseed in Sniffin' Kids test.[6,7,10-13] These items are either not local products or unavailable in Thailand, so we cannot expect reliable familiarity from Thai children. Additionally, some odors in commercial test kits, such as gasoline, paint thinner, smoke, and turpentine, may cause irritation or discomfort in young children.[3,14-17]

The aim of this study was to develop an odor identification test kit suitable for Thai children, with potential applicability in other Southeast Asian countries. We selected local fruits, foods, and flavoring agents

familiar to target population. Additionally, we aimed to evaluate the kit's reliability in distinguishing between children with and without olfactory dysfunction.

MATERIALS AND METHODS

The study was divided into three phases. Phase 1 involved picture identification and odor selection for designing the test kit. Phase 2 focused on validating the test kit with subjects who had no reported smell dysfunction. In Phase 3, the test was used to differentiate between children with and without olfactory dysfunction. Participants were children aged 5 to 12 years. Children above 12 years old were not recruited because a previous study found no significant differences in smell scores among children aged 9-11, 12-14, and 15-17 years.[18] The study was approved by The Institutional Review Board of Siriraj Hospital, Faculty of Medicine Siriraj Hospital, Mahidol University (COA Si 148/2021). Informed consent was obtained from parents or legal guardians, and informed assent was obtained from the children. The research was conducted in accordance with the World Medical Association's Code of Ethics (Helsinki Declaration).

Power analysis

The sample size for Phase 1 was calculated based on the study by Dżaman, which reported that 76% of children aged 5-7 years and 90% of children over 7 years could correctly identify odors.[19] The total sample size for Phase 1 was 53 participants.

For Phases 2 and 3, the sample sizes were calculated using the study by *Grossman et al*[20], which compared olfactory function between patients with repaired cleft palate (CP) and a control group. With 80% power for a two-sided test and a 1:1 ratio, the required number of participants per group was 31.

Phase 1: Picture identification and Odor selection

Fifty-three children, divided into three age groups — 5 to 7 years, over 7 to 10 years, and over 10 to 12 years — were recruited after a thorough history-taking and complete ear, nose and throat (ENT) examination. History was obtained from both the children and their guardians regarding nasal symptoms and the child's ability to smell food. The inclusion criteria for Phases 1 and 2 were children without nasal symptoms, such as rhinorrhea, nasal congestion, or loss of smell. Exclusion criteria included children with: (1) structural lesions in the nasal cavity, such as nasal polyps or sinonasal tumors; (2) a history of head trauma; (3) previous sinonasal surgery; (4) recent upper respiratory tract infection within the past week; and (5) speech or language impairment.

We selected 11 common odorants from previous studies, including banana, orange, lime, grape, strawberry, cinnamon, garlic, chocolate, coffee, mint, and rose.[5-7,10-12,19,21-24] To these, we added six new odorants familiar to Thai children, including dried squid, cheese, tomato sauce, caramel, jasmine and salak-flavored syrup. Salak, a tropical fruit native to Southeast Asia, is used in the concentrated syrup. In total, 17 odorants were selected from categories such as fruit, food, beverage, and flavoring agents.

The odor-producing substances in the study were chosen from synthetic materials or commercial form of the natural products, as well as from fresh preparations, depending on the strong smell, resemblance of the odor to the natural products, and the convenience for use. For example, the smell of grape juice and strawberry jam were stronger, easier to be recognized and more convenient for use than fresh grapes or strawberries. The same consideration was applied for the use of salak-flavored syrup instead of fresh salak. Finally, ten odorants from synthetic materials and seven odorants from fresh preparations were used in this study. We controlled the consistency of the odor quality by using fresh preparations from the same source or commercial brands, with meticulous measurement of the substance and the weight was recorded before use in every session of the test. The source materials included the following: banana (Cavendish banana), 0.7 g of both peel and flesh from the brand Dole; orange (Mandarin orange), 0.3 g of peel from the brand 2.P.H.; lime (Key Lime), 0.6 g of peel from the brand Pakbangpun; strawberries (strawberry jam), 0.9 g from the brand Best Foods; mint (mint powder), 0.1 g from the brand Royal Project; rose (Wine & Roses Anti-Aging Body Oil), 0.1 g from the brand Erb; jasmine, 0.2 g from jasmine flowers; caramel (Vanilla Flavour), 0.2 g from the brand Winner's; tomato sauce, 0.5 g from the brand Prego; garlic (peeled garlic), 0.2 g from the brand My Choice; cinnamon (Chinese Five-Spice or Pae-Lo powder), 0.1 g from the brand Home Fresh Mart; dried squid, 0.2 g from the brand Taotong; cheese (Parmesan cheese), 0.2 g from the brand Imperial; chocolate, 0.2 g from the brand Van Houten; coffee, 0.2 g from the brand Nescafe Red Cup; salak-flavored syrup, 0.8 g from the brand Hale's Blue Boy; and grape (red grape juice), 0.8 g from the brand Tipco. Each odorant was placed in an opaque bottle and covered with an aluminum lid featuring small holes punched in it to release the odor without revealing the contents. The pattern for the holes was designed on the aluminum foil, and the same template was used for every bottle (Fig 1). The lid was disposed of after each use. The entire

Fig 1. An odorant presented in an opaque bottle, covered with an aluminum foil lid with holes. The bottles were placed 1 cm below and parallel to the child's nostril.

set was replaced every three days, and the odorants made from fresh substances were stored at 5-7°C to prevent deodorization.[25]

We selected pictures that were easy for children to clearly identify. These pictures were validated by children during a preliminary phase to ensure their suitability. For picture identification, the researcher showed a picture of each substance and asked the child to identify it. A score of 1 or 0 was assigned for each item. After showing each picture, two bottles with different odorants were presented — one containing the same substance as in the picture and the other containing an odor from a different category, including fruit, food, beverage, and flavoring agents. The bottles were placed 1 cm below and parallel to the child's nostril, and the child smelled both, choosing the one that matched the picture. The child could smell as many times or for as long as needed until they answered the question. This approach was consistent across all phases. The number and percentage of correct picture identifications and odor selections were recorded.

We planned to select 6-8 odorants for the test kit based on their high ranking and correct identification, both by picture and odor, in more than 90% of participants. A 4-alternative forced-choice (AFC) paradigm was designed for each odorant to help children recognize the correct scent and reduce guessing. In the 4-AFC, four labelled pictures were presented: one for the target odorant and three for distractors. Two distractors were "related", belonging to the same category as the target odor, such as fruits, while one distractor was unrelated, from a different category of odor source.

Phase 2: Validation of the odor identification test and Test-retest reliability

Thirty-one children without nasal symptoms were recruited and divided into three age groups as in Phase 1. Each child was tested using the odor identification test kit from Phase 1, and smell scores were recorded. For each item from the test kit, only one bottle was presented, and the children responded using the 4-AFC method. A correct answer was scored as 1, while an incorrect answer as 0. To assess test-retest reliability, the same participants were tested again at least one week later, following the protocol of previous studies.[6,7,18]

Phase 3: Odor identification test in children with olfactory dysfunction or repaired CP

Participants in Phase 3 included children with symptoms of reduced smell due to upper respiratory tract infection or congenital anosmia, such as Kallmann syndrome. Children with repaired CP were also recruited, as evidence suggests factors like anatomical abnormalities, reduced nasal airflow, and smaller nasal volume may contribute to olfactory dysfunction.[20,26,27] History of subjective olfactory issues and ENT examinations were conducted for all participants. The odor identification test was administered as in Phase 2, and smell scores were recorded.

Outcome measure and statistical analyses

The outcomes for Phase 1 included demographic data and the results of picture identification and odor selection, which were analyzed using descriptive statistics. Ranking of the results was used to select the odorants for the test kit.

In Phases 2 and 3, the outcomes were the smell scores and the time taken to complete the test kit. Analysis of variance (ANOVA) or Kruskal-Wallis tests were used to compare the outcomes among the three age groups. Test-retest reliability was analyzed using paired T-tests or Wilcoxon's test. A receiver operator characteristic curve (ROC) and the Youden index were used to determine the cut-off score that could differentiate between normal and abnormal olfactory function. Additionally, the 10th percentile of smell scores across all age groups was used as an alternative cut-off point between normal and abnormal olfactory function. Statistical analyses were performed using SPSS version 22.0.

RESULTS

The demographic data and time taken for all tests are presented in Table 1. The mean age across all phases was approximately 8 years. In Phase 1, the mean test

TABLE 1. Demographic data.

	Phase 1 (n=53)	Phase 2 (n=31)	Phase 3 (n=36)	
			Cleft palate	Olfactory dysfunction
Sex				
Male	27 (50.9%)	16 (51.6%)	18 (72.0%)	7 (63.6%)
Female	26 (49.1%)	15 (48.4%)	7 (28.0%)	4 (36.4%)
Age (year, mean ± SD)	8.4 ± 2.3	8.4 ± 2.2	8.0 ± 2.3	8.0 ± 2.1
Age stratification				
5-7 yrs.	18 (34.0%)	10 (32.3%)	10 (40.0%)	5 (45.5%)
7.01-10 yrs.	17 (32.1%)	11 (35.3%)	10 (40.0%)	4 (36.4%)
10.01-12 yrs.	18 (34.0%)	10 (32.3%)	5 (20.0%)	2 (18.2%)
Duration of the test (minute, mean ± SD)	8.9 ± 3.5	3.2 ± 1.4	3.3 ± 1.7	5.0 ± 1.8

Abbreviation: SD, standard deviation

duration was 8.9 minutes (SD 3.5), which included both picture identification and odor selection. The odor identification test kit was developed by integrating these two components, as shown in Table 2. The top seven odorants that met the criteria were lime, orange, banana, grape, salak-flavored syrup, coffee, and chocolate. The total score ranged from 0 to 7.

The results of Phase 2 are shown in Table 3. The mean test duration for normal subjects was 3.2 ± 1.4 minutes, with no significant difference in the amount of time spent across all three age groups ($p = 0.066$). The overall mean smell score for children was 6.7 (SD 0.7), with no significant difference between males and females ($p = 0.143$). However, there was a significant difference in scores among the three age groups, with the youngest group scoring the lowest ($p = 0.036$). A significant positive correlation was found between smell scores and the age of the children ($r = 0.512$, $p = 0.003$). Coffee was the only odor with a significant difference in identification between age groups ($p = 0.03$). We tested the effect of the season by performing the test with fresh fruit in different seasons, the results showed no difference of the mean scores ($p = 0.868$). Test-retest reliability showed a strong correlation ($r = 0.932$, $p < 0.001$).

In Phase 3, we conducted the test on children with smell loss or those at risk for olfactory dysfunction. This group included 11 children with complaints of smell loss, consisting of 10 cases of acute rhinosinusitis (ARS) and one case of Kallmann syndrome. Additionally, 25 children with repaired CP participated, however, none of the CP group reported subjective olfactory dysfunction prior to the test.

The mean duration to complete the test for the 11 children with olfactory dysfunction was 5.0 minutes (SD 1.8), which was significantly longer than the time taken by normal subjects in phase 2 ($p = 0.005$). The mean score for children with olfactory dysfunction was 3.8 (SD 1.6), which was significantly lower than that of the normal subjects in Phase 2 ($p < 0.001$). The smell scores for all groups are presented in Table 4. There was no significant difference between the CP group and the normal group in either time spent ($p = 0.991$) or smell scores ($p = 0.946$). Therefore, the children with repaired CP were excluded from the validation and ROC analysis.

Finally, we validated the test by including 31 children with normal olfaction from Phase 2 and 11 children with olfactory dysfunction from Phase 3. The test demonstrated a sensitivity of 90.32%, specificity of 90.91%, and accuracy of 90.5% in detecting smell loss. The positive predictive value was 96.6%, with a negative predictive value of 76.9%, a positive likelihood ratio (LR) of 9.93, a negative LR of 0.1, and a diagnostic odds ratio of 93.3. The ROC curve, shown in Fig 2, indicated a cut-off score of 5.5 for diagnosing normal odor identification. The area under curve (AUC) was 0.928. The mean score at the 10^{th} percentile in normal subjects was 5.2.

TABLE 2. Phase 1: Picture identification and Odor selection.

Substance	Percent of correct picture identification					Percent of correct odor selection				
	All age (n=53)	5-7 yrs. (n=18)	7.01-10 yrs. (n=17)	10.01-12 yrs. (n=18)	Age group comparison (*p*-value)**	All age (n=53)	5-7 yrs. (n=18)	7.01-10 yrs. (n=17)	10.01-12 yrs. (n=18)	Age group comparison (*p*-value)**
Cheese	96.2%	88.9%	100%	100%	0.321	96.2%	94.4%	100%	94.4%	1.0
Squid	75.5%	50.0%	82.4%	94.4%	0.010*	98.1%	94.4%	100%	100%	1.0
Strawberry	100%	100%	100%	100%	-	96.2%	100%	100%	88.9%	0.321
Cinnamon	81.1%	72.2%	76.5%	94.4%	0.211	92.5%	88.9%	100%	88.9%	0.530
Lime	**98.1%**	100%	94.1%	100%	**0.321**	**98.1%**	94.4%	100%	100%	**1.0**
Garlic	79.2%	55.6%	82.4%	100%	0.002*	94.3%	88.9%	94.1%	100%	0.530
Orange	**98.1%**	100%	94.1%	100%	**0.321**	**98.1%**	94.4%	100%	100%	**1.0**
Tomato sauce	100%	100%	100%	100%	-	96.2%	88.9%	100%	100%	0.321
Banana	**100%**	100%	100%	100%	-	**100%**	100%	100%	100%	-
Grape	**98.1%**	94.4%	100%	100%	**1.0**	**100%**	100%	100%	100%	-
Caramel	90.6%	88.9%	82.4%	100%	0.185	86.8%	83.3%	82.4%	94.4%	0.603
Salak flavored syrup	**98.1%**	94.4%	100%	100%	**1.0**	**100%**	100%	100%	100%	-
Jasmine	81.1%	55.6%	88.2%	100%	0.001*	100%	100%	100%	100%	-
Coffee	**100%**	100%	100%	100%	-	**98.1%**	94.4%	100%	100%	**1.0**
Rose	86.8%	66.7%	94.1%	100%	0.008*	94.3%	88.9%	94.1%	100%	0.530
Chocolate	**100%**	100%	100%	100%	-	**98.1%**	94.4%	100%	100%	**1.0**
Mint	84.9%	72.2%	88.2%	94.4%	0.190	92.5%	83.3%	94.1%	100%	0.207

Note: Level of significant *$p < 0.05$; ** By the chi-square test; The selected odors in the odor identification test kit are presented in bold.
Abbreviation: yrs, years

TABLE 3. Phase 2: Odor identification scores and duration of the test in normal subjects.

	All age (n=31)	5-7 years (n=10)	7.01-10 years (n=11)	10.01-12 years (n=10)	Age group comparison (p-value)
Duration of the test** (minute, mean ± SD)	3.2 ± 1.4	4.0 ± 1.2	3.1 ± 1.8	2.6 ± 0.5	0.066
Score** (mean ± SD)	6.7 ± 0.7	6.2 ± 1.1	6.8 ± 0.4	7.0 ± 0.0	0.036*
Target odor***					
Chocolate	96.8%	90.0%	100%	100%	0.913
Orange	96.8%	100%	90.9%	100%	0.639
Salak flavored syrup	93.5%	90.0%	90.9%	100%	0.588
Banana	96.8%	90.0%	100%	100%	0.053
Coffee	93.5%	80.0%	100%	100%	0.030*
Lime	93.5%	80.0%	100%	100%	0.375
Grape	96.8%	90.0%	100%	100%	0.576

Note: Level of significant *$p < 0.05$; ** By the one-way ANOVA; *** By the chi-square test or Fisher's exact test
Abbreviation: SD, standard deviation

TABLE 4. Comparison of odor identification scores among 3 groups.

	Normal children (n = 31)	Repaired CP (n = 25)	Olfactory dysfunction (n = 11)	*p*-value
Score**	6.7 ± 0.7	6.6 ± 0.6	3.8 ± 1.6	<0.001*
Post Hoc Tests***				
	6.7 ± 0.7	6.6 ± 0.6		0.946
	6.7 ± 0.7		3.8 ± 1.6	<0.001*
		6.6 ± 0.6	3.8 ± 1.6	<0.001*

Note: Level of significant *$p < 0.05$; ** By the one-way ANOVA; *** By the Tukey Post Hoc Tests
Abbreviation: CP, cleft palate

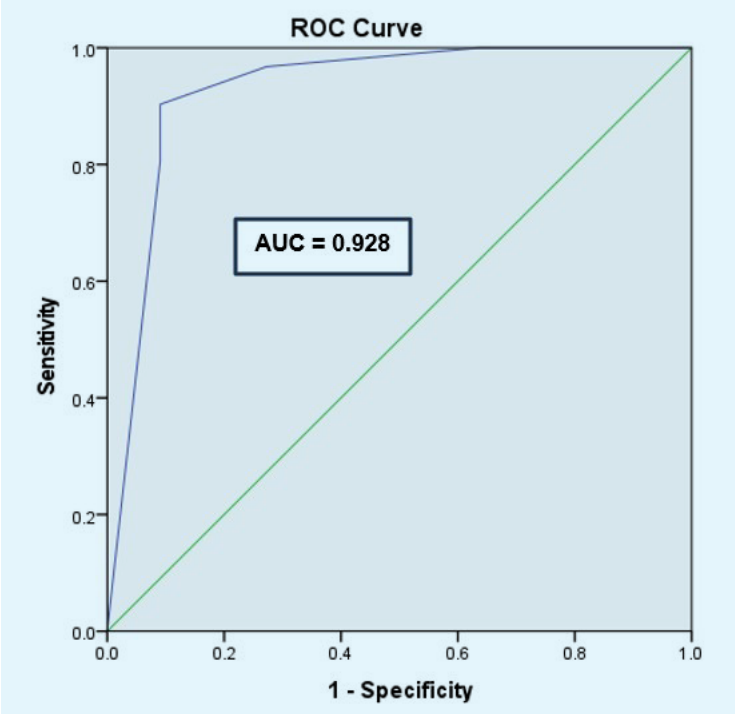


Fig 2. Receiver operator characteristic (ROC) curve showing the cut-off smell score at 5.5 for children across all age groups. The area under curve (AUC) is 0.928.

DISCUSSION

Olfactory dysfunction can lead to decreased appetite, malnutrition and a reduced quality of life.[1,28,29] Smell tests may be beneficial for children with low appetite or poor weight gain. We developed an odor identification test kit using food and flavoring agents because they are closely associated with appetite. Some of the odorants in this study were used in previous studies,[5-7,10-13,19,21-24] combined with new substances commonly found local foods and desserts. Children aged 5-12 years should be familiar with these common fruits and flavoring agents. However, studies have also been conducted on children below 5 years old.[7,10,11,23] *Schriever et al.*[10] found that U-sniff test was reliable and valid for children aged ≥ 4 years. *Cavazzana et al.*[23] reported that children aged 3-4.99 years could successfully complete the Sniffin' Kids test using 11 odors instead of 14. In our study, four substances (squid, garlic, jasmine, and rose) showed that picture identification was influenced by age ($p < 0.05$). However, age did not affect odor selection for any of the 17 odorants in Phase 1. Odor identification performance improves with age, likely due to enhanced cognitive function and verbal abilities[7,30,31] In our study, all children in the 10.01 to 12-year-old group correctly identified all seven odorants in the test kit. *Schriever et al.*[13] also found a positive correlation scores on the Sniffin' Kids test and age ($p < 0.001$; $r = 0.29$). *Dżaman et al.*[19] reported a statistically significant influence of age on odor identification ability ($r = 0.676$, $p < 0.001$). The average time taken to complete the test for seven odors

was 3 to 5 minutes. The longest duration recorded for a correct answer to a single odor test was 30 seconds. Participants who took more than 30 seconds either failed to provide an answer or gave an incorrect one. The test-retest reliability of our test ($r = 0.932$, $p < 0.001$) was comparable to other smell tests for children.[6,7,11,13,32,33]

Eleven children who were able to express their loss of smell had either ARS or Kallmann syndrome. These children took longer to complete the test and had significantly lower smell scores. ARS causes mucosal swelling, impaired mucociliary transport, and increased secretion, all of which prevent odors from reaching the olfactory epithelium.[34,35] Kallmann syndrome is a genetic disease characterized by hypogonadotropic hypogonadism and anosmia.[36] Magnetic resonance imaging of patients with Kallmann syndrome reveals aplasia or hypoplasia of the olfactory bulbs and tracts.[37]

Children with CP have anatomical abnormalities that can contribute to olfactory dysfunction. *Grossmann et al.*[20] found that patients with unilateral CP had a higher smell threshold on the cleft side. In their study, the mean number of identified odors was 1.8 in the cleft group and 2.7 in the control group, though the subjective sense of smell was similar in both groups.[20] *Mani et al.*[26] demonstrated a significant reduction in smell scores in unilateral CP patients when compared to participants without cleft ($p = 0.005$). Despite this evidence, the results of our study showed no difference between children with and without CP in terms of smell scores or the time taken to complete the test. *Roosenboom et al.*[38] also reported no significant difference in Sniffin' Sticks scores between participants with non-syndromic cleft lip and/or CP and participants without cleft.

In our study, the discriminant scores between normal and abnormal olfactory function were determined using the ROC curve and the 10th percentile of the smell score, with both methods yielding similar results. Since the scores were recorded as integers, a smell score of 6 and 7 was considered indicative of normal olfactory function based on both approaches.

The limitations of our study were the small number of children with olfactory dysfunction. Kallmann syndrome is a rare genetic disorder, and we had only one case in this study. Additionally, there were 31 participants with normal smell in Phase 2, but we were unable to provide normative data for subgroup due to the limited number of participants within each subgroup. Furthermore, we did not use a tool to measure the concentration of the odors, which might be the limitation of the study. However, we had normal adult subjects who smelled the substance at the concentrations used in the test kit before using them for the first time to ensure the optimal level of the odor. After that, we controlled the amount of the substance in every use with precise measurements. For future development, an odor meter should be used to test the consistency of odors, ensuring a uniform scent every time a test is conducted.

We developed the test kit using odor sources from widely available food products that can be found in convenience stores across Thailand and countries in Southeast Asia. With only seven items, the test kit is easy to prepare, accurate regardless of the season, and suitable for use in a variety of hospital settings. The test duration was no more than 5 minutes, making it suitable for incorporation into regular outpatient visits. The selection of odors could be applied to children in other Southeast Asian countries where commercial test kits are either unavailable or unfamiliar. In the future, we plan to transform these odorants into pen-like sticks to enhance their clinical use and facilitate future research. Further studies are needed to evaluate the test in children with specific causes of smell loss and those with failure to thrive, where olfactory dysfunction may be a cause.

CONCLUSION

The odor identification test kit consisted of seven odorants derived from fresh fruit, food and flavoring agents that are familiar to Thai children. Children aged 5-12 years with no symptoms of smell loss were able to complete the test in an average of 3.2 minutes. The test effectively distinguished children with reported smell loss from those with normal olfactory function. The cut-off scores were 5.5 based on the ROC analysis and 5.2 from the 10th percentile. Children with repaired CP showed no significant difference in smell scores compared to children without cleft.

Data Availability Statement

The data that support the findings of this study are available upon request from the corresponding author, [K.U.]. The data are not publicly available due to containing information that could compromise the privacy of research participants.

ACKNOWLEDGEMENTS

The authors would like to express appreciation to Miss Jeerapa Kerdnopakun and Miss Ngamrat Treerassapanich, our research assistants, for their help with reference searches, manuscript formatting, and reference management.

DECLARATION

Grants and Funding Information

This research project was supported by the Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, grant number (IO) R016431041.

Conflict of Interest

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

Registration Number of Clinical Trial

TCTR20250207001

Author Contributions

T.W: Conceptualization, Methodology, Validation, Investigation, Resources, Writing – Original Draft, Visualization, Supervision; S.W.: Validation, Investigation, Resources, Writing – Original Draft.; T.S.: Methodology, Formal analysis.; A.T.: Investigation, Resources, Writing – Original Draft.; V.V.: Investigation, Resources, Writing – Original Draft.; K.U.: Conceptualization, Methodology, Validation, Formal analysis, Data Curation, Writing – Review & Editing, Project administration.

Use of Artificial Intelligence

None

Ethics Approval

This study was approved by the Institutional Review Board of Siriraj Hospital, Faculty of Medicine Siriraj Hospital, Mahidol University (COA Si 148/2021).

REFERENCES

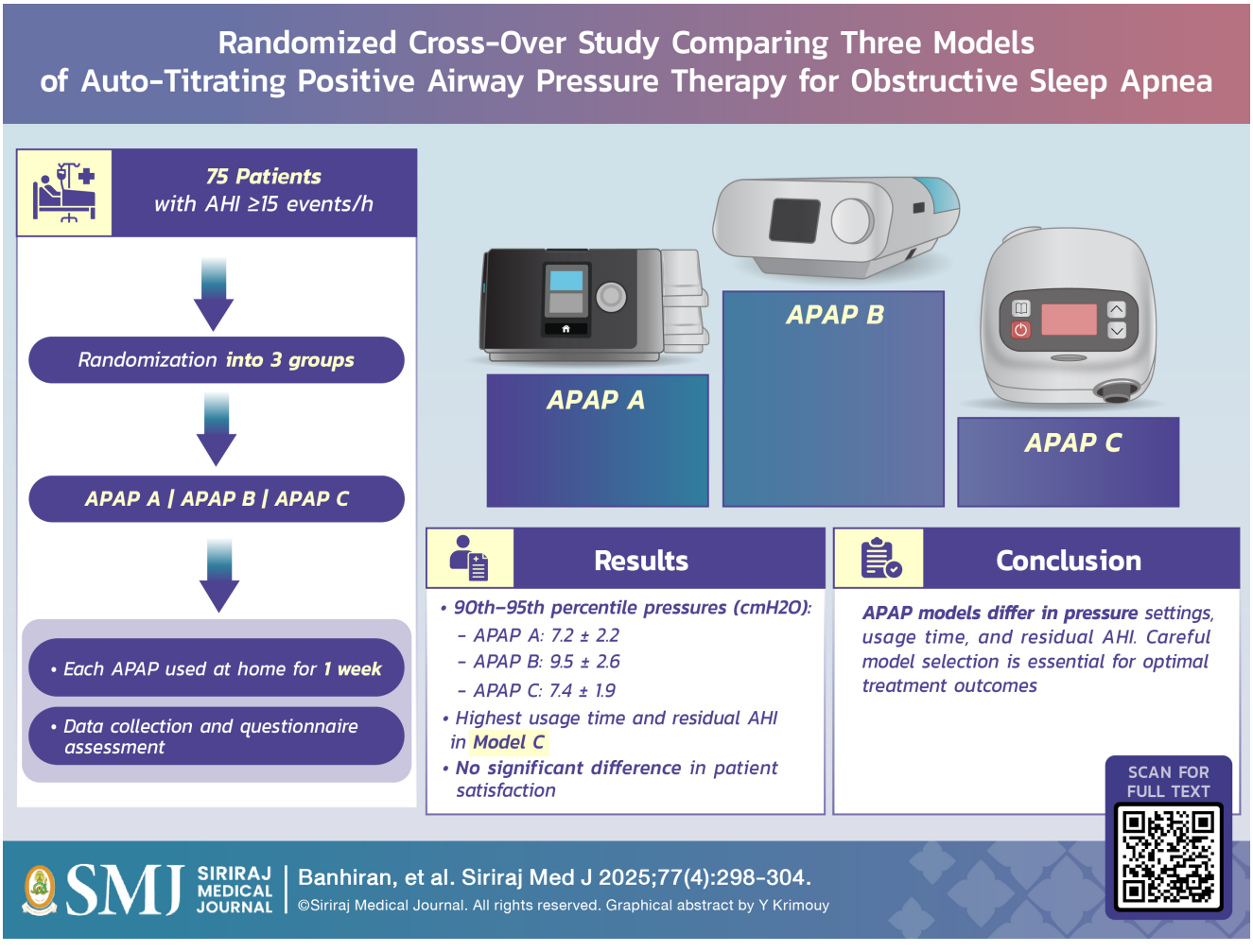
1. Pinkaew B, Assanasen P, Michel O, Talek K, Phonmanee T, Jeerapa Kerdnoppakhun J x. Impact Assessment of Smell and Taste Disorders on Quality of Life in Thais Using the SF-36 Health Survey (Thai version). Siriraj Med J. 2019;71(2):102-9.
2. Li W, Luxenberg E, Parrish T, Gottfried JA. Learning to smell the roses: experience-dependent neural plasticity in human piriform and orbitofrontal cortices. Neuron. 2006;52(6):1097-108.
3. Whitcroft KL, Altundag A, Balungwe P, Boscolo-Rizzo P, Douglas R, Bnecilla MLB, et al. Position paper on olfactory dysfunction: 2023. Rhinology. 2023;61(33):1-108.
4. Ottaviano G, Cantone E, D'Errico A, Salvalaggio A, Citton V, Scarpa B, et al. Sniffin' Sticks and olfactory system imaging in patients with Kallmann syndrome. Int Forum Allergy Rhinol. 2015;5(9):855-61.
5. Cameron EL. Olfactory perception in children. World J Otorhinolaryngol Head Neck Surg. 2018;4(1):57-66.
6. Schriever VA, Agosin E, Altundag A, Avni H, Van HC, Cornejo C, et al. Development of an International Odor Identification Test for Children: The Universal Sniff Test. J Pediatr. 2018;198:265-272.e3.
7. Cameron EL, Doty RL. Odor identification testing in children and young adults using the smell wheel. Int J Pediatr Otorhinolaryngol. 2013;77(3):346-50.
8. Fornazieri MA, Ebara LK, Araujo RG, Lima JVF, Favareto FB, Pinna FR, et al. Adaptation of the Pediatric Smell Wheel(TM) to evaluate olfactory function in Brazilian children. Braz J Otorhinolaryngol. 2022;88(Suppl 5):S47-S51.
9. Ayabe-Kanamura S, Schicker I, Laska M, Hudson R, Distel H, Kobayakawa T, et al. Differences in perception of everyday odors: a Japanese-German cross-cultural study. Chem Senses. 1998;23(1):31-8.
10. Schriever VA, Zscheile L, Gellrich J, Hummel T. Odor identification performance in children aged 3–6 years. Pediatr Res. 2021;89(5):1304-9.
11. Dalton P, Doty RL, Murphy C, Frank R, Hoffman HJ, Maute C, et al. Olfactory assessment using the NIH Toolbox. Neurology. 2013;80(11 Suppl 3):S32-S36.
12. Murphy C, Anderson JA, Markison S. Psychophysical Assessment of Chemosensory Disorders in Clinical Populations. In: Kurihara K, Suzuki N, Ogawa H, eds. Olfaction and Taste XI. Springer Japan; 1994.p.609-13.
13. Schriever VA, Mori E, Petters W, Boerner C, Smitka M, Hummel T. The "Sniffin' Kids" test--a 14-item odor identification test for children. PLoS One. 2014;9(6):e101086.
14. Smith WM, Davidson TM, Murphy C. Toxin-induced chemosensory dysfunction: a case series and review. Am J Rhinol Allergy. 2009;23(6):578-81.
15. Agin K, Hassanian-Moghaddam H, Shadnia S, Rahimi HR. Characteristic manifestations of acute paint thinner-intoxicated children. Environ Toxicol Pharmacol. 2016;45:15-9.
16. Milton LA, White AR. The potential impact of bushfire smoke on brain health. Neurochem Int. 2020;139:104796.
17. Filipsson AF. Short term inhalation exposure to turpentine: toxicokinetics and acute effects in men. Occup Environ Med. 1996;53(2):100-5.
18. Gellrich J, Sparing-Paschke LM, Thieme T, Schwabe K, Dworschak A, Hummel T, et al. Normative data for olfactory threshold and odor identification in children and adolescents. Int J Pediatr Otorhinolaryngol. 2019;123:5-9.
19. Dżaman K, Zielnik-Jurkiewicz B, Jurkiewicz D, Molińska-Glura M. Test for screening olfactory function in children. Int J Pediatr Otorhinolaryngol. 2013;77(3):418-423.
20. Grossmann N, Brin I, Aizenbud D, Sichel JY, Gross-Isseroff R, Steiner J. Nasal airflow and olfactory function after the repair of cleft palate (with and without cleft lip). Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2005;100(5):539-44.
21. Doty R, Marcus A, Lee W. Development of the 12-item Cross-Cultural Smell Identification Test (CC-SIT). Laryngoscope. 1996;106(3 Pt 1):353-6.
22. Bastos LOD, Guerreiro MM, Lees AJ, Warner TT, Silveira-Moriyama L. Effects of Age and Cognition on a Cross-Cultural Paediatric Adaptation of the Sniffin' Sticks Identification Test. PLoS One. 2015;10(8):e0131641.
23. Cavazzana A, Wesarg C, Schriever VA, Hummel T, Lundstrom JN, Parma V. A Cross-Cultural Adaptation of the Sniffin' Sticks Olfactory Identification Test for US children. Chem Senses. 2017;42(2):133-40.

24. Dalton P, Mennella JA, Maute C, Castor SM, Silva-Garcia A, Slotkin J, et al. Development of a test to evaluate olfactory function in a pediatric population. Laryngoscope. 2011;121(9):1843-50.
25. Chen D, Zhang Y, Zhao J, Liu L, Zhao L. Research Progress on Physical Preservation Technology of Fresh-Cut Fruits and Vegetables. Horticulturae. 2024;10(10):1098.
26. Mani M, Moren S, Thorvardsson O, Jakobsson O, Skoog V, Holmstrom M. EDITOR'S CHOICE: objective assessment of the nasal airway in unilateral cleft lip and palate--a long-term study. Cleft Palate Craniofac J. 2010;47(3):217-24.
27. Warren DW, Drake AF. Cleft nose. Form and function. Clin Plast Surg. 1993;20(4):769-79.
28. DeVere R. Disorders of Taste and Smell. Continuum (Minneap Minn). 2017;23(2, Selected Topics in Outpatient Neurology):421-46.
29. Welge-Lüssen A, Hummel T. Management of Smell and Taste Disorders : A Practical Guide for Clinicians. Georg Thieme Verlag; 2014.
30. Monnery-Patris S, Rouby C, Nicklaus S, Issanchou S. Development of olfactory ability in children: sensitivity and identification. Dev Psychobiol. 2009;51(3):268-76.
31. Gellrich J, Sparing-Paschke LM, Hummel T, Schriever VA. The Influence of Cognitive Parameters on Olfactory Assessment in Healthy Children and Adolescents. Chem Senses. 2021;46.
32. Mariño-Sánchez F, Valls-Mateus M, Fragola C, Los Santos G, Aguirre A, Alonso J, et al. Pediatric Barcelona Olfactory Test -6 (pBOT-6): Validation of a Combined Odor Identification and Threshold Screening Test in Healthy Spanish Children and Adolescents. J Investig Allergol Clin Immunol. 2020;30(6):439-47.
33. Krantz EM, Schubert CR, Dalton DS, Zhong W, Huang GH, Klein BEK, et al. Test-retest reliability of the San Diego Odor Identification Test and comparison with the brief smell identification test. Chem Senses. 2009;34(5):435-40.
34. Ciofalo A, de Vincentiis M, Zambetti G, Altissimi G, Fusconi M, Greco A, et al. Olfactory dysfunction in acute rhinosinusitis: intranasal sodium hyaluronate as adjuvant treatment. Eur Arch Otorhinolaryngol. 2017;274(2):803-8.
35. Dalton P. Olfaction and anosmia in rhinosinusitis. Curr Allergy Asthma Rep. 2004;4(3):230-6.
36. Dodé C, Hardelin J-P. Kallmann syndrome. Eur J Hum Genet. 2009;17(2):139-46.
37. Klingmuller D, Dewes W, Krahe T, Brecht G, Schweikert HU. Magnetic resonance imaging of the brain in patients with anosmia and hypothalamic hypogonadism (Kallmann's syndrome). J Clin Endocrinol Metab. 1987;65(3):581-4.
38. Roosenboom J, Hermans R, Lammens F, Samain JL, Devriendt K, Poorten VV, et al. Olfactory function in patients with nonsyndromic orofacial clefts and their unaffected relatives. Am J Med Genet A. 2018;176(11):2375-81.

Randomized Cross-Over Study Comparing Three Models of Auto-Titrating Positive Airway Pressure Therapy for Obstructive Sleep Apnea

Wish Banhiran, M.D.[1], Sirithorn Kormongkolkul, M.D.[1], Phawin Keskool, M.D.[1,*], Wattanachai Chotinaiwattarakul, M.D.[2,3], Sarin Rungmanee, M.D.[3]

[1]Department of Otorhinolaryngology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [2]Division of Neurology, Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [3]Siriraj Sleep Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.



**Corresponding Author: Phawin Keskool*
E-mail: sientpk@hotmail.com. Phawin.kes@mahidol.ac.th
Received 3 December 2024 Revised 16 February 2025 Accepted 16 February 2025
ORCID ID:http://orcid.org/0000-0003-2743-4688
https://doi.org/10.33192/smj.v77i4.272529

ABSTRACT

Objective: To compare pressure, average usage time, residual apnea-hypopnea index (AHI), and patient satisfaction across three models of auto-titrating positive airway pressure (APAP) (A, B, and C refers to APAP devices — APEX, Philips, and Hoffrichter, respectively).

Materials and Methods: Seventy-five adult patients with an AHI of ≥15 events/h who were willing to use APAP were included and randomly assigned to three groups with different APAP sequences generated by computer randomization. After using each model at home for a week, patients returned to the clinic for data collection and switched to the next model. They completed questionnaires regarding their symptoms before and after the therapeutic session.

Results: Data from 62 patients (43 males, 19 females) who completed the research protocol were analyzed. The average 90th–95th percentile pressures for APAP models A, B and C were 7.2 ± 2.2, 9.5 ± 2.6 and 7.4 ± 1.9 cmH_2O, respectively ($p < 0.001$), with an intra-class correlation coefficient (ICC) of 0.52 (95% CI 0.22-0.70). In addition, average usage time and residual AHI differed significantly, with the highest values in model C. However, no significant differences were found in mean pressure or patient satisfaction across the three APAP models.

Conclusion: The 90th–95th percentile pressures, average time usage, and residual AHI varied among APAP models, showing only moderate consistency. These findings suggest that careful consideration is required when selecting an APAP model for home use, as it may affect pressure determination and treatment outcomes.

Keywords: Auto-titrating positive airway pressure; APAP; obstructive sleep apnea; pressure titration; continuous positive airway pressure; CPAP (Siriraj Med J 2025; 77: 298-304)

INTRODUCTION

Obstructive sleep apnea (OSA) is a common and potentially serious disease characterized by recurrent episodes of partial or complete upper-airway obstruction during sleep, leading to hypoxemia, hypercarbia, and interrupted sleep.[1-8] Patients with OSA often present with symptoms such as loud habitual snoring, excessive daytime sleepiness (EDS), irritability, reduced concentration, memory decline, diminished quality of life, an increased risk of motor vehicle or occupational accidents, and coexisting cardiovascular disease.[9-13]

The first-line treatments for OSA currently include sleep hygiene, weight reduction for overweight or obese patients, and continuous positive airway pressure (CPAP) therapy, particularly for those with moderate to severe OSA.[14-20] The primary mechanism of CPAP is to act as a pneumatic splint, keeping the upper airway open by applying positive pressure to the pharynx via the device's interface. When used regularly and with optimal airway pressure, CPAP can alleviate symptoms and improve patients' quality of life.[14,21,22]

There are several methods to determine the optimal airway pressure for OSA patients, including in-lab polysomnography (PSG) with CPAP titration, either full-night or split-night protocols;[14,23] auto-titrating positive airway pressure (APAP) at home;[20,24,25] and predictive calculated formulas. However, PSG is expensive and cumbersome, and predictive formulas are usually inaccurate. Home APAP titration allows patients to use APAP at home for a week, after which the 90th -95th percentile pressure that reduces the respiratory event index (REI) to 5 events/h or less is selected as the optimal fixed CPAP pressure. This method is currently considered the most practical.[20,24,25] However, there are various models of APAP available, and limited data exists comparing clinically relevant differences and the optimal pressure settings between them.[26,27] At Siriraj Hospital, three models of APAP are used for OSA patients, but no data regarding the differences in their effectiveness is available. Therefore, the primary objective of this study is to compare the 90th–95th percentile pressure recorded from the three models of APAP used at Siriraj Hospital.

The secondary objective is to compare the three APAP models in terms of mean pressure, maximum pressure, average usage time, residual AHI, and patient satisfaction.

MATERIALS AND METHODS

Study design

This randomized crossover study was conducted on three models of APAP at the snoring clinic in the Department of Otorhinolaryngology and the Siriraj Sleep Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, between September 2018 and December 2019. The study was approved by the Siriraj Institutional Review Board (SIRB) (COA no.

Si 418/2018) and written consent was obtained from all participants before they enrolled.

Subjects

The inclusion criteria were patients aged ≥ 18 years with an apnea-hypopnea index (AHI) ≥15 events/h, as determined by diagnostic PSG. Exclusion criteria included patients with severe or unstable medical conditions, such as recent myocardial infarction or stroke, and those who declined CPAP therapy.

Intervention

During a routine follow-up visit after PSG, patients were randomly assigned to one of three groups according to APAP sequences generated by a randomization program by a statistician not involved in data collection, patients were then asked to use all three APAP models at home, each for one week. After using each model, the patients returned to the clinic for a follow-up, where their symptoms were reviewed, APAP data downloaded, and they switched to the next model (Fig 1). All patients completed questionnaires regarding their symptoms before treatment and after each therapeutic session. Treatment intolerance or failure was recorded if participants withdrew from the study due to adverse effects or if they failed to follow up for any reason.

Auto-titrating positive airway pressure

Three models of APAP were used in this study: (A) XT Auto (Apex Medical Corporation, New Taipei City, Taiwan), (B) REMstar Auto A-Flex System One 60 Series (Philips Respironics Murrysville, USA), and (C) Hoffrichter Point 2 AutoCPAP (Hoffrichter GmbH, Schwerin, Germany). The pressure for all devices was set between 5-15 cmH_2O. Each patient received education on APAP use before treatment and was instructed to use the devices with a properly selected nasal mask every night or as much as they could tolerate.

Sample size calculation

The sample size of this study was calculated by using data from a previous study by *Nolan, et al.*[26] The 90^{th}–95^{th} percentile pressure comparison among the three APAP models was reported as an intraclass correlation coefficient (ICC). Based on an ICC of 0.8 and an acceptable discrepancy within a 95% confidence interval of 0.15, the calculated sample size was 62. To account for an estimated dropout rate of about 20%, the total sample size for this study was set at 75 participants.

Randomization, allocation, and concealment

Patients were randomly assigned to one of three treatment groups using sequential numbers from a block-

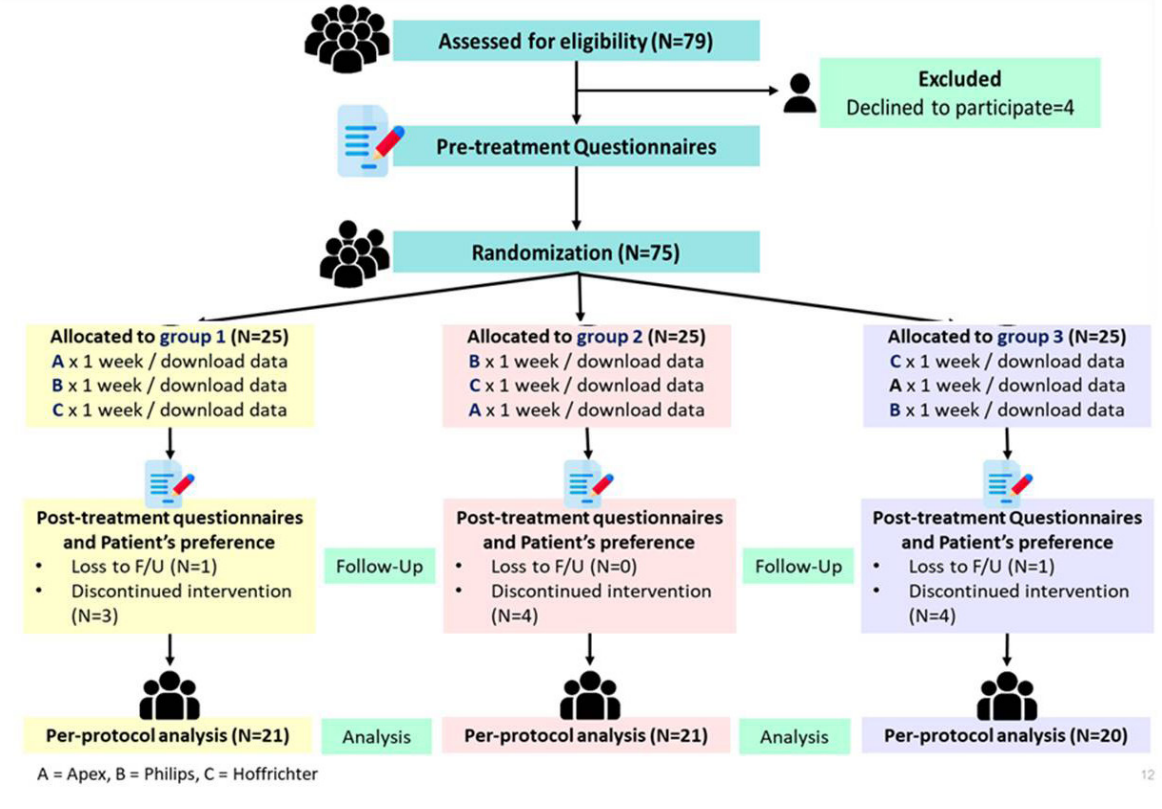


Fig 1. Consolidated Standards of Reporting Trials (CONSORT) flowchart of patients.
Abbreviations: APAP = auto- titrating positive airway pressure; A = APEX; B = Philips; C = Hoffrichte

computerized randomization program managed by a statistician not involved in data collection. The patients in each group were assigned to use the APAP models in the following order: group 1 (25 patients) — A, B, and C; group 2 (25 patients) — B, C, and A; and group 3 (25 patients) — C, A, and B (Fig 1). Patients and researchers were blinded to the sequential numbers prior to group assignment but not during the interventions.

Statistical analysis

Categorical data were presented as numbers and percentages, while continuous data were presented as means ± standard deviation (SD). The 90th–95th percentile pressure comparison among the three APAP models was reported using the intraclass correlation coefficient (ICC), which was interpreted as follows: ICC < 0.4 indicated a poor level of consistency; ICC 0.4-0.74 indicated a moderate level of consistency; and ICC ≥ 0.75 indicated an excellent level of consistency. Other variables were compared using one-way repeated measures analysis of variance (ANOVA) and the Bonferroni post hoc test, with results reported at a 95% confidence interval (CI). A *p* value of < 0.05 was considered statistically significant. All statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) software, version 22.0.

RESULTS

Seventy-nine patients were initially enrolled, but four were excluded (patients declined PAP therapy), leaving 75 patients for randomized. During the follow-up period, 13 patients withdrew from the study (two were lost to follow-up and 11 discontinued the intervention due to intolerance of APAP therapy). As a result, data from 62 patients (43 males and 19 females) with a mean age of 50.2 ± 12.3 years (range 25-72) were analyzed. The mean body mass index (BMI) and mean AHI of all patients were 28.6 ± 5.0 kg/m^2 and 53.0 ± 27.7 events/h, respectively. There were no statistically significant differences in age, gender, body mass index (BMI), AHI, 90th–95th percentile pressure, or mean and minimum oxygen saturation among the three patient groups (Table 1).

Primary outcome

The average 90th–95th percentile pressures for APAP models A, B and C were 7.2 ± 2.2, 9.5 ± 2.6 and 7.4 ± 1.9 cmH_2O, respectively, with an ICC of 0.52 (95% CI 0.22-0.70), indicating a moderate level of consistency.

Secondary outcomes

There were no significant differences in the mean pressure or patient satisfaction among the three APAP models. However, the average usage time and residual AHI differed significantly, shown in Table 2.

TABLE 1. Demographics of the three groups of participants.

	Group 1 (ABC, N = 21)	Group 2 (BCA, N = 21)	Group 3 (CAB, N = 20)	*P*-value
Male, N (%)	14	14	15	0.44
Age (y)	51.8 ± 12.2	51.4 ± 13.6	47.5 ± 11.0	0.47
BMI (kg/m^2)	28.2 ± 4.1	27.9 ± 5.0	29.6 ± 6.1	0.55
AHI, events/h	49.9 ± 26.8	55.4 ± 31.6	53.7 ± 24.6	0.80
Mean O_2 saturation (%)	94.1 ±2.9	93.9 ± 2.9	93.6 ± 2.7	0.87
Lowest O_2 saturation (%)	80.1 ± 7.9	80.3 ± 10.5	79.1 ± 6.5	0.89
90th – 95th pressure of A (cmH_2O)	7.2 ± 1.9	6.9 ± 2.4	7.5 ± 2.5	0.84
90th – 95th pressure of B (cmH_2O)	9.0 ± 1.9	10.1 ± 3.2	9.6 ± 2.6	0.41
90th – 95th pressure of C (cmH_2O)	7.4 ± 2.0	7.5 ± 2.1	7.2 ± 1.8	0.82

Abbreviations: BMI = body mass index, AHI = apnea-hypopnea index, APAP = auto-titrating positive airway pressure, 90th – 95th pressure = 90th – 95th percentile pressure of APAP

TABLE 2. Recorded data of the three models of APAP.

	A	B	C	*P*- value
90th–95th pressure (cmH_2O)	7.2 ± 2.2	9.5 ± 2.6	7.4 ± 1.9	<0.001*
Mean pressure (cmH_2O)	6.6 ± 1.7	7.5 ± 1.8	5.6 ± 0.8	0.058
Average time usage (h/night)	5.5 ± 1.7	5.2 ± 1.9	5.7 ± 1.6	0.011*
Residual AHI (events/h)	2.8 ± 2.8	3.2 ± 2.2	7.7 ± 6.1	<0.001*

Abbreviations: AHI = apnea-hypopnea index, APAP = auto-titrating positive airway pressure, 90th – 95th pressure = 90th – 95th percentile pressure of APAP

*The statistical significant was accepted as *p*-value < 0.05.

Adverse side effects and treatment intolerance

The common side effects of APAP therapy were dry mouth, discomfort, burden, a feeling of burden, and nasal obstruction. However, these side effects were mild in most patients.

DISCUSSION

Although there are several methods to determine the effective airway pressure for OSA patients, home APAP titration currently appears to be the most practical method.[20,24,25] However, there is limited data comparing the optimal generated pressure across different APAP models. After randomization with no significant differences in baseline characteristics among the three APAP sequence groups, the results showed that the 90th–95th percentile pressure among the three APAP models were moderately consistent (ICC of 0.52, 95% CI 0.22-0.70), with the highest pressure in model B. A subsequent analysis comparing the 90th–95th percentile APAP pressure and the optimal pressure derived from split-night PSG in 30 patients demonstrated that model B had pressures closest to those obtained from split-night PSG (ICC of 0.76, 95% CI 0.50-0.89), which is consistent with previous studies.[27] Despite this, there were no significant differences in mean pressure or patient satisfaction among the three APAP models.

Although the average time of the APAP models was statistically different, the difference was slight (no more than 30 min/night), and likely not clinically significant. In terms of residual AHI, model C had the highest residual AHI compared to models A and B, which were nearly identical. This difference in residual AHI could have implications for follow-up in assessing treatment effectiveness.

This study had several limitations. First, the optimal pressure derived from APAP was not compared with the gold standard, i.e. PSG with CPAP titration, in all cases. Only 30 patients had data available from split-night PSG. Second, the three APAP models used in our hospital may not represent other models used elsewhere. Nevertheless, the results of this study suggest that using different APAP models to determine optimal pressure at home could result in varying treatment outcomes due to differences in the 90th-95th pressure and residual AHI recorded by each device. We recommend that future studies explore the long-term clinical impacts of different APAP models on treatment outcomes.

CONCLUSION

Although no significant differences were found in mean pressure or patient satisfaction, the 90th–95th percentile pressures differed among the three APAP models, with only moderate consistency. APAP model B had the highest 90th–95th percentile pressure, closely aligning with the optimal pressure derived from split-night PSG, compared to the other two models. Additionally, differences in average usage time and residual AHI were observed among APAP models, particularly with model C. These findings highlight the importance of carefully selecting an APAP model for home use to ensure optimal pressure determination, as different models may lead to varying treatment outcomes. Further studies are needed to explore these differences.

Data Availability Statement

The data that support the findings of this study are available upon request from the corresponding author, [P.K.]. The data are not publicly available due

to containing information that could compromise the privacy of research participants.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to Mr.Suthipol Udompanturak for his contribution as a consulting statistician, as well as to Miss Jeerapa Kerdnoppakhun and Miss Ngamrat Treerassapanich for their invaluable assistance as research assistants. The authors also extend their appreciation to the staff of the Department of Otorhinolaryngology, Siriraj Sleep Center, and to all patients involved in this study

DECLARATION

Grants and Funding Information

This research project was supported by the Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, grant number (IO) R016232008.

Conflict of Interest

The authors declare they have no conflict of interest.

Registration Number of Clinical Trial

TCTR20250223004

Author Contributions

W.B., S.K.; Conceptualization W.B.; Writing - Original Draft. S.K., W.C., and S.R.; Data Curation, Methodology, Investigation. P.K; cWriting - Review & Editing. All authors have accepted responsibility for the entire content of this manuscript and have approved its submission.

Use of Artificial Intelligence

None

Compliance with Ethical Standards

Ethical approval: All procedures performed in studies involving human participants were conducted in accordance with the ethical standards of the institutional and/or national research committee and the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

REFERENCES

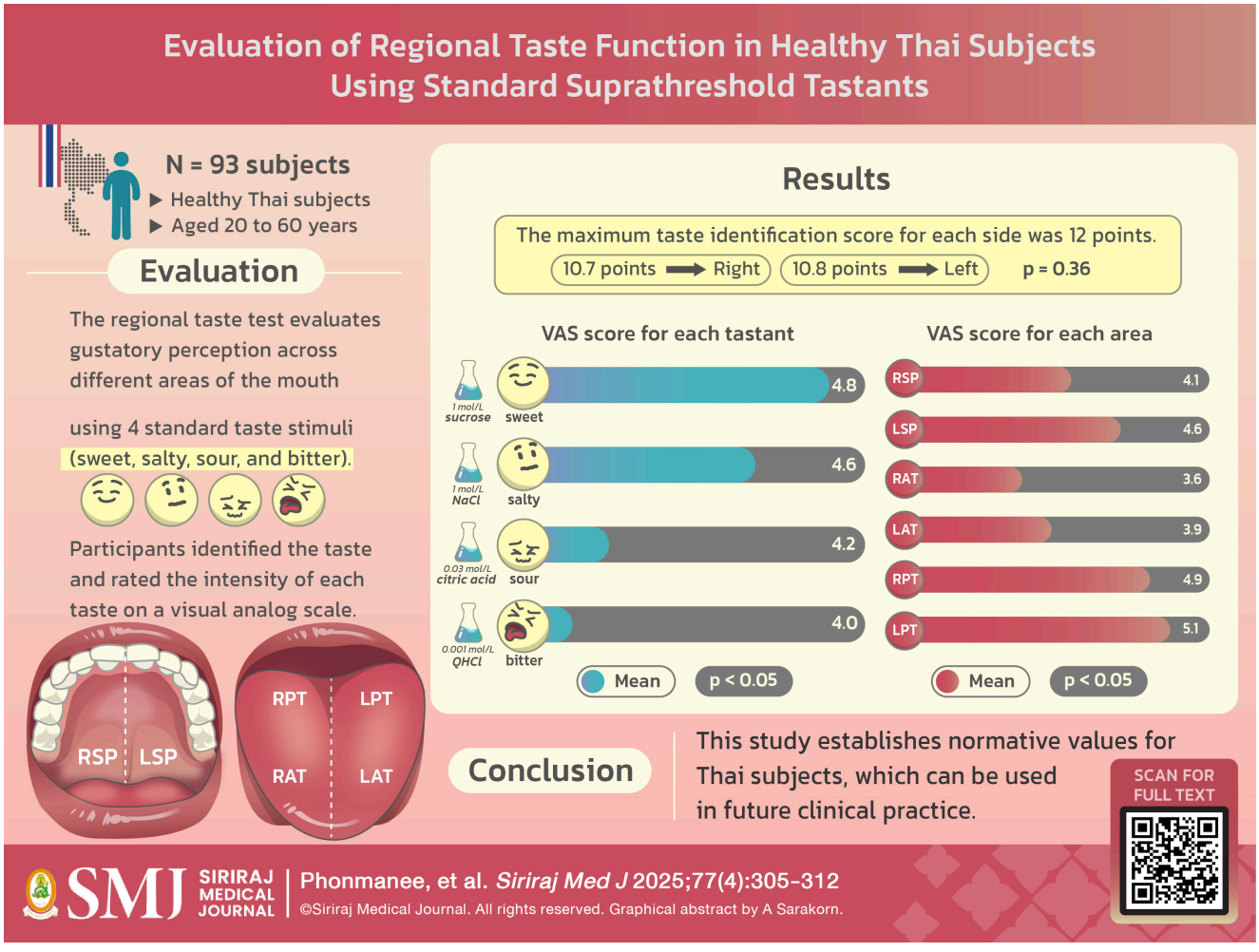
1. Mehra R, Moul DE, Strohl KP. Sleep Breathing Disorder: Clinical Overview. In: Kryger M, Roth T, editors. Principles and Practice of Sleep Medicine. 6th ed. Philadelphia, PA: Elsevier; 2017. p. 167-73.
2. Dempsey JA, Veasey SC, Morgan BJ, O'Donnell CP. Pathophysiology of sleep apnea. Physiol Rev. 2010;90(1):47-112.
3. Rungmanee S, Banhiran W, Keskool P, Assanasen P, Chotinaiwattarakul W, Nujchanart N. Characteristics and clinical presentations of patients at the Siriraj snoring clinic. Siriraj Med J. 2020;72(3):202-8.
4. Franklin KA, Lindberg E. Obstructive sleep apnea is a common disorder in the population-a review on the epidemiology of sleep apnea. J Thorac Dis. 2015;7(8):1311-22.
5. Sompradeekul S, Maranetra N, Naruman C, Nana A. The Prevalence of Obstructive Sleep Apnea Among Professional Hospital Drivers. Siriraj Med J. 2002;54(10):625-33.
6. Ungkanont K, Suraseranivongse S, Laohaprasitiporn D, Vichyanond P. Pediatric Obstructive Sleep Apnea. Siriraj Med J. 2001;53(10):750-3.
7. Soonthornkes N, Sakornwattananon O, Jirapitakkul J, Nuallaong W. Burnout and Associated Factors among Thai Anesthesiology Residents. Siriraj Med J. 2023;75(10):736-43.
8. Nantsupawat N, Khamchet S, Buawongpong N, Pinyopornpanish K, Nantsupawat T, Angkurawaranon C, et al. Sleep Quality and Associated Factors in Elderly Patients with Type-2 Diabetes Mellitus. Siriraj Med J. 2023;75(5):392-8.
9. Dong JY, Zhang YH, Qin LQ. Obstructive sleep apnea and cardiovascular risk: meta-analysis of prospective cohort studies. Atherosclerosis. 2013;229(2):489-95.
10. Banhiran W, Assanasen P, Metheetrairut C, Nopmaneejumruslers C, Chotinaiwattarakul W, Kerdnoppakhun J. Functional outcomes of sleep in Thai patients with obstructive sleep-disordered breathing. Sleep Breath. 2012;16(3):663-75.
11. Banhiran W, Assanasen P, Metheetrairut C, Chotinaiwattarakul W. Health-related quality of life in Thai patients with obstructive sleep disordered breathing. J Med Assoc Thai. 2013;96(2): 209-16.
12. Tregear S, Reston J, Schoelles K, Phillips B. Obstructive sleep apnea and risk of motor vehicle crash: systematic review and meta-analysis. J Clin Sleep Med. 2009;5(6):573-81.
13. Rodenstein D. Sleep apnea: traffic and occupational accidents--individual risks, socioeconomic and legal implications. Respiration. 2009;78(3):241-8.
14. Epstein LJ, Kristo D, Strollo PJ, Jr., Friedman N, Malhotra A, Patil SP, et al. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. J Clin Sleep Med. 2009;5(3):263-76.
15. Lee SA, Paek JH, Han SH. Sleep hygiene and its association with daytime sleepiness, depressive symptoms, and quality of life in patients with mild obstructive sleep apnea. J Neurol Sci. 2015;359(1-2):445-9.
16. Tuomilehto H, Seppa J, Uusitupa M. Obesity and obstructive sleep apnea--clinical significance of weight loss. Sleep Med Rev. 2013;17(5):321-9.
17. Aloia MS, Stanchina M, Arnedt JT, Malhotra A, Millman RP. Treatment adherence and outcomes in flexible vs standard continuous positive airway pressure therapy. Chest. 2005; 127(6):2085-93.
18. Juhasz J, Becker H, Cassel W, Rostig S, Peter JH. Proportional positive airway pressure: a new concept to treat obstructive sleep apnoea. Eur Respir J. 2001;17(3):467-73.
19. Berry RB, Sriram P. Auto-adjusting positive airway pressure treatment for sleep apnea diagnosed by home sleep testing. J Clin Sleep Med. 2014;10(12):1269-75.
20. Wongsritrang K, Fueangkamloon S. Comparison of the clinical

outcomes between unattended home APAP and polysomnography manual titration in obstructive sleep apnea patients. J Med Assoc Thai. 2013;96(9):1159-63.
21. Kim Y, Koo YS, Lee HY, Lee SY. Can Continuous Positive Airway Pressure Reduce the Risk of Stroke in Obstructive Sleep Apnea Patients? A Systematic Review and Meta-Analysis. PLoS One. 2016;11(1):e0146317.
22. Kushida CA, Littner MR, Hirshkowitz M, Morgenthaler TI, Alessi CA, Bailey D, et al. Practice parameters for the use of continuous and bilevel positive airway pressure devices to treat adult patients with sleep-related breathing disorders. Sleep. 2006;29(3):375-80.
23. Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K, et al. Clinical Practice Guideline for Diagnostic Testing for Adult Obstructive Sleep Apnea: An American Academy of Sleep Medicine Clinical Practice Guideline. J Clin Sleep Med. 2017;13(3):479-504.
24. Hui DS, Ng SS, To KW, Ko FW, Ngai J, Chan KK, et al. A randomized controlled trial of an ambulatory approach versus the hospital-based approach in managing suspected obstructive sleep apnea syndrome. Sci Rep. 2017;8:45901.
25. Guerrero A, Embid C, Isetta V, Farre R, Duran-Cantolla J, Parra O, et al. Management of sleep apnea without high pretest probability or with comorbidities by three nights of portable sleep monitoring. Sleep. 2014;37(8):1363-73.
26. Nolan GM, Ryan S, O'Connor T M, McNicholas WT. Comparison of three auto-adjusting positive pressure devices in patients with sleep apnoea. Eur Respir J. 2006;28(1):159-64.
27. Gagnadoux F, Pevernagie D, Jennum P, Lon N, Loiodice C, Tamisier R, et al. Validation of the System One RemStar Auto A-Flex for Obstructive Sleep Apnea Treatment and Detection of Residual Apnea-Hypopnea Index: A European Randomized Trial. J Clin Sleep Med. 2017;13(2):283-90.

Evaluation of Regional Taste Function in Healthy Thai Subjects Using Standard Suprathreshold Tastants

Tharatham Phonmanee, B.Ed., Naruphon Sastranuruk, M.D., Paraya Assanasen, M.D. , Kawita Atipas, M.D.*

Department of Otorhinolaryngology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.



*Corresponding author: Kawita Atipas
E-mail: kawita.ati@mahidol.ac.th
Received 3 December 2024 Revised 22 January 2025 Accepted 22 January 2025
ORCID ID:http://orcid.org/0000-0001-6784-9525
https://doi.org/10.33192/smj.v77i4.272535

ABSTRACT

Objective: To evaluate regional taste function in normal Thai subjects using standard suprathreshold tastants.

Materials and Methods : A total of 93 healthy Thai subjects (41 males and 52 females), aged 20 to 60 years, were recruited. Regional taste function was assessed using four standard suprathreshold tastants, with both identification scores and VAS scores recorded.

Results: The mean total taste identification score was 10.7 on the right side and 10.8 on the left side, with no statistically significant differences between sides across all subgroups. The mean VAS score was highest for sweet tastants and for the posterior tongue as the tested area. No significant differences were observed between sexes or among age groups when analyzed by tastant and area.

Conclusion: The regional (spatial) taste test is a useful gustatory test. This study establishes normative values for Thai subjects, which can be used in future clinical practice for patient evaluations.

Keywords: Regional taste test; spatial taste test; gustatory test; gustatory function (Siriraj Med J 2025; 77: 305-312)

INTRODUCTION

Taste is an important sensory function, as it helps stimulate appetite, identify food components, and detect toxic substances. Several factors can lead to taste dysfunction, including infections, systemic health conditions, poor oral hygiene, prior surgeries, radiation exposure, head injuries, or aging.[1,2] The four basic taste categories are salty, sweet, sour, and bitter, with umami (savory flavor) also proposed as a distinct taste.[1] The prevalence of taste dysfunction has been reported to range from 0.6% to 20%.[3] Loss of taste can have numerous adverse effects, including reduced appetite, weight loss, and psychological issues such as depression.[4,5] Studies have shown that taste dysfunction can significantly impact a patient's quality of life.[4,6] However, patients often cannot identify taste loss and frequently confused olfactory loss with gustatory loss.[4] Furthermore, self-reported taste assessments via questionnaires tend to have low sensitivity and specificity compared to objective taste tests.[7] Localized taste loss is especially hard to detect, as it may lack clear symptoms and is often compensated for by other areas.[8]

Assessing taste function is technically challenging and time-consuming. The two primary types of chemical taste testing are detection threshold tests and regional (spatial) taste tests. This study focuses on regional taste tests in normal Thai subjects, targeting areas innervated by the chorda tympani branch of the facial nerve (anterior tongue), the glossopharyngeal nerve (posterior tongue), and the greater superficial petrosal nerve (soft palate).[6] Data on taste testing in Thailand and Southeast Asian countries are limited, particularly regarding regional taste assessments. The results of this study will provide a valuable reference for regional taste testing in Thai clinical settings.

MATERIALS AND METHODS

A total of 93 healthy Thai subjects (41 males, 52 females) aged between 20 and 60 were recruited. Subjects were required to refrain from eating food, except water, for at least one hour before testing. Subjects with conditions that could affect taste and smell function, such as tongue lesions, a history of tongue surgery or cranial nerve-related surgeries, sinonasal diseases, smoking, or those taking medications that could affect taste function, were excluded from the study. Alcohol consumption status and dietary habits were not part of the exclusion criteria.

The study protocol and consent procedures received ethical approval from the Siriraj Institutional Review Board (COA No. Si 676/2012). All subjects provided documented informed consent before participating in the study.

Test procedures

The regional (spatial) taste function test was adapted from the test developed by the Connecticut Chemosensory Clinical Research Center (CCCRC).[8] The test evaluated six areas: the right and left anterior tongue, posterior tongue, and soft palate, as shown in Fig 1. The liquid taste solutions and concentrations used were 1 mol/L sodium chloride (NaCl) for salty, 1 mol/L sucrose for sweet, 0.03 mol/L citric acid for sour, and 0.001 mol/L quinine hydrochloride (QHCl) for bitter,[8] all at suprathreshold concentrations.

Each taste solution was applied to the six areas using a sterile cotton swab in a random order, resulting in 24 tests. Participants identified the taste as salty, sweet, sour, or bitter, earning a score of 1 for each correct response, with a maximum score of 12 per side of the tongue. Subjects also rated the intensity of each taste on

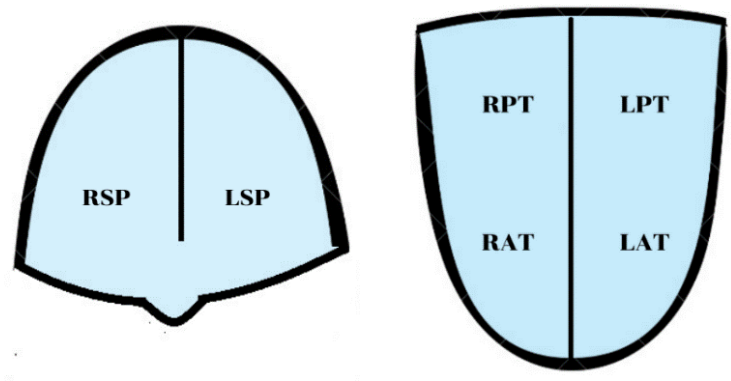


Fig 1. Images showing the tested palate and tongue areas. The labeled areas are where the four tastants were applied.
RSP=right soft palate, LSP=left soft palate, RPT= right posterior tongue, LPT=left posterior tongue, RAT=right anterior tongue, LAT=left anterior tongue (Drawing by Tharatham Phonmanee)

a visual analog scale from 0 to 10, with a higher score reflecting stronger intensity. Between tests, participants rinsed their mouths to minimize carryover effects. All tests were conducted by a single investigator to eliminate inter-rater variability.

Statistical analysis

All statistical analyses were performed using SPSS Statistics version 15.0 (SPSS Inc., Chicago, IL, USA). Paired t-tests were conducted to compare differences between the right and left sides of the mouth. Repeated measures ANOVA was used to analyze differences in VAS scores among taste stimuli and regions. The chi-square test was used to assess statistical differences in the percentage of acceptable taste identification between sexes and across age groups. A p-value <0.05 was considered statistically significant.

RESULTS

A total of 93 subjects participated in the study (mean age 41.6 ± 11.6), comprising 41 males (mean age 40.8 ± 11.3 years) and 52 females (mean age 42.1 ± 11.9 years). Regional taste function testing was performed in six areas using the four basic taste solutions as described earlier. The maximum taste identification score for each side was 12 points. Mean scores for each sex and age group are presented in Table 1. No statistically significant differences were found between the mean scores for each side of the mouth.

Taste identification was deemed acceptable if participants correctly identified the taste in at least five out of six areas of the tongue. The proportion of acceptable taste identification for each taste is shown in Table 2. No statistically significant differences in taste identification ability were observed based on sex and age group.

The mean VAS score for each taste across all six areas is shown in Table 3. Differences in VAS scores among the tastes were statistically significant for the total participant group as well as the male and female subgroups (p <0.05). The mean VAS score was highest for sweet, followed by salty, sour, and bitter, respectively.

Taste identification by area was analyzed, with results shown in Table 4. Acceptable taste identification was defined as correctly identifying at least three of the four basic taste stimuli in each area of the mouth. No statistically significant differences in taste identification ability by sex or age group were observed for any area. The mean VAS score differed significantly among the areas, with the posterior tongue recording the highest mean VAS score, followed by the soft palate and anterior tongue, respectively (Table 5 and Fig 2).

DISCUSSION

Clinical evaluation of taste function is not a routinely performed test. Various methods exist for assessing taste, including self-ratings of taste function, chemical stimuli tests, and electrical stimuli tests (also known as electrogustometry).[9,10] Chemical stimuli tests are commonly divided into whole-mouth and regional testing. These tests typically involve the four basic taste qualities: salty, sweet, sour and bitter. However, specific taste substances and concentrations vary among institutions. Taste tests are generally categorized as threshold tests and suprathreshold tests based on the concentrations of tastants. The application methods depend on the target area being tested, such as the right and left sides of the tongue, the tip/anterior and posterior sections, the four quadrants of the tongue, or areas of the tongue in conjunction with the palatal region. Different tools are used to deliver tastants, including cotton swabs, paper discs, edible wafers, taste strips, and solutions.[9-12]

This study focuses on regional taste testing and provides normative data for Thai subjects. The technique and concentrations used at Siriraj's Smell & Taste Clinic

TABLE 1. Total taste identification scores for each side of the mouth.

	Number of subjects (%)	Mean ± SD		p-value
		Right	Left	
Total	93 (100.0)	10.7±1.7	10.8±1.5	0.36
Sex:				
Male	41 (44.1)	10.6±1.7	10.7±1.6	0.44
Female	52 (55.9)	10.8±1.7	10.9±1.5	0.58
Age:				
20-29	19 (20.4)	11.0±1.2	10.9±1.3	0.87
30-39	23 (24.8)	10.8±1.7	11.0±1.3	0.55
40-49	20 (21.5)	10.4±2.0	10.5±1.8	0.51
50-59	31 (33.3)	10.8±1.8	10.9±1.7	0.42

TABLE 2. Proportion of acceptable taste identification by each taste stimulus (the taste is correctly identified in at least five out of six areas).

	Number of subjects (%)			
	Salty	Sweet	Sour	Bitter
Total	81 (87.1)	90 (96.8)	71 (76.3)	75 (80.6)
Sex:				
Male	36 (87.8)	41 (100.0)	30 (73.2)	33 (80.5)
Female	45 (86.5)	49 (94.2)	41 (78.8)	42 (80.8)
p-value	0.97	0.25	0.52	0.86
Age:				
20-29	18 (94.7)	18 (94.7)	16 (84.2)	14 (73.7)
30-39	20 (87.0)	22 (95.7)	17 (73.9)	20 (87.0)
40-49	16 (80.0)	20 (100.0)	14 (70.0)	14 (70.0)
50-59	27 (87.1)	30 (96.8)	24 (77.4)	27 (87.1)
p-value	0.34	0.89	0.76	0.62

TABLE 3. Mean VAS score for the four tastants (salty, sweet, sour, and bitter).

	Mean ± SD				p-value
	Salty	Sweet	Sour	Bitter	
Total	4.6±2.1	4.8±2.0	4.2±1.9	4.0±1.9	<0.05*
Sex:					
Male	4.3±2.1	4.4±2.1	3.7±1.9	3.4±1.6	<0.05*
Female	4.9±2.1	5.0±1.8	4.6±1.7	4.3±2.1	<0.05*
Age:					
20-29	5.9±2.2	5.1±2.4	4.5±1.9	4.5±2.3	<0.05*
30-39	5.1±2.1	5.3±1.9	4.6±2.2	4.6±1.8	0.16
40-49	4.4±1.9	4.7±1.8	4.1±1.6	3.5±1.4	<0.05*
50-59	3.6±1.8	4.1±1.7	3.8±1.8	3.4±1.9	0.06

TABLE 4. Proportion of acceptable taste identification in each area of the mouth (at least three out of four basic tastes are correctly identified in that area).

	Number of subjects (%)					
	RSP	LSP	RAT	LAT	RPT	LPT
Total	87 (93.5)	87 (93.5)	80 (86.0)	77 (82.8)	88 (94.6)	88 (94.6)
Sex:						
Male	36 (87.8)	38 (92.7)	36 (87.8)	34 (82.9)	39 (95.1)	39 (95.1)
Female	51 (98.1)	49 (94.2)	44 (84.6)	43 (82.7)	49 (94.2)	49 (94.2)
p-value	0.08	1.00	0.66	0.98	1.00	1.00
Age:						
20-29	18 (94.7)	18 (94.7)	17 (89.5)	16 (84.2)	19 (100.0)	19 (100.0)
30-39	22 (95.7)	22 (95.7)	19 (82.6)	21 (91.3)	23 (100.0)	22 (95.7)
40-49	17 (85.0)	19 (95.0)	17 (85.0)	14 (70.0)	18 (90.0)	18 (90.0)
50-59	30 (96.8)	28 (90.3)	27 (87.1)	26 (83.9)	28 (90.3)	29 (93.5)
p-value	0.44	0.94	0.92	0.36	0.26	0.64

Abbreviations: RSP=right soft palate, LSP=left soft palate, RAT=right anterior tongue, LAT=left anterior tongue, RPT= right posterior tongue, LPT=left posterior tongue

TABLE 5. Mean VAS score for each area of the mouth.

	Mean ± SD						p-value
	RSP	LSP	RAT	LAT	RPT	LPT	
Total	4.1±1.8	4.6±1.9	3.6±1.8	3.9±1.7	4.9±2.2	5.1±2.1	<0.05*
Sex:							
Male	3.6±1.8	4.2±2.0	3.2±1.4	3.5±1.6	4.5±2.2	4.7±2.2	<0.05*
Female	4.5±1.7	4.9±1.8	4.0±2.0	4.2±1.7	5.3±2.2	5.5±2.0	<0.05*
Age:							
20-29	4.5±1.8	5.4±2.3	4.3±2.3	4.5±2.0	5.6±2.4	5.8±2.2	<0.05*
30-39	4.7±1.9	5.1±1.8	4.2±1.5	4.5±1.8	5.1±2.1	5.8±2.2	<0.05*
40-49	3.7±1.6	4.3±1.7	3.4±1.1	3.8±1.2	5.1±2.1	4.9±1.6	<0.05*
50-59	3.8±1.7	4.0±1.8	2.9±1.7	3.1±1.3	4.3±2.3	4.4±2.1	<0.05*

Abbreviations: RSP=right soft palate, LSP=left soft palate, RAT=right anterior tongue, LAT=left anterior tongue, RPT= right posterior tongue, LPT=left posterior tongue

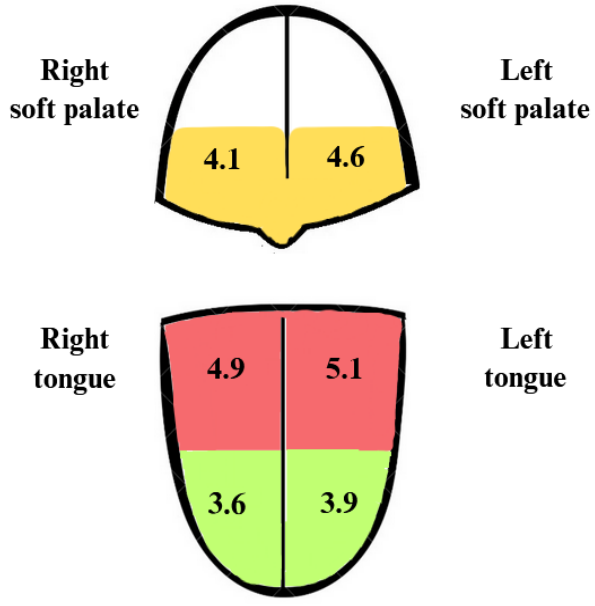


Fig 2. Images showing the mean VAS score for each tested palate and tongue area.
(Drawing by Tharatham Phonmanee)

are adapted from the CCCRC spatial taste test.[8] The main advantage of regional taste testing over whole-mouth tests is its ability to localize specific lesions. Taste buds are primarily located on the tongue but are also found on the soft palate, pharynx, and epiglottis, with taste receptors distributed across other organ systems.[13,14] The nerves that innervate each area of the tongue vary based on the location and types of taste papillae; the anterior tongue contains fungiform papillae and is innervated by the chorda tympani nerve, a branch of the facial nerve; the posterior tongue contains circumvallate and foliate papillae and is innervated by the glossopharyngeal nerve; and the soft palate is innervated by the greater superficial petrosal nerve, which is also a branch of the facial nerve.[13] Regional taste tests can detect specific nerve dysfunctions such as chorda tympani nerve injury after middle ear surgery, which whole-mouth tests may miss. A study by *Matsuda et al.*[15] demonstrated regional taste dysfunction in elderly subjects for NaCl.

As there is no standard cut-off point for correct taste identification, we used a threshold of five out of six areas as the passing level for each taste. With these concentrations, over 70% of subjects passed for each taste stimulus. Subjects rated sweet as the strongest taste overall on the VAS. While intensity ratings for sour tend to be the highest in other studies, comparisons with our findings are not possible due to differences in testing techniques and concentrations.[4,16] The trend of high VAS scores for sweet may influence the consumption of sweet foods and drinks among Thai people; however, further studies on this topic are needed.

By area, over 80% of subjects could identify at least three out of the four basic tastes. The posterior tongue recorded the highest average VAS scores, followed by the soft palate and anterior tongue. The glossopharyngeal nerve carries special visceral afferent fibers from the posterior third of the tongue. The neural pathway passes through the inferior ganglion, then through the jugular foramen to the nucleus solitarius, and continues to other brain areas.[17] The difference in taste sensation across various areas is suggested to be caused by the number of taste buds and nerves that are stimulated.[18] Previous studies on regional taste sensitivity have yielded conflicting findings, with many reporting taste sensitivity varies by stimulus across different areas. *Nilsson et al.*[19] reported higher taste thresholds on the soft palate compared to the tongue. *Collings et al.*[20] found that NaCl had the lowest threshold on the anterior tongue and the highest on the soft palate. Similarly, sucrose had the lowest threshold on the anterior tongue but the highest on the sides, while quinine had the lowest threshold on the soft palate. Another study by *Feeney et al.*[21] found higher intensity ratings for bitter and umami on the posterior tongue compared to the anterior tongue. *Sato et al.*[18] reported no differences in thresholds for sweet, salty, and bitter across regions, but sour thresholds were higher on the soft palate than on the tip of the tongue.

Previous studies have shown that older age is associated with decreased taste perception.[16,22,23] However, our study found no significant differences in taste identification ability across age groups when analyzed by individual taste stimuli and areas of the tongue. This may be because the inclusion criteria limit the maximum age to 60 years, which excludes the elderly population. While some studies suggest that females have better taste sensitivity than males[23,24], this study found no significant differences between male and female subjects.

Genetic variations and ethnicity have been shown to influence taste perception. Studies conducted in the United States have demonstrated differences in taste intensity ratings among various ethnic groups, such as non-Hispanic black and non-Hispanic white adults.[25,26] Another study by *Yang et al.*[27] also showed differences in taste perception between Asians and Caucasians. Therefore, our results are more applicable for use with Thai patients.

CONCLUSION

The regional (spatial) taste test is a valuable tool for assessing gustatory function, complementing threshold tests and electrogustometry. The technique used in this study, adapted from the CCCRC spatial taste test, is also used in clinical practice at Siriraj's Smell & Taste

Clinic. This method effectively identifies taste function for specific taste stimuli in particular regions of the tongue and soft palate. Our findings revealed no significant difference in taste ability between the two sides of the mouth. The mean VAS score was highest for sweet as a tastant and for the posterior tongue as a tested area. Additionally, no significant differences were observed between sexes or among age groups when analyzed by tastant and area. This study also establishes normative values for Thai subjects, which can serve as a reference in future clinical practice. Further studies that include a broader patient age range, all smoking statuses, and patients with taste dysfunction would be valuable to enhance the generalizability and applicability of the data.

Data Availability Statement

The data that support the findings of this study are available upon request from the corresponding author, [K.A.]. The data are not publicly available due to containing information that could compromise the privacy of research participants.

ACKNOWLEDGEMENTS

The authors acknowledge Pimrapat Tengtrakulchareon for her assistance with statistical analysis and Jeerapa Kerdnoppakhun for her support in manuscript preparation.

DECLARATION

Grants and Funding Information

This research project was funded by the Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand, grant number (IO) R015631041.

Conflict of Interest

The authors declare no conflict of interest.

Registration Number of Clinical Trial

TCTR20250211007

Author Contributions

Conceptualization and methodology, N.S. and P.A.; Investigation, N.S. and P.A.; Formal analysis, T.P., N.S. and K.A.; Visualization and writing – original draft, T.P., N.S., P.A. and K.A.; Writing – review and editing, T.P. and K.A.; Funding acquisition, P.A.; Supervision, P.A.; All authors have read and agreed to the final version of the manuscript.

Use of Artificial Intelligence

The manuscript was not produced using artificial intelligence.

Human Ethics Approval Declaration

This study was conducted according to the Declaration of Helsinki and was approved by the Siriraj Institutional Review Board (approval No. Si 676/2012). Informed consent was obtained from all individual participants involved in the study.

REFERENCES

1. Daly BP, Daly MP, Minniti N, Daly JM. Sense of Taste (Effect on Behavior). In: Ramachandran VS, editor. Encyclopedia of Human Behavior (Second Edition). San Diego: Academic Press; 2012. p. 373-8.
2. Gonzales GM, Cook MJ. CHAPTER 13 - DISORDERS OF SMELL AND TASTE. In: Schapira AHV, Byrne E, DiMauro S, Frackowiak RSJ, Johnson RT, Mizuno Y, et al., editors. Neurology and Clinical Neuroscience. Philadelphia: Mosby; 2007. p. 171-7.
3. Liu G, Zong G, Doty RL, Sun Q. Prevalence and risk factors of taste and smell impairment in a nationwide representative sample of the US population: a cross-sectional study. BMJ Open. 2016;6(11):e013246.
4. Deems DA, Doty RL, Settle RG, Moore-Gillon V, Shaman P, Mester AF, et al. Smell and taste disorders, a study of 750 patients from the University of Pennsylvania Smell and Taste Center. Arch Otolaryngol Head Neck Surg. 1991;117(5):519-28.
5. Hur K, Choi JS, Zheng M, Shen J, Wrobel B. Association of alterations in smell and taste with depression in older adults. Laryngoscope Investig Otolaryngol. 2018;3(2):94-9.
6. Pinkaew B, Assanasen P, Michel O, Talek K, Phonmanee T, Kerdnoppakhun J. Impact Assessment of Smell and Taste Disorders on Quality of Life in Thais Using the SF-36 Health Survey (Thai version). Siriraj Med J. 2019;71(2):102-9.
7. Soter A, Kim J, Jackman A, Tourbier I, Kaul A, Doty RL. Accuracy of self-report in detecting taste dysfunction. Laryngoscope. 2008;118(4):611-7.
8. Bartoshuk L. Clinical evaluation of the sense of taste. Ear Nose Throat J. 1989;68(4):331-7.
9. Ikui A. A review of objective measures of gustatory function. Acta Otolaryngol Suppl. 2002(546):60-8.
10. Tomita H, Ikeda M, Okuda Y. Basis and practice of clinical taste examinations. Auris Nasus Larynx. 1986;13 Suppl 1:S1-15.
11. Mueller C, Kallert S, Renner B, Stiassny K, Temmel AF, Hummel T, et al. Quantitative assessment of gustatory function in a clinical context using impregnated "taste strips". Rhinology. 2003;41(1):2-6.
12. Ahne G, Erras A, Hummel T, Kobal G. Assessment of gustatory function by means of tasting tablets. Laryngoscope. 2000;110(8):1396-401.
13. Spence C. The tongue map and the spatial modulation of taste perception. Curr Res Food Sci. 2022;5:598-610.
14. Finger TE, Kinnamon SC. Taste isn't just for taste buds anymore. F1000 Biol Rep. 2011;3:20.
15. Matsuda T, Doty RL. Regional taste sensitivity to NaCl: relationship to subject age, tongue locus and area of stimulation. Chem Senses. 1995;20(3):283-90.
16. Barragán R, Coltell O, Portolés O, Asensio EM, Sorlí JV,

Ortega-Azorín C, et al. Bitter, Sweet, Salty, Sour and Umami Taste Perception Decreases with Age: Sex-Specific Analysis, Modulation by Genetic Variants and Taste-Preference Associations in 18 to 80 Year-Old Subjects. Nutrients. 2018;10(10).
17. Rea P. Chapter 2 - Head. In: Rea P, editor. Essential Clinically Applied Anatomy of the Peripheral Nervous System in the Head and Neck: Academic Press; 2016. p. 21-130.
18. Sato K, Endo S, Tomita H. Sensitivity of three loci on the tongue and soft palate to four basic tastes in smokers and non-smokers. Acta Otolaryngol Suppl. 2002(546):74-82.
19. Nilsson B. Taste acuity of the human palate. III. Studies with taste solutions on subjects in different age groups. Acta Odontol Scand. 1979;37(4):235-52.
20. Collings VB. Human taste response as a function of locus of stimulation on the tongue and soft palate. Perception & Psychophysics. 1974;16(1):169-74.
21. Feeney EL, Hayes JE. Regional differences in suprathreshold intensity for bitter and umami stimuli. Chemosens Percept. 2014;7(3-4):147-57.
22. Mojet J, Christ-Hazelhof E, Heidema J. Taste perception with age: generic or specific losses in threshold sensitivity to the five basic tastes? Chem Senses. 2001;26(7):845-60.
23. da Silva LA, Lin SM, Teixeira MJ, de Siqueira JT, Jacob Filho W, de Siqueira S. Sensorial differences according to sex and ages. Oral Dis. 2014;20(3):e103-10.
24. Bartoshuk LM, Duffy VB, Miller IJ. PTC/PROP tasting: anatomy, psychophysics, and sex effects. Physiol Behav. 1994;56(6):1165-71.
25. Williams JA, Bartoshuk LM, Fillingim RB, Dotson CD. Exploring Ethnic Differences in Taste Perception. Chem Senses. 2016;41(5):449-56.
26. Bowser S, Farnsworth N, Russell K, Schlechter H, Bernstein S, Courville AB, et al. Sweet taste perception is greater in non-Hispanic black than in non-Hispanic white adults. Nutrition. 2019;59:103-7.
27. Yang Q, Williamson A-M, Hasted A, Hort J. Exploring the relationships between taste phenotypes, genotypes, ethnicity, gender and taste perception using Chi-square and regression tree analysis. Food Quality and Preference. 2020;83:103928.