Siriraj Medical Journal
is deeply sadden by the passing of

Her Majesty Queen Sirikit
The Queen Mother

We pay tribute to Her Majesty's
lifelong devotion to the nation
and the people of Thailand

PMAC
PRINCE MAHIDOL AWARD CONFERENCE 2025

E-ISSN 2228-8082
Volume 77, Number 11, November 2025

SMJ

Siriraj Medical Journal

The world-leading biomedical science of Thailand

ORIGINAL ARTICLE REVIEW ARTICLE MONTHLY

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 11
November 2025

ORIGINAL ARTICLE

REVIEW ARTICLE

SIRIRAJ MEDICAL JOURNAL

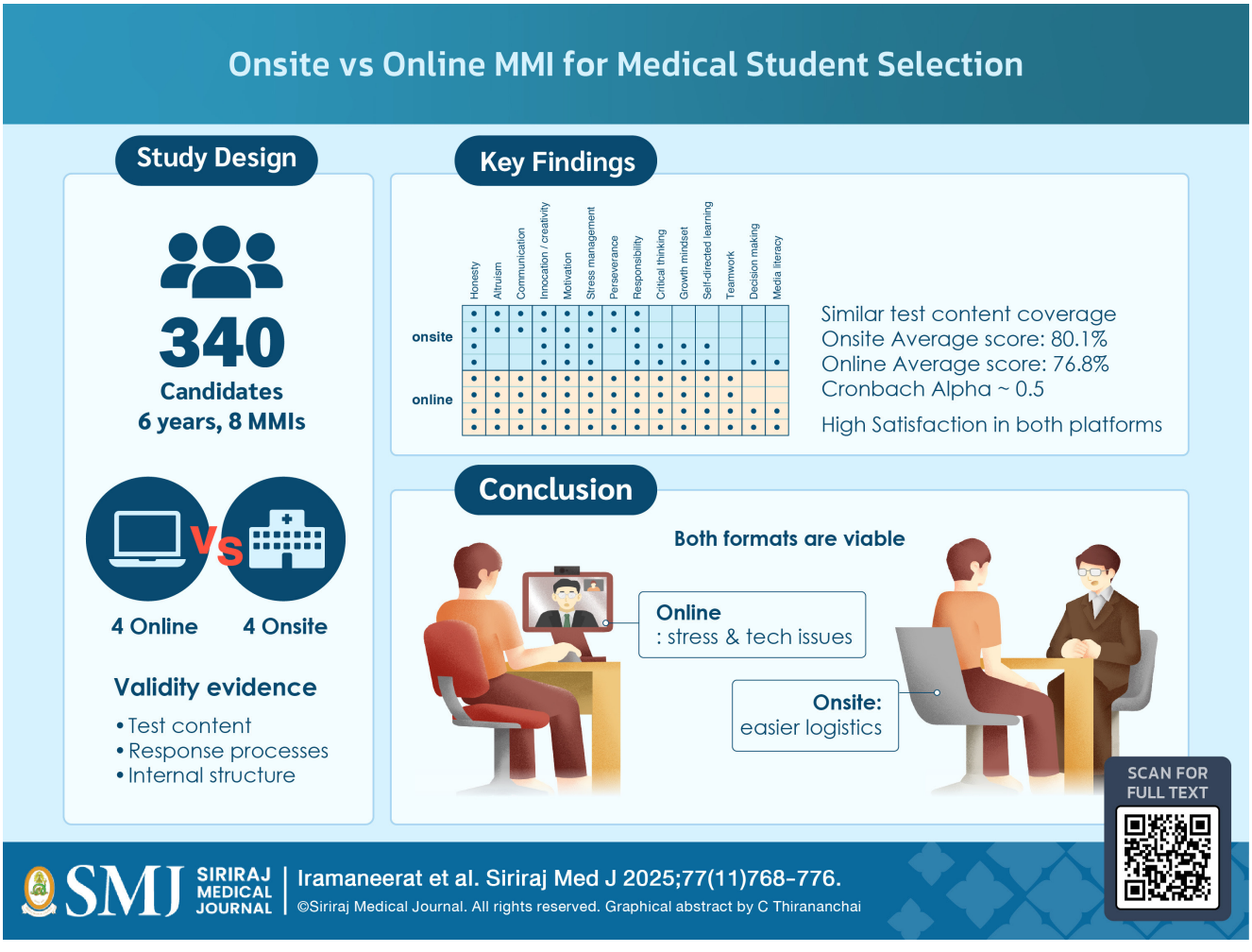
https://he02.tci-thaijo.org/index.php/sirirajmedj/index

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

Validity and Reliability of the Multiple-Mini Interview (MMI) in Medical Student Selection: A Comparison of Onsite and Online Platforms

Cherdsak Iramaneerat, M.D., MHPE, Ph.D.[1,*], Punyapat Maprapho, M.D.[2], Orawan Uakit, M.Ed.[3]

[1]Department of Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [2]Siriraj Health Science Education Excellence Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [3]Undergraduate Education Unit, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.



**Corresponding author: Cherdsak Iramaneerat*
E-mail: cherdsak.ira@mahidol.ac.th
Received 30 July 2025 Revised 17 August 2025 Accepted 18 August 2025
ORCID ID:http://orcid.org/0000-0002-2750-3981
https://doi.org/10.33192/smj.v77i11.276829

ABSTRACT

Objective: The objective of this study was to compare the administration of Multiple-Mini Interview (MMI) between the online and onsite platforms, in terms of validity evidence.

Materials and Methods: We retrospectively reviewed records of past MMI scores, examination materials, and participant questionnaire responses over a six-year period at the Faculty of Medicine Siriraj Hospital, Mahidol University. A total of eight MMI administrations were included, four onsite and four online. Validity evidence was assessed based on three key sources: test content, response processes, and internal structure.

Results: Over six years, eight MMIs were conducted, with 237 out of 340 candidates admitted to a medical school. Content analysis of the test specification tables indicated that both onsite and online platforms adequately addressed similar objectives. Participant satisfaction ratings were comparable between onsite and online MMIs. Qualitative analysis revealed minor issues in a few stations regarding clarity of instructions and scoring criteria. Additionally, some online MMI stations showed discrepancies between task time requirements and allotted time, and more technical issues were also reported. Score analysis showed that both highest and average scores from online MMIs were slightly lower than those from onsite MMIs. However, both formats yielded moderately reliable test scores (Cronbach's Alpha of 0.49 - 0.51).

Conclusion: The online MMI is a viable alternative to the traditional onsite MMI. Both platforms effectively covered the same assessment objectives, yielded comparable reliability and participant satisfaction.

Keywords: Medical student selection; multiple mini-interview; validity; reliability (Siriraj Med J 2025; 77: 768-776)

INTRODUCTION

Interviewing plays a central role in the selection process of medical students. While various methods are used during admissions, such as pre-admission grades, academic records, aptitude tests, reference letters, and interviews, the interview remains a key tool for evaluating non-cognitive attributes that medical schools seek as a foundation for producing high-quality doctors. Traditionally, personal interviews involved a small number of interviewers engaging candidates in unstructured interview questions to evaluate attitudes. However, this format has been criticized for its low reliability, lack of standardization, susceptibility to interviewer bias, and limited generalizability across different contexts.[1-3] To address these limitations, the Multiple-Mini Interview (MMI) was developed by McMaster University as an alternative to traditional interviews.[2]

In the MMI, candidates rotate through a series of stations, each designed to assess specific personal attributes. These assessments are based on candidates' performance in a task or response to questions. Trained interviewers observe and give scores using a standardized rating scale.[2,3] The MMI has been widely adopted by health professional schools. A systematic review revealed that undergraduate health professional schools typically use five to 12 stations (average = 9.2 stations), with each station lasting five to ten minutes (average = 7.3 minutes), and staffed by one to two interviewers per station.[3] Psychometric studies have shown that MMI scores demonstrate high internal consistency reliability (intra-station inter-item correlation = 0.96), and moderate to high inter-station reliability (ranging from 0.59 – 0.87).[3-5] In general, MMIs with a greater number of stations yielded higher reliability.[2,6] Importantly, MMI scores show low correlations with past academic performance and personal interview scores,[2,7] but are strongly correlated with judgment and decision-making abilities (r = 0.75).[3] MMI scores have also been shown to predict Objective Structured Clinical Examination (OSCE) scores, licensing examination performance, clinical decision-making performance, and clerkship performance measures.[8,9] Although the MMI process is often perceived as more stressful than traditional interviews, the majority of candidates report preferring the MMI to traditional interviews.[10,11]

With increasing evidence suggesting that problems among medical students during their study might relate to non-academic issues, such as burnout[12], our medical school looked for innovative approach in medical student selection besides academic readiness. With the intention to bring in students who had personal characteristics that fit with medical school environment, Faculty of Medicine Siriraj Hospital introduced MMI into medical student selection in 2018. We initially administered MMI in an onsite format. Due to the COVID-19 pandemic, the MMI transitioned to an online format during 2021–2022, before returning to onsite administration in 2023 as public health conditions improved.

Although internet-based MMIs (iMMIs) have been

described in literature since 2013, often involving multiple sessions with questions drawn from large item banks to reduce costs for international candidates[13], our online approach differed. At our institution, a single-session online MMI was implemented because multiple sessions could increase the risk of station content being shared between candidates participating in different sessions, potentially compromising the fairness of this highly-competitive, high-stakes process. To ensure comparability and fairness in candidate evaluation, all applicants completed the same set of tasks and interview questions within a single session. The objective of this study was to compare the administration of MMI between the onsite and online platforms, in terms of validity evidence.

MATERIALS AND METHODS

According to Messick's framework, validity refers to the degree to which evidence supports the intended interpretation and use of test scores.[14] A comprehensive validity study would demonstrate evidence from five sources including test content, response processes, internal structure, relations to other variables, and consequences. However, some sources of validity evidence are rarely reported in the literature, namely response processes and consequences.[15] This study focused on three sources of validity evidence: test content, response processes, and internal structure.[14] Test content refers to the extent to which MMI scenarios reflect the competencies expected of future medical professionals. Response processes examine whether candidates approach the tasks in ways that align with the intended constructs. Internal structure involves analyzing the relationships among scores across stations to assess consistency and ensure whether the MMI reliably measures the targeted attributes. We employed a mixed-methods approach, incorporating both quantitative and qualitative data to gather validity evidence. We did not include the analysis related to relation to other variables and consequences in this study due to the difficulty in making valid inferences that those two aspects in medical students that obtained later for years were solely caused by the MMI administration platforms, as many other factors in the curriculum could influence them more.

After obtaining ethical approval from the Institutional Review Board at our medical school (COA no. Si 160/2023), we retrospectively reviewed past MMI score records, examination documents, and participants' questionnaire responses over a period of six years. The data included eight MMI administrations, of which four were conducted onsite and four online. In Thailand, multiple rounds of student selection are held for admission into medical schools. The Faculty of Medicine Siriraj Hospital employs the MMI in selection rounds aimed at students with specific achievements or those who had already obtained a bachelor's degree. These rounds are highly competitive and allow the faculty to select candidates who best align with the program's criteria. MMIs were conducted once annually from 2018 to 2020 and twice annually from 2021 onward. However, in 2023, only the first MMI session was included in this study, as the second session occurred after data analysis had been completed. Over the six-year study period, 340 candidates participated in MMIs, and 237 were admitted to the MD program at the Faculty of Medicine Siriraj Hospital.

Quantitative data analysis

We retrospectively reviewed test scores from all eight MMIs, comparing onsite and online MMIs in terms of score distribution, average score, and internal consistency reliability. All statistical analyses were conducted in PASW Statistics 18.0.[16]

Qualitative data analysis

We conducted a document analysis of the test specification blueprints for all MMI sessions to compare content coverage. Additionally, we analyzed responses from the open-ended sections of post-test surveys completed by candidates, interviewers, and administrative staff. Comments and feedback were summarized and categorized into themes.

RESULTS

The characteristics of the eight MMI sessions included in this review are summarized in Table 1. Each session consisted of 8–12 stations, with 1–6 rest stations, and each active station lasted 8 minutes. The total duration of each MMI ranged from 1.5 to 2.5 hours. Across cohorts, around 51%–81% of candidates were selected for admission. However, these selection rates were influenced by a fixed admission quota for each round and therefore do not necessarily reflect differences in candidate competency across cohorts.

Test content

Content analysis of the test specification tables revealed that both onsite and online platforms covered similar objectives. The candidate attributes targeted across all stations included honesty, altruism, creativity, motivation for medical study, responsibility, self-directed learning, social media literacy, critical thinking skills, teamwork skills, communication skills, perseverance, stress management, growth mindset, situational awareness and decision-making. Table 2 showed the comparison of

TABLE 1. Characteristics of the MMIs included in the analysis.

Exam	Platform	Candidates	Passing candidates	Number of stations
2018	onsite	29	21 (72.41%)	8
2019	onsite	24	21 (87.5%)	8
2020	onsite	49	25 (51.02%)	8
2021 (round 1)	online	31	25 (80.65%)	8
2021 (round 2)	online	14	7 (50%)	8
2022 (round 1)	online	53	40 (75.47%)	12
2022 (round 2)	online	13	8 (61.54%)	12
2023	onsite	127	90 (70.86%)	10

TABLE 2. Comparison of test specification between two MMI platforms.

Exam	Platform	Honesty	Altruism	Communication	Innovation/ creativity	Motivation	Stress management	Perseverance	Responsibility	Critical thinking	Growth mindset	Self-directed learning	Teamwork	Decision making	Media literacy
2018	onsite	●	●	●	●	●	●	●	●						
2019	onsite	●	●	●	●	●	●	●	●						
2020	onsite	●			●	●	●		●	●	●	●			
2023	onsite	●			●	●	●		●	●	●	●		●	●
2021 (round 1)	online	●	●	●	●	●	●	●	●	●	●	●	●		
2021 (round 2)	online	●	●	●	●	●	●	●	●	●	●	●	●		
2022 (round 1)	online	●	●	●	●	●	●	●	●	●	●	●	●	●	●
2022 (round 2)	online	●	●	●	●	●	●	●	●	●	●	●	●	●	●

test content coverage between the two platforms. With more experience in conducting MMI, administrators tended to increase content coverage to more domains. Almost all objectives were covered by both platforms, except only teamwork which was not covered in an onsite platform, due to the decision to decrease the number of stations when changing from online MMI back to onsite MMI. Although both platforms evaluated similar competencies, the nature of tasks differed. Onsite MMIs allowed interviewers to observe candidates interacting with standardized patients and/or use onsite instruments and tools (e.g., video clips, part-task trainers, mannequins, puzzles, or cameras). In contrast, online MMIs had limited access to such resources. Therefore, online station design required greater creativity and utilized video clips, pen-and-paper tasks, and online tools.

Response processes

Validity evidence related to response processes can be categorized into two types: **direct methods**, which capture real-time thought processes, and **indirect methods**, which infer cognitive engagement through participant reflection.[17] Due to feasibility constraints, we employed an indirect approach using a post-test satisfaction surveys from candidates, interviewers, and administrative staff as indirect evidence for how effectively the MMI assessed intended competencies.

Quantitative analysis of the survey data demonstrated overall satisfaction with the assessment procedure. Two key aspects were evaluated: (1) satisfaction with test content and (2) satisfaction with test administration. Content satisfaction was measured using four items of the relevance and appropriateness of the MMI content. Satisfaction with test administration was assessed using six items in online MMIs and three items in onsite MMIs, focusing on organization, equipment, station arrangement, assistance, and internet connectivity. All items were rated on a five-point Likert scale, where five represented excellent and one indicated improvement required. A summary of survey rating data is summarized in Table 3.

Qualitative analysis of open-ended survey responses revealed three key themes: test instructions, time management, and technical issues.

Theme 1: Test Instruction

A few MMI stations in each administration received feedback suggesting improvements in the clarity of instructions or scoring criteria. The frequency of this issue was similar between the online and onsite formats. One interviewer noted that a candidate appeared to misunderstand the general instruction for an online MMI. On the other hand, a significant number of comments from candidates expressed appreciation for the MMI format, stating that it effectively showcased their potential to become competent medical students.

Theme 2: Time Management

Many candidates and interviewers commented on a mismatch between task demands, and the time allotted per station. Most felt the time was insufficient to complete the tasks, while a few reported having too much time. A unique challenge in the online MMI was the delivery of the one-minute warning. Unlike the onsite format, which used a shared audible signal, the online version relied on a small pop-up notification, some of which were reportedly missed. Nearly all time-related concerns were associated with the online MMIs, except for a single issue reported during the initial onsite implementation in 2018.

Theme 3: Technical Issues

Numerous comments highlighted how technical issues could hinder candidates' ability to demonstrate their true potential. Reported issues included unfamiliarity with software, unstable internet connection, difficulty handling digital station instruments, and suboptimal performance by standardized patients (SPs). Online MMIs, in particular, received more feedback about these

TABLE 3. Satisfaction ratings of the MMIs.

Onsite MMI			Online MMI		
Exam	Test content Mean (SD)	Test administration Mean (SD)	Exam	Test content Mean (SD)	Test administration Mean (SD)
2018	3.81 (0.34)	4.44 (0.13)	2021 (round 1)	3.72 (0.32)	4.19 (0.15)
2019	4.33 (0.09)	4.47 (0.13)	2021 (round 2)	4.28 (0.22)	4.80 (0.17)
2020	4.32 (0.23)	4.70 (0.14)	2022 (round 1)	4.23 (0.15)	4.48 (0.09)
2023	4.32 (0.17)	4.64 (0.07)	2022 (round 2)	4.41 (0.11)	4.61 (0.14)
Average	4.20 (0.21)	4.56 (0.12)	Average	4.16 (0.20)	4.52 (0.14)

technical issues. Both interviewers and candidates reported experiencing higher stress levels during online MMIs due to the added challenge of managing multiple tasks during online interview.

Internal structure

Three core components of internal structure are dimensionality, measurement invariance, and reliability.[18] This study focused on the measurement invariance and reliability. For measurement invariance, we examined score distributions and average score across various MMI administrations. We also calculated internal consistency reliability for each round of MMI scores.

Table 4 presents the lowest, highest, and average scores all eight MMIs. For onsite MMIs, the aggregated data showed a minimum score of 63.5%, a maximum of 90.9%, and an average of 80.1%. For online MMIs, the lowest score was 64.1%, the highest 86.3%, and the average score was 76.8%. The highest and average scores obtained from online MMIs seemed to be slightly lower than onsite MMIs.

The adjusted Cronbach's alpha values for all eight MMI administrations are presented in Table 5. The average Cronbach's alpha for the onsite MMIs was 0.49, while the average for online MMIs was 0.51. Line graph of adjusted Cronbach's alpha of all eight administrations revealed that both platforms yielded similar levels of reliability, ranging between 0.36 – 0.61 (Fig 1). These results indicate that both platforms showed a similar level of internal consistency reliability.

TABLE 4. Score distribution and average scores.

Onsite MMI				Online MMI			
Exam	Score (%)			Exam	Score (%)		
	Lowest	Highest	Average		Lowest	Highest	Average
2018	66.1	92.3	83.2	2021 (round 1)	67.7	88.9	79.6
2019	67.3	86.9	78.8	2021 (round 2)	69.8	84.2	75.7
2020	60.7	93.5	80.6	2022 (round 1)	60.9	89.0	78.7
2023	60.0	90.8	77.7	2022 (round 2)	57.8	83.2	73.0
Average	63.5	90.9	80.1	Average	64.1	86.3	76.8

TABLE 5. Internal consistency reliability of MMI scores.

Onsite MMI			Online MMI		
Exam	Number of stations	Cronbach's Alpha (adjusted)	Exam	Number of stations	Cronbach's Alpha (adjusted)
2018	8	0.458	2021 (round 1)	8	0.492
2019	8	0.363	2021 (round 2)	8	0.437
2020	8	0.604	2022 (round 1)	12	0.612
2023	10	0.547	2022 (round 2)	12	0.504
Average		0.49	Average		0.51

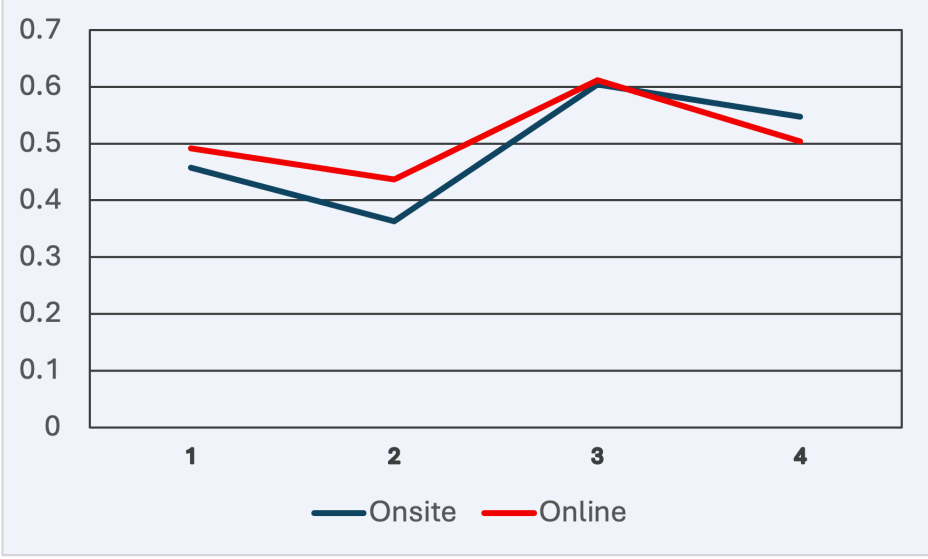


Fig 1. Internal consistency reliability of MMI scores, comparing between onsite and online platforms.

DISCUSSION

This study compared the traditional onsite format and the newer online format of the MMI to assess their validity and reliability in the context of medical student selection. We focused on three sources of validity evidence, including test content, response processes, and internal structure.[14]

Test content

Analysis of the test specification tables revealed that both MMI platforms were comparable in their coverage of objectives, targeting a range of personal characteristics important in medical education. Onsite MMI implementation was relatively straightforward as medical teachers were already familiar with the OSCE format, which closely resembles onsite MMI. In contrast, implementing an online MMI presented unique challenges. Each station required careful task design and scoring criteria to meet assessment objectives within a virtual environment. Observing candidates remotely demanded creativity and extensive reliance on digital technologies, such as teleconferencing platforms, internet connectivity, file-sharing tools, and video recording capabilities. Candidates were limited to basic tools such as pens, pencils, and paper, unlike onsite MMI, which could use mannequins, part-task trainers, and standardized patients. While the online MMI was able to meet test content coverage, it required more effort to do so. Additionally, the online format inadvertently assessed candidates' technological proficiency, with those more adept with digital tools appearing less stressed during the process, likely due to the heavy dependence on digital technology.

Response processes

We employed both quantitative and qualitative data to evaluate how effectively each platform allowed candidates to demonstrate their competencies, free from interference by administrative factors. Quantitative analysis of satisfaction ratings from participants indicated high satisfaction with both test content and administration. However, ratings tended to dip when a new interview format was first introduced. This trend was observed in the 2018 cohort (initial onsite MMI) and in the first round of the 2021 cohort (initial online MMI), likely due to anxiety associated with unfamiliar testing environments. Despite this, no significant impact on test scores was observed (Table 4). Satisfaction ratings were comparable between platforms, with onsite MMI scoring revealing a mean of 4.38 (SD = 0.16) and online MMI a mean of 4.34 (SD = 0.17). Despite comparable satisfaction for the overall testing processes, qualitative analysis revealed important insights into the response processes. Three major themes emerged from participants' qualitative responses: test instruction, time management, and technical issues.

Although MMIs have been discussed extensively in literature for many years[2], they remain a relatively new and unfamiliar to candidates and interviewers in Thailand, necessitating significant communication efforts for implementation. For each cohort, detailed instructions were provided through both written and verbal formats, along with opportunities for participants to seek clarification. Despite these measures, a few participants still reported confusion, particularly regarding task-specific instructions at individual stations. This underscores the need for thorough item review and the importance

of clarity in all task instructions and scoring criteria to support response process validity.

Feedback on time management revealed two important lessons. First, pilot testing is essential for detecting mismatches between task demands and time allocation, allowing adjustments before actual administration. Second, managing timing across geographically dispersed participants posed a unique challenge for the online format. Unlike onsite MMIs, where a single shared signal minimizes distraction, the online platform relied on individual notifications. This calls for administrators to develop carefully planned communication strategies to preserve validity in the response process.

Technical issues emerged as a third key theme and potential threat to response process validity. While onsite MMIs occasionally involved equipment and computers, technical problems were minimal and manageable within a controlled environment. In contrast, online MMIs were more vulnerable to external factors such as varying internet connectivity and access to suitable equipment in remote locations. Managing these issues required flexibility on the part of the administrators, including rescheduling. Consequently, both candidates and interviewers reported greater stress during online MMIs.

Internal structure

Evaluation of MMI test scores revealed slightly lower scores in the online format compared to the onsite format. Given that both formats were designed to address similar competencies and administered to candidates of comparable educational levels, we hypothesized that the lower scores in the online MMIs may be attributed to validity issues related to response processes, such as minor technical difficulties during online administration, which may have increased candidate stress and perceived task difficulty.

Analysis of internal consistency reliability supported findings from previous research regarding the relationship between the number of stations and test reliability. Our study showed average Cronbach's alpha values of 0.47, 0.55, and 0.56 for MMIs with 8, 10, and 12 stations, respectively. These findings are consistent with prior literature suggesting reliability improves with an increasing number of stations, typically plateauing around 10 stations.[12,19] However, the reliability observed in our context was lower than that reported in other studies, which reported Cronbach's alpha values ranging from 0.69-0.98.[3,20] We hypothesize that enhancing the design of the scoring rubric and improving rater training could improve reliability in our context. Notably, internal consistency reliability was comparable between onsite and online MMIs.

In summary, our findings suggest that online MMIs, with thoughtful adaptation of tasks at some stations, can effectively assess the same test objectives as traditional onsite MMIs. Both platforms demonstrated comparable levels of reliability and participant satisfaction. However, the online MMI required more technical resources and posed greater logistical challenges, contributing to higher participant stress and slightly lower average scores.

Our experience indicates that while the online MMI is a viable alternative when in-person testing is not feasible, under normal circumstances, onsite MMIs remain easier to manage and are less likely to cause stress among participants. Future efforts to enhance MMI quality should focus on improving scoring rubrics and rater training to strengthen internal consistency reliability. Based on these findings, we recommended transitioning back to onsite MMI as soon as pandemic restrictions allowed (beginning in 2023).

There are several limitations to the generalizability of the findings from this study. First, the study was conducted at a single medical school, evaluating a specific set of candidate characteristics. Therefore, the test specifications, station designs, and evaluation criteria may differ from those used at other institutions. Second, the online MMI was implemented rapidly in response to the COVID-19 pandemic, with limited preparation time and minimal rater training. These constraints may have influenced the outcomes. It is possible that with more extensive planning, better technological infrastructure, and enhanced training, the results of online MMIs under more prepared conditions and with improved technology could yield different results. Additionally, although the study benefited from a large sample size (340 candidates) collected over six years across eight MMI cohorts, this extended period may have introduced variability. Changes in student characteristics, preparedness, technological familiarity, and participant attitudes over time could have influenced the results. Furthermore, we acknowledged that we did not provide a complete list of validity evidence as suggested by Messick's framework. We only provided limited validity evidence in three aspects. There were two more sources of validity evidence, relations to other variables, and consequences, waiting to be explored in future studies.

CONCLUSION

Online MMI is a viable alternative to onsite administration when in-person interviews requiring candidates and interviewers to come together are difficult. With appropriate task modifications, the online format can fulfill the same test specifications and achieve comparable score reliability. Participants reported overall satisfaction

with both MMI formats, though some raised minor concerns in online MMIs related to technical issues during the online sessions, which appeared to slightly increase the perceived difficulty and resulted in marginally lower average scores. In summary, both online and onsite MMIs showed no significant difference in their validity evidence related to test content, response processes, and internal structure. However, the MMI administrators have to deal with some difficulty in setting up an online MMI.

Data Availability Statement

The datasets generated and analyzed in this study are not publicly available. However, they can be accessed upon reasonable request made to a corresponding author.

ACKNOWLEDGEMENT

We thank Siriraj Health Science Education Excellence Center for their support.

DECLARATIONS

Grants and Funding Information

This study received financial support from an educational research grant from the Faculty of Medicine Siriraj Hospital, Mahidol University.

Conflict of Interest

All the authors confirm that they have no personal or professional conflicts of interest to declare relating to any aspect of this research study.

Registration Number of Clinical Trial

Not applicable.

Author Contributions

Conceptualization and methodology, C.I.; Data collection O.U.; Data analysis, C.I.; Manuscript preparation, C.I.; Critical review and editing, C.I. and P.M. All authors reviewed the results and approved the final version of the manuscript.

Use of Artificial Intelligence

Artificial intelligence was not used in the preparation of the manuscript. All study concepts, analysis, interpretation, and writing were carried out by the authors.

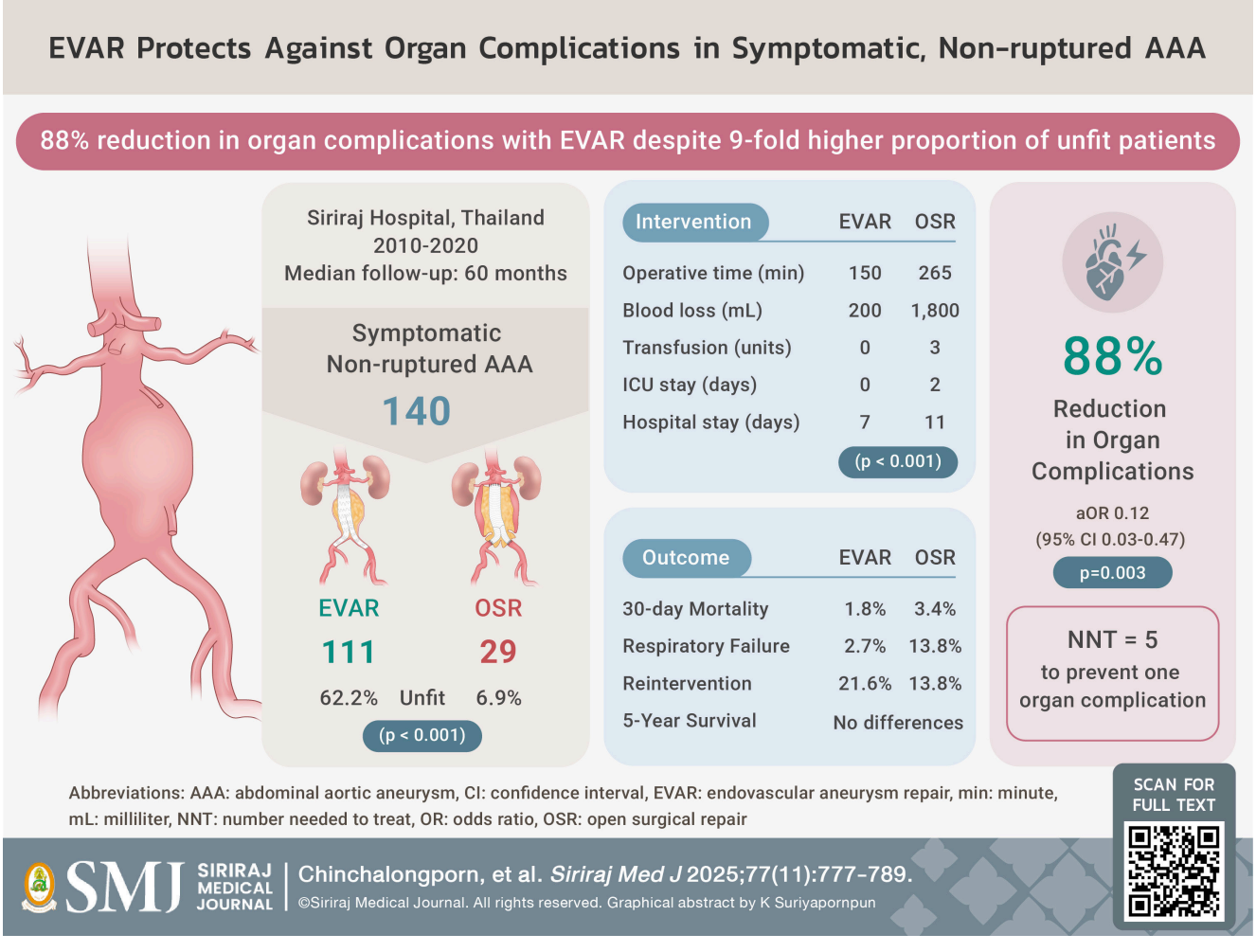
REFERENCES

1. Eva KW, Macala C, Fleming B. Twelve tips for constructing a multiple mini-interview. Med Teach. 2019;41(5):510-16.
2. Eva KW, Rosenfeld J, Reiter HI, Norman GR. An admissions OSCE: the multiple mini-interview. Med Educ. 2004;38(3): 314-26.
3. Rees EL, Hawarden AW, Dent G, Hays R, Bates Hassell AB. Evidence regarding the utility of multiple mini-interview (MMI) for selection to undergraduate health programs: A BEME systematic review: BEME Guide No. 37. Med Teach. 2016;38(5): 443-55.
4. Hecker K, Donnon T, Fuentealba C, Hall D, Illanes O, Morck DW, et al. Assessment of applicants to the veterinary curriculum using a multiple mini-interview method. J Vet Med Educ. 2009; 36(2):166-73.
5. Uijtdehaage S, Doyle L, Parker N. Enhancing the reliability of the multiple mini-interview for selecting prospective health care leaders. Acad Med. 2011;86(8):1032-9.
6. Sebok SS, Luu K, Klinger DA. Psychometric properties of the multiple mini-interview used for medical admissions: findings from generalizability and Rasch analyses. Adv Health Sci Educ Theory Pract. 2014;19(1):71-84.
7. Eva KW, Reiter HI, Rosenfeld J, Trinh K, Wood TJ, Norman GR. Association between a medical school admission process using the multiple mini-interview and national licensing examination scores. JAMA. 2012;308(21):2233-40.
8. Eva KW, Reiter HI, Rosenfeld J, Norman GR. The ability of the multiple mini-interview to predict preclerkship performance in medical school. Acad Med. 2004;79(10 Suppl):S40-2.
9. Reiter HI, Eva KW, Rosenfeld J, Norman GR. Multiple mini-interviews predict clerkship and licensing examination performance. Med Educ. 2007;41(4):378-84.
10. Dowell J, Lynch B, Till H, Kumwenda B, Husbands A. The multiple mini-interview in the U.K. context: 3 years of experience at Dundee. Med Teach. 2012;34(4):297-304.
11. McAndrew R, Ellis J. An evaluation of the multiple mini-interview as a selection tool for dental students. Br Dent J. 2012;212(7):331-5.
12. Thamwiriyakul N, Thamissarakul S, Wannapaschaiyong P. Association between grit and burnout among clinical medical students. Siriraj Med J. 2025; 77(2):175–82.
13. Tiller D, O'Mara D, Rothnie I, Dunn S, Lee L, Roberts C. Internet-based multiple mini-interviews for candidate selection for graduate entry programmes. Med Educ. 2013;47(8):801-10.
14. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Educational Research Association 2014.
15. Cizek GJ. Sources of validity evidence for educational and psychological tests: A follow-up study. Paper presented at the Annual meeting of the National Council on Measurement in Education, Denver, CO, May 2010.
16. PASW Statistics for Windows, version 18.0 [program]. Chicago: SPSS Inc., 2009.
17. Padilla JL, Benítez I. Validity evidence based on response processes. Psicothema. 2014;26(1):136-44.
18. Rios J, Wells C. Validity evidence based on internal structure. Psicothema. 2014;26(1):108-16.
19. Eva KW, Reiter HI, Trinh K, Wasi P, Rosenfeld J, Norman GR. Predictive validity of the multiple mini-interview for selecting medical trainees. Med Educ. 2009;43(8):767-75.
20. Pau A, Jeevaratnam K, Chen YS, Fall AA, Khoo C, Nadarajah VD. The Multiple Mini-Interview (MMI) for student selection in health professions training - a systematic review. Med Teach. 2013;35(12):1027-41.

Endovascular Versus Open Repair for Symptomatic, Non-Ruptured Abdominal Aortic Aneurysms: A Retrospective Cohort Study

Wanchai Chinchalongporn, M.D., Ph.D.[1], Pitchaya Wisantanon, M.D.[1], Khamin Chinsakchai, M.D.[1,*], Chanean Ruangsetakit, M.D.[1], Chumpol Wongwanit, M.D.[1], Kiattisak Hongku, M.D.[1], Sasima Tongsai, Ph.D.[2], Nuttawut Sermsathanasawadi, M.D., Ph.D.[1], Suteekhanit Hahtapornsawan, M.D.[1], Nuttawut Pungpunngam, M.D.[1], Tossapol Prapassaro, M.D.[1], Kanin Pruekprasert, M.D.[1]

[1]Division of Vascular Surgery, Department of Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand, [2]Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.



*Corresponding author: Khamin Chinsakchai
E-mail: khamin.chi@mahidol.edu
Received 18 August 2025 Revised 7 September 2025 Accepted 8 September 2025
ORCID ID:http://orcid.org/0000-0002-5420-3204
https://doi.org/10.33192/smj.v77i11.277132

ABSTRACT

Objective: To compare perioperative organ morbidity, 30-day mortality, and 5-year overall survival following endovascular aneurysm repair (EVAR) versus open surgical repair (OSR) for symptomatic, non-ruptured abdominal aortic aneurysm (SnAAA).

Materials and Methods: We retrospectively reviewed 140 consecutive patients treated for SnAAA between 2010 and 2020 (EVAR n=111, OSR n=29). The primary endpoint was 30-day all-cause mortality; secondary endpoints included perioperative complications and 5-year survival. Due to extreme baseline imbalances (62.2% unfit in EVAR vs 6.9% in OSR) and limited sample size, multivariable adjustment was used instead of propensity score methods.

Results: Thirty-day mortality was 1.8% in the EVAR group and 3.4% in the OSR group (p=0.504, Fisher's exact test). EVAR was associated with an 88% reduction in perioperative organ complications (adjusted OR 0.12, 95% CI 0.03-0.47, p=0.003), including a significant reduction in respiratory failure (2.7% vs 13.8%, OR 0.17, 95% CI 0.04-0.83, p=0.034). EVAR also resulted in shorter operative time (median 150 vs 265 minutes, p<0.001), reduced blood loss (200 vs 1,800 mL, p<0.001), and shorter hospital stay (7 vs 11 days, p<0.001). Five-year survival analysis revealed no significant difference between groups (log-rank p=0.193; adjusted HR 1.09, 95% CI 0.42-2.85, p=0.857).

Conclusions: Despite being performed in a higher-risk cohort, EVAR provided significant protection against organ complications without compromising long-term survival compared to OSR. These findings support EVAR as the preferred approach for unfit patients with SnAAA, while both strategies remain appropriate options for fit patients.

Keywords: Abdominal aortic aneurysm; symptomatic; non-ruptured; endovascular repair; open repair; perioperative outcomes; survival (Siriraj Med J 2025; 77: 777-789)

INTRODUCTION

Abdominal aortic aneurysm (AAA) remains a significant cause of morbidity and mortality worldwide, with its prevalence increasing in aging populations.[1] While asymptomatic AAA has been extensively studied, the optimal management of symptomatic, non-ruptured AAA (SnAAA) remains uncertain in vascular surgery.[2]

Endovascular aneurysm repair (EVAR) has revolutionized AAA treatment since its introduction in the early 1990s. Multiple randomized controlled trials (EVAR 1, DREAM, OVER) have demonstrated the short-term perioperative benefits of EVAR over open surgical repair (OSR), although long-term survival outcomes remain similar.[3-6] More recent evidence suggests improved durability with newer-generation devices[7], and the 2024 European Society for Vascular Surgery guidelines emphasize anatomical suitability and patient fitness as key factors in treatment selection.[8]

While the outcomes of EVAR and OSR have been well-documented for asymptomatic and ruptured AAAs, data specific to SnAAAs remains limited.[9,10] This subgroup presents a unique clinical challenge: the urgency of intervention limits extensive optimization, yet the absence of rupture allows for potential consideration of either approach. Existing studies have shown mixed results, with some suggesting short-term benefits of EVAR and others reporting no significant difference in outcomes.[9]

Management of SnAAA is particularly crucial, as these patients face an estimated 3-month rupture risk of up to 50% if left untreated.[11] Furthermore, the physiological stress of symptomatic presentation may influence perioperative outcomes differently compared with elective cases. While concerns remain regarding EVAR durability and higher reintervention rates in asymptomatic AAAs[12], the applicability of these findings to urgent symptomatic cases is unclear. In these patients, the reduced physiologic impact of minimally invasive approaches may offer meaningful advantages.[13]

Given these knowledge gaps, there is a clear need for comprehensive studies comparing outcomes of EVAR and OSR specifically in SnAAA patients. Understanding the comparative effectiveness of these approaches in this high-risk population is essential for optimizing clinical decision-making and resource allocation. Therefore, we conducted a retrospective cohort analysis comparing short- and long-term outcomes of EVAR versus OSR in patients with SnAAA. We hypothesized that EVAR would demonstrate non-inferior perioperative mortality and superior recovery outcomes compared to OSR, despite being used in a potentially higher-risk patient population. Our primary objective was to compare 30-day all-cause mortality, with secondary objectives including perioperative organ complications, long-term survival, and reintervention rates.

MATERIALS AND METHODS

Study design and patient selection

This retrospective cohort study was conducted at Siriraj Hospital, Thailand, between January 2010 and December 2020. Medical records of patients diagnosed with SnAAA who underwent either EVAR or OSR were reviewed. The study protocol was approved by the Siriraj Institutional Review Board on November 22, 2021 (SIRB; COA no. Si 921/2021).

Symptomatic, non-ruptured AAA was defined as an intact AAA with:

- Abdominal and/or back pain attributable to the aneurysm, AND
- Tenderness over the aneurysm on physical examination, AND
- No evidence of rupture on computed tomography (CT) imaging (absence of retroperitoneal hematoma, contrast extravasation, or other signs of rupture)

Inclusion criteria:

1. Patients diagnosed with SnAAA as defined above
2. Patients who underwent either EVAR or OSR

Exclusion criteria:

1. Infective native AAA
2. Isolated internal or external iliac artery aneurysm
3. Ruptured AAA (evidence of blood outside the aortic wall)
4. Asymptomatic AAA undergoing elective repair

Of the 230 patients with SnAAA screened during the study period, 89 were excluded (87 with infected AAA and 2 with isolated iliac artery aneurysms). Of the 141 eligible patients, one OSR patient was lost to follow-up, leaving 140 patients in the final analysis (111 EVAR, 29 OSR) (Fig 1).

Data collection

We collected data on patient demographics, comorbidities, AAA morphology, operative details, and postoperative outcomes. Surgical fitness was assessed according to the criteria described by *Brown et al.*[14] Patients were classified as unfit for surgery based on the EVAR trial criteria, which included: severe cardiac disease (recent myocardial infarction or angina onset within the past 3 months, unstable angina, severe valvular disease, significant arrhythmia, or uncontrolled heart failure); poor respiratory status (inability to climb one flight of stairs without dyspnea, FEV_1 <1.0 L, PaO_2 <8.0 kPa, or $PaCO_2$ >6.5 kPa); or renal impairment (serum creatinine >200 µmol/L (2.26 mg/dL)).

All patients underwent multidetector computed tomography angiography (CTA) with 1-mm slice thickness. Measurements included neck length and diameter, presence of neck calcification or thrombus, neck angulation, maximum AAA diameter, and total aneurysm length.

Surgical procedures

The choice between EVAR and OSR was based on anatomical suitability for EVAR, overall patient fitness, surgeon preference, and device availability. Anatomical suitability was assessed using standard criteria, including a proximal neck length ≥15 mm, neck angulation <60° and adequate iliac access vessels.

All EVAR procedures were performed in operating rooms equipped with fluoroscopic guidance, using commercially available stent graft systems. OSR was performed via a transperitoneal approach, with infrarenal aortic clamping when possible. Both EVAR and OSR procedures were performed in accordance with established standard practices.[8]

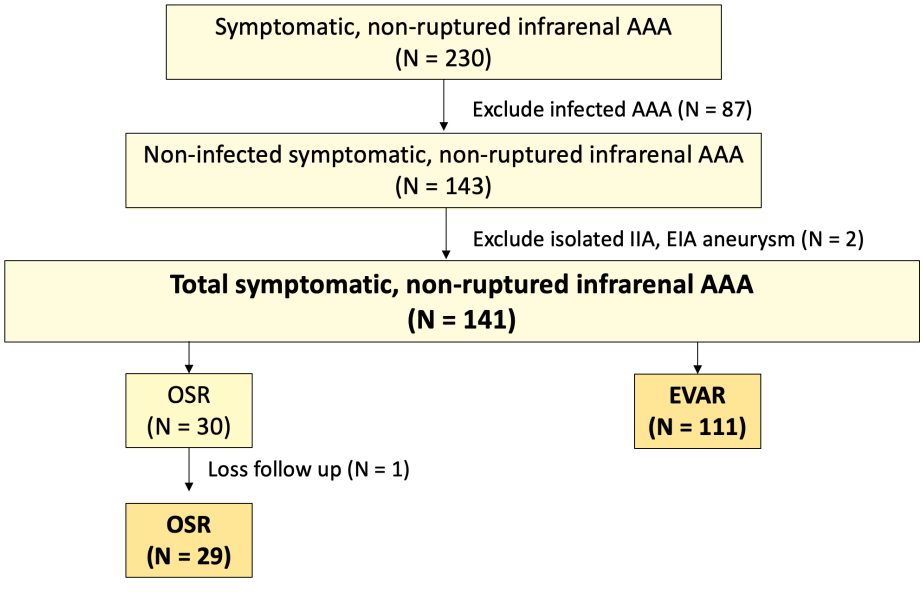


Fig 1. STROBE flow diagram showing screening, exclusions, and final cohorts.
Abbreviations: AAA: Abdominal aortic aneurysm, EIA: External iliac artery, EVAR: Endovascular aneurysm repair, IIA: Internal iliac artery, OSR: Open surgical repair

Outcome measures

The primary endpoint was perioperative mortality, defined as death from any cause within 30 days of the procedure or during the same hospital admission.

Secondary endpoints included:

1. Perioperative complications (within 30 days):
   - Organ complications: cardiac, respiratory, renal, or gastrointestinal dysfunction
   - Infection complications: surgical site, graft, or systemic infections
   - Technical complications: procedure-specific issues requiring reintervention
2. Operative details: operative time, estimated blood loss, and blood transfusion requirements
3. Length of stay: intensive care unit (ICU) and total hospital duration
4. Long-term survival: all-cause mortality during follow-up
5. Reintervention rates: any aneurysm-related reintervention

Perioperative complications were defined according to standardized criteria[15] (Supplementary Table S1).

Follow-up protocol

Patients were followed at 1 month, 6 months, and annually thereafter. The EVAR group underwent CTA at each interval, while the OSR group underwent CTA at 2 and 5 years postoperatively. In patients with advanced chronic kidney disease (stage III or IV), duplex ultrasound was substituted for CTA.[16] Survival data were obtained from hospital records and the Thai national death registry.

Statistical analysis

Sample size and power: A post-hoc power analysis revealed 80% power to detect a 20% absolute difference in organ complications, but only 15% power to detect differences in perioperative mortality due to the low event rate.

Statistical methods: Continuous variables were expressed as mean ± standard deviation or median (interquartile range), depending on normality (assessed by the Shapiro-Wilk test). Categorical variables were presented as numbers and percentages. Comparisons between the EVAR and OSR groups were made using Student's t-test or Mann-Whitney U test for continuous variables, and Pearson's chi-square or Fisher's exact test for categorical variables, as appropriate.

Approach to confounding: Given the extreme selection bias evident in baseline characteristics (62.2% unfit in EVAR vs 6.9% in OSR) and a limited OSR sample size (n=29), propensity score analysis was considered but ultimately deemed inappropriate due to poor overlap between groups. Instead, multivariable models directly adjusted for age and fitness status as the primary confounders based on clinical relevance and baseline imbalances were used.

Multivariable analysis: Due to the low event rate for perioperative mortality (three deaths in total), multivariable analysis was not performed, in accordance with the conventional guideline requiring at least 10 events per predictor variable. Therefore, only unadjusted comparisons using Fisher's exact test were conducted for this outcome.

For secondary outcomes with adequate events (≥10), multivariable logistic regression was conducted, adjusting for age and fitness status. Model fit was assessed using the Hosmer-Lemeshow test, with a p-value >0.05 indicating adequate fit. Covariates were selected based on clinical relevance and baseline imbalance magnitude. Given 21 organ complication events, we limited primary models to two variables following established guidelines. Sensitivity analyses with additional anatomical variables were considered but not performed to avoid violating statistical constraints.

Survival analysis: Kaplan-Meier curves were generated for overall survival and compared using the log-rank test. Cox proportional hazards models were used to calculate adjusted hazard ratios, controlling for age and fitness status. The proportional hazards assumption was tested via Schoenfeld residuals and graphical assessment.

Missing data: Missing data was minimal (<5%) for key variables. A complete case analysis was used for multivariable models. Patients with missing outcome data were excluded from relevant analyses.

A two-tailed p-value <0.05 was considered statistically significant. All statistical analyses were performed using SPSS software (version 21, IBM Corp., Armonk, NY, USA).

RESULTS

Patient selection and baseline characteristics

During the study period, 140 patients with SnAAA were identified and included in the analysis: 111 (79.3%) underwent EVAR and 29 (20.7%) underwent OSR. Baseline demographic and clinical characteristics are presented in Table 1.

The mean age was similar between groups (76.6 ± 8.1 years for EVAR vs 74.2 ± 8.8 years for OSR, p=0.155). Males comprised 73.9% of the EVAR group and 65.5% of the OSR group (p=0.371). No significant differences were found in the prevalence of comorbidities, including

TABLE 1. Baseline characteristics and AAA morphology in patients undergoing OSR and EVAR for symptomatic AAA.

Characteristic	OSR (n=29)	EVAR (n=111)	P-value
Demographics			
Age (years), mean ± SD	74.2 ± 8.8	76.6 ± 8.1	0.155
Male sex, n (%)	19 (65.5)	82 (73.9)	0.371
Comorbidities, n (%)			
Coronary arterial disease	7 (24.1)	26 (23.4)	0.936
COPD	3 (10.3)	14 (12.6)	1.000
Hypertension	21 (72.4)	81 (73.0)	0.952
Dyslipidemia	8 (27.6)	31 (27.9)	0.971
Chronic kidney disease	2 (6.9)	15 (13.5)	0.525
Diabetes mellitus	6 (20.7)	10 (9.0)	0.100
Cerebrovascular disease	0 (0.0)	9 (8.1)	0.204
Current smoking	4 (13.8)	17 (15.3)	1.000
Risk Assessment			
ASA class III-IV, n (%)	14 (48.3)	74 (66.7)	0.292
Unfit for surgery, n (%)	2 (6.9)	69 (62.2)	<0.001
AAA Morphology			
Neck length (mm), mean ± SD	22.8 ± 14.6	26.5 ± 15.3	0.434
Neck diameter (mm), mean ± SD	21.0 ± 5.3	22.0 ± 3.6	0.343
Neck calcification ≥40%, n (%)	0 (0.0)	3 (2.7)	1.000
Neck thrombus ≥40%, n (%)	3 (10.3)	9 (8.1)	0.118
Neck angle (degrees), median (IQR)	70 (20, 119)	40 (10, 76)	0.119
AAA diameter (mm), mean ± SD	64.4 ± 20.5	60.5 ± 15.6	0.298
AAA length (mm), mean ± SD	126.9 ± 30.3	119.2 ± 26.0	0.958
Anesthesia, n (%)			<0.001
General anesthesia	29 (100)	60 (54.1)	
Regional anesthesia	0 (0)	43 (38.7)	
Local anesthesia	0 (0)	8 (7.2)	

Abbreviations: AAA: abdominal aortic aneurysm; ASA: American Society of Anesthesiologists; CKD: chronic kidney disease; COPD: chronic obstructive pulmonary disease; Cr: creatinine; EVAR: endovascular aneurysm repair; IQR: interquartile range; OSR: open surgical repair; SD: standard deviation

coronary artery disease, chronic obstructive pulmonary disease, hypertension, dyslipidemia, chronic kidney disease, diabetes, or cerebrovascular disease. Current smoking rates were also similar (15.3% in the EVAR group vs 13.8% in the OSR group (p=1.000)).

While not reaching statistical significance, several anatomical variables showed numerical imbalances between groups. Neck thrombus ≥40% was present in 10.3% of OSR patients versus 8.1% of EVAR patients (p=0.118), and median neck angle differed substantially (70° in OSR vs 40° in EVAR, p=0.119). These differences may indicate anatomical selection bias.

Notably, the percentage of patients deemed unfit for surgery was significantly higher in the EVAR group compared to the OSR group (62.2% vs 6.9%, p<0.001), reflecting real-world selection bias that favors EVAR in higher-risk patients. Despite this imbalance, American Society of Anesthesiologists (ASA) classifications did not differ significantly between groups, with 66.7% of EVAR and 48.3% of OSR patients classified as ASA III or IV (p=0.292).

AAA morphology

Anatomical characteristics were similar between groups (Table 1). No significant differences were observed in neck length (26.5 ± 15.3 mm for EVAR vs 22.8 ± 14.6 mm for OSR, p=0.434), neck diameter, neck calcification or thrombus, neck angulation, AAA diameter (60.5 ± 15.6 mm for EVAR vs 64.4 ± 20.5 mm for OSR, p=0.298), or total aneurysm length.

Admission-to-procedure timing

The time from admission to procedure did not differ significantly between groups (Mann-Whitney U test, p=0.574). The EVAR group had a mean preoperative duration of 1.70 ± 2.53 days (median 1.00 day, range 0-16 days) compared to 1.31 ± 1.14 days for the OSR group (median 1.00 day, range 0-4 days). Extended intervals (>7 days) occurred in 8 EVAR patients (7.2%) due to staged bilateral internal iliac artery embolization, following institutional protocol requiring 1-2 week intervals between procedures. No OSR patients experienced delays beyond 4 days.

Operative details and early outcomes

Operative characteristics demonstrated significant advantages for EVAR (Table 2). The median operative time was significantly shorter for EVAR (150 minutes [IQR 120-187] vs 265 minutes [IQR 225-365], p<0.001). EVAR also resulted in significantly less blood loss (median

TABLE 2. Perioperative outcomes in OSR versus EVAR groups.

Outcome	OSR (n=29)	EVAR (n=111)	P-value
Primary outcome			
30-day mortality, n (%)	1 (3.4)	2 (1.8)	0.504*
Operative details			
Operative time (min), median (IQR)	265 (225, 365)	150 (120, 187)	<0.001
Blood loss (mL), median (IQR)	1800 (850, 2250)	200 (100, 400)	<0.001
Blood transfusion (units), median (IQR)	3 (2, 4)	0 (0, 1)	<0.001
Postoperative recovery			
ICU stay (days), median (IQR)	2 (1, 6)	0 (0, 1)	<0.001
Hospital stay (days), median (IQR)	11 (8, 21)	7 (4, 11)	<0.001
Reinterventions			
Any reintervention, n (%)	4 (13.8)	24 (21.6)	0.348
Early (≤30 days), n (%)	2 (6.9)	10 (9.0)	1.000
Late (>30 days), n (%)	2 (6.9)	14 (12.6)	0.524

Abbreviations: EVAR: endovascular aneurysm repair; ICU: intensive care unit; IQR: interquartile range; OSR: open surgical repair
* *Fisher's exact test used for perioperative mortality due to low event rate; all other P-values from Mann-Whitney U or Chi-square tests as appropriate.

200 mL [IQR 100-400] vs 1,800 mL [IQR 850-2250], p<0.001) and lower transfusion requirements (0 units [IQR 0-1] vs 3 units [IQR 2-4], p<0.001).

Postoperative recovery was faster with EVAR, reflected in shorter ICU stays (median 0 days [IQR 0-1] vs 2 days [IQR 1-6], p<0.001) and total hospital stays (median 7 days [IQR 4-11] vs 11 days [IQR 8-21], p<0.001).

Anesthesia type differed significantly between treatment groups. All OSR patients received general anesthesia due to surgical necessities. In contrast, EVAR procedures permitted anesthetic flexibility: 54.1% (60/111) general anesthesia, 38.7% (43/111) regional anesthesia (spinal or epidural), and 7.2% (8/111) local anesthesia with conscious sedation. The ability to use regional or local anesthesia in 45.9% of EVAR cases (51/111) contributed to reduced perioperative physiologic stress, particularly in high-risk patients.

Perioperative mortality

The perioperative mortality rate was 1.8% (2/111) in the EVAR group compared to 3.4% (1/29) in the OSR group (p=0.504, Fisher's exact test). Due to the limited number of events (n=3), multivariable adjustment for mortality was not feasible. The single death in the OSR group resulted from sepsis, while in the EVAR group, one death was due to multiple organ failure and the other from stroke.

Perioperative complications

Overall perioperative complications occurred in 29.7% (33/111) of EVAR patients and 34.5% (10/29) of OSR patients (p=0.621). The distribution of complications is detailed in Table 3.

Adjusted analyses for perioperative complications are shown in Table 5. After adjustment for age and fitness

TABLE 3. Perioperative complications in OSR versus EVAR groups.

Complication type	OSR (n=29)	EVAR (n=111)	P-value
Any complication, n (%)	10 (34.5)	33 (29.7)	0.621
Organ complications, n (%) *	10 (34.5)	11 (9.9)	0.002
Myocardial infarction	2 (6.9)	3 (2.7)	0.276
Congestive heart failure	3 (10.3)	4 (3.6)	0.156
Respiratory failure	4 (13.8)	3 (2.7)	0.034
Renal failure	3 (10.3)	4 (3.6)	0.156
Ischemic colitis	1 (3.4)	2 (1.8)	0.504
Infection complications, n (%)	6 (20.7)	20 (18.0)	0.742
Chest infection	3 (10.3)	8 (7.2)	0.697
Fever of unknown origin	0 (0.0)	5 (4.5)	0.583
Wound infection	0 (0.0)	3 (2.7)	1.000
Intraabdominal infection	1 (3.4)	0 (0.0)	0.207
Graft infection	1 (3.4)	1 (0.9)	0.373
Septicemia	3 (10.3)	8 (7.2)	0.697
Wound dehiscence	0 (0.0)	1 (0.9)	1.000
Urinary tract infection	2 (6.9)	5 (4.5)	0.634
Technical complications, n (%)	2 (6.9)	9 (8.1)	1.000
Type IA endoleak	-	4 (3.6)	-
Graft limb occlusion	1 (3.4)	2 (1.8)	0.504
Groin hematoma	0 (0)	1 (0.9)	1.000
Groin seroma	0 (0)	1 (0.9)	1.000
Bilateral IIA occlusion	0 (0)	1 (0.9)	1.000
IIA anastomosis stenosis	1 (3.4)	-	-

* Some patients experienced multiple complications. P-values for individual technical complications were not calculated as many are procedure-specific.
Abbreviations: EVAR: endovascular aneurysm repair; OSR: open surgical repair

status, EVAR was associated with significantly lower odds of organ complications compared to OSR (adjusted OR 0.12, 95% CI 0.03-0.47, p=0.003), representing an 88% reduction in odds. This protective effect was primarily driven by a significant reduction in respiratory failure (OR 0.17, 95% CI 0.04-0.83, p=0.034), representing an 83% reduction in odds. The number needed to treat (NNT) with EVAR to prevent one organ complication was 5, and to prevent one episode of respiratory failure was 10.

No significant differences were observed in adjusted analyses for overall perioperative complications (aOR 0.80, 95% CI 0.29-2.15, p=0.652), infectious complications (aOR 0.76, 95% CI 0.23-2.45, p=0.640), or technical complications (aOR 1.47, 95% CI 0.26-8.25, p=0.664).

Exploratory stratified analysis by fitness status was performed (Supplementary Table S2), though interpretation was limited by extreme imbalance in the unfit subgroup. Among fit patients (n=69), organ complications occurred in 1 of 42 (2.4%) EVAR patients versus 10 of 27 (37.0%) OSR patients (p<0.001, Fisher's exact test). Among unfit patients (n=71), meaningful comparison was precluded by having only 2 OSR patients (both without complications) versus 69 EVAR patients with a 14.5% complication rate.

Reinterventions

Early reinterventions (≤30 days) were required in 9.0% (10/111) of EVAR patients and 6.9% (2/29) of OSR patients (p=1.000). EVAR-specific early complications included Type IA endoleaks requiring additional procedures (n=4) and graft limb occlusion (n=2). OSR-related complications requiring early reintervention included internal iliac artery anastomosis stenosis and external iliac artery dissection (Table 4).

Late reinterventions (>30 days) were performed in 12.6% (14/111) of EVAR patients and 6.9% (2/29) of OSR patients (p=0.524). Details of late complications and reinterventions are detailed in Table 4.

Long-term survival

The median follow-up was 60 months (range 1-120 months). Kaplan-Meier analysis revealed no significant difference in overall survival between treatment groups (log-rank p=0.193, Fig 2). The mean survival time was 8.35 years (95% CI 6.82-9.88) for OSR and 7.66 years (95% CI 6.63-8.70) for EVAR. Median survival was not reached in either group during the follow-up period.

Cox proportional hazards regression (Table 6) confirmed no significant difference in mortality risk between EVAR and OSR after adjusting for age and fitness status (adjusted HR 1.09, 95% CI 0.42-2.85, p=0.857). Age was identified as a significant independent predictor of mortality (HR 1.06 per year, 95% CI 1.01-1.12, p=0.020), while fitness status showed a borderline association with survival (HR 2.24 for unfit patients, 95% CI 0.97-5.16, p=0.058).

The 5-year overall survival rate was 59.7% in the OSR group and 41.5% in the EVAR group. The 5-year reintervention-free survival rate (Fig 3) was 83.8% in the OSR group and 73.5% in the EVAR group (log-rank p=0.355), indicating no significant difference in freedom from reintervention.

DISCUSSION

In this retrospective cohort of 140 patients with SnAAA, EVAR was associated with an 88% reduction in perioperative organ complications (adjusted OR 0.12, 95% CI 0.03-0.47, p=0.003) compared to OSR, despite being performed in a significantly higher proportion of unfit patients (62.2% vs 6.9%). This protective effect against organ dysfunction, primarily driven by reduced respiratory failure, remained significant after adjustment for baseline differences. Importantly, these perioperative advantages did not come at the expense of long-term survival (adjusted HR 1.09, 95% CI 0.42-2.85, p=0.857) or increased perioperative mortality (1.8% vs 3.4%, p=0.504).

The extreme difference in fitness status between groups represents substantial selection bias that challenges and strengthens our analysis. The 62.2% of unfit patients in the EVAR group compared with the 6.9% in the OSR group reflects real-world practice where EVAR is often selected for patients deemed high risk for open surgery. This selection bias limited methods such as propensity score matching due to poor overlap, necessitating direct multivariable adjustment.

Paradoxically, this selection bias strengthens our conclusions about EVAR's protective effect against organ complications. Despite being performed in patients with significantly higher baseline risk, EVAR was still associated with superior organ protection. This finding suggests that the true benefit of EVAR in a more balanced or comparable population may exceed the effect estimated in our adjusted analysis.

Our findings align with previous studies examining symptomatic, non-ruptured AAAs. *De Martino et al.*[9] reported similar perioperative advantages of EVAR in symptomatic AAAs, although their analysis did not adjust for fitness status. *Soden et al.*[17] found comparable benefits using data from the American College of Surgeons National Surgical Quality Improvement Program, with EVAR associated with reduced cardiac and pulmonary complications.

TABLE 4. Complications requiring reintervention in OSR and EVAR groups.

Time Period	Group	Complication	Reintervention	Occurrence (n)	Time to Reintervention
Early	**OSR**	EIA dissection	Thrombo-embolectomy and EIA stenting	1	Day 1
(≤30 days)		IIA anastomosis stenosis	IIA stenting	1	Day 1
	EVAR	Type IA endoleak	Aortic cuff extension	3	Days 3, 5, 12
		Type IA endoleak	Palmaz stent	1	Day 19
		Ischemic colitis	Hartmann's procedure	2	Day 2
		Graft limb occlusion	Thrombo-embolectomy	2	Days 0, 12
		Groin complication	Explore wound	2	Days 0, 14
Late	**OSR**	Type II endoleak	Embolization	1	2.5 years
(>30 days)		Adhesive SBO	EL with lysis adhesion	1	7 years
	EVAR	Type IA endoleak	Palmaz stent	1	2 months
		Type IB endoleak	Iliac stent	1	2.5 years
		Type IA, IB endoleak	Palmaz stent and iliac extension	1	5 months
		Type II endoleak	Embolization	4	2.5, 3, 4, 6 years
		Type III endoleak	Relining both iliac limb (Kissing stent)	1	5 months
		Type III endoleak	Iliac extension	1	23 months
		Infected graft	Graft excision	1	2 years
		Suprarenal aortic aneurysm	3 fenestrated grafts	1	2.5 years
		Iliac limb occlusion	Femoro-femoral bypass graft	2	4 months, 1 year
		Iliac limb migration	Iliac extension	1	5 years

Abbreviations: EIA: external iliac artery; EL: exploratory laparotomy; EVAR: endovascular aneurysm repair; IIA: internal iliac artery; OSR: open surgical repair; SBO: small bowel obstruction

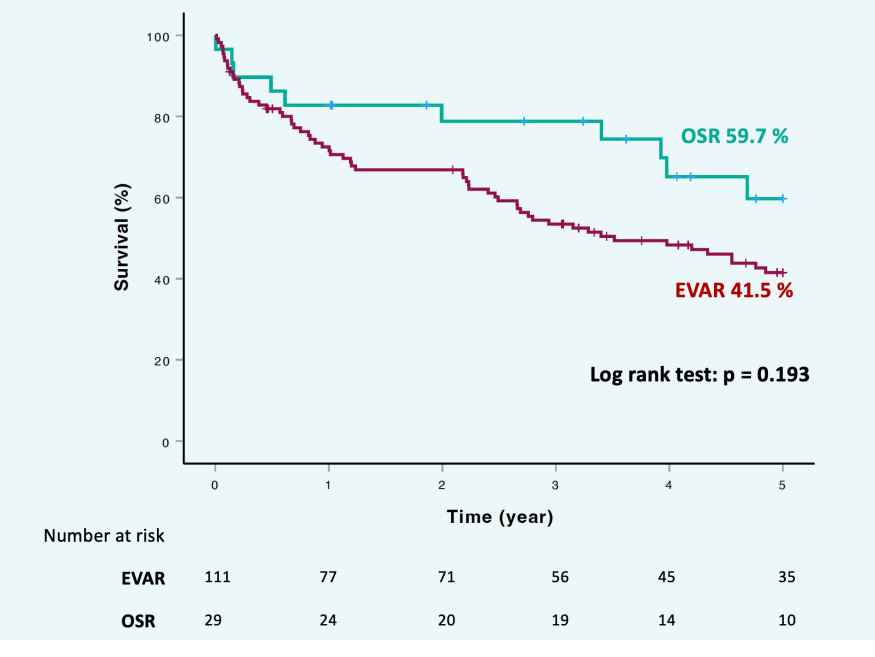


Fig 2. Kaplan–Meier curves comparing overall survival in patients undergoing Endovascular Aneurysm Repair (EVAR) and Open Surgical Repair (OSR).

TABLE 5. Multivariable Logistic Regression Analysis of Perioperative Outcomes*.

Outcome	Unadjusted OR	Adjusted OR*	95% CI	P-value
Any complication	0.80	0.80	0.29-2.15	0.652
Organ complications	0.21	0.12	0.03-0.47	0.003
Infection complications	0.84	0.76	0.23-2.45	0.640
Technical complications	1.19	1.47	0.26-8.25	0.664

*Adjusted for age and fitness status

TABLE 6. Summary of survival analysis results.

Analysis	Estimate	95% CI	P-value
Kaplan-Meier Analysis			
Mean survival, years			
OSR	8.35	6.82-9.88	-
EVAR	7.66	6.63-8.70	-
Log-rank test	-	-	0.193
Cox Regression Analysis*			
EVAR vs OSR (HR)	1.09	0.42-2.85	0.857
Age per year (HR)	1.06	1.01-1.12	0.020
Unfit vs Fit (HR)	2.24	0.97-5.16	0.058

*Adjusted model including treatment, age, and fitness status
Abbreviations: CI: confidence interval; HR: hazard ratio

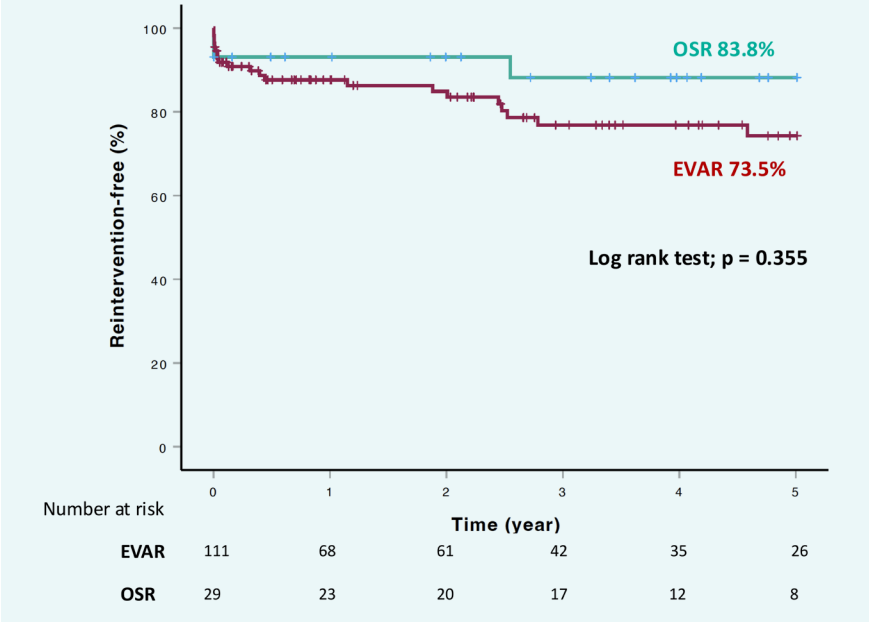


Fig 3. Kaplan–Meier curves comparing reintervention-free survival in patients undergoing Endovascular Aneurysm Repair (EVAR) and Open Surgical Repair (OSR).

The 88% reduction in organ complications observed in our study exceeds rates reported in studies of asymptomatic AAAs. This greater benefit likely reflects the heightened vulnerability of symptomatic AAA patients to physiologic stress. The urgency of repair in these cases often precludes extensive preoperative optimization, making the minimally invasive nature of EVAR particularly advantageous. Our finding that respiratory failure was the primary driver of organ complications (13.8% OSR vs 2.7% EVAR) supports this hypothesis, as avoiding laparotomy and general anesthesia substantially reduces pulmonary morbidity. The number needed to treat analysis revealed clinically meaningful benefits, with an NNT of 5 for preventing organ complications and 10 for preventing respiratory failure. These values suggest substantial clinical benefit, with only 5 patients requiring EVAR instead of OSR to prevent one organ complication. However, these observational estimates should be interpreted cautiously given the selection bias inherent in our retrospective design.

In terms of long-term survival, our findings are consistent with those of the EVAR-1 trial and subsequent meta-analyses[6], which showed no significant difference between EVAR and OSR after adjustment for baseline characteristics. However, our study uniquely demonstrates this equivalence in the context of SnAAA and under conditions of extreme selection bias.

These findings have important implications for clinical decision-making in SnAAA. For patients deemed unfit for major surgery, EVAR should be considered the preferred approach when anatomically feasible, given the 88% reduction in perioperative organ complications. The operational advantages of EVAR demonstrated in our study (Table 2) have important healthcare system implications. The substantial reductions in operative time, blood product utilization, and ICU requirements are particularly valuable in resource-constrained settings where OR availability, blood supplies, and ICU capacity are limited. For symptomatic AAA requiring urgent intervention, these operational efficiencies may facilitate more timely treatment and better resource allocation. This is particularly relevant for patients with limited physiologic reserve who may not tolerate postoperative organ dysfunction. For fit surgical candidates, both EVAR and open surgical repair remain reasonable options, and the choice should be individualized based on anatomical suitability, surgeon expertise, and patient preference. The absence of a survival difference allows these additional factors to guide decision-making. In urgent clinical settings, EVAR offers practical advantages, including shorter operative time (150 vs 265 minutes), reduced blood loss (200 vs 1,800 mL), and lower ICU utilization, all of which facilitate rapid patient stabilization.

Our study has several strengths. First, it focuses specifically on SnAAA, a population with limited representation in existing evidence. Second, we applied robust multivariable adjustment to address confounding, despite challenges posed by extreme selection bias. Third, our long-term follow-up (median 60 months) allowed assessment of both early and late outcomes. Lastly, we provided transparent reporting of analytical limitations, enhancing the interpretability of our findings.

Nonetheless, key limitations must be acknowledged. Treatment allocation in our study followed institutional protocols based on anatomical suitability, fitness status, and device availability. However, the retrospective design precluded detailed analysis of specific decision-making rationales for individual cases. This represents an important limitation as unmeasured selection factors may have influenced outcomes. The predominance of EVAR (79.3%) likely reflects both contemporary practice trends and appropriate patient selection for minimally invasive approaches in high-risk patients. The substantial selection bias, reflected by the extreme imbalance in fitness status (62.2% vs 6.9% unfit), mirrors real-world practice but limits causal inference despite statistical adjustment. The degree of baseline imbalance prevented the use of propensity score methods, and residual confounding likely persists despite multivariable adjustment. The overall perioperative mortality of 2.1%, while favorable for patients, limited our ability to perform adjusted mortality comparisons. With only three deaths, we could not perform multivariable analysis without violating established statistical principles regarding events per variable — a limitation increasingly common in contemporary vascular surgery studies as outcomes improve. The small open surgical repair group sample size (n=29), though indicative of contemporary practice trends favoring EVAR, reduced statistical power for some comparisons, particularly subgroup analyses. Residual confounding from anatomical factors (neck angle, neck thrombus) likely persists despite adjustment. While sensitivity analyses including these variables were considered, they could not be appropriately performed given our event rate. This represents an important limitation, though the magnitude of EVAR's protective effect (88% reduction) suggests findings are clinically meaningful despite potential residual confounding. As a single-center retrospective study, our findings may not generalize to other settings or populations, and minor technical complications not requiring reintervention may have been underreported. Additionally, we were unable to assess quality of life outcomes or patient-reported

measures, nor fully disentangle anatomical selection bias as some patients may have been directed to open repair due to unsuitable anatomy for EVAR rather than fitness considerations alone.

Future research should focus on several areas. Prospective registries with standardized data collection could better characterize patient selection criteria and outcomes. Although randomized trials in the urgent SnAAA setting may be challenging, they would provide more definitive evidence. Additionally, development of risk prediction models specific to SnAAA could guide patient selection. Finally, incorporation of patient-reported outcome measures and quality of life would provide a more comprehensive view of treatment effectiveness.

CONCLUSIONS

In patients with symptomatic non-ruptured AAAs, EVAR provided significant protection against perioperative organ complications, with an 88% reduction in odds compared to OSR. This benefit was observed despite EVAR being predominantly used in patients deemed unfit for open surgery. Long-term survival was equivalent between approaches after adjustment for baseline differences. These findings support EVAR as the preferred option for unfit patients with suitable anatomy, while both strategies remain reasonable choices for fit patients. The marked reduction in organ morbidity without survival compromise positions EVAR as an effective first-line therapy for symptomatic AAAs in appropriately selected patients.

Data Availability Statement

The data that support the findings of this study are not publicly available due to privacy and ethical restrictions related to the Personal Data Protection Act (PDPA) of Thailand. Data may be available from the corresponding author (Khamin Chinsakchai) upon reasonable request and with permission from the Siriraj Institutional Review Board.

ACKNOWLEDGEMENTS

The authors wish to express their sincere gratitude to the entire team at the Siriraj Aortic Center. We especially thank the dedicated nurses and nursing assistants for their invaluable assistance and compassionate patient care.

DECLARATIONS

Grants and Funding Information

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

Conflict of Interest

The authors declare no conflict of interest.

Registration Number of Clinical Trial

Not applicable. This study was a retrospective review of existing medical records and was not registered as a clinical trial.

Author Contributions

Conceptualization and methodology: K.C., P.W., W.C.; Investigation: K.C., P.W., W.C.; Formal analysis: S.T.; Visualization and writing – original draft: W.C.; Writing – review and editing: K.C., W.C., C.R., C.W., K.H., N.S., S.H., N.P., T.P., K.P. All authors have read and agreed to the published version of the manuscript.

Use of Artificial Intelligence

The authors used a generative AI language model to improve the grammar, clarity, and style of the manuscript text. AI was not used for data collection, statistical analysis, interpretation of results, or the generation of figures and tables. All final content, including all scientific claims and conclusions, was reviewed and edited by the authors, who are fully responsible for the integrity of the work.

REFERENCES

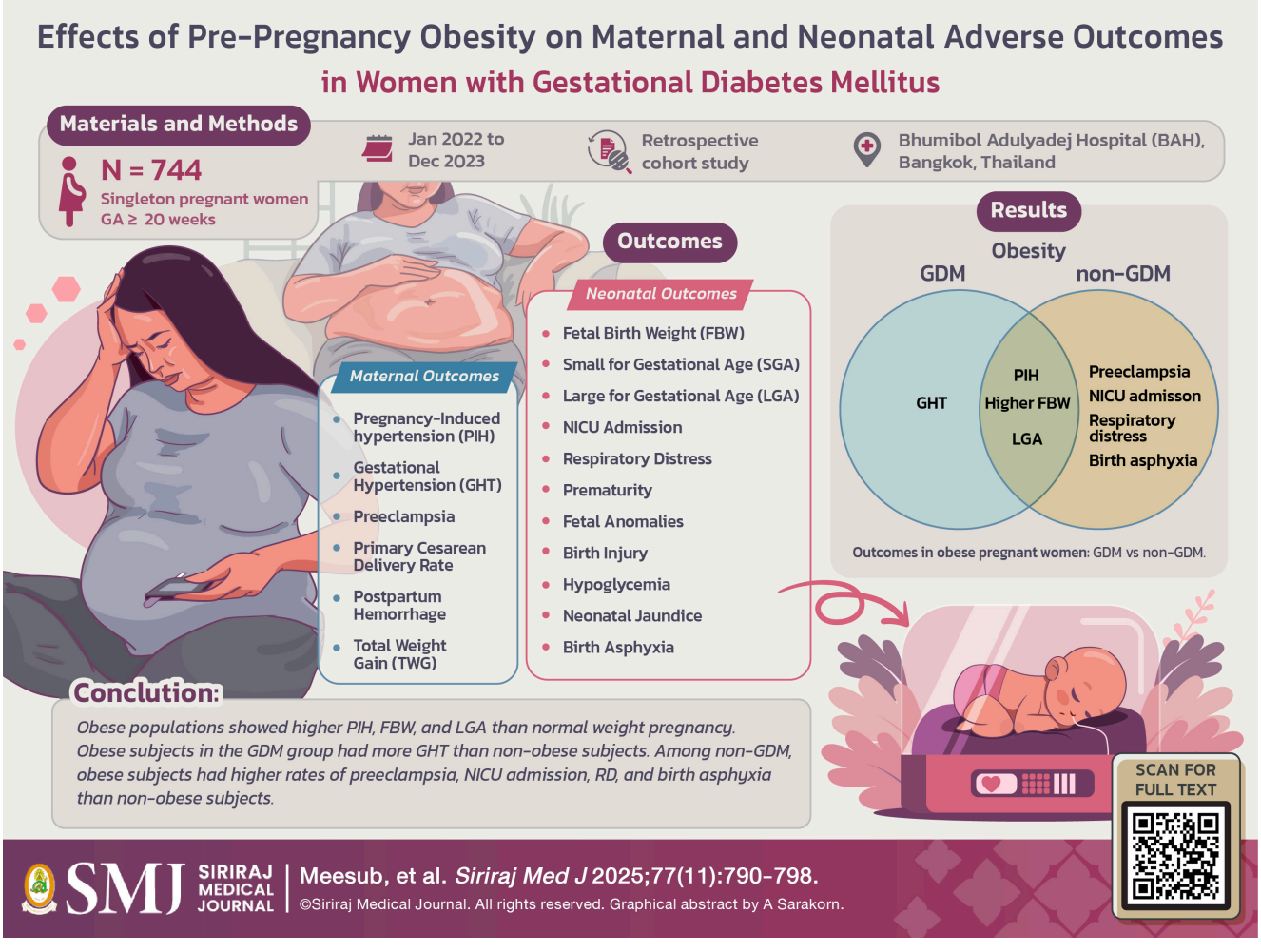
1. Sampson UK, Norman PE, Fowkes FG, Aboyans V, Song Y, Harrell FE, et al. Global and regional burden of aortic dissection and aneurysms: mortality trends in 21 world regions, 1990 to 2010. Glob Heart. 2014;9(1):171-80.e10.
2. Kontopodis N, Galanakis N, Antoniou SA, Tsetis D, Ioannou CV, Veith FJ, et al. Meta-Analysis and Meta-Regression Analysis of Outcomes of Endovascular and Open Repair for Ruptured Abdominal Aortic Aneurysm. Eur J Vasc Endovasc Surg. 2020; 59(3):399-410.
3. Prinssen M, Verhoeven EL, Buth J, Cuypers PWM, van Sambeek MRHM, Balm R, et al. A randomized trial comparing conventional and endovascular repair of abdominal aortic aneurysms. N Engl J Med. 2004;351(16):1607-18.
4. EVAR trial participants. Endovascular aneurysm repair versus open repair in patients with abdominal aortic aneurysm (EVAR trial 1): randomised controlled trial. Lancet. 2005;365(9478): 2179-86.
5. Lederle FA, Freischlag JA, Kyriakides TC, Padberg FT, Matsumura JS, Kohler TR, et al. Outcomes following endovascular vs open repair of abdominal aortic aneurysm: a randomized trial. JAMA. 2009;302(14):1535-42.
6. Powell JT, Sweeting MJ, Ulug P, Blankensteijn JD, Lederle FA, Becquemin JP, et al. Meta-analysis of individual-patient data from EVAR-1, DREAM, OVER and ACE trials comparing outcomes of endovascular or open repair for abdominal aortic aneurysm over 5 years. Br J Surg. 2017;104(3):166-78.
7. Antoniou GA, Antoniou SA, Torella F. Editor's Choice - Endovascular vs. Open Repair for Abdominal Aortic Aneurysm: Systematic Review and Meta-analysis of Updated Peri-operative and Long Term Data of Randomised Controlled Trials. Eur J

Vasc Endovasc Surg. 2020;59(3):385-97.
8. Wanhainen A, Van Herzeele I, Goncalves FB, Montoya SB, Berard X, Boyle JR, et al. Editor's Choice -- European Society for Vascular Surgery (ESVS) 2024 Clinical Practice Guidelines on the Management of Abdominal Aorto-Iliac Artery Aneurysms. Eur J Vasc Endovasc Surg. 2024;67(2):192-331.
9. De Martino RR, Nolan BW, Goodney PP, Chang CK, Schanzer A, Cambria R, et al. Outcomes of symptomatic abdominal aortic aneurysm repair. J Vasc Surg. 2010;52(1):5-12.e1.
10. Prapassaro T, Chinsakchai K, Techarattanaprasert S, Wongwanit C, Ruansetakit C, Hongku K, et al. Determining Perioperative Mortality in Patients with Ruptured Abdominal Aortic Aneurysm: Insights from a Retrospective Cohort Study. Siriraj Med J. 2024;76(8):480-7.
11. Darling RC, Messina CR, Brewster DC, Ottinger LW. Autopsy study of unoperated abdominal aortic aneurysms. The case for early resection. Circulation. 1977;56(3 Suppl):II161-4.
12. Patel R, Sweeting MJ, Powell JT, Greenhalgh RM. Endovascular versus open repair of abdominal aortic aneurysm in 15-years' follow-up of the UK endovascular aneurysm repair trial 1 (EVAR trial 1): a randomised controlled trial. Lancet. 2016;388(10058):2366-74.
13. Wongwanit C, Mutirangura P, Chierakul N, Chaiyasoot W, Phongraweewan O. Rapidly Enlarging and Asymptomatic Abdominal Aortic Aneurysm in a Male Patient with Chronic Obstructive Pulmonary Disease: A Case Report of Endovascular Aortic Aneurysm Repair (EVAR). Siriraj Med J. 2006;58(5):812-8.
14. Brown LC, Epstein D, Manca A, Beard JD, Powell JT, Greenhalgh RM. The UK Endovascular Aneurysm Repair (EVAR) trials: design, methodology and progress. Eur J Vasc Endovasc Surg. 2004;27(4):372-81.
15. Chinsakchai K, Thorthititum D, Hongku K, Wongwanit C, Tongsai S, Sermsathanasawadi N, et al. Endovascular Versus Open Repair for Asymptomatic Abdominal Aortic Aneurysms: A 12-Year Retrospective Cohort Analysis. Ann Vasc Surg. 2025;112:363-72.
16. Uthoff H, Peña C, Katzen BT, Gandhi R, West J, Benenati JF, et al. Current clinical practice in postoperative endovascular aneurysm repair imaging surveillance. J Vasc Interv Radiol. 2012;23(9):1152-9.e6.
17. Soden PA, Zettervall SL, Ultee KH, Karling JD, Buck DB, Hile CN, et al. Outcomes for symptomatic abdominal aortic aneurysms in the American College of Surgeons National Surgical Quality Improvement Program. J Vasc Surg. 2016;64(2):297-305.

Effects of Pre-Pregnancy Obesity on Maternal and Neonatal Adverse Outcomes in Women with Gestational Diabetes Mellitus

Patcharaporn Meesub, M.D.[1], Piyawan Pariyawateekul, M.D.[1,*], Wiyada Luangdansakul, M.D.[1], Monyada Pleankong, M.D.[1], Buppa Smanchat, M.D.[1], Konkarn Bhamarapravatana, Ph.D.[2], Komsun Suwannarurk, M.D.[3]

[1]Department of Obstetrics and Gynecology, Bhumibol Adulyadej Hospital, Royal Thai Air Force, Bangkok 10220, Thailand, [2]Department of Preclinical Sciences, Faculty of Medicine, Thammasat University, Pathumthani 12120, Thailand, [3]Department of Obstetrics and Gynecology, Faculty of Medicine, Thammasat University, Pathumthani 12120, Thailand.



*Corresponding author: Piyawan Pariyawateekul
E-mail: nootpiya@hotmail.com
Received 26 September 2024 Revised 30 July 2025 Accepted 30 July 2025
ORCID ID:http://orcid.org/0000-0001-7705-6537
https://doi.org/10.33192/smj.v77i11.271329

ABSTRACT

Objective: To explore the effects of pre-pregnancy obesity and GDM on pregnancy outcomes.

Materials and Methods: This retrospective cohort study was conducted between January 2022 and December 2023 at Bhumibol Adulyadej Hospital (BAH), Bangkok, Thailand. Participants were singleton pregnant women aged between 16 and 45 years old and delivered at BAH. Exclusion criteria were prepregnant medical diseases. All subjects underwent a 75-g oral glucose tolerance test for diagnosis of GDM. There were GDM and the non-GDM groups. Data was reviewed from electronic medical records, including demographic and clinical characteristics; namely: total weight gain (TWG), pregnancy induced hypertension (PIH), fetal birth weight (FBW), LGA neonates, neonatal intensive care unit (NICU) admission, mild respiratory distress (RD), and birth asphyxia.

Results: A total of 744 participants were included and randomly divided equally into GDM and non-GDM groups, with 372 cases each. PIH, FBW and LGA of obese subjects in both GDM and non-GDM groups were significantly higher than those in non-obese subjects. Subjects with obesity in both groups had significantly less TWG than those in non-obese subjects. GDM group participants with obesity had a significantly higher prevalence of Gestational hypertension (GHT) than non-obese subjects. Among the non-GDM group, obese subjects had significantly higher prevalence of preeclampsia, NICU admission, mild RD, and birth asphyxia than non-obese subjects.

Conclusion: Obese populations showed higher PIH, FBW, and LGA than normal weight pregnancy. Obese subjects in the GDM group had more GHT than non-obese subjects. Among non-GDM, obese subjects had higher rates of preeclampsia, NICU admission, RD, and birth asphyxia than non-obese subjects.

Keywords: Obesity; pregnancy outcomes; gestational diabetes mellitus (Siriraj Med J 2025; 77: 790-798)

INTRODUCTION

Gestational diabetes mellitus (GDM) is a temporally abnormal glucose metabolism during pregnancy.[1] Insulin resistance and inadequate insulin response occur among GDM subjects.[2] Moreover, fetal hyperglycemia develops following maternal hyperglycemia. Sustained high blood sugar in the fetus induces hyperinsulinemia and stimulates excessive somatic fetal growth, causing macrosomia and large-for-gestational age (LGA) neonates. Fetal macrosomia is defined as fetal birth weight exceeding 4,000 g or 4,500.[3] LGA neonates are defined as fetal birth weight of more than the 90th percentile. Fetal macrosomia and LGA lead to various consequences, such as an increased rate of cesarean delivery and shoulder dystocia at birth.[4]

GDM and obesity among pregnant women are the most common metabolic complications in pregnancy. Obese pregnant women have increased insulin resistance and uncontrolled weight gain.[5,6] Increased risk of macrosomia, higher rate of cesarean delivery due to labor dystocia[7], maternal perioperative complications and neonatal sepsis were reported among obese pregnant women with normoglycemia.[8-10] Obesity is defined according to the definition by the World Health Organization Asian Pacific region as a body mass index (BMI) threshold of more than or equal to 25 kg/m^2,[11] while in the U.S., BMI equal or more than 30 kg/m^2 is used.[12]

Previous studies reported that maternal obesity at the first antenatal care (ANC) was associated with a higher neonatal birth weight, macrosomia's risk, and cesarean delivery rate than those with GDM.[13-16]

The aim of the study was to investigate the effects of pre-pregnancy obesity and GDM on neonatal birth weight, macrosomia, cesarean delivery rate, low Apgar score, neonatal intensive care unit (NICU) admission, and other adverse pregnancy outcomes among pregnant Thai women.

MATERIALS AND METHODS

Study design and setting

The retrospective cohort study was conducted at the Department of Obstetrics and Gynecology, Bhumibol Adulyadej Hospital (BAH), Bangkok, Thailand. It was approved by the Institutional Review Board of BAH (IRB 77/66), Royal Thai Air Force.

Study participants and data collection

The subjects included pregnant women who visited the antenatal clinic and delivered their children at BAH. Data were reviewed from medical records kept in BAH computerized databases between January 2022 and December 2023. The inclusion criteria included healthy women carrying singleton pregnancies with gestational age (GA) ≥ 20 weeks. Exclusion criteria included women who had pre-existing underlying diseases, namely Diabetes mellitus (DM), Chronic hypertension (CHT), and antiphospholipid syndrome (APS).

Demographic characteristics, height, and weight were gathered during the first antenatal visit. Maternal BMI was calculated as weight (kg) divided by height squared (m^2). According to the Asia-Pacific criteria, BMI over 25.0 kg/m^2 was diagnosed as maternal obesity.[11] All participants underwent GDM screening between 24 and 28 weeks of gestation.

A 75-g oral glucose tolerance test (OGTT) was the first step for screening GDM. From IADPSG criteria, GDM was diagnosed when any blood glucose level was above the cut-off values of fasting plasma glucose (92 mg/dL), 1-hr glucose (180 mg/dL), or 2-h glucose (153 mg/dL).[17]

All pregnant women were followed up by obstetricians until delivery. The non-GDM pregnant women were monthly appointed in the first and second trimester. Appointments every two weeks were applied during the third trimester of pregnancy. Weekly antenatal visits were appointed for women with GDM. In addition, consultation to endocrinologists, nutritionists, and dieticians was applied for pregnant women with GDM.

To ensure adequacy of glycemic control, Self-monitoring of blood sugar was regularly performed among pregnant women with GDM. When the glycemic target was not achieved, insulin was prescribed. Targets for glycemic control were recommended for fasting 1-hr and 2-hr postprandial glucose values were less than 95 mg/dL, 140 mg/dL and 120 mg/dL, respectively.[18] Women in the non-GDM group were followed up the same in either the obesity group or non-obesity group.

Subjects were divided into the study (GDM) and the control (non-GDM) groups. Comparisons were made between obese and non-obese subjects in both groups regarding clinical and adverse pregnancy outcomes. Maternal outcomes consisted of pregnancy-induced hypertension (PIH), gestational hypertension (GHT), preeclampsia, primary cesarean delivery rate, and postpartum hemorrhage.

The fetal outcomes included fetal birth weight (FBW), large-for-gestational age (LGA) neonate (weight over the 90th percentile for gestational age), and small-for-gestational age (SGA) neonates (weight below 10th percentile for gestational age), fetal macrosomia, preterm birth (below 37 weeks of gestation), NICU admission, respiratory complication, fetal anomaly, hypoglycemia (serum glucose concentration < 40 mg/dL in symptomatic, or < 45 mg/dL in asymptomatic term neonates), neonatal jaundice and birth asphyxia (an Apgar score of <7).

Statistical analysis

All statistical analyses were performed using Stata Statistical Software: Release Version 16. All statistical analysis involved the chi-square, independent T-test. A p-value of < 0.05 was classified as statistically significant.

RESULTS

Medical records of 4,694 pregnant women who visited the antenatal clinic and delivered their children at BAH between January 2022 and December 2023 were retrieved. Inclusion criteria were singleton pregnant women, 20 weeks of gestation or more and delivery at BAH. Exclusion criteria were maternal medical conditions that affected pregnancy outcomes, namely chronic hypertension, diabetes mellitus (DM) and antiphospholipid syndrome (APS).

As shown in Fig 1, 744 subjects were recruited for analysis. All subjects were divided into GDM and non-GDM groups, each with 372 cases, respectively. There were 186 obese and 186 non-obese women in the GDM group. There were 152 obese and 220 non-obese women in the non-GDM group.

The mean age of all subjects was 30.6 years old. Obese subjects in both groups had higher BMI and pre-pregnancy weights (PPW) than non-obese counterparts with statistical significance as shown in Table 1. Only the TWG of obese subjects were lower than those non-obese subjects with statistical significance. The average GA of obese and non-obese subjects among the non-GDM group was 38.9 and 38.6 weeks (p=0.026), respectively. Average GA among obese and non-obese members in the GDM group were comparable (38.1/38.1 weeks).

From Table 2, mean fetal birth weight (FBW), LGA and pregnancy induced hypertension (PIH) prevalent in obese pregnant women in both GDM and non-GDM groups were significantly higher than in normal-weight pregnant women. Among GDM pregnancies, obese pregnant women had a higher rate of GHT development than non-obese subjects with statistical significance. However, among the non-GDM group, obese pregnant women had higher rates of preeclampsia, respiratory distress syndrome (RDS), NICU admission, and birth asphyxia than those of normal-weight pregnant women as presented in Table 2.

DISCUSSION

Obesity was diagnosed using BMI values. According to WHO guidelines, BMI of equal to or more than 30 kg/m^2 was classified as obesity. Data from the Asia-Pacific study, cut points of BMI for obesity of people in the region was lowered to 25 kg/m^2.[19] From the current study, TWG of obese pregnant women was significantly less than those of non-obese pregnant subjects. Ke and

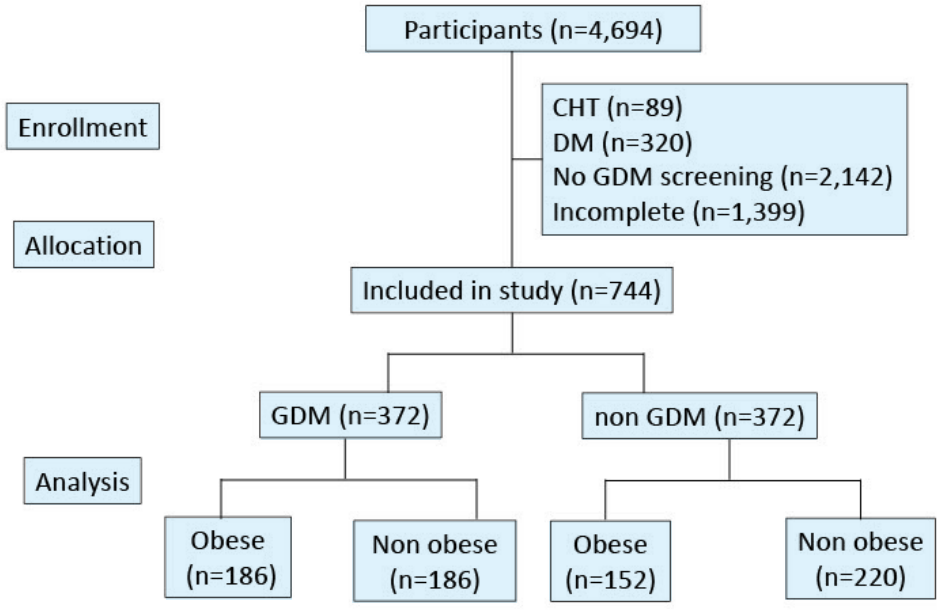


Fig 1. Flow chart of study.
Abbreviations: GDM: gestational diabetes mellitus, CHT: chronic hypertension

TABLE 1. Maternal baseline characteristics and maternal and fetal outcomes by gestational diabetes mellitus (GDM) and obesity status (n=744).

	GDM (n=372)*		*p*-value	Non-GDM (n=372)*		*p*-value
	Obesity (186)	Non (186)		Obesity (152)	Non (220)	
Age (yr)**	31.4±5.1	31.7±5.9	0.533	30.5±5.5	29.1±5.8	0.017
PPW (kg)**	75.6±11.5	52.6±7.5	<0.001	76.0±11.6	53.6±7.4	<0.001
Height (cm)**	158.2±6.2	157.5±6.3	0.243	159.9±5.94	159.7±5.8	0.750
PPBMI (kg/m^2)**	30.1±3.79	21.2±2.47	<0.001	29.7±3.94	21.0±2.3	<0.001
TWG (kg)	10.3±5.8	12.8±5.7	<0.001	13.6±7.2	15.3±5.8	0.015
Smoking	0 (0)	3 (1.6)	0.082	1 (0.7)	1 (0.5)	0.792
GA (weeks)**	38.1±1.7	38.1±1.2	0.859	38.9±1.2	38.6±1.4	0.026
Underlying disease						
Thyroid disease	3 (1.6)	3 (1.6)	1.000	2 (1.3)	3 (1.4)	0.969
Thalassemia	41 (22.0)	38 (20.4)	0.704	31 (20.4)	36 (16.4)	0.320
Heart	1 (0.5)	0 (0)	0.317	1 (0.7)	0 (0)	0.228
Multiparity	135 (72.6)	129 (69.4)	0.493	122 (80.3)	139 (63.2)	<0.001
History of macrosomia	4 (2.2)	4 (2.2)	1.000	3 (1.9)	2 (0.9)	0.381

*n (%), **Mean ± standard deviation (SD)
Abbreviations: GDM: gestational diabetes mellitus, PPW: pre-pregnancy weight, BMI: body mass index, PPBMI: pre-pregnancy BMI, TWG: total weight gain, GA: gestational age at birth.

TABLE 2. Analysis of maternal and fetal outcomes of GDM (n=372) and non-GDM (n=372).

	GDM*		*p*-value	Non-GDM *		*p*-value
	Obesity (186)	Non (186)		Obesity (152)	Non (220)	
PIH	26 (14.0)	7 (3.8)	0.001	14 (9.2)	6 (2.7)	0.006
GHT	13 (7.0)	1 (0.5)	0.001	6 (4)	3 (1.4)	0.111
Preeclampsia	13 (7.0)	6 (3.2)	0.099	8 (5.3)	3 (1.4)	0.029
C/S	64 (34.4)	57 (30.7)	0.439	58 (38.2)	64 (29.1)	0.067
PPH	6 (3.2)	5 (2.7)	0.760	3 (2.0)	6 (2.7)	0.642
FBW (Kg)**	3.3±0.5	3.1±5.1	0.002	3.6±0.4	3.3±0.5	<0.001
SGA	3 (1.6)	9 (4.8)	0.078	1 (0.7)	3 (1.4)	0.516
LGA	75 (40.3)	54 (29.0)	0.022	121 (79.6)	101 (46.0)	<0.001
NICU	23 (12.4)	20 (10.8)	0.627	21 (13.8)	10 (4.6)	0.
RD	25 (13.4)	21 (11.3)	0.529	16 (10.5)	13 (5.9)	0.103
Mild	21 (11.3)	20 (10.8)	0.869	16 (10.5)	11 (5.0)	0.043
Moderate	4 (2.2)	1 (0.5)	0.177	0	2 (0.9)	0.239
Preterm	13 (7.0)	24 (12.9)	0.057	2 (1.3)	11 (5.0)	0.057
Fetal anomaly	21 (11.3)	23 (12.4)	0.748	12 (7.9)	16 (7.3)	0.823
Birth injury	1 (0.5)	3 (1.6)	0.315	4 (2.6)	5 (2.3)	0.825
Hypoglycemia	8 (4.3)	5 (2.7)	0.397	2 (1.3)	1 (0.5)	0.361
Neonatal Jaundice	16 (8.6)	19 (10.2)	0.594	12 (7.9)	11 (5.0)	0.254
Birth asphyxia	4 (2.2)	1 (0.5)	0.177	3 (2.0)	0	0.036

*n (%), **Mean ± standard deviation (SD)

Abbreviations: GDM: gestational diabetes mellitus, PIH: pregnancy induced hypertension, GHT: gestational hypertension, C/S: primary cesarean delivery, PPH: postpartum hemorrhage, FBW: fetal birth weight, SGA: small for gestational age, LGA: large for gestational age, NICU: neonatal intensive care unit admission, RD: respiratory distress.

Lyu reported from China in year 2023 and 2024 that obese mothers had lower TWG than their non-obese counterparts.[14,15] TWG during pregnancy came from maternal and fetal components. Maternal factors, namely insulin level and glucose level, played important roles during the first half of pregnancy. Fetal factors, namely fetal growth and fluid expansion, caused the weight gain in the latter half of pregnancy.[20] From Lyu's study, obesity and GDM were independent factors for TWG. The findings from the current study supported Ke's and Lyu's findings.

Previous literature reported that pre-pregnancy overweight/obesity, GDM and excessive weight gain were independent factors of LGA risk.[14,15] Theoretically, obesity results from food overconsumption. Pregnant women tended to increase their food consumption due to high levels of progesterone during pregnancy.[21] However, pregnant mothers with obesity were subjects of concern to their attending physicians. Diet control and GDM screening were major interventions.

From the current study, obesity was a risk factor for PIH development among pregnant women. Insulin resistance and low-grade inflammation of endothelium and increased sodium resorption were the explanation of PIH among obese subjects.[5] Chinese studies by Sun and Ke reported that obesity among GDM subjects exhibited higher risk of PIH than those with normal weight.[14,22] A report from the US by Kim in 2016 said participants with obesity had four-fold increased risk of GHT than those with normal-weight pregnancies.[23]

From the present study, obesity brought an increased risk of fetal birth weight consequences to heavy weight babies (LGA or macrosomia) in both GDM and non-GDM groups. This finding supported results from previous studies that obesity was associated with higher neonatal birth weight and macrosomia.[13-16,22,23] Risk of higher neonatal birth weight from maternal obesity ranged between 1.7 and 2.8 fold.[13-16,22,23] Excessive nutritional environment, hyperlipidemia and high blood glucose levels were common findings among obese women. Fetuses in obese pregnant women were also exposed to high levels of glucose leading to higher fetal insulin secretion, fetal overgrowth and adipose tissue deposition.[5,15]

From the current study, obese subjects had a higher risk of preeclampsia than normal-weight pregnant mothers among non-GDM groups. However, there were no significant differences between obese and non-obese subjects in pre-eclampsia development among the GDM group.

Lyu and colleagues reported in 2024 that obesity and GDM were not associated with pre-eclampsia occurrence.[15] However, Lyu used the BMI cut point at a level equal to or more than 28 kg/m^2, while the current study used the BMI cut point at a level equal to or more than 25kg/m^2 according to the South and Southeast Asia consensus suggested for the Thai population.[19] Data from Sweden and China in 2021 said that risk of severe preeclampsia from obesity in Swedish and Chinese subjects were 2.1 and 4.5 fold, respectively. Severe pre-eclampsia's risk from GDM in both populations were 2.3 and 2.1 folds, respectively.[24] Ke's study reported from China that high weight and obesity were prognostic factors for pre-eclampsia development with OR at 2.7.[14] The association of obesity and GDM to pre-eclampsia was inconsistent among previous studies. Prevalence of pre-eclampsia in the current study, Lyu's, Ke's, Yang's and Swedish studies were 2.6, 1.6, 5.5, 1.6 and 0.9 percent, respectively.[14,15,24]

Pathophysiology of pre-eclampsia was a consequence of vascular factors (endovascular trophoblast invasion of uterine spiral arterioles), immunological factors (dysfunction between maternal, placenta, and fetus tissue), inflammatory and genetic factors.[25] Obesity and GDM might be associated with pre-eclampsia. Most obese subjects had metabolic syndromes consisting of insulin resistance.[5] Consequences of insulin resistance were low-grade inflammation, vascular endothelial activation and increased sodium reabsorption.[25]

From the current study, subjects with obesity in the non-GDM group had an increased risk of NICU admission, mild RDS and birth asphyxia in neonates. Addicott from the US reported in 2024 that increased maternal BMI brought a higher risk of adverse neonatal outcomes, namely preterm, low Apgar score, prolonged neonatal length of stay and NICU admission than those with normal-BMI.[26] Lutsiv from Canada reported in 2015 that morbidly obese patients (Class III, BMI ≥ 40kg/m^2) had higher risk of adverse neonatal outcomes than patients with class I and class II obesity (BMI 30 to 39.9kg/m^2).[27] Data from China and the Philippines in 2016 and 2017 showed that high BMI of patients was associated with poor adverse neonatal outcomes.[23,28]

Overconsumption of food resulted in obesity. Increased adipokines and cytokines in obese subjects led to chronic inflammation. This might be an explanation of adverse neonatal outcomes with obese mothers.[5] The findings from the current study supported existing literature.[23,26-28]

Among GDM groups, obese subjects had higher risk of GHT than non-obese subjects from the present study with statistical significance as shown in Table 2. From US and Chinese studies, overweight and obese participants had significantly higher risk of GHT than their normal-weight counterparts.[22,23] The findings from the current study are in lieu with US and Chinese studies only in GDM subjects. Another explanation was that the average BMI of the GDM group was higher than the non-GDM group (25.7 vs 24.5, p-value 0.004). This indicated that subjects with GDM were more likely to be obese than their non-GDM counterparts. Data from the previous studies were summarized and presented in Table 3.

The study's strengths include its large sample size, robust methodology, and comparison between GDM and non-GDM groups. The retrospective nature is a notable limitation, as the study relies on historical medical records, which may have incomplete data. The lack of consideration for lifestyle factors (diet, exercise) and genetic predisposition to obesity is also a limitation. To prevent the risk of adverse pregnancy outcomes caused by obesity, the following suggestions should be considered namely weight reduction before pregnancy, proper pre-pregnancy weight, pre-conceptional counseling, early ANC and suitable plans should be made and agreed with pregnant women.

CONCLUSION

Obesity in pregnant women (BMI equal or more than 25 kg/m^2) increased the occurrence of PIH, FBW and LGA when compared to those with normal weight. Moreover, obese pregnant women concurrent with GDM had higher risk of GHT. Whereas obese pregnant women with non-GDM had increased risk of preeclampsia, NICU admission, RD, and birth asphyxia.

Gestational diabetes mellitus (GDM) is a common

TABLE 3. Comparison adverse maternal and fetal outcomes of obese women.*

Authors	Present	Yue S.	Alfadhli	Ke JF.	Lyu Y.	Kim SS.	Sun Y.	Yang Y.	Yang Y.	Addicott	Kim T.	Lutsiv
Years	2024	2022	2021	2023	2024	2016	2020	2021	2021	2024	2017	2015
Country	Thailand	Vietnam	Saudi Arabia	China	China	US	China	Sweden	China	USA	US	Canada
Total case (N)	744	4,970	531	746	802	112,309	3,172	555,446	79,243	858	338	-
Total Obese	338(45.4)	267(5.4)	254(47.8)	218(29.2)	71(8.9)	19,482	59	63,846	2,693	472	338	-
GDM	186(55)	78(29.2)	170(66.9)	218	23(32.4)	1,928(1.7)	12	-	-	59	44(13)	-
cut-off BMI	≥25	≥27.5	≥30	≥24	≥28	≥30	≥28	-	-	≥30	≥30	≥30
PPBMI(kg/m^2)**	29.9	-	35.6	25.8	-	18.5	-	24.6	21.9	40.1		-
Age (yr)**	31	30.3	32.9	31	30.6	27.8	18-44	30.9	28.6	27.1	29	-
GA (weeks)**	38.5	38.6	-	38.6	48	-	-	-	-	38.9	38.5	-
Multiparity	257(76)	167(62.5)	-	99	42	13,473	-	-	-	39	250(74)	-
Hx.macrosomia	7(2.1)	12(4.5)	-	-	-	-	-	-	-	-	-	-
PPH	NS	-	-	-	NS	NS	-	-	-	-	S	S
GDM	-	-	-	-	-	S	S	-	-	S	S	S
TWG (kg)	S	S	-	S	-	-	-	-	-	S	-	S
PIH	S		-	S	-	-	-	-	S	-	-	S
GHT	S	-	-	-	-	S	S	-	-	-	NS	S
Preeclampsia	S	-	-	NS	NS	-	-	S	-	S	NS	S
C/S	NS	S	S	S	S	S	S	S	S	S	S	S
FBW (Kg)	S	-	S	S	S	S	S	-	-	-	-	S
SGA	NS	-	-	NS	NS	-	NS	S	-	-	-	NS
LGA	S	S	S	S	S	S	S	S	S	-	S	S
NICU	S	-	S	-	-	S	-	-	S	S	S	S
RDS	S	-	-	-	-	S	-	-	S	-	S	-
Preterm	NS	S	-	S	NS	S	NS	S	-	S	NS	S
Birth asphyxia	S	-	S	-	-	NS	-	-	-	NS	S	NS

*n (%), **Mean ± standard deviation (SD)

Abbreviations: GDM: gestational diabetes mellitus, PPBMI: pre-pregnancy BMI, GA: gestational age at birth, Hx.macrosomia: History of macrosomia, PPH: postpartum hemorrhage, TWG: total weight grain, PIH: pregnancy induced hypertension, GHT: gestational hypertension, C/S: primary cesarean delivery, FBW: fetal birth weight, SGA: small for gestational age, LGA: large of gestational age, NICU: neonatal intensive care unit admission, RDS: respiratory distress syndrome.

complication during pregnancy. Maternal hyperglycemia leads to fetal hyperglycemia and stimulates excessive somatic fetal growth causing macrosomia and large-for-gestational age (LGA) neonates. Maternal obesity is also the most common metabolic complication in pregnancy. Normoglycemic obese women had an increased risk of macrosomia and higher rate of cesarean delivery.

Obesity in pregnancy increased higher occurrence of PIH, FBW and LGA than those with normal weight among pregnant women. Obese pregnant women with GDM had more GHT than non-obese pregnant women. Obesity with non-GDM pregnant women increases the risk of pre-eclampsia, NICU admission, RDS, and birth asphyxia

Data Availability Statement

Data about individual identified samples of this research will be available from the corresponding author Piyawan Pariyawateekul upon reasonable request after the main results of the research have been published.

ACKNOWLEDGEMENT

The present study was funded by Bhumibol Adulyadej Hospital Research Fund. The authors would like to thank all subjects who participated in this study.

DECLARATIONS

Grants and Funding Information

This research has been funded by the Bhumibol Adulyadej Hospital Fund (Funding No. 77/66).

Conflict of Interest

There was no conflict of interest.

Registration Number of Clinical Trial

Not applicable.

Author Contributions

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

Use of Artificial Intelligence

No content was generated using AI.

Institutional Review Board Statement

This retrospective review has been approved by the local Institutional Review board (reference number: 77/66).

REFERENCES

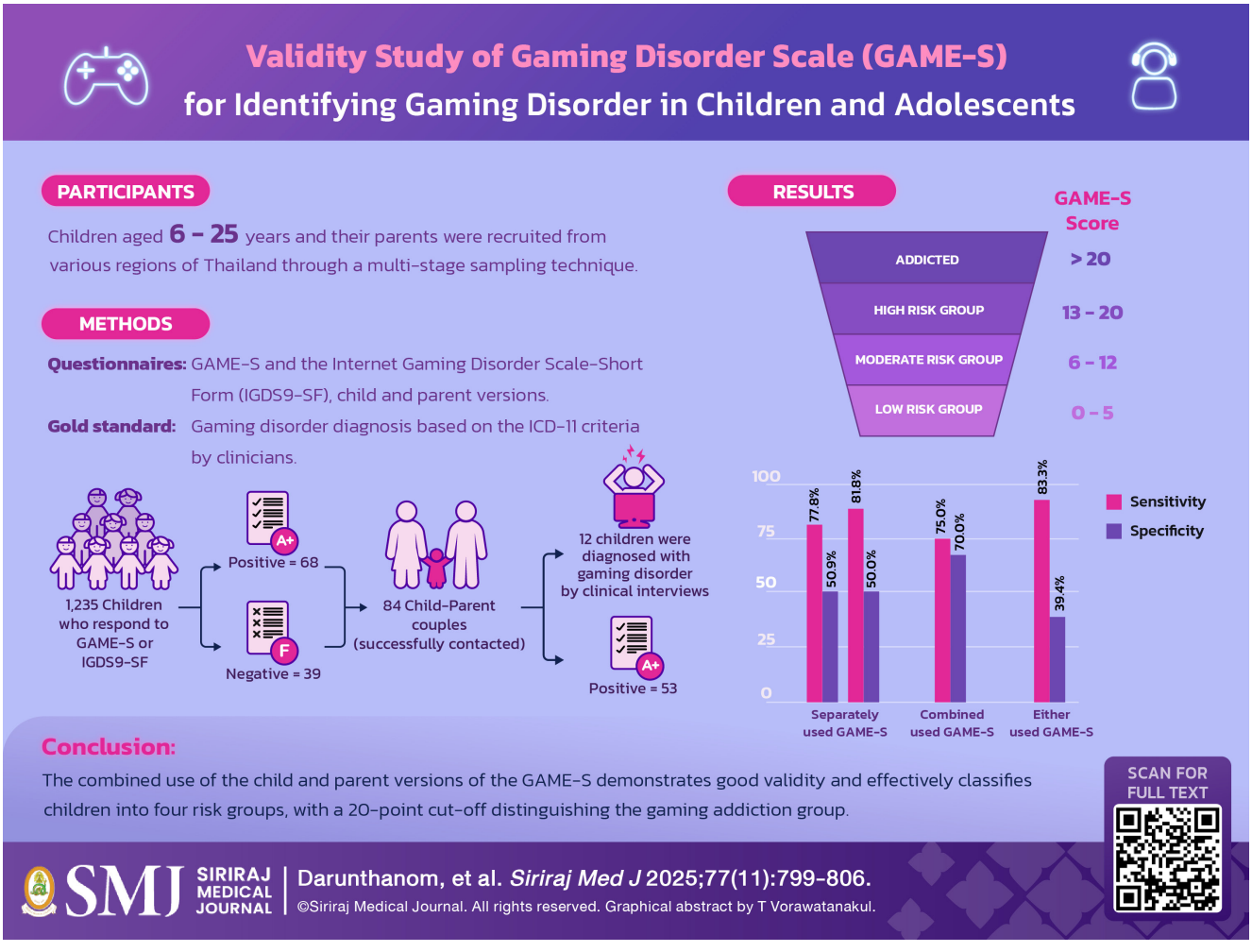
1. Gestational Diabetes Mellitus: ACOG Practice Bulletin, Number 190. Obstet Gynecol. 2018;131:e49-64.
2. Cunningham FG, Leveno KJ. Diabetes Mellitus. In: Cunningham FG, Leveno KJ, editors. Williams Obstetrics. 26th ed. New York: McGraw Hill Education; 2022. p. 1068–88.
3. Macrosomia: ACOG Practice Bulletin, Number 216. Obstet Gynecol. 2020;135(1):e18-e35.
4. Cunningham FG, Leveno KJ. Fetal-growth disorders. In: Cunningham FG, Leveno KJ, editors. Williams Obstetrics. 26th ed. New York: McGraw Hill Education; 2022. p. 823–37.
5. Cunningham FG, Leveno KJ. Obesity. In: Cunningham FG, Leveno KJ, editors. Williams obstetrics. 26th ed. New York: McGraw Hill Education; 2022. p. 902–14.
6. Sunsaneevitayakul P, Sompagdee N, Asad-Dehghan M, Talungchit P. Effect of gestational weight gain on overweight and obese pregnant women. Siriraj Med J. 2022;74(6):364–70.
7. Wadeng W, Onwatanasarikul W, Khongthon N, Pleankong M, Prommas S, Bhamarapravatana K, et al. Sonographic assessment of fetal biometric parameters and maternal height ratio for prediction of cesarean delivery due to labor dystocia. Siriraj Med J. 2025;77(5):436–45.
8. Ijäs H, Koivunen S, Raudaskoski T, Kajantie E, Gissler M, Vääräsmäki M. Independent and concomitant associations of gestational diabetes and maternal obesity to perinatal outcome: a register-based study. PLoS One. 2019;14(8):e0221549.
9. Asanathong N, Jiamjira-anon N, Eiamcharoenwit J, Mantaga S, Thanakiattiwibun C, Siriussawakul A, et al. Incidence of adverse perioperative airway complications in obese non-pregnant and pregnant patients undergoing general anesthesia. Siriraj Med J. 2022; 74(3):178-84.
10. Sirivunnabood T, Wanitpongpan P, Yapan P. Incidence and risk factors of neonatal sepsis in preterm premature rupture of membranes before 34 weeks of gestation. Siriraj Med J. 2022; 74(3):169-77.
11. World Health Organization. The Asia-Pacific perspective: redefining obesity and its treatment. Geneva: WHO Western Pacific Region; 2000. p.15–21.
12. Castaneda C, Marsden K, Maxwell T, Ten Eyck P, Kuwaye D, Kenne KA, et al. Prevalence of maternal obesity at delivery and association with maternal and neonatal outcomes. J Matern Fetal Neonatal Med. 2022;35(25):8544-51.
13. Alfadhli EM. Maternal obesity influences birth weight more than gestational diabetes. BMC Pregnancy Childbirth. 2021; 21(1):111.
14. Ke JF, Liu S, Ge RL, Ma L, Li MF. Associations of maternal pre-pregnancy BMI and gestational weight gain with the risks of adverse pregnancy outcomes in Chinese women with gestational diabetes mellitus. BMC Pregnancy Childbirth. 2023;23(1):414.
15. Lyu Y, Cui M, Zhang L, Zheng G, Zuo H, Xiu Q, et al. Pre-pregnancy body mass index, gestational diabetes mellitus and gestational weight gain: individual and combined effects on fetal growth. Front Public Health. 2024;12:1354355.
16. Yue S, Thi VTK, Dung LP, Nhu BTH, Kestelyn E, Thuan DT, et al. Clinical consequences of gestational diabetes mellitus and maternal obesity as defined by Asian BMI thresholds in Viet

Nam: a prospective, hospital-based, cohort study. BMC Pregnancy Childbirth. 2022;22(1):195.
17. Weinert LS. International Association of Diabetes and Pregnancy Study Groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy: comment to the International Association of Diabetes and Pregnancy Study Groups Consensus Panel. Diabetes Care. 2010;33(7):e97.
18. American Diabetes Association. Standards of medical care in diabetes. Diabetes Care. 2010;33(Suppl 1):S11-61.
19. Tham KW, Abdul Ghani R, Cua SC, Deerochanawong C, Fojas M, Hocking S, et al. Obesity in South and Southeast Asia-A new consensus on care and management. Obes Rev. 2023;24(2):e13520.
20. Cunningham FG, Leveno KJ. Maternal physiology. In: Cunningham FG, Leveno KJ, editors. Williams obstetrics. 26th ed. New York: McGraw Hill Education; 2022. p. 50–79.
21. Waclawek T, Park SQ. Potential mechanisms and modulators of food intake during pregnancy. Front Nutr. 2023;10:1032430.
22. Sun Y, Shen Z, Zhan Y, Wang Y, Ma S, Zhang S, et al. Effects of pre-pregnancy body mass index and gestational weight gain on maternal and infant complications. BMC Pregnancy Childbirth. 2020;20(1):390.
23. Kim SS, Zhu Y, Grantz KL, Hinkle SN, Chen Z, Wallace ME, et al. Obstetric and neonatal risks among obese women without chronic disease. Obstet Gynecol. 2016;128(1):104-12.
24. Yang Y, Le Ray I, Zhu J, Zhang J, Hua J, Reilly M. Preeclampsia prevalence, risk factors, and pregnancy outcomes in Sweden and China. JAMA Netw Open. 2021;4(5):e218401.
25. Cunningham FG, Leveno KJ. Preeclampsia syndrome. In: Cunningham FG, Leveno KJ, editors. Williams obstetrics. 26th ed. New York: McGraw Hill Education; 2022. p. 688–710.
26. Addicott K, Nudelman M, Putty K, Prasher P, Preston D, Yoost JL, et al. Adverse perinatal outcomes associated with increasing maternal obesity. Am J Perinatol. 2024;41(9):1275-81.
27. Lutsiv O, Mah J, Beyene J, McDonald SD. The effects of morbid obesity on maternal and neonatal health outcomes: a systematic review and meta-analyses. Obes Rev. 2015;16(7):531-46.
28. Kim T, Burn SC, Bangdiwala A, Pace S, Rauk P. Neonatal morbidity and maternal complication rates in women with a delivery body mass index of 60 or higher. Obstet Gynecol. 2017;130(5):988-93.

Validity of the Gaming Disorder Scale (GAME-S) for Identifying Gaming Disorder in Children and Adolescents

Winudda Darunthanom, M.D., Chanvit Pornnoppadol, M.D., Wanlop Atsariyasing, M.D.*

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



*Corresponding author: Wanlop Atsariyasing
E-mail: wanlop.atr@mahidol.ac.th
Received 10 April 2025 Revised 16 May 2025 Accepted 17 May 2025
ORCID ID:http://orcid.org/0000-0001-8868-8302
https://doi.org/10.33192/smj.v77i11.274309

ABSTRACT

Objective: This study aimed to evaluate the validity of the Gaming Disorder Scale (GAME-S) for diagnosing gaming disorder among children and adolescents in Thailand.

Materials and Methods: A total of 1,235 individuals, aged 6-25 years, and their parents were recruited from various regions of Thailand through a multi-stage sampling technique. Data were collected using both the child and parent versions of the GAME-S. The results were compared with ICD-11 clinical diagnoses of gaming disorder, conducted by trained clinicians. Concurrent validity was evaluated through the Thai version of the Internet Gaming Disorder Scale-Short Form (IGDS9-SF).

Results: A total of 1,235 children and adolescents and their parents participated to assess the correlation between the GAME-S and IGDS9-SF. For GAME-S diagnostic accuracy analysis, 84 child–parent pairs, balanced by gender (mean age = 13.48 years), were included. Sensitivities of the child and parent versions were 77.8% and 81.8%, respectively, while specificity was 50.9% and 50.0%. Combining both versions increased sensitivity to 75.0% and specificity to 70.0%. The GAME-S correlated strongly with IGDS9-SF (child: r=0.823, p=0.01; parent: r=0.82, p=0.01). Latent class analysis identified a four-class model as the best fit. The gaming addiction group was effectively characterized by a cut-off score of 20.

Conclusion: The combined use of both the child and parent versions of the GAME-S demonstrates good validity and is a reliable tool for screening children and adolescents for gaming disorder. Additionally, it can categorize children and adolescents into four risk groups, helping identify individuals and provide appropriate interventions for each group.

Keywords: Gaming disorder; sensitivity and specificity; questionnaire; children; adolescent (Siriraj Med J 2025; 77: 799-806)

INTRODUCTION

Gaming disorder is abnormal gaming behavior characterized by three core groups of symptoms, based on the *International Classification of Diseases, 11th Revision* (ICD-11) diagnostic criteria[1], including impaired control over gaming, prioritization of gaming over other activities, and continued gaming despite negative consequences. These behaviors must persist for at least 12 months and significantly impact daily life, including health, education, work, and relationships. The global prevalence of gaming addiction is estimated at 3.3%, with the prevalence in adolescents and young adults around 6.3%.[2] In Thailand, a nationwide survey in 2006 found gaming addiction prevalence among children and adolescents to be between 4.1% and 5.5%.[3]

Gaming disorder has been linked to various physical and psychological issues. Physically, prolonged gaming can lead to eye strain, back pain, irregular eating habits, and inadequate sleep due to prolonged gaming sessions. Psychologically, gaming disorder is associated with emotional instability, irritability, aggressive behavior, and difficulties with responsibility and social interaction. Children with gaming addiction may disengage from school, exhibit behavioral problems, or even engage in gambling and theft. Families of affected individuals often experience increased conflict and a reduced quality of life, both physically and mentally. If left untreated, gaming disorder can lead to long-term developmental and behavioral problems in children and adolescents.[3-5]

Early screening and intervention are crucial for effective treatment of gaming disorder.[6] Several assessment tools have been developed internationally and in Thailand to screen and assess gaming disorder.[7] Globally recognized tools include the Game Addiction Scale for Adolescents (GASA)[8], the Internet Gaming Disorder Scale–Short Form (IGDS9-SF)[9], and the Ten-Item Internet Gaming Disorder Test (IGDT-10).[10] In Thailand, two key tools have been developed: the Game Addiction Screening Test (GAST)[3], and the Gaming Disorder Scale (GAME-S).[11] Additionally, the Social-Media Addiction Screening Scale (S-MASS)[21] is another tool in Thailand developed to screen addiction. However, it focuses on social media, not gaming directly.

The GAME-S consists of nine items based on the ICD-11 criteria and was specifically designed for the Thai context, targeting individuals aged 10–24 years. It has demonstrated good reliability in both the self-report and parent-report versions.[11] In contrast, although the Game Addiction Screening Test (GAST) has shown reasonable utility, it was developed prior to the establishment of standardized diagnostic criteria for gaming disorder, limiting its diagnostic alignment. Some of its items—

such as "I have skipped class to play games"—may no longer reflect current gaming patterns, especially as online gaming has become more accessible without the need to skip school. Additionally, the GAST includes a relatively large number of items, which can increase respondent burden. Compared to GAST, the GAME-S is shorter, more relevant to contemporary behaviors, aligned with international diagnostic standards, and thus more practical and valid for both research and clinical screening.[11]

However, *Sangkhaphan's* study did not evaluate the criterion validity of the GAME-S or examine its cut-off point, both of which are important for clinical use and epidemiological studies. Therefore, this study aims to assess the sensitivity and specificity of the GAME-S by comparing its results with clinical interview diagnoses. Additionally, it will evaluate the concurrent validity of the GAME-S against the IGDS9-SF (Thai-version) in children and adolescents. Finally, we seek to classify risk groups for gaming disorder.

MATERIALS AND METHODS

Population and sample

This study was conducted as a questionnaire-based cross-sectional study. Data were collected between July and November 2023. The study population consisted of individuals aged 6 to 25 years in Thailand.

The sample included children and adolescents and their parents or guardians living in communities from 15 provinces, representing different country regions. A multi-stage stratified random sampling method was used to ensure proportional representation. The number of participants from each region was determined based on the population distribution of children and adolescents aged 6-25 years, as reported by the Department of Local Administration, Ministry of Interior. The selected regions included Bangkok, Central region (three provinces), Northern region (three provinces), Northeastern region (five provinces), and Southern region (three provinces), or a total of 15 provinces. Within each region, a province was purposively selected based on the availability of child and adolescent psychiatrists to ensure that participants identified as at high risk for gaming addiction could access appropriate care.

One district was randomly chosen within each selected province through simple random sampling. Participants were then randomly selected from urban and rural areas in equal numbers. For each area, two participants (one male and one female) were selected from each age group (6 to 25 years), totaling 80 participants per province. In total, 80 participants were recruited from each of the 14 regional provinces and 160 participants from Bangkok, for a planned total of 1,280 participants.

Sample size calculation

The sample size was determined using the n4Studies program and the formula for a finite population proportion.[12] The sample size calculation was as follows:

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

Thailand's estimated population of children and adolescents aged 6-25 years is 16,503,799 (N = 16,503,799).[13] The prevalence of gaming addiction was assumed to be 15%[14] with an acceptable margin of error of 2% (d = 0.02) and a 95% confidence level (Alpha (α) = 0.05, Z (0.975) = 1.96). Based on this, the minimum required sample size was 1,225 participants. To account for potential sample loss and incomplete data, the sample size was increased by 5%, resulting in a final sample size of 1,280 participants.

To evaluate the sensitivity and specificity of the GAME-S, participants who screened negative on both the GAME-S and IGDS9-SF (Thai version) were randomly selected and matched by gender, age, and province. The number of negative-screened participants was half the number of positive-screened participants. A 2:1 ratio of positive to negative cases was used to improve sensitivity estimation, in line with the screening objective of the GAME-S. A higher proportion of positive cases ensured sufficient power for analysis, as recommended in validity studies.[19,20]

Inclusion and exclusion criteria

The inclusion criteria included individuals aged 6 to 25 years and their parents or guardians who could communicate effectively in Thai. The exclusion criteria included individuals who were unable to complete the questionnaire.

All participants provided informed consent before participation. Written informed consent was obtained from parents or guardians, and written assent was obtained from children and adolescents, using signed paper forms. The study was conducted following ethical standards and approved by the Siriraj Institutional Review Board (SiRB), Faculty of Medicine Siriraj Hospital, Mahidol University (COA no. Si 209/2022).

Instruments

General Information Questionnaire: This questionnaire collected demographic details, including gender, age,

education level, academic performance, and socioeconomic status.

Gaming Disorder Scale (GAME-S): The GAME-S consists of the self-version and the parent-report version. Each version includes nine items, assessing three categories of gaming disorder symptoms based on the ICD-11 criteria: three items related to difficulty in controlling gaming, three items focused on prioritizing gaming over other important activities, and three items related to continuing gaming despite experiencing negative consequences. Responses were scored from 0 (not at all) to 3 (definitely). Gaming disorder was considered present if at least one item from each of the three symptom groups was rated at level three. The reliability of the GAME-S was validated, with Cronbach's alpha coefficients of 0.92 (self-version) and 0.96 (parent-version), and factor loadings ranging from 0.66 to 0.86. The Pearson correlation coefficients between the GAME-S and GAST were 0.87 for the self-version and 0.92 for the parent-version.

Internet Gaming Disorder Scale-Short Form (IGDS9-SF): The original IGDS9-SF was developed by *Pontes and Griffiths*[9] and the Thai version was translated by *Pornnapadol*.[15] It assesses gaming addiction in children and adolescents, with separate versions for children and adolescents, and parents. The scale consists of nine items based on the diagnostic criteria of the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5). The responses for each item are 'Yes' (one point) and 'No' (0 points). A total score of five or higher indicates gaming disorder. The Thai version of IGDS9-SF has a Cronbach's alpha coefficient of 0.913 (self-report version) and 0.922 (parent-report version).

Diagnosis of Gaming Disorder: Diagnosis was conducted through interviews with physicians and psychologists trained in assessing gaming disorder based on the ICD-11 criteria. The interrater reliability (Kappa) for the diagnosis of gaming disorder was 1.00.

Data collection process

The researchers collaborated with the Children and Youth Council of Thailand and local networks in each province to disseminate information within communities. Individuals who met the inclusion criteria and agreed to participate were provided with questionnaires.

Participants who screened positive for gaming disorder based on either the GAME-S or IGDS9-SF (Thai version) (either self-report or parent-report) were invited for clinical interviews to confirm the diagnosis. It should be noted that interviewers were aware of the participants' screening test scores; therefore, they were not blinded assessors.

Data analysis

Data analysis was conducted using SPSS version 18. Descriptive statistics were employed to present general information. To assess validity, sensitivity and specificity were calculated. Concurrent validity was assessed using Pearson's correlation coefficient. Additionally, latent class analysis (LCA) was performed to examine the underlying structure of the GAME-S and assess its ability to classify individuals based on response patterns. Model selection was guided by statistical fit indices and clinical interpretability.

RESULTS

1,058 parents and 1,235 children completed the GAME-S and IGDS9-SF (Thai version) questionnaires, including both child and parent versions. Among the children, 68 met the screening criteria for gaming disorder based on the GAME-S or IGDS9-SF (Thai version) and were eligible for diagnostic interviews. An additional 39 children who did not meet the criteria were randomly selected, bringing the total number of planned interviews to 107. Of these, 84 were successfully conducted, identifying 12 cases of gaming disorder.

Demographic characteristics of interviewed participants

Most parents were female (86.4%), with a mean age of 42.77 years, and 60.6% were married. Most parents had an educational level below a bachelor's degree, and household income was below 10,000 Thai baht (Table 1). Most children were male (59.4%), with a mean age of 13.48 years. Additionally, 47.4% of children were enrolled in grades 1-6, and 40.3% were in grades 7-12. Also, most children had a GPA between 3.00 and 4.00 (57.8%) (Table 2).

Sensitivity and specificity of the GAME-S

Sensitivity and specificity analysis showed that the child version of the GAME-S had a sensitivity of 77.8% and a specificity of 50.9%, while the parent version had a sensitivity of 81.8% and a specificity of 50.0%. When both versions were combined, the specificity improved, with a sensitivity of 75.0% and a specificity of 70.0%. Using either the child or parent version alone yielded the highest sensitivity (83.3%) but the lowest specificity (39.4%) (Table 3).

Concurrent validity of the GAME-S

Concurrent validity was evaluated using Pearson's correlation coefficient, which revealed a strong correlation between the GAME-S and IGDS9-SF (Thai version) in both the child ($r=0.82$, $p=0.01$) and parent ($r=0.81$, $p=0.01$) versions.

TABLE 1. Demographic data of parents.

Demographic data of parents (n=81)	n	%
Gender		
Male	12	14.8
Female	69	85.2
Age (mean ± S.D.)		
Range = 20-65 years	41.81 ± 9.69	
Education level		
Below bachelor's degree	52	64.2
Bachelor's degree	26	32.1
Higher bachelor's degree	3	3.7
Marital status		
Single	12	14.8
Living with spouse	51	63.0
Separated/widowed/divorced	18	22.2
Family monthly income (THB)		
< 10,000	30	37.0
10,000-20,000	29	35.8
Greater than 20,000	22	27.2

Note: S.D.=standard deviation. THB=Thai Baht

TABLE 2. Demographic data of children.

Demographic data of children (n=84)	n	%
Gender		
Male	50	59.5
Female	34	40.5
Age (mean ± S.D.)		
Range = 6-24 years	13.44 ± 4.79	
Education level		
Year 1-6	35	41.7
Year 7-12	32	38.1
University	8	9.5
Unreported	9	10.7
Grade Point Average		
0.00-2.00	6	7.2
2.01-2.50	13	15.5
2.51-3.00	18	21.4
3.01-3.50	27	32.1
3.51-4.00	20	23.8

Note: S.D. = standard deviation. THB = Thai Baht

TABLE 3. Sensitivity and specificity of the GAME-S.

Version	Sensitivity (%)	Specificity (%)
Child	77.8	50.9
Parent	81.8	50.0
Child + Parent Combined	75.0	70.0
Either Child or Parent	83.3	39.4

Latent Class Analysis (LCA)

LCA identified a 4-class model as the best fit for both parent-reported (AIC=6814.43, BIC=6854.15, entropy=0.836) and child-reported (AIC=7751.87, BIC=7792.82, entropy=0.854) GAME-S. The cut-off points derived from the 4-class LCA model are presented in Table 4.

DISCUSSION

This study assessed the validity of the GAME-S in diagnosing gaming disorder among children and adolescents. The results indicate that the GAME-S demonstrates acceptable sensitivity and specificity for detecting gaming disorder, with improved specificity and maintained sensitivity when combining reports from children and parents. Additionally, a strong correlation was found between the GAME-S and IGDS9-SF (Thai version) for both child and parent versions.

When compared to the IGDS9-SF (Thai version), a self-report measure for diagnosing gaming disorder based on ICD-11 criteria, the GAME-S child version showed higher sensitivity (77.8%) but lower specificity (50.9%) than the IGDS9-SF (50.0% and 97.6%, respectively). Combining the child and parent versions in the GAME-S maintained higher sensitivity (75.0% vs 50.0%), while improving specificity, though it remained lower (70.0% and 97.6%).

A similar pattern was observed when comparing the GAME-S to the IGDS9-SF (Thai version), which is used to diagnose gaming disorder based on the DSM-5 criteria. The GAME-S, whether the child-only version or the combined child and parent versions, consistently demonstrated higher sensitivity and lower specificity. These differences may be attributed to variations in the sample populations. In this study, participants were drawn from a community sample with balanced gender distribution, aged between 6 and 25 years, and predominantly from lower socioeconomic status. In contrast, the study by *Boonyapasert* et al. (2021)[16] used a clinical sample composed mainly of males aged 11-18 years, with higher socioeconomic status.

When compared to the GASA (Arabic version), a self-report measure, the GAME-S (in both the child-only and combined versions) showed comparable sensitivity to the GASA (Arabic version) (77.8%, 75.0%, and 76.9%, respectively). However, the GAME-S demonstrated lower specificity than the GASA (Arabic version) (50.9%, 70.0%, and 84.2%). These discrepancies may be due to

TABLE 4. Cut-off points from the four-class Latent Class Analysis model of the GAME-S.

Score (Total 0-27)	Parent-version (n=1,058)		Child-version (n=1,235)	
	n	%	n	%
0-5	451	42.6	602	48.7
6-12	327	30.9	356	28.8
13-20	220	20.8	233	18.9
>20	60	5.7	44	3.6

variations in the sample population and the criteria used as the diagnostic gold standard. The study by *Abolfotouh et al.* (2024)[17] involved adolescents recruited online who had higher parental education levels. Furthermore, their diagnostic approach was based on affirmative responses to items related to relapse, withdrawal, conflict, and problems, rather than the DSM-5 or ICD-11 criteria, and did not involve clinical interviews.

When compared to the IGDT-10 (Japanese version), a self-report measure, the GAME-S (in both the child-only and combined child-parent versions) demonstrated lower sensitivity (77.8%, 75.0%, and 87.5%) and lower specificity (50.9%, 70.0%, and 85.2%). These differences may be attributed to variations in the sample populations. In the study by *Mihara et al.* (2022)[18], participants were gamers aged 10-29 years, and clinical interviews based on DSM-5 criteria were used as the diagnostic gold standard for internet gaming disorder.

For risk classification, based on latent class analysis (LCA) results and prevalence data from the GAME-S, participants were categorized into four risk groups based on their gaming disorder severity scores:

1) Low-risk group: Participants with GAME-S scores ranging from 0 to 5 (child: 48.7%, parent: 42.6%), corresponding to the lowest-risk LCA group. These individuals show no significant gaming-related issues and are unlikely to fulfil the criteria for gaming disorder.

2) Moderate-risk group: Participants with GAME-S scores ranging from 6 to 12 (child: 28.8%, parent: 30.9%) represent an elevated but subclinical level of gaming behavior. While they do not meet the diagnostic criteria, they may engage in prolonged gaming that causes some functional impact.

3) High-risk group: Participants with scores ranging from 13 to 20 (child: 18.9%, parent: 20.8%), corresponding to the high-risk group identified by LCA. These individuals display problematic gaming behaviors, with a higher likelihood of functional impairment, warranting clinical attention.

4) Addicted group (gaming disorder diagnosis): Participants scoring above 20 (child: 3.6%, parent: 5.7%) closely correspond to the prevalence of gaming disorder as assessed by GAME-S (child: 3.2%, parent: 4.3%). This group is most likely to meet the diagnostic criteria for gaming disorder and may require further psychiatric evaluation and intervention.

This classification offers a structured approach to the early identification and intervention, ensuring that at-risk individuals receive appropriate monitoring and clinical care.

The study has several strengths, including a diverse, community-based sample drawn from all regions of Thailand and a broad age range of participants (6-25 years), which enhances the generalizability of the findings. The use of clinical interviews based on the ICD-11 diagnostic criteria further enhances the study's criterion validity. Additionally, it collects data from both child and parent perspectives, which improves the reliability of the responses. However, some limitations should be acknowledged. The sample consisted of individuals from non-clinical backgrounds with predominantly low socioeconomic status, which may limit the generalizability of the results to clinical populations or those with different socioeconomic backgrounds. Additionally, not all eligible participants attended clinical interviews, potentially introducing attribution bias. Interviewers were informed of the participants' screening test scores and thus were not blinded during the diagnostic assessment. Lastly, test-retest reliability was not assessed, leaving the stability of responses over time undetermined.

Future research should aim to improve specificity by refining scoring thresholds or incorporating behavioral indicators, such as gaming duration and psychosocial impact. Longitudinal studies are also needed to assess the predictive validity of GAME-S over time. Additionally, cross-cultural validation studies could expand their applicability beyond Thailand, particularly in Southeast Asian populations, where gaming disorder is an emerging public health concern.

CONCLUSION

The GAME-S, particularly when using both the child and parent versions, is an effective screening tool for gaming disorder among children and adolescents. Additionally, the GAME-S can categorize children and adolescents into four risk groups (low, moderate, high, and addicted groups) to help identify individuals and provide appropriate interventions for each group.

Data Availability Statement

The data underlying the results of this study are not accessible to the public due to ethical and confidentiality constraints. However, de-identified data can be provided upon request to the corresponding author, pending approval from the institutional ethics review board. Relevant secondary data sources are referenced in the bibliography.

ACKNOWLEDGEMENTS

The research team would like to express sincere gratitude to the Child and Youth Council of Thailand and its network across 15 provinces for their support

in coordinating and collecting data. We also thank Ms. Natthaporn Kangsawiwat for her coordination throughout the research project, and Ms. Narathip Sanguanpanit from the Faculty of Medicine Siriraj Hospital for her statistical consultation. Lastly, we would like to thank Mr. Aditya Rana from the Faculty of Medicine Siriraj Hospital for his assistance with language editing of this manuscript.

DECLARATIONS

Grants and Funding Information

This research was funded by the National Research Council of Thailand for the fiscal year 2023. National Research Council of Thailand (NRCT)/N32A660126.

Conflict of Interest

The authors have no conflicts of interest to disclose regarding this study.

Registration Number of Clinical Trial

Not applicable.

Author Contributions

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

Use of Artificial Intelligence

Not applicable.

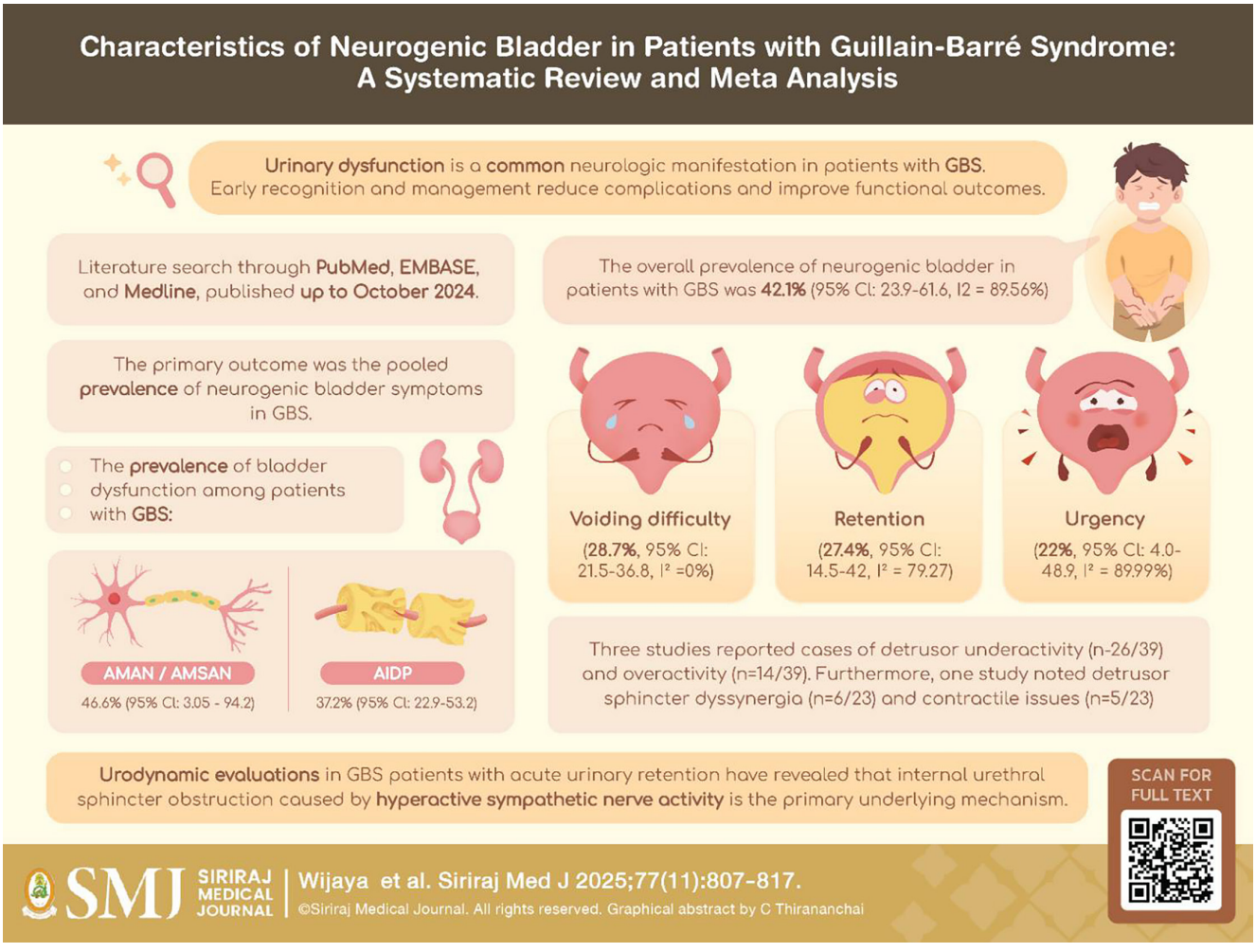
REFERENCES

1. World Health Organization. Gaming disorder. 2018 [cited 2023 Dec 27]. Available from: https://www.who.int/standards/classifications/frequently-asked-questions/gaming-disorder.
2. Kim HS, Son G, Roh EB, Ahn WY, Kim J, Shin SH, et al. Prevalence of gaming disorder: A meta-analysis. Addict Behav. 2022;126:107183.
3. Pornnoppadol C, Sornpaisarn B, Kumkiang K. The development of Game Addiction Screening Test (GAST). J Psychiatr Assoc Thailand. 2014;59(1):3–14.
4. Pornnoppadol C, Ladawan na Ayudhaya S, Phoasavasdi C, Surapongphiwattana T. Development of Game Addiction Protection Scale (GAME-P). J Psychiatr Assoc Thailand. 2017;62(1):3-16.
5. Phengpan K. Preventing health problems of childhood addicted to online games. J Ratchathani Innov Health Sci. 2017;1(1):16-30.
6. Saunders JB, Hao W, Long J, King DL, Mann K, Fauth-Bühler M, et al. Gaming disorder: Its delineation as an important condition for diagnosis, management, and prevention. J Behav Addict. 2017;6(3):271-9.
7. Yoon S, Yang Y, Ro E, Ahn W-Y, Kim J, Shin S-H, et al. Reliability, convergent, and discriminant validity of gaming disorder scales: A meta-analysis. Front Psychol. 2021;12:764209
8. Lemmens J, Valkenburg P, Peter J. Development and validation of a game addiction scale for adolescents. Media Psychol. 2009;12(1):77-95.
9. Pontes HM, Griffiths MD. Measuring DSM-5 internet gaming disorder: Development and validation of a short psychometric scale. Comput Human Behav. 2015;45:137-43.
10. Király O, Sleczka P, Pontes HM, Urbán R, Griffiths MD, Demetrovics Z. Validation of the Ten-Item Internet Gaming Disorder Test (IGDT-10) and evaluation of the nine DSM-5 Internet Gaming Disorder criteria. Addict Behav. 2017;64:253-60.
11. Sangkhaphan T, Pornnoppadol C, Hataiyusuk S. Development of the Gaming Disorder Scale (GAME-S). J Psychiatr Assoc Thailand. 2023;68(1):50–61.
12. Ngamjarus C, Chongsuvivatwong V. n4Studies: Sample size and power calculations for iOS. 2014. /Daniel WW. Biostatistics: A foundation of analysis in the health sciences. 6th ed: John Wiley&Sons, Inc; 1995.
13. Department of Provincial Administration. Ministry of Interior. Official statistics registration systems 2022 [cited 2024 November 14]. Available from: https://stat.bora.dopa.go.th/stat/statnew/statMONTH/statmonth/#/displayData.
14. Kolkijkovin V, Wisitpongaree C, Tachakasem P, Pornnoppadol C, Supawattanabodee B. Computer Game Addiction: Risk and Protective Factors in Students in Dusit District, Bangkok. Vajira Med J. 2015;59(3):1-13.
15. Pornnoppadol C, Pariwatcharakul P. The Development of Internet Gaming Disorder Scale–Short-Form (IGDS9-SF). 2021.
16. Boonyaprasert U, Kiatrungrit K. Validity and Reliability of the Thai Internet Gaming Disorder Scale 9-item (IGD Scale-9 TH) in adolescents in the psychiatry out-patient unit. J Psychiatr Assoc Thailand. 2021;66(3):243–60.
17. Abolfotouh MA, Barnawi NA. Prevalence and prediction of video gaming addiction among Saudi adolescents, using the Game Addiction Scale for Adolescents (GASA). Psychol Res Behav Manag. 2024;17:3889–903.
18. Mihara S, Osaki Y, Kinjo A, Matsuzaki T, Nakayama H, Kitayuguchi T, et al. Validation of the Ten-Item Internet Gaming Disorder Test (IGDT-10) based on the clinical diagnosis of IGD in Japan. J Behav Addict. 2022;11(4):1024–34.
19. Bujang MA, Adnan TH. Requirements for minimum sample size for sensitivity and specificity analysis. J Clin Diagn Res. 2016;10(10):YE01–6.
20. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. J Biomed Inform. 2014;48:193–204.
21. Chanpen S, Pornnoppadol C, Vasupanrajit A, Dejatiwongse Na Ayudhya Q. Assessment of the validity and reliability of the Social-Media Addiction Screening Scale (S-MASS). Siriraj Med J [Internet]. 2023 Mar 1 [cited 2025 May 16];75(3):167-80. Available from: https://he02.tci-thaijo.org/index.php/sirirajmedj/article/view/261044

Characteristics of Neurogenic Bladder in Patients with Guillain-Barré Syndrome: A Systematic Review and Meta-analysis

Vincent Ongko Wijaya, M.D.[1], Desak Made Cittarasmi Saraswati Seputra, M.D.[1], Ni Made Dwita Pratiwi, M.D.[2,*], Komang Arimbawa, M.D.[2]

[1]Department of Neurology, Faculty of Medicine, Universitas Udayana, Denpasar, Bali, Indonesia, [2]I.G.N.G Ngoerah General Hospital, Denpasar, Bali, Indonesia.



**Corresponding author: Ni Made Dwita Pratiwi*
E-mail: dwtpratiwi@gmail.com
Received 18 June 2025 Revised 6 September 2025 Accepted 6 September 2025
ORCID ID:http://orcid.org/0000-0002-3731-3663
https://doi.org/10.33192/smj.v77i11.276071

ABSTRACT

Objective: This review aims to determine the prevalence and characteristics of neurogenic bladder issues in patients diagnosed with Guillain-Barré Syndrome.

Materials and Methods: The authors conducted a literature search through PubMed, EMBASE, and Medline, published up to October 2024. Moreover, supplementary sources were obtained through the examination related references. Studies presenting the urinary dysfunction features of GBS patients in their data were included. The primary outcome was the pooled prevalence of neurogenic bladder symptoms in GBS. The quality of the included studies was evaluated using the Oxford Centre for Evidence-Based Medicine guidelines. Selected studies were included in the meta-analysis of proportion and heterogeneity test.

Results: From 257 identified studies, 6 observational studies were included in the final analysis, with 375 participants included. The overall prevalence of neurogenic bladder in patients with GBS was 42.1% (95% CI: 23.9-61.6, I^2 = 89.56%). Voiding difficulty (28.7%, 95% CI: 21.5-36.8, I^2 =0%), retention (27.4%, 95% CI: 14.5–42, I^2 = 79.27), and urgency (22%, 95% CI: 4.0–48.9, I^2 = 89.99%) was commonly reported bladder symptoms, respectively. Acute Axonal Pattern (AMAN/AMSAN) was more common (46.6%, 95% CI: 3.05-94.2, I^2 =95.31%) than Acute Demyelinating Pattern (AIDP) (37.2%, 95% CI: 22.9-53.2, I^2 = 0%). Detrusor underactivity (n=26/39) was mostly reported findings based on urodynamic tests.

Conclusion: Urinary dysfunction is a common neurologic manifestation in patients with GBS. Early recognition and management reduce complications and improve functional outcomes.

Keywords: Neurogenic bladder; GBS; urinary dysfunction (Siriraj Med J 2025; 77: 807-817)

INTRODUCTION

Guillain-Barré syndrome (GBS) is defined as an acute immune-mediated neurological disorder characterized by the demyelination of nerve cells, which manifests as polyradiculoneuropathy. A common presentation includes a gradual ascending weakness in the extremities, as well as bulbar or ocular muscle involvement.[1] Although less typical, autonomic dysfunction, particularly lower urinary tract issues, may present as the initial symptom of GBS.[2,3] Therefore, the prompt recognition of this condition is crucial to prevent further complications.

A previous study involving 65 GBS patients found that 27.7% experienced bladder dysfunction, with urine retention reported in 9.2% of cases. In a longitudinal cohort study, over one-third of patients with GBS reported experiencing urinary urgency and frequency, with nearly 50% noting that these issues disrupted their daily activities. Additional urodynamic abnormalities included underactive and overactive detrusor, as well as, to a lesser extent, hyperactive sphincter has been linked as underlying mechanism. Urinary complications were more common in the GBS subtype of acute inflammatory demyelinating polyneuropathy (AIDP) compared to acute motor axonal neuropathy (AMAN), and these issues were associated with severity of disease and coexisting of bowel impairment.[4,5]

Recent studies indicate that bladder dysfunction affects more than a quarter of individuals with GBS; however, it often goes underdiagnosed, suggesting that the actual prevalence may be higher than reported. Consequently, the complete spectrum of this syndrome remains unclear. Furthermore, only a limited number of investigations have documented neurogenic bladder symptoms in GBS. This review aims to determine the prevalence and characteristics of neurologic bladder issues in patients diagnosed with Guillain-Barré Syndrome. Consequently, this meta-analysis is necessary to clarify the discrepancies of urinary disturbances prevalence in GBS.

MATERIALS AND METHODS

Search strategy and selection criteria

A systematic review followed by a meta-analysis was conducted according to Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines.[6] A comprehensive literature search was conducted using PubMed, EMBASE, and Medline databases, focusing on publications up to 31 October 2024. The literature search utilized the terms "neurogenic bladder" OR "bladder" OR "urinary" AND "Guillain-Barré Syndrome," as well as their relevant derivatives from selected articles. Additionally, we reviewed the reference lists of these papers and their subsequent citations to identify further relevant studies. We included observational studies that reported prevalence data concerning bladder dysfunction among patients with Guillain-Barré Syndrome. Moreover,

we reviewed the reference lists of the articles through our search strategy to uncover any additional studies that may have been overlooked.

Studies were included if they reported bladder symptoms in patients over 18 years, with a diagnosis of GBS based on clinical electrophysiology findings. The definition of "neurogenic bladder" in our study was considered as symptoms of urinary retention, urgency, incontinence, and voiding difficulty caused by damage in nervous system without any other potential causes (e.g., obstructive, urinary tract infection). The exclusion criteria were 1) non-original studies (e.g. review articles); 2) were written in non-English languages; 3) non-human subjects; 4) conference papers, conference presentations, books, graduate degree theses, and another non-peer-reviewed articles. In line with our primary objective, participants which had pre-existing urinary tract symptoms before the onset of GBS were also excluded. Patients with comorbidity of urinary infection as a complication of hospital stay were also excluded in our search strategy.

The eligibility of studies was initially assessed through independent screening of titles and abstracts by two authors. Full-text articles that passed the initial screening were subsequently reviewed by two reviewers to confirm their compliance with the established inclusion and exclusion criteria. Data were extracted for each eligible study into a customized spreadsheet by one reviewer. Additionally, articles referenced in the collected studies were included and manually screened. Any discrepancies were resolved in consultation with the third author. From each reviewed article, the following data were extracted and organized: author names, study design, country of origin, study population, participant age, bladder symptoms, and main findings.

Assessment of quality and risk of bias

The quality of each study was assessed using the Oxford Centre for Evidence-Based Medicine quality ratings, which range from Level 1 to 5. Level 1 indicating a well-powered and appropriately conducted randomized controlled trial (RCT), and Level 5 reflecting expert opinion and case reports.[7]

Statistical analysis

The primary outcome of the study was the pooled prevalence of neurogenic bladder manifestations in GBS, utilizing point prevalence estimates when available. The I^2 tests were employed to assess heterogeneity among the studies, with studies exhibiting an I^2 value greater than 50% classified as having high heterogeneity. A fixed-effects model was applied in situations where no significant heterogeneity was detected among the studies; conversely, a random-effects model was utilized when the data were determined to be heterogeneous. Forest plots were used to illustrate the proportions reported in each study, along with the combined estimated prevalence plots accompanied by 95% confidence intervals. Substantial heterogeneity was defined as $I^2 > 50\%$ and statistical significance was established at $p < 0.05$. The analysis was performed using MedCalc V.19.2.0 software.

RESULTS

An initial search yielded 257 potentially relevant manuscripts based on the applied search strategy. Following the application of exclusion criteria, 66 articles were screened for eligibility, of which six studies fulfilled the inclusion criteria and were included in the quantitative synthesis. The search and selection process are illustrated in the PRISMA flow diagram (Fig 1).

Study characteristics

This review incorporated two prospective studies[5,11], two retrospective studies[10,12], and two cross-sectional studies[4,9] encompassing a total of 275 patients (with sample sizes ranging from 7 to 171) who had a previous history of GBS. Patient data for the included studies originated from multiple countries, including the United States, Japan, India, and Australia. The characteristics and quality ratings of the studies included in this meta-analysis are detailed in Table 1.

The overall prevalence of neurogenic bladder among patients with GBS was found to be 42.1% (95% CI: 23.9-61.6) across four studies[4,5,11,12], which comprised a total of 240 subjects and exhibited a high level of heterogeneity ($I^2 = 89.56\%$), as shown in Table 2. Additionally, the forest plot depicting the prevalence (%) of neurogenic bladder in GBS is presented in Fig 2.

According to electrodiagnostic findings, the prevalence of bladder dysfunction among patients with Guillain-Barré Syndrome (GBS) is documented at 46.6% (95% CI: 3.05-94.2) for the Acute Axonal Pattern (AMAN/AMSAN) and 37.2% (95% CI: 22.9-53.2) for the Acute Demyelinating Pattern (AIDP), as detailed in Table 3. Forest plots illustrating the prevalence (%) of patients with both the Acute Axonal Pattern and the Acute Demyelinating Pattern are presented in Fig 3.

Voiding difficulty is the most commonly reported bladder symptom among patients with GBS, with an overall pooled prevalence of 28.7% (95% CI: 21.5-36.8) derived from a total of 141 subjects, displaying a low level of heterogeneity ($I^2 = 0\%$). A forest plot depicting the prevalence (%) of voiding difficulty is provided in

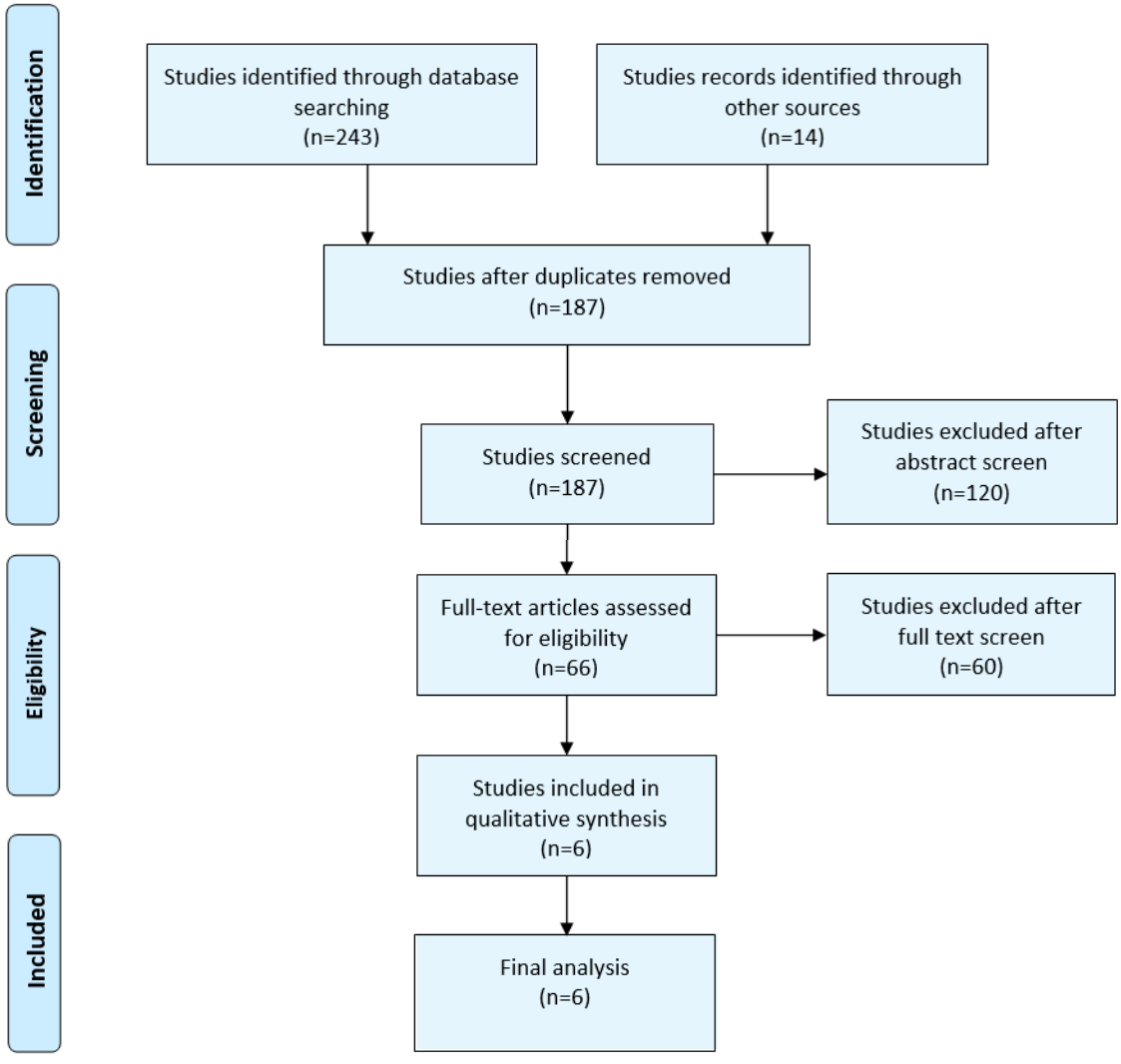


Fig 1. Study selection process.

Fig 4. The random-effects analysis also revealed the proportions of other symptoms, such as retention, which was found in 27.4% of patients (95% CI: 14.5–42.7; Fig 5; 219 subjects), and urgency, reported at 22% (95% CI: 4.0–48.9; Fig 6; 141 subjects), as summarized in Table 4.

Additionally, we included studies that examined bladder dysfunction in GBS patients through urodynamic tests, as outlined in Table 5. Three studies reported cases of detrusor underactivity (n=26/39) and overactivity (n=14/39). Furthermore, one study noted detrusor sphincter dyssynergia (n=6/23) and contractile issues (n=5/23).

DISCUSSION

The analysis included a total of 275 patients, of whom 42.1% exhibited signs of bladder dysfunction. Previous studies have reported a prevalence of dysautonomia as high as 67%, with bladder dysfunction observed in 11% to 30% of cases.[9] Additionally, a past cohort study indicated that more than half of subjects with GBS experienced urinary dysfunction, which can significantly impact daily living.[9,10]

The variability in reported prevalence among studies may be attributed to the diagnostic challenges faced, particularly in acute settings. The diagnostic criteria for GBS specify that bladder dysfunction is a diagnosis of exclusion, likely because it is essential to distinguish GBS patients from those with acute myelopathy, where urinary retention is frequently present.[9] However, the identification of upper motor neuron signs might aid in differentiating these conditions.[4,11,12] Bladder dysfunction occurring at symptom onset, coupled with a deficit in sensory level, can complicate the diagnosis of GBS and suggest acute myelopathy. Additionally, the presence of lost reflexes or areflexia may help differentiate these conditions during physical examinations. While unusual, sensory level changes and bladder dysfunction may manifest as initial symptoms of GBS.[12]

TABLE 1. Studies included in systematic review.

No	Author	Study Type	Country	Study Group	Median (range) or Mean (SD) Age	Follow-Up Duration	Neurogenic Bladder Diagnosis	Bladder Symptoms and/ or % of Total	Key Findings/ Summary	Study Quality Level
1	Sakakibara et al, 2009[4]	Cross Sectional	Japan	65 GBS patients	41 (13-81)	Not applicable	Urodynamic Study	Urinary symptoms observed in 27.7% patient: • Urinary retention (9.2%) • Voiding difficulty (24.6%) • Urinary urgency (7.7%)	Urinary symptoms more common in woman (37.5%), older than 35 years (31.4%) and, AIDP type of GBS (39%).	4
2	Amatya et al, 2013[5]	Prospective Cohort	Australia patients	66 GBS	55.6 (18.1)	3-6 weeks	Questionnaire	• Incomplete emptying (19.7%) • Frequency (33.3%) • Intermittency (27.3%) • Urgency (39.4%) • Weak stream (33.3%) • Straining (12.1%) • Nocturia (59.1%)	• Urinary symptoms interfered with daily life in 49% of cases, while 10.6% reported a negative impact on quality of life (QoL). • A significant relationship was observed between bladder symptoms and both the level of urogenital distress ($p < 0.001$) and the impact of urinary problems ($p < 0.001$). Higher scores on bladder symptom scales were also significantly correlated with psychological distress, functional limitations, and reduced participation.	2

TABLE 1. Studies included in systematic review. (Continue)

No	Author	Study Type	Country	Study Group	Median (range) or Mean (SD) Age	Follow-Up Duration	Neurogenic Bladder Diagnosis	Bladder Symptoms and/ or % of Total	Key Findings/ Summary	Study Quality Level
3	Wheeler et al, 1984[9]	Cross Sectional	Massachusetts	7 GBS patients	51 (32-60)	Not applicable	Urodynamic Study	• Detrusor areflexia with appropriate sphincter relaxation (57.1%) • Detrusor hyperreflexia (42.9%)	• Urological changes began after the neurologic changes • Urodynamic evaluation was performed at average 9 weeks (range 3-24 weeks) of voiding dysfunction	4
4	Sakikibara et al, 1997[10]	Retrospective study	Japan	28 GBS patients	37 (8-69)	Not applicable	Urodynamic Study	• Micturition disturbance in 7 of 28 patient (25%) • The major symptoms of micturition disturbance were: - Voiding difficulty (6 of 7 patient) - Transient urinary retention (3 of 7 patient) - Nocturnal urinary frequency (3 of 7 patient) - Urinary urgency (3 of 7 patient) - Diurnal urinary frequency (2 of 7 patient) - Urge urinary incontinence (2 of 7 patient) - Stress incontinence (2 of 7 patient)	Micturition symptoms emerged following the onset of weakness and gradually improved in parallel with other neurological signs.	3

TABLE 1. Studies included in systematic review. (Continue)

No	Author	Study Type	Country	Study Group	Median (range) or Mean (SD) Age	Follow-Up Duration	Neurogenic Bladder Diagnosis	Bladder Symptoms and/ or % of Total	Key Findings/ Summary	Study Quality Level
5	Naphade et al, 2012[11]	Prospective Cohort	India	38 GBS patients	28 (9-65)	2 months	Urodynamic Study	Urodynamic abnormality (60.53%) - Detrusor underactivity (65%) - Detrusor sphincter dyssynergia (26%) - Detrusor overactivity (13%)	• Urodynamic abnormalities were more commonly found in axonal variant of GBS • Patients with urodynamic abnormality significantly had more disability based on Hughes motor grade. • Patients with normal urodynamic at baseline, continued to show normal results after 2 months of follow-up	2
6	Chakraborty et al, 2020[12]	Case control	United States	171 GBS patients	55 (16)	Not applicable	Clinical Findings	• Urinary Retention, 17 (23.9%)	• The study included 171 patients with GBS, of which 71 (38%) exhibited signs of dysautonomia. • 72% of patients had demyelinating type and 36% of this patients had dysautonomia.	3

TABLE 2. Overall proportion of neurogenic bladder in patients with GBS.

Studies	Sample size	Prevalence (%)	95% CI	Weight (%)		I^2	p value
				Fixed	Random		
Sakakibara 2009	65	27.7	17.3 to 40.2	27.1	25.3		
Naphade 2012	38	60.5	43.4 to 76	15.9	23.8		
Amatya 2013	66	59.1	40.3 to 71	27.5	25.3		
Chakraborty 2020	71	23.9	14.6 to 35.5	29.5	25.5		
Total (random effects)	240	42.1	23.9 to 61.6	100	100	89.56 %	< 0,0001

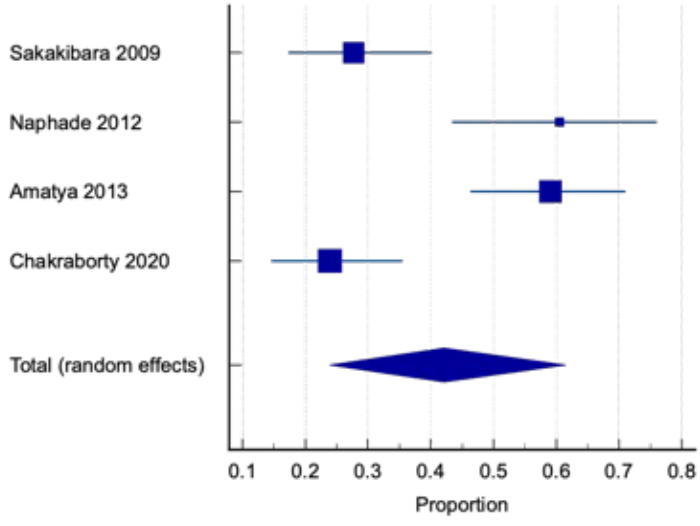


Fig 2. Forest plot demonstrating the pooled prevalence (%) of neurogenic bladder in patients with GBS across included studies, with corresponding 95% confidence intervals.

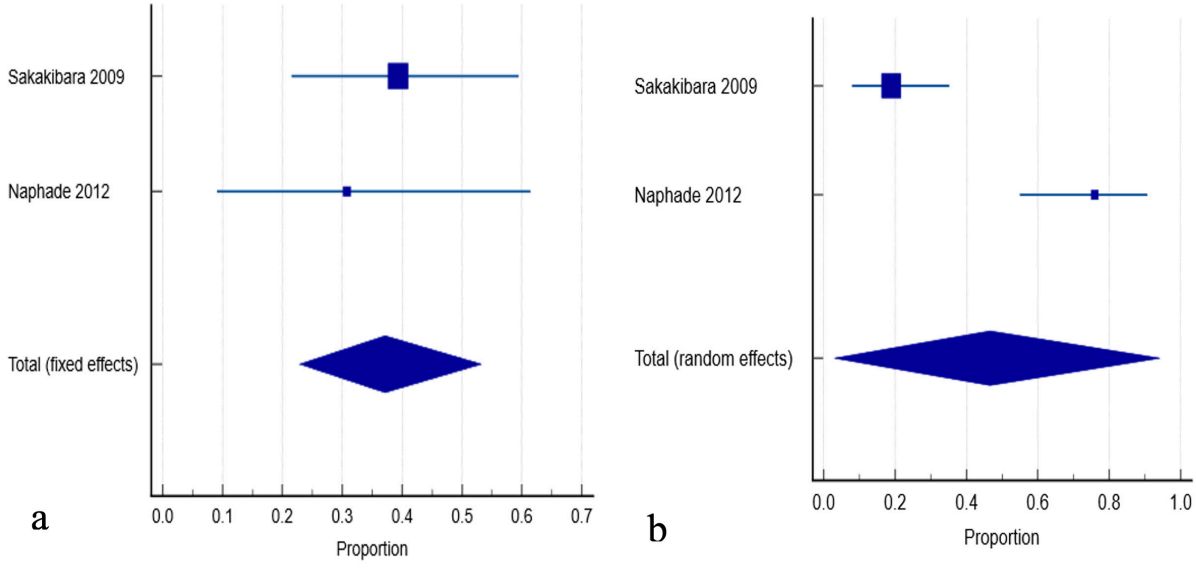


Fig 3. Forest plot demonstrating proportion of neurogenic bladder in patients with (a) Acute Demyelinating Pattern (AIDP) and (b) Acute Axonal Pattern (AMSAN) across included studies, with corresponding 95% confidence intervals.

TABLE 3. Results of meta-analysis of neurogenic bladder proportion based on GBS patterns.

Symptoms	Number of studies	Pooled sample size	Pooled prevalence (%)	95% CI	I^2	p value
Acute Demyelinating Pattern (AIDP)	2	41	37.2	22.9 to 53.2	0%	0,627
Acute Axonal Pattern (AMAN/AMSAN)	2	62	46.6	3.05 to 94.2	95,31%	< 0,0001

TABLE 4. Results of meta-analysis of proportion based on bladder symptoms in patients with GBS.

Symptoms	Number of studies	Pooled sample size	Pooled prevalence (%)	95% CI	I^2	p value
Retention	5	219	27.4	14.5 to 42.7	79,27%	0,0007
Voiding difficulty	3	141	28.7	21.5 to 36.8	0%	0,4905
Urgency	3	141	22.0	4.0 to 48.9	89,99%	< 0,0001

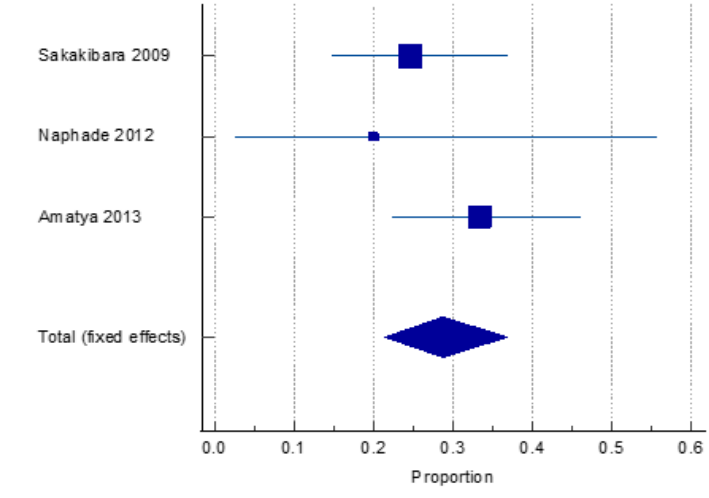


Fig 4. Forest plot demonstrating the proportion estimates (%) of voiding difficulty in patients with GBS across included studies, with corresponding 95% confidence intervals.

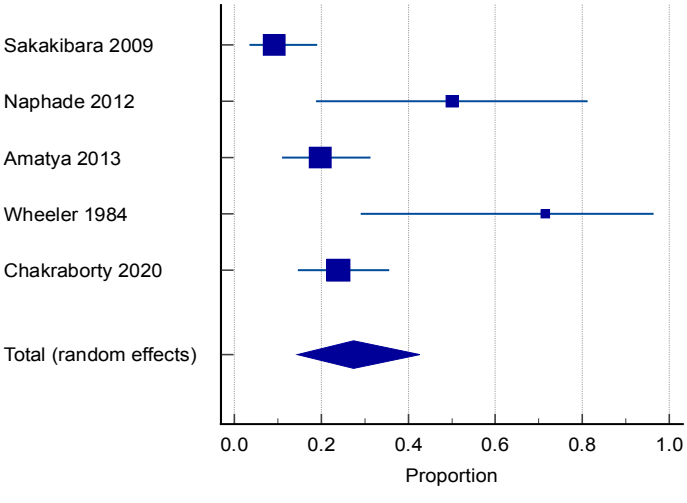


Fig 5. Forest plot demonstrating the proportion estimates (%) of urinary retention in patients with GBS across included studies, with corresponding 95% confidence intervals.

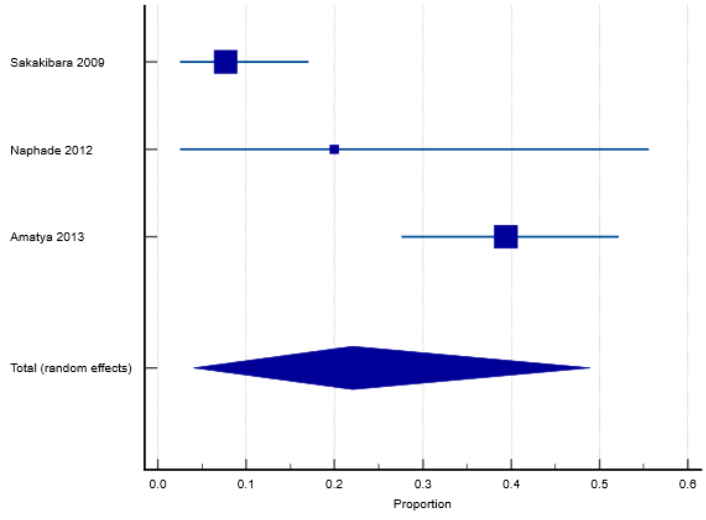


Fig 6. Forest plot demonstrating the proportion estimates (%) of urinary urgency in patients with GBS across included studies, with corresponding 95% confidence intervals.

TABLE 5. Proportion of bladder dysfunction in patients with GBS based on urodynamic findings.

Studies	Detrusor Underactivity (n)	Detrusor Sphincter Dyssynergia (n)	A contractile (n)	Overactivity (n)	Total sample (n)
Wheeler 1984	7	NR	NR	8	9
Sakakibara 2009	15	6	5	3	23
Naphade 2012	4	0	0	3	7

Studies have investigated the autonomic mechanisms involved in bowel and bladder function in GBS. Urodynamic evaluations in GBS patients with acute urinary retention have revealed that internal urethral sphincter obstruction caused by hyperactive sympathetic nerve activity is the primary underlying mechanism, rather than bladder paralysis due to parasympathetic failure. Other potential mechanisms include non-relaxation of the urethral sphincter during voiding (sphincter dysfunction), reduced bladder sensation from afferent fibers damage in the bladder wall or possible overdistension injuries in the early phase of urinary retention, along with detrusor overactivity.[12,13]

In our GBS patients, the most common urinary symptoms were voiding difficulties (24.6%) and urinary retention (9.2%). Previous case series of GBS have also reported a predominance of voiding symptoms. The lower urinary tract alternates between storage and voiding functions through specific autonomic innervation. Consequently, neuropathic involvement of the bladder may contribute to both voiding and storage symptoms.[14]

Functional evaluations such as uroflowmetry, post-void residual measurement, and urodynamic studies are essential to detect detrusor underactivity in patients who present with suspected neurogenic bladder, including patient with GBS.[15] However, data on urodynamic studies performed during the acute phase for participants is quite limited. In a recent study, these assessments were conducted within 8 weeks following the onset of the disease. Bladder underactivity, specifically areflexia, is more prevalent than overactivity. An underactive detrusor muscle is commonly observed in cases of peripheral neuropathy, suggesting a lesion in the parasympathetic postganglionic cholinergic nerves. The inflammation or immune dysfunction affecting the lumbosacral autonomic fibers may provide insight into the underlying causes.[9,11,16]

This review has some limitations. As this review excluded studies not published in English, there is a potential for language bias, and relevant evidence published in other languages may have been overlooked. Additionally, most of the studies included in this review were inherent to its retrospective design, thus the manifestation of urinary dysfunction was made based on medical records making it difficult to track the trajectory of the disease. Prevalence of urinary dysfunction may have been underreported because milder forms might have gone untreated, the use of urinary catheterization may have occurred before urinary retention became evident especially due to pre-existing comorbidities, and medications may have contributed to their occurrence. In a 6-year cohort study, reported that autonomic dysfunction was a predictor of mechanical ventilation use in severe GBS.[17]

In this review, subgroup analysis was not performed since neurogenic bladder was not the primary outcome of most of included studies, therefore only prevalence estimates of bladder dysfunction could be extracted. Most of the studies also had limited duration of follow up, therefore the prognosis of GBS patients with urinary dysfunction was not concluded. Until today, the optimal treatment of dysautonomia in which urinary dysfunction in GBS has not been well studied. The treatment goals include achieving social continence, alleviating symptoms, promoting consistent and complete bladder emptying at suitable intervals, preventing infections, and preserving renal function. Further well-designed prospective cohort studies with a longer follow-up period and standardized urodynamic testing are required to support the finding of this study. This review highlights the need for increased clinical awareness and may contribute to the development of practical approaches for screening and managing bladder dysfunction in patients with GBS.

CONCLUSIONS

Our study emphasizes that urinary dysfunction is a common manifestation of autonomic dysfunction in hospitalized patients admitted for GBS. Early recognition and management can prevent further deterioration of the disease and improve functional outcomes.

Data Availability Statement

The data supporting the findings of this review are available within the article.

ACKNOWLEDGEMENT

None.

DECLARATIONS

Grants and Funding Information

This research did not receive any external funding.

Conflict of Interest

The authors declare no competing interests.

Registration Number of Clinical Trial

PROSPERO CRD420251018021.

Available from https://www.crd.york.ac.uk/PROSPERO/view/CRD420251018021.

Author Contributions

V.O., D.M., D.P., K.A.; conceptualization. V.O., D.M.; methodology and formal analysis: V.O., D.P.; Original draft of the manuscript. K.A.; visualization. V.O., D.M.; validation, review and editing of the manuscript. All authors read and approved the final manuscript.

Use of Artificial Intelligence

Not applicable.

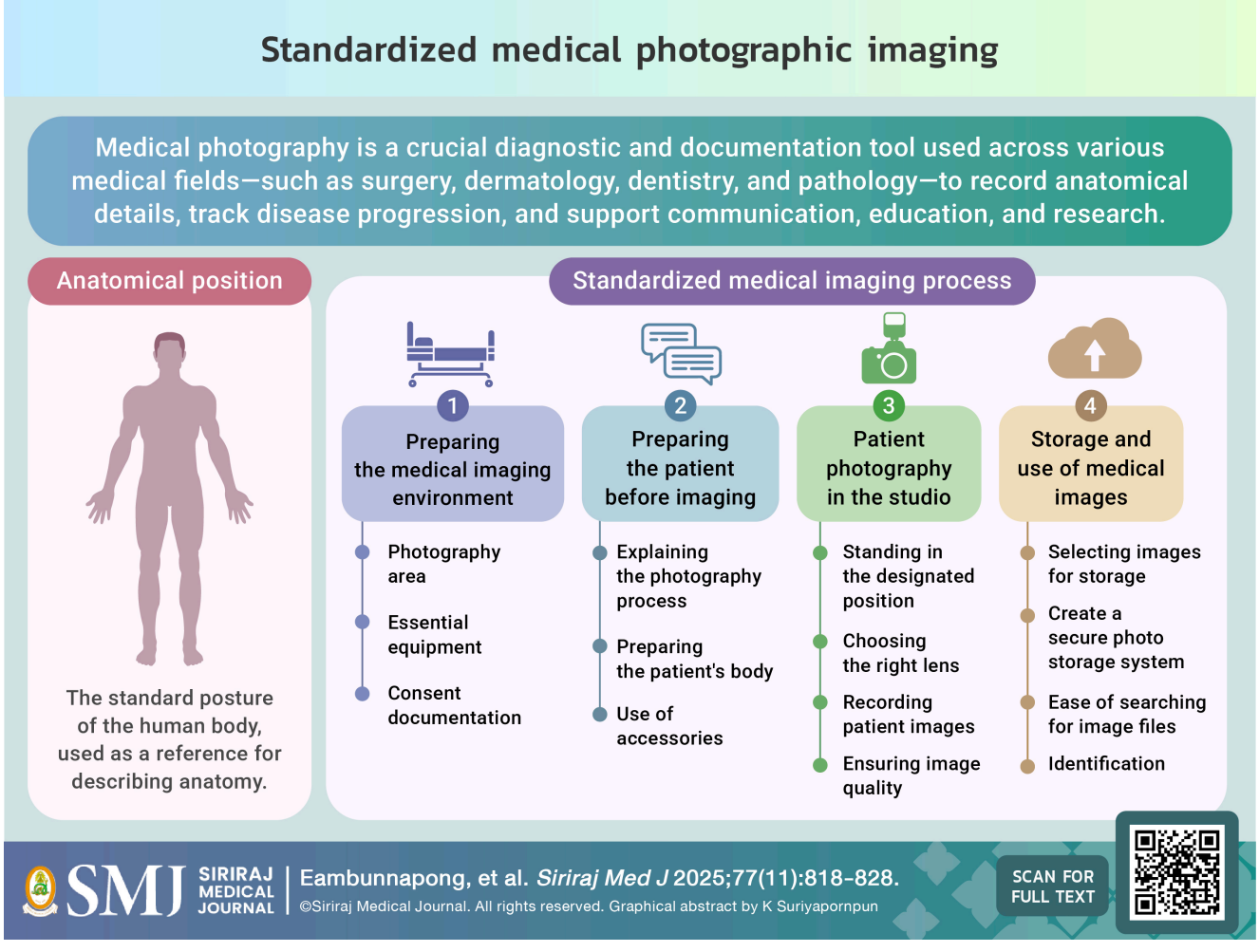
REFERENCES

1. Zaeem Z, Siddiqi ZA, Zochodne DW. Autonomic involvement in Guillain-Barré syndrome: an update. Clin Auton Res. 2019;29(3):289-99.
2. Burakgazi AZ, Alsowaity B, Burakgazi ZA, Unal D, Kelly JJ. Bladder dysfunction in peripheral neuropathies. Muscle Nerve. 2012;45(1):2-8.
3. Khoo CS, Ali AH, Remli R, Tan HJ. A case of Guillain-Barré syndrome (GBS) presenting with acute urinary retention and T6 sensory level. Clin Med (Lond). 2018;18(4):308-10.
4. Sakakibara R, Uchiyama T, Kuwabara S, Mori M, Ito T, Yamamoto T, Awa Y, Yamaguchi C, Yuki N, Vernino S, Kishi M, Shirai K. Prevalence and mechanism of bladder dysfunction in Guillain-Barré Syndrome. Neurourol Urodyn. 2009;28(5):432-7.
5. Amatya B, Khan F, Whishaw M, Pallant JF. Guillain-Barré syndrome: prevalence and long-term factors impacting bladder function in an Australian community cohort. J Clin Neurol. 2013;9(3):144-50.
6. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA- P) 2015 statement. Syst Rev. 2015;4(1):1.
7. OCEBM Levels of Evidence Working Group. The Oxford Levels of Evidence 2. Oxford: Oxford Centre for Evidence-Based Medicine; 2011.
8. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60.
9. Wheeler JS Jr, Siroky MB, Pavlakis A, Krane RJ. The urodynamic aspects of the Guillain-Barré syndrome. J Urol. 1984;131(5):917-9.
10. Sakakibara R, Hattori T, Kuwabara S, Yamanishi T, Yasuda K. Micturitional disturbance in patients with Guillain-Barré syndrome. J Neurol Neurosurg Psychiatry. 1997;63(5):649-53.
11. Naphade PU, Verma R, Garg RK, Singh M, Malhotra HS, Shankwar SN. Prevalence of bladder dysfunction, urodynamic findings, and their correlation with outcome in Guillain-Barre syndrome. Neurourol Urodyn. 2012;31(7):1135-40.
12. Chakraborty T, Kramer CL, Wijdicks EFM, Rabinstein AA. Dysautonomia in Guillain-Barré Syndrome: Prevalence, Clinical Spectrum, and Outcomes. Neurocrit Care. 2020;32(1):113-20.
13. Singh NK, Jaiswal AK, Misra S, Srivastava PK. Assessment of autonomic dysfunction in Guillain– Barre syndrome and its prognostic implications. Acta Neurol Scand. 1987;75(2):101–5.
14. Sakakibara R, Uchiyama T, Tamura N, Kuwabara S, Asahina M, Hattori T. Urinary retention and sympathetic sphincter obstruction in axonal Guillain-Barré syndrome. Muscle Nerve. 2007;35(1):111–5.
15. Ramart P, Ackerman AL. Recurrent Urinary Tract Infection in Women from a Urologist's Perspective. Siriraj Med J. 2023;75(1):55-61.
16. Tagaloa E, Venter F, Liang L, Bhaika J, Aguirre D, Patel J, et al. A Rare Case of Guillain-Barré Syndrome with Severe Pandysautonomia. J Investig Med High Impact Case Rep. 2021;9:23247096211019558.
17. Wen P, Wang L, Liu H, Gong L, Ji H, Wu H, et al. Risk factors for the severity of Guillain-Barré syndrome and predictors of short-term prognosis of severe Guillain-Barré syndrome. Sci Rep. 2021;11(1):11578.

Standardized Medical Photographic Imaging of Patients: A Proposed Guideline

Kawitsara Eambunnapong, Ph.D.[1], Wisit Sombatthawornkun, M.Sc.[2,*], Prachyanun Nilsook, Ph.D.[3]

[1]Department of Orthopaedic Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [2]Vejnitatphattana School, Medical Education Technology Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [3]Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Thailand.



*Corresponding author: Wisit Sombatthawornkun
E-mail: wisit.som@mahidol.ac.th
Received 14 September 2025 Revised 10 October 2025 Accepted 10 October 2025
ORCID ID:http://orcid.org/0009-0004-9750-7285
https://doi.org/10.33192/smj.v77i11.277671

ABSTRACT

Medical photographic imaging plays an essential role in the diagnosis, treatment planning, and follow-up care of patients. Medical photographic imaging also plays a vital role in supporting academic/instructional and research documentation and materials. Therefore, to ensure high-quality, reliable, and consistent medical photographic imaging, standardized guidelines for photographing medical patients are needed. Based on a synthesis of relevant published documents, we hypothesized that medical photographic imaging standards can be divided into the following three key domains: 1) characteristics of; 2) essential elements of; and, 3) standardized process of producing good medical photographic imaging. Establishing and consolidating these standards into a medical photographic imaging guideline for medical clinicians, researchers, and educators will improve the usefulness and consistency of photographic communication of medical patient information. Accordingly, the aim of this study was to establish a medical photographic imaging guideline for use by any subset involved in human patient healthcare, including clinicians, medical imaging specialists, researchers, teachers/students, human healthcare manufacturers/publishers, and government health policy organizations. Moreover, acceptance and establishment of the proposed medical photographic imaging guideline will facilitate and promote improved communication of medical information internationally, making comparisons of treatment and research outcomes more accurate, consistent, and systematic. The proposed medical photographic imaging guideline represents an important advancement in medical knowledge, including how medical photographic imaging is learned, used, and communicated in clinical practice.

Keywords: Standardized medical photographic imaging; medical patients; proposed guideline (Siriraj Med J 2025; 77: 818-828)

INTRODUCTION

Photography is characterized as the art and science of capturing light to create an image that captures the essence of a moment using both technical tools and the creative expression of the photographer.[1] Photography originated in ancient times, but was first used in medicine in 1852.[1,2] With the advancement of imaging technology, photography has become an essential diagnostic tool for medical doctors across various specialties and subspecialties.[1,3-12] In addition, medical photography is being increasingly applied in medical education.[5,7,9-11,13,14] Medical photographs, whether taken in a medical or accident/crime setting, capture specific patient characteristics that have helped to shape clinical standards of practice.[15] Medical photographs also help to document a patient's journey from admission to recovery.[10,11,16] However, a significant challenge is the limited availability of professional medical photographers.[17] It is, therefore, necessary to provide healthcare personnel with knowledge/training in standard medical photographic imaging techniques so they can produce higher quality medical photographic images, which, as shown by related research, yield significant benefits for both medical personnel and patients.

Medical photographic imaging

Medical photography is an important diagnostic tool that provides detailed anatomical information for physicians to perceive and analyze.[1,3-5,7,9,11,15,17-20] It is widely applied across various fields for documenting anatomical abnormalities and tracking disease progression. Examples include dentistry.[4,8] plastic and cosmetic surgery[2,5,18,21-23], plastic surgery[1,2,11,15,24-27], veterinary medicine[8], dermatology for skin lesions[2,3,6,8,10,16,24,28-33], emergency care such as open fractures[17], forensic medicine[11,13], anatomy, pathology, osteopathology[2,7,24] and in patients with cleft lip and palate.[18] Medical photography also enhances communication between doctors[1,15,17,24,34], serves as a valuable tool in medical education[7,17-20,34], and plays an important role in academic presentations and publications.[5]

Characteristics of good medical photographic imaging

The essential characteristics of high-quality medical photography include: 1. Clarity & Resolution: Images must be sharp and detailed to minimize misinterpretation[1,3,15,17,21,35]; 2. Correct Scale & Magnification: Standardized magnification allows accurate comparisons before and after treatment, ensuring consistency across cases[15,21,32]; 3. Consistency & Reproducibility: The camera must always be held at the same distance from the object to facilitate comparison of data or results[5,15,21,29,35]; 4. Proper Lighting & Exposure: Uniform and appropriate lighting must be maintained[5,21,29,32,35]; 5. Correct Positioning & Framing: Patients and anatomical structures should be consistently positioned[21,32,35]; 6. Minimizing Artifacts: Non-medical elements must be excluded.[5,15,36] This includes covering the patient's eyes with black bars to conceal their identity[7,33] and avoiding

jewelry, glasses, cosmetics, clothes, accessories, or unkempt hair[5,18,35]; and 7. Compliance with International Standards: Medical photography must be done after obtaining written informed consent.[1,5-7,9,11,23,24,29,35-37] (Fig 1)

Important elements in medical photographic imaging

The key components of medical photography include: 1. Medical photographer[4,11,18,35], 2. Imaging environment[11,18,21,22,28,35,38], 3. Equipment for image recording[4,5,7,8,10,11,17,18,21,22,24,26-29,31,33,35,38] 4. Standardized angles and image techniques[2,11,15,18,21,25,28,35,38,39] 5. Standardized lighting[4,5,10,11,13, 15,21,22,26-28,35], 6. Patient preparation and positioning[4,11,21,28,35], 7. Software[6,8,10,11,24,29,31,33,38], 8. Imaging phase[11,13,18], 9. Stance[18,27,35], 10. Background[2,10,11,18,19,25,26,28,35,38], 11. Accessories[4,11,28,35], 12. Patient Consent[1,5-7,11,19,23,24,28-30,32-37,39], and 13. Dissemination of medical photographs[6,10,19,23] (Fig 2).

Medical photographer

In many countries, medical photographs are taken by professionals skilled in positioning patients with various medical conditions to capture the most accurate images. The characteristics of good medical photography therefore depends largely on the skill of the photographer.[11,17] Anonymous photographs must be taken without compromising image quality.[1] Key characteristics of a good medical photographer include: 1. Medical Knowledge: A basic understanding of human anatomy, medical photography perspectives, medical terminology, and image documentation to communicate the meaning correctly; 2. Specialized Photography Skills: Proficiency in controlling light, shadow and composition, and the ability to produce images suitable for clinical analysis; 3. Technological Proficiency: Skilled use of cameras, accessories, image-editing software, and database management systems. Training in proper patient imaging and secure data storage is essential; 4. Ethics: Ability to respect patient privacy and maintain confidentiality, ensuring that images are not altered in ways that misrepresent medical information. In some sensitive medical conditions requiring patients to remove clothing, their privacy must be respected and concerns acknowledged; 5. Communication Skills: Effective collaboration with physicians, nurses, researchers, and patients, translating medical needs into clear and easy-to-understand images; 6. Creativity: Ability to compose and present images in a visually effective and meaningful way; and 7. Time Management: Capacity to work under tight timelines and plan photoshoots both in the studio and at various hospitals.[1,8,11,17] Since technology evolves rapidly, medical photographers must be able to continually update their skills and knowledge.

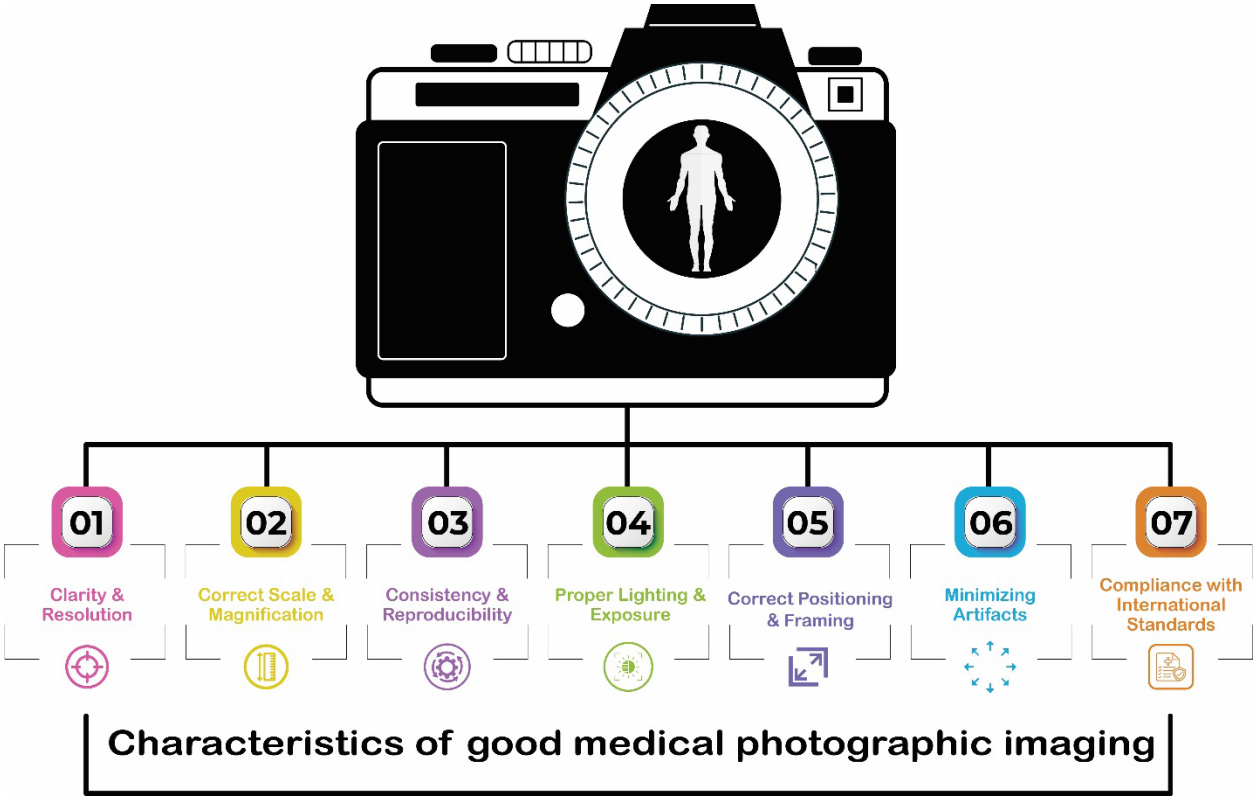


Fig 1. Characteristics of good medical photographic imaging.

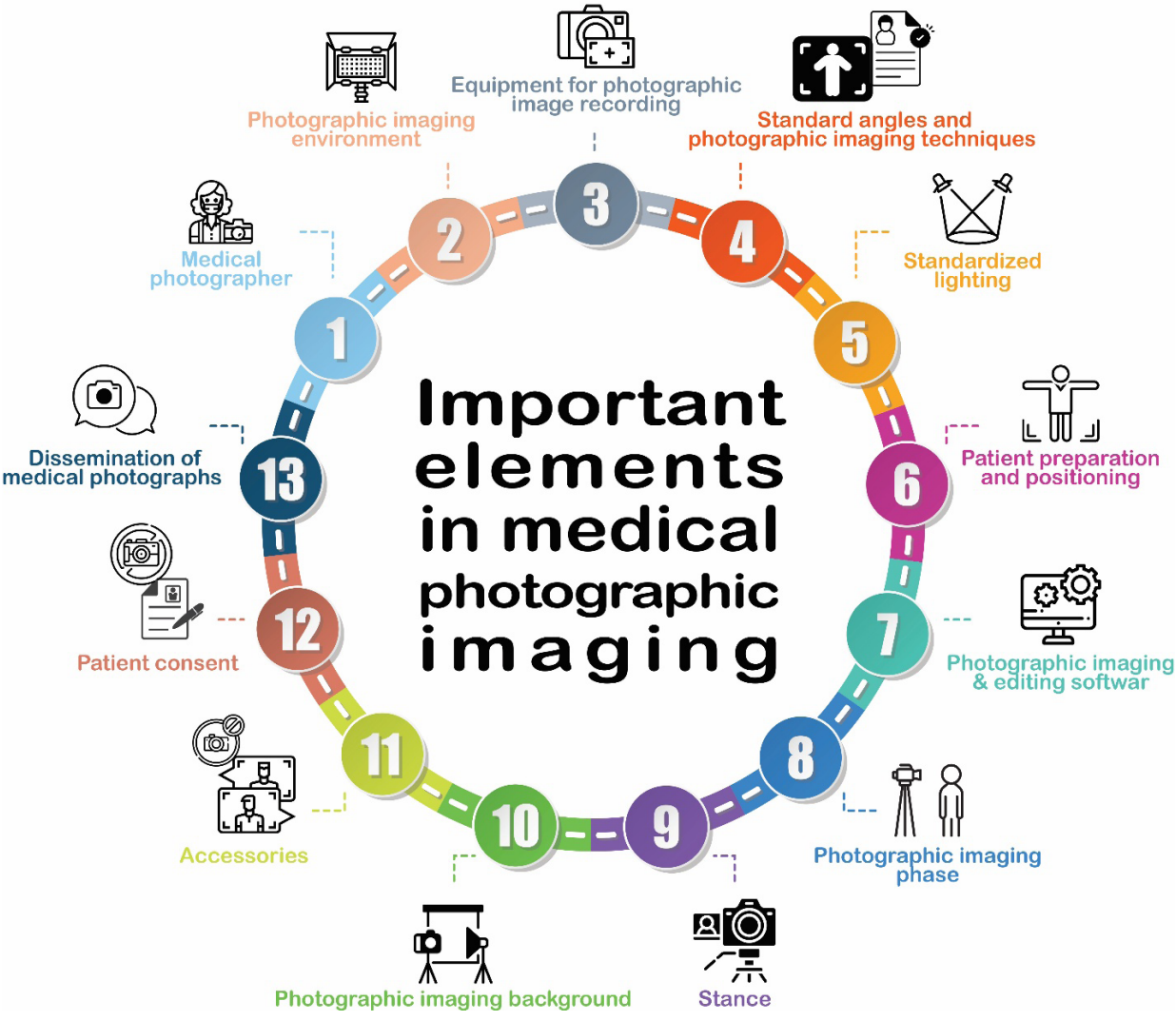


Fig 2. Important elements in medical photographic imaging.

Photographic imaging environment

Medical photography should be conducted in a dedicated, quiet area that ensures privacy and comfort, so that the angle and direction of photographs can be adjusted.[18,26] Ideally, the area should be divided into three sections: Image Display Area: For immediate access and reviews; Dressing Room: To provide patient comfort and privacy; and Photography Studio: A sufficiently large space equipped with controlled studio lighting.[11] Having a fixed location for medical photography with the same background helps maintain image consistency.[2,21,32] If windows are present, curtains should be drawn to block external light.[11] Establishing a permanent and well-equipped photography environment enhances is therefore an important part of achieving high-quality results and minimizing required time to achieve acceptable results.

Equipment for photographic image recording

The equipment used to record medical images must be reliable since there is only one opportunity to record a lesion. When patients undergo intraoperative procedures, images cannot be retaken because lesions will change, making pre-treatment and intraoperative images critical for documenting progression and treatment outcomes.[8,10,13] Equipment selection should emphasize ease of use.[40] Based on a review of relevant literature, the equipment used for recording medical images includes the following:

Camera

Digital single-lens reflex cameras are the most widely used by professional photographers in medical imaging due to their ability to review, store, upload, or delete images as they are taken.[4,9-11,13,18,22,26, 27,32,35,40] Cameras equipped with overhead flash reduce glare and reflection.[28] A good understanding of lens aperture and shutter speed is necessary to adjust image quality and troubleshoot errors when something goes wrong.[21] Image quality depends on the quality of the camera device.[17] Essential features of medical cameras include the ability to

change lenses and adjust settings that control aperture size and shutter speed, through-the-lens focusing, metering system, exposure settings, and compatibility with external flashes.[13,32] The amount of detail a camera can capture is called its resolution and measured in pixels. Most modern digital cameras feature higher resolutions and allow for large, detailed images without loss of clarity.[28] In wound photography, high-resolution cameras provide precise visualization of wound margins and enable enlargement while maintaining clarity.[40]

Lens

The lens is one of the most important components of medical photography equipment.[8,9,11,18,27] A high-resolution lens with an appropriate focal length for the area being photographed is essential. A high depth of field is recommended to ensure that the entire area of interest, or larger anatomical structures, remain in focus. Variable focal length (zoom) lenses should be avoided, as they make it difficult to standardize clinical photography.[11,13] Lens selection is therefore critical in clinical documentation.[8,9,11,22,26] The focal length should always match the clinical application and the size of the area being photographed.[11] For consistency and accurate comparison, it is important to use the same lens when capturing the same view.

Memory card

All digital cameras store images on memory cards, which are available in various types. The specific type of memory card is less important than ensuring sufficient storage capacity and keeping multiple spare cards available.[8]

Battery

At least one or two fully charged rechargeable battery packs should always be available along with a reliable charger.[8]

Tripod

A tripod is a valuable accessory in medical photography.[8,28] It allows precise control of camera angles[2,22,35], reduces image blur when photographing without flash[28,35], and helps maintain consistency in serial image documentation, such as pre- and post-surgical comparisons.[22,35] With the growing use of mobile phones in medical photography, additional accessories such as memory cards, batteries, and tripods, are also recommended to ensure image quality and consistency.

Standardized angles and photographic imaging techniques

Accurate patient positioning is essential for clear comparison between preoperative and postoperative images.[18] Anatomical landmarks should be aligned in standardized positions to record patient symptoms in a reproducible manner.[2,5,15,21] Six standard views are commonly used in medical imaging (Fig 3): 1) frontal and posterior views, 2) right and left oblique views, and 3) right and left side angles.[9,26] These standardized views must be applied consistently before and after surgery. Postoperative photographs are typically taken about one year following surgery once healing has occurred and lesions have diminished. Maintaining the same preoperative standards is essential to achieve the best possible results.

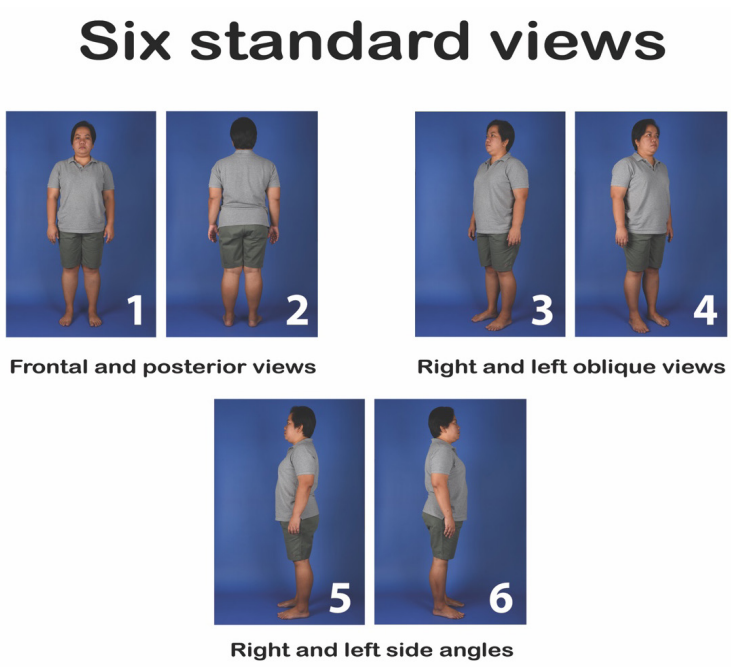


Fig 3. Six standard views are commonly used in medical imaging.

Standard lighting

Natural light is unsuitable for medical photography because it varies by time of day and weather conditions. Instead, controlled artificial lighting should be used.[9-11] The aim is to capture clinical symptoms, not cosmetic images, so accessories such as taillights or reflectors are unnecessary.[18,28] Therefore, standardized lighting is essential to minimize shadows, provide adequate depth of field[9,11,13,35], and ensure consistent image quality.[5,9,11,21,35] The light source should never be placed behind the subject.[28] For portraits, illumination from the same direction allows reliable comparison of features, expressions, and treatment outcomes.[5,9,15,35] The recommended setup involves two electronic light sources positioned 1 to 1.5 meters from the patient at 45-degree angles. Once established, lighting should not be altered, as changes lead to inconsistencies in image quality.

Patient preparation and positioning

Informed consent must be obtained before any formal photography, particularly if the images will be published or displayed. The purpose of the photograph must be clearly explained to the patient.[5,9,11,18,32] Patients should remove all distracting elements, such as visible rings, watches, and bracelets, when arms or legs are photographed.[2,5,21,22] Female patients with long hair should tie it back so the face, forehead and both ears are visible, using accessories such as headbands, clips, or elastic hair bands.[5,9,11,32] Glasses, visible earrings, and nose bridges should be avoided. Simple makeup is acceptable, but lipstick and foundation on the nose or lips should be avoided, as they may conceal scars or relevant features.[5,11,18,30] In dermatology, makeup should be avoided altogether to reveal skin irregularities and blemishes that inform diagnosis and treatment.[9,11,22,32] When photographing the scalp, excess oil may cause glare, so the hair should be wiped with tissue to reduce shine before imaging.[28] Proper patient positioning is crucial for standardization across different views, though it is often difficult to control and a common cause of non-standardized images.[9]

Photographic imaging & editing software

Image editing software is an important tool in medical photography, as it enhances image quality to support clarity and allows for more precise medical analysis. The key aspects are as follows:

Image editing

When photographing patients, if the face is visible, black bars should be placed over the eyes. Tattoos, birthmarks, and other unique features that may reveal a patient's identity should be avoided or edited out.[8,30,33] The face is the most identifiable anatomical feature; scars, melanocytic nevi, other birthmarks, and tattoos may still allow recognition by acquaintances. Image editing helps preserve patient privacy, and identifiable features should be removed as necessary.[6,37] Automatic camera corrections, such as red-eye removal, blemish correction, or cosmetic filters, may obscure clinical findings and must be disabled in medical photography.[10]

Image file format

With the increased use of digital cameras, image file sizes have grown dramatically.[8,10] Each file format has different properties relevant to accuracy and usage: JPEG (Joint Photographic Experts Group): Small file size, suitable for web use and sharing on social media, TIF (Tagged Image File Format): Large file size, high quality, suitable for professional printing, and RAW: Unprocessed files from which brightness, white balance, and sharpness can be adjusted later without loss of quality.[10] The choice of format depends on the medical photographer's expertise and the institution's imaging system.

Access to image data

Storage of patient images must comply with strict security standards to prevent unauthorized access.[1,29,35] Each image should contain information such as patient name, date of photography, patient status, and clinical status (preoperative, intraoperative, or postoperative). For follow-up, postoperative photographs should be taken during routine visits to ensure consistency comparability.[11,35,40]

Storage area

Adequate and secure storage is essential for data retrieval and security.[11,26,29,40] Identifiable tattoos or body art should be separated from clinical records whenever possible.[11,26] Proper storage allows longitudinal tracking of wound healing and disease progression.[17,40] File format guidelines are as follows: For hospital archival, original images should be stored in TIFF or RAW, while JPEG copies may be generated for general use, such as for sending images to doctors.

Photographic imaging phase

A constant distance between the camera and the patient is critical. Images of the same anatomical region should always be captured from the same distance.[11,13,18] Pre-set distances provide consistency which improves readability. Patients should be positioned at least one foot away from the background as an appropriate distance

between the object and background reduces shadows.[9,11,32] Pre- and postoperative images must be taken from equal distances; otherwise, lesions may be missed or inconsistencies introduced.[11,21] If excessive background shadow occurs, increasing the distance slightly beyond 1 foot or more usually resolves it. This principle applies to both full-body and partial body photography.

Stance

Consistent patient positioning is one of the one of the most challenging but essential aspects of clinical photography and requires consistent practice.[18,21] Specific body parts should be framed to emphasize anonymity.[15] Precise patient positioning is essential for professional photographs.[11] Standard positioning is important because it facilitates recognition of abnormal conditions, and annotating the patient's position t on the image can aid in future reference.[10]

Photographic imaging background

The purpose of the background is to isolate the subject.[9,11,13] The ideal background for medical photography is smooth, non-reflective, and sky-blue, which complements skin tones and is easy on the eyes.[2,9,11,13,18,22,25,26,32,35,38] When printed in grayscale, blue appears as neutral gray, preserving detail.[11,35] It also allows for greater depth of field and adjusts shadows without overexposing objects.[9,11] The background color should also be consistent with appropriate contrast.[2,11,26,28] Also, patient clothing and background should remain consistent across sessions.[11,21] In addition to the background, the walls and ceiling should be off-white.[11] Once chosen, background color should not be altered.

Accessories

Useful accessories include measuring tapes and a marker.[28] Additional positioning aids may be employed to help patients maintain consistent posture across multiple views.[35]

Patient consent

The consent process should always be clearly authorized and agreed upon by the patient. Patient privacy and dignity must be never be violated, and patients retain the right to withdraw consent at any time.[1,7,9,11,19,23,24,28,30,32-36,39] Written consent should be obtained prior to any photography, as this provides legal and ethical protection for health professionals and researchers.[6,7,9,11,19,23,28-30,35,39] Compliance with data protection laws and privacy regulations is essential to ensure that healthcare data are collected, stored, and used securely and lawfully.[17] Images may include a hospital card, tag or patient number to facilitate later identification.[28] Personal information in electronic form must be safeguarded by appropriate security measures to prevent unauthorized access. Respect for human autonomy and privacy is a fundamental principle of medical ethics.[1,30,36,39,41] Patients must be informed that photographs are a tool for treatment planning. If photographs are used for educational, academic, exhibition, or publication purposes, additional statements regarding patient confidentiality are required.[6,7,9,35] If photographs are to be used for non-medical purposes, such as marketing, a separate consent form describing the intended use must be provided.[11,35] Patient consent should clearly specify the purpose of the photographs, and patients should have the option to decide how their images may be used.

Dissemination of medical photographs

The dissemination of medical photographs across platforms including print, websites, personal devices, television, and social media must always protect patient privacy and rights. Written informed consent remains the practice.[23,28-30] Consent forms must explicitly address issues related to patient identity, as public disclosure requires particular care.[23] Therefore, protecting patient data is of utmost importance.[17] The increasing availability of the internet and digital publishing has contributed to an increase in the number of photographs in circulation, underscoring the need for careful ethical consideration in the dissemination of medical photographs.[19]

Standardized medical imaging process

The creation of standardized medical photographs requires adherence to strict tools, technologies, and protocols to ensure high-quality images. The process can be divided into four main steps. (Fig 4)

1. Preparing the medical imaging environment

The imaging environment consists of:

1.1 Photography area – The area should be large enough to accommodate all photographic equipment. If patients must undress, a fully enclosed dressing room should be available. A display area should also be provided so patients can review their images.

1.2 Essential equipment – Includes camera, lens, memory card, battery, studio light, backdrop and tripod, all of which should be prepared in advance.

1.3 Consent documentation – Patients must sign informed consent forms to protect their rights and privacy. Establishing a permanent medical photography environment reduces workflow time and improves

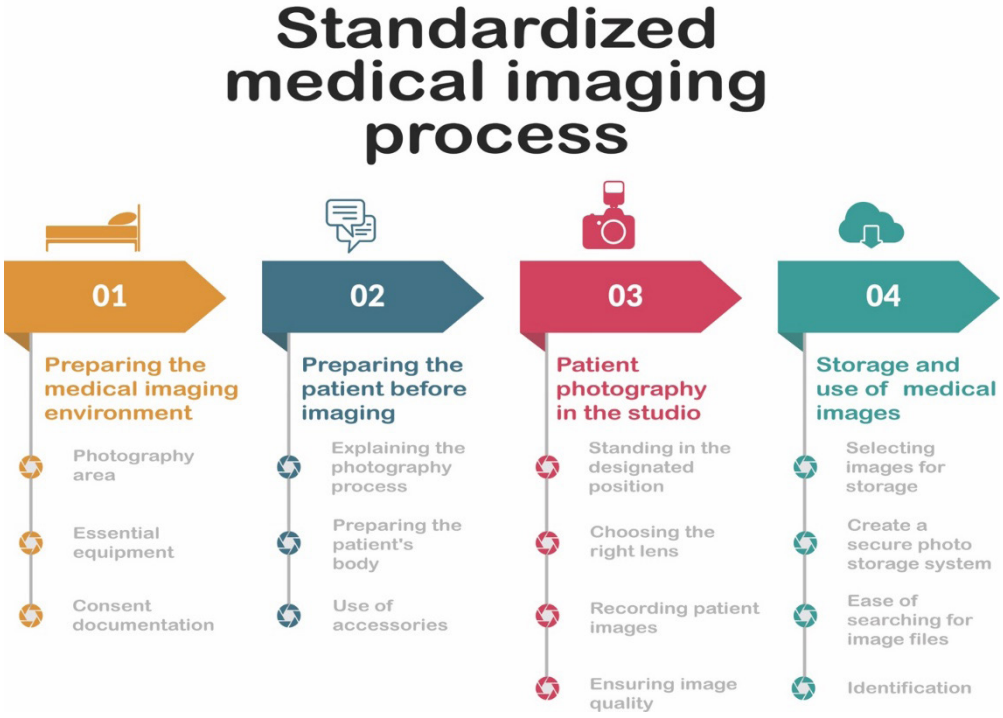


Fig 4. Standardized medical imaging process.

consistency. Importantly, patients retain the right to withdraw consent at any time.

2. Preparing the patient before imaging

Patient preparation is a critical step before beginning any imaging procedure. The preparation process includes the following steps:

2.1 Explaining the photography process: Medical photographers should carefully review and understand the doctor's request for photography r to ensure accuracy. If the patient agrees to being photographed, they must sign a consent form to protect their rights and privacy before any imaging begins.

2.2 Preparing the patient's body: Preparation varies depending on the type of disease being documented. Some procedures may require the patient to remove certain clothing items or all clothing. The photographers should clearly explain the purpose of the images, demonstrate the required poses and angles, and assess the patient's readiness beforehand. Careful planning is essential, as clinical images are cannot be taken again.

2.3 Use of accessories: Preparation for female patients often takes longer than male patients. For example, distracting jewelry must be removed, and long hair should be adjusted if it obscures the neck or ears, particularly when photographing the face. Photographers must always keep extra equipment available. After the session, the medical photographer should review the images to ensure they meet professional standards.

3. Patient photography in the studio

For every patient photoshoot conducted in a studio, a relative or an assistant photographer should remain nearby to help prevent accidents. If the patient experiences any adverse symptoms, assistance can be provided immediately assisted. The patient photoshoot process is as follows:

3.1 Standing in the designated position: The patient should be positioned exactly as instructed by the photographer. All necessary equipment must be prepared in advance. For example, a seated photo may be taken first, followed by a standing full-length image. Any repositioning, such as moving a chair to take a full-length standing photo, should be handled by the photographer's assistant. A tripod is essential to maintain consistent positioning and to reproduce image accurately. Standardized anatomical positioning must be followed: the patient should stand upright with the head and eyes facing forward, arms resting close to the body with palms facing forward, and feet slightly apart to distribute body weight evenly.

3.2 Choosing the right lens: The same lens type should be used across different angles to maintain consistency in distance and perspective. This ensures comparability of images when monitoring disease progression.

3.3 Recording patient images: Clinical photoshoots are not intended just for aesthetics; their primary purpose is to document conditions clearly and realistically. Lighting must remain consistent in direction and intensity. Patient

positioning should follow the doctor's instructions and safety must always be the main concern.

3.4 Ensuring image quality: Image quality depends heavily on the skill on the medical photographer's skill. Clinical photographs must be realistic, not distorted, and show signs of disease, such as skin color, shallow or deep wounds. Most importantly, the images must be sharp to help doctors make accurate diagnoses.

4. Storage and use of medical images

After recording, images should be carefully selected and stored in a secure, standardized, and easily accessible system. When editing is necessary, identifying features should be removed as necessary. Each photo should include key information such as the patient's name, doctor's name, date of imaging, and patient status before, during, or after surgery. For patients requiring follow-up, postoperative photographs should be taken in coordination with scheduled visits to ensure consistency and comparability. Access to medical photographs must be restricted to physicians or authorized personnel only. Properly organized and secure image storage not only safeguards patient data but also enhances efficiency in patient care and reduces risk of data loss.

CONCLUSION

Medical imaging plays a vital role in diagnosis, treatment planning, treatment follow-up, and in academic or research papers. To achieve high-quality accurate, and reliable images, clear and standardized guidelines are essential. From the synthesis of related documents, standardized medical photography can be summarized into three key components: 1. Characteristics of high-quality medical imaging: This includes clarity and resolution, correct scale and magnification, consistency and reproducibility, proper lighting and exposure, accurate positioning and framing, minimization of artifacts, and compliance with medical guidelines. 2. Essential elements of medical photography: These encompass the role of trained medical photographers, the imaging environment, recording equipment, standardized angles and techniques, standardized lighting, patient preparation and positioning, supporting software, imaging phases and stance, background, use of accessories, patient consent, and appropriate dissemination of images. 3. Standardized Medical Imaging Process: This involves preparing the imaging environment, preparing the patient before imaging, following a structured studio photography process, and ensuring secure and systematic storage and use of medical image files. By integrating these three elements, medical imaging can achieve the highest standards of quality. This not only enhances diagnostic accuracy and treatment effectiveness but also provides a strong foundation for developing reliable medical image databases in line with international standards in the digital age.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to all advisors and experts in medical photography for their valuable guidance and constructive suggestions throughout the development of the proposed standardized guidelines for medical photographic imaging of patients. Appreciation is also extended to the affiliated institutions for their support in providing essential data and facilitating the research process, as well as to the medical professionals and specialists who contributed their insights and experiences, which were instrumental in shaping this proposal. Finally, the authors wish to thank all volunteer patients for their kind cooperation and trust in participating in this study, which aims to enhance the quality and ethical standards of medical photographic imaging in Thailand.

DECLARATIONS

Grants and Funding Information

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

Conflict of Interest

The authors declare that they have no conflicts of interest.

Registration number of clinical trial

This is an academic article with no human trials.

Author Contributions

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

Use of Artificial Intelligence

The authors declare no use of artificial intelligence.

REFERENCES

1. Lau CK, Schumacher HH, Irwin MS. Patients' perception of medical photography. J Plast Reconstr Aesthet Surg. 2010;63(6): e507-11.
2. Uzun M, Bülbül M, Toker S, Beksaç B, Kara A. Medical

photography: principles for orthopedics. J Orthop Surg Res. 2014;9:23.
3. Ali Z, Sørensen JA, Zhang DG, Ghazanfar MN, Allerup JAC, Maurer M, et al. Smartphone Photographs of Chronic Urticaria Taken by Patients Are of Good Quality and Useful in the Clinic. Dermatology. 2024;240(3):357-61.
4. Ferreira CdAP, Pereira FdAV, Rodrigues MVB, Cabral JLdOA, de Campos Tuña IT. Art and science of dental photography: suggested photographic protocol with cellular device. Observatório De La Economía Latinoamericana. 2024;22(8):e6167-e.
5. Henderson JL, Larrabee WF, Krieger BD. Photographic Standards for Facial Plastic Surgery. Arch Facial Plast Surg. 2005;7(5):331-3.
6. Lakdawala N, Fontanella D, Grant-Kels JM. Ethical considerations in dermatologic photography. Clin Dermatol. 2012;30(5):486-91.
7. Macintosh T. Ethical considerations for clinical photography in the global South. Dev World Bioeth. 2006;6(2):81-8.
8. Mutalik S. Digital clinical photography: practical tips. J Cutan Aesthet Surg. 2010;3(1):48-51.
9. Swamy RS, Most SP. Pre- and postoperative portrait photography: standardized photos for various procedures. Facial Plast Surg Clin North Am. 2010;18(2):245-52, Table of Contents.
10. Towbin AJ, Roth CJ, Bronkalla M, Cram D. Workflow Challenges of Enterprise Imaging: HIMSS-SIIM Collaborative White Paper. J Digit Imaging. 2016;29(5):574-82.
11. Yavuzer R, Smirnes S, Jackson IT. Guidelines for standard photography in plastic surgery. Ann Plast Surg. 2001;46(3):293-300.
12. Jirakulsawat A, Noocharoen N, Srinimit B, Anakmeteeprugsa S. Assessment on knowledge and satisfaction level of delirium video for education in geriatric patients undergoing elective noncardiac surgery. Siriraj Med J. 2024;76(7):415-21.
13. Barut C, Ertilav H. Guidelines for standard photography in gross and clinical anatomy. Anat Sci Educ. 2011;4(6):348-56.
14. Bunyaratavej S, Jirawattanadon P, Sereeaphinan C, Wongdama S, Sombatmaithai S, Leeyaphan C. Mobile Device Digital Photography for Teledermatology Consultation: Real-Life Situations. Siriraj Med J. 2023;75(12):871-9.
15. Han Lee J, Vercler CJ. Foucault and Clinical Photography: An Analysis and Mitigation of the Clinical Gaze in Pediatric Craniofacial Plastic Surgery. FACE. 2024;5(2):292-9.
16. Daltveit S, Kleppe L, Petterteig MO, Moi AL. Photographs in burn patient diaries: A qualitative study of patients' and nurses' experiences. Intensive Crit Care Nurs. 2024;82:103619.
17. Radotra I, Azimi DY, Maamoun W. The use of smartphone-application based medical photography for open fractures: a national survey of orthoplastic affiliated major trauma centres in England. Injury. 2022;53(6):2028-34.
18. Chen X, Wang Z, Song T. Basic and Informative Photographic Standards for Cleft Lip Patients. J Craniofac Surg. 2024;35(7):e608-e12.
19. Devakumar D, Brotherton H, Halbert J, Clarke A, Prost A, Hall J. Taking ethical photos of children for medical and research purposes in low-resource settings: an exploratory qualitative study. BMC Medical Ethics. 2013;14(1):27.
20. Nguyen TT, VanderWalde L, Bellavance E, Eisenhauer T, Hieken T, Johnson N, et al. Ethical Considerations of Medical Photography in the Management of Breast Disease. Ann Surg Oncol. 2018;25(10):2801-6.
21. Hagan KF. Clinical Photography for the Plastic Surgery Practice—The Basics. Plast Surg Nurs. 2008;28(4):188-92.
22. Parsa S, Basagaoglu B, Mackley K, Aitson P, Kenkel J, Amirlak B. Current and Future Photography Techniques in Aesthetic Surgery. Aesthet Surg J Open Forum. 2022;4:ojab050.
23. Patmon D, Sandhu H, Girotto J, Ford R. Legal Ramifications of Publishing Patient Photographs: A Review of Legal Cases. Plast Reconstr Surg Glob Open. 2023;11(8):e5162.
24. Cura M, Alves H, Andrade JP. Medical photography usage amongst doctors at a Portuguese hospital. Int J Environ Res Public Health. 2022;19(12):7304.
25. Joumblat NR, Chim J, Aguirre Sanchez PG, Bedolla E, Salgado CJ. Guidelines for the standardization of genital photography. Aesthet Surg J. 2018;38(10):1124-30.
26. Thornton SM, Attaluri PK, Wirth PJ, Shaffrey EC, George RE, Carbullido MK, et al. Picture Perfect: Standardizing and Safekeeping Clinical Photography in Plastic Surgery. Aesthet Surg J Open Forum. 2024;6:ojae012.
27. Wong MS, Vinyard WJ. Photographic standards for the massive weight loss patient. Ann Plast Surg. 2014;73:S82-S7.
28. Ashique K, Kaliyadan F. Clinical photography for trichology practice: Tips and tricks. Int J Trichology. 2011;3(1):7-13.
29. Bradshaw LM, Gergar ME, Holko GA. Collaboration in wound photography competency development: a unique approach. Adv Skin Wound Care. 2011;24(2):85-92.
30. Creighton S, Alderson J, Brown S, Minto C. Medical photography: ethics, consent and the intersex patient. BJU Int. 2002;89(1):67-71.
31. AlSuhaymi S, Alghubaywi F, AlHarthi R, Al AlSheikh A, AlJasser MI. Perception and acceptance of medical photography among Arab dermatology patients: a cross-sectional analysis. Dermatol Reports. 2024;16(3):9925.
32. Nair AG, Santhanam A. Clinical photography for periorbital and facial aesthetic practice. J Cutan Aesthet Surg. 2016;9(2):115-21.
33. Rimoin L, Haberle S, Aspey LD, Grant-Kels JM, Stoff B. Informed consent, use, and storage of digital photography among Mohs surgeons in the United States. Dermatol Surg. 2016;42(3):305-9.
34. Mansouri ABA, Charfi R, Litaiem N, Chakroun A, Mzoughi K, Zakhama L, et al. A Smartphone use in Medical Photography. International Journal of Research and Scientific Innovation. 2024;10(12):744-7.
35. DiBernardo BE, Adams LR, Krause J, Fiorillo MA, Gheradini G. Photographic Standards in Plastic Surgery. Plast Reconstr Surg. 1998;102(2):559-68.
36. Rahman A, Abdus S. Clinical Photography of Patients in the Era of Social Media: An Ethical Issue. Pakistan Journal of Ethics. 2024;4(1):43-6.
37. Nettrour JF, Burch MB, Bal BS. Patients, pictures, and privacy: managing clinical photographs in the smartphone era. Arthroplast Today. 2019;5(1):57-60.
38. Golinko MS, Abou-Ghanem O, Simard M, Caudill M, Pontell ME, O'Sick N, et al. ImageAssist: A Novel Smartphone Application for Standardized Clinical Photography That Implements and Automates American Society of Plastic Surgery/Plastic Surgery Foundation Photographic Guidelines With Background Deletion. Ann Plast Surg. 2024;92(4):367-72.
39. Petersilge CA, McDonald J, Bishop M, Yudkovitch L, Treuting C, Towbin AJ. Visible Light Imaging: Clinical Aspects with

an Emphasis on Medical Photography-a HIMSS-SIIM Enterprise Imaging Community Whitepaper. J Digit Imaging. 2022;35(3): 385-95.
40. Rennert R, Golinko M, Kaplan D, Flattau A, Brem H. Standardization of wound photography using the Wound Electronic Medical Record. Adv Skin Wound Care. 2009;22(1):32-8.
41. Troy A, Keane K, Cullivan O, Khalid M, Afridi I, Gerraghty K, et al. AB010. SOH24AB_137. The use of smartphones for medical photography in Irish urology: convenient or concerning? Mesentery and Peritoneum. 2024;8.