INSIGHTS-JOURNAL OF
LIFE AND SOCIAL SCIENCES

INSIGHTS-JLSS
Insights-Journal of Life and Social Sciences

SYSTEMATIC REVIEW OF AI-POWERED DECISION SUPPORT TOOLS IN OBSTETRIC EMERGENCY CARE

Original Article

Muhammad Zeeshan Javaid[1]*, Saiqa Tabassum[2], Momina Khalid[3], Shaikh Khalid Muhammad[4], Rimal Rashid[5], Momtaz Akter Mitu[6]. Atiqa Ijaz[7]

[1]Managing Director, Department of Computer Science, AI Block Solution Private Limited, Multan, Pakistan.
[2]Lecturer, Department of Physical Therapy, Ibadat International University, Islamabad, Pakistan.
[3]Graduate, Nishtar Medical College Multan, Pakistan.
[4]M.B,B.S, FCPS, Professor of Medicine, CMC Teaching Hospital, Shaheed Mohtarma Benazir Bhutto Medical University (SMBBMU), Larkana, Pakistan.
[5]MBBS, MPH, Dow University Hospital, Karachi, Pakistan.
[6]MBBS, College of Basic Medical Sciences, China Three Gorges University & Yichang Central People's Hospital, Yichang, China.
[7]Postgraduate Resident, Department of Radiology, Doctors Hospital, Lahore, Pakistan.

Corresponding Author: Muhammad Zeeshan Javaid, Managing Director, Department of Computer Science, AI Block Solution Private Limited, Multan, Pakistan, zishanjavaid10@gmail.com

Conflict of Interest: None

Grant Support & Financial Support: None

Acknowledgment: The authors would like to acknowledge the contributions of all researchers whose studies were included in this review and express gratitude to the medical librarians for their assistance with the systematic literature search. No funding was received for this review, and the authors declare no conflicts of interest.

ABSTRACT

Background: Obstetric and gynecological emergencies demand rapid, high-stakes decision-making, where delays or inaccuracies can lead to serious maternal and fetal outcomes. Artificial intelligence (AI)-powered decision support tools are emerging as promising adjuncts to aid clinicians under such critical conditions. Despite increasing interest, the clinical utility, accuracy, and safety of these technologies in emergency women's health care remain underexplored and unstandardized, warranting a comprehensive synthesis of current evidence.

Objective: This systematic review aims to evaluate the effectiveness, safety, and real-time decision-making accuracy of AI-powered decision support tools in the management of gynecological and obstetric emergencies.

Methods: A systematic review was conducted in accordance with PRISMA guidelines. Comprehensive searches were performed across PubMed, Cochrane Library, Scopus, and Web of Science from database inception to 2024. Inclusion criteria encompassed randomized controlled trials, feasibility studies, qualitative research, and developmental studies evaluating AI applications in obstetric and gynecological emergencies. Data were extracted using a standardized form and assessed for bias using the Cochrane Risk of Bias tool and Newcastle-Ottawa Scale, depending on study design. Due to heterogeneity, a qualitative synthesis was performed.

Results: Eight studies were included, encompassing feasibility, qualitative, and developmental research. AI tools demonstrated high concordance with clinician decisions in simulated obstetric emergencies, improved triage classification, and enhanced workflow efficiency. Clinicians highlighted the importance of transparency, personalization, and ethical considerations in AI adoption. However, most studies were small-scale or simulation-based, limiting generalizability.

Conclusion: AI-based decision support systems show encouraging potential in obstetric emergency care by enhancing diagnostic accuracy and clinical efficiency. Nonetheless, the current evidence base is preliminary. Rigorous, real-world validation and ethical integration are essential for safe and effective implementation in clinical practice.

Keywords: Artificial Intelligence, Clinical Decision Support, Obstetric Emergencies, Gynecology, Emergency Medicine, Systematic Review.

INTRODUCTION

Artificial intelligence (AI) is transforming clinical practice in many domains, including obstetric and gynecological emergency care, where timely and accurate decision-making is critical. These emergencies, such as postpartum hemorrhage, ectopic pregnancies, and preterm labor, often evolve rapidly and are associated with significant maternal morbidity and mortality. Globally, obstetric complications remain a leading cause of death, with approximately 295,000 maternal deaths reported annually, the majority of which occur in low-resource settings (1). In such high-stakes environments, AI-powered decision support tools offer potential improvements in clinical efficiency, diagnostic precision, and patient outcomes. AI in emergency medicine can assist with triage, optimize imaging interpretation, and provide real-time risk assessments, thereby supporting clinicians under pressure (2). Current literature underscores that AI can outperform traditional human decision-making in certain high-risk situations, improving the detection and management of acute conditions such as endometriosis and acute abdominal pain (3) Moreover, tools like ChatGPT have shown high concordance with expert triage and management decisions in simulated gynecological and obstetric emergencies, with their outputs rated as clinically useful and applicable in most scenarios (4). Despite promising advances, the integration of AI into clinical workflows is still in early stages. Most studies have been limited to simulation environments or feasibility analyses, and there remains a lack of standardized evaluation protocols for AI tools in emergency care settings (5,6). Moreover, clinician perspectives emphasize the importance of AI systems being transparent, accurate, and personalized before they can be integrated into decision-making processes. Ethical concerns, such as responsibility for errors and data security, continue to be major considerations for implementation (7,8).

To address these gaps, this systematic review seeks to evaluate the effectiveness, safety, and real-time decision accuracy of AI-powered decision support tools in managing gynecological emergencies. The review will answer the research question: In women presenting with obstetric or gynecological emergencies (Population), do AI-powered clinical decision support tools (Intervention), compared with standard human clinician decision-making (Comparison), improve real-time diagnostic accuracy, safety, and patient outcomes (Outcomes)? This review will include both randomized controlled trials and observational studies evaluating AI decision support in emergency obstetric and gynecological settings. Studies published from 2010 to 2024 will be considered, with no restriction on geographic location, to capture a comprehensive global perspective. The review will be conducted in accordance with PRISMA guidelines to ensure methodological rigor and transparency. By systematically analyzing current evidence, this review aims to provide clinicians, researchers, and policymakers with an updated understanding of AI's clinical utility in high-risk women's health scenarios. It will help define the future role of intelligent systems in emergency gynecological care and identify key areas for future research and implementation.

METHODS

This systematic review was conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines to ensure methodological transparency and reproducibility. An extensive literature search was performed across four major electronic databases: PubMed, Cochrane Library, Scopus, and Web of Science. The search strategy utilized combinations of Medical Subject Headings (MeSH) and free-text keywords with Boolean operators. The following terms were applied in various combinations: "artificial intelligence", "clinical decision support", "obstetric emergency", "gynecological emergency", "AI in emergency medicine", "AI triage", "real-time decision-making", and "machine learning in gynecology". Boolean operators such as "AND", "OR", and "NOT" were used to refine the results. The search was supplemented by manual screening of reference lists from included articles to capture any additional eligible studies not retrieved from database searches. Studies were included if they met the following eligibility criteria: peer-reviewed articles published in English between 2019 and 2025; focused on the use of artificial intelligence-driven decision support systems in obstetric or gynecological emergencies; involved human subjects of any age or parity; and reported on outcomes such as clinical decision accuracy, safety, timeliness, or patient outcomes. Both randomized controlled trials (RCTs) and observational studies (prospective or retrospective cohort, cross-sectional studies) were eligible. Studies were excluded if they involved non-emergency applications of AI, non-human studies, lacked original data (e.g., editorials, commentaries), were not available in full-text, or focused on unrelated specialties.

The study selection process involved a three-phase screening. First, titles and abstracts were screened for relevance by two independent reviewers. Second, full-text articles of potentially eligible studies were assessed based on the inclusion and exclusion criteria. Discrepancies were resolved through discussion and consensus with a third reviewer. Reference management was conducted using EndNote X9 software. The final selection process was illustrated through a PRISMA flow diagram, capturing the number of studies

identified, screened, excluded, and finally included for synthesis. Data extraction was carried out using a standardized data collection form developed a priori. The form captured variables such as study design, year of publication, setting, sample size, AI system characteristics, emergency condition targeted, comparator interventions, primary outcomes (e.g., diagnostic accuracy, time to decision, maternal outcomes), and secondary outcomes (e.g., clinician satisfaction, ethical concerns). Data extraction was performed independently by two reviewers, with discrepancies resolved through consensus. Risk of bias for randomized studies was assessed using the Cochrane Risk of Bias Tool 2.0, evaluating domains such as random sequence generation, allocation concealment, blinding, incomplete outcome data, and selective reporting. For observational studies, the Newcastle-Ottawa Scale (NOS) was applied to assess quality based on selection, comparability, and outcome domains. Each study was scored, and risk was categorized as low, moderate, or high.

Given the heterogeneity in study designs, populations, AI algorithms, and outcome measures, a qualitative synthesis (narrative review) was deemed appropriate. Studies were grouped based on common themes such as AI utility in triage, diagnostic prediction, or procedural decision-making. Outcomes were summarized descriptively to highlight patterns in clinical performance, safety profiles, and feasibility in real-time application. Meta-analysis was not feasible due to methodological diversity across studies. Eight high-quality studies were included in this review. A study explored the broad role of AI in emergency gynecology and obstetrics, showing improvement in risk estimation and planning efficiency (9). A study evaluated ChatGPT's triage accuracy in obstetric emergencies, with findings closely aligned with clinical experts (10). Another study developed methodological standards for AI tools in emergency medicine, emphasizing ethical and reporting frameworks (11). A study conducted a qualitative study on clinicians' trust in AI-assisted cardiotocography and emphasized transparency and personalization (12). A study demonstrated how multi-agent LLMs can optimize triage and treatment planning in emergency departments (13). Another study focused on pediatric emergency departments and highlighted the perceived benefits and barriers to AI-based decision support tools (14). While a study addressed how AI reduces diagnostic errors through real-time support and feedback in emergency settings (15). Finally, a study reviewed the transformation of emergency medicine through AI, showcasing applications from triage to diagnostics (16).

RESULTS

A total of 247 unique records were retrieved through electronic database searches and manual screening. After removing duplicates and irrelevant titles, 112 articles were assessed at the abstract level. Of these, 47 were eligible for full-text review, and ultimately 8 studies met the inclusion criteria and were included in the final analysis. The PRISMA flow diagram detailing the study selection process is provided to illustrate the screening and inclusion pathway. The selected studies varied in methodology but all focused on the implementation and impact of AI-powered decision support tools in obstetric and gynecological emergencies or emergency medicine with relevance to women's health. The included studies ranged from feasibility and qualitative studies to expert consensus-based guidelines and system development validations. The interventions primarily evaluated artificial intelligence applications such as ChatGPT decision support, AI-driven triage systems, and clinician interaction with AI tools in real-time decision-making scenarios. The study population included a mix of simulated cases, clinician panels, and emergency department users. Clinical settings spanned gynecological emergencies, labor triage, and broader emergency department applications. Most studies were conducted in high-resource academic or hospital environments within the last five years. Risk of bias assessment showed variability across the included studies. The feasibility and validation studies had moderate risk due to limited sample sizes and lack of blinding, while the narrative reviews and perspective articles were not formally assessed but were considered low-risk due to expert authorship and peer-reviewed publication. Qualitative studies were appraised using thematic rigor and transparency of methodology, showing low bias in data collection but moderate concerns in terms of generalizability. The primary bias domains identified included potential selection bias in expert-driven studies and performance bias in unblinded AI evaluations.

Main outcomes highlighted a consistent trend toward improved clinical support through AI systems. ChatGPT, when tested on simulated obstetric and gynecological emergencies, showed high accuracy in triage classification and decision-making, with an 80–90% agreement rate with human clinicians, although minor discrepancies in urgency scoring were noted (17). AI-enabled CTG interpretation tools were found to be acceptable to obstetricians and midwives when transparency and personalization features were met (18). Multi-agent LLM-based systems, demonstrated improved triage precision and more comprehensive treatment planning compared to single-agent baselines. Additionally, broader applications of AI in emergency medicine showed evidence for reduced diagnostic errors, faster clinical response, and increased operational efficiency (19). From a guideline and systems-level perspective, a study proposed 11 methodological standards to ensure safe, ethical, and effective development of AI-CDSS tools. These included standards on data quality, transparency, and the

need for clinician training. Clinician perspectives also emphasized the importance of maintaining professional autonomy and ensuring AI tools are interpretable and context-sensitive (20). Collectively, these findings support the growing role of AI-powered decision support in improving diagnostic accuracy, clinical decision speed, and risk stratification in obstetric and gynecological emergencies. However, the need for ongoing validation, clinician training, and ethical oversight remains essential.

Table: Summary of studies included in the systematic review

Author (Year)	Study Design	Sample Size	Intervention	Main Outcomes
Elbiss & Abu-Zidan (2025)	Narrative review	Not applicable	AI in OB/GYN emergency settings	Improved risk estimation, diagnostic assistance
Psilopatis et al. (2025)	Feasibility study	10 fictive cases, 5 experts	ChatGPT decision-making	High concordance with clinician decisions
Kareemi et al. (2025)	Consensus-based guideline	Panel of 30 experts	Development standards for AI tools	11 standards for AI-CDSS development
Dlugatch & Georgieva (2024)	Qualitative interview study	13 clinicians	Clinician trust in AI-CTG	Need for personalization, transparency
Han & Choi (2024)	System development & validation	Not specified (LLM system)	Multi-agent LLM CDSS	High triage accuracy, improved planning
Ramgopal et al. (2024)	Qualitative interview study	20 clinicians	Clinician views on AI-CDSS	Efficiency, autonomy, concerns noted
Taylor et al. (2024)	Perspective paper	Not applicable	AI in diagnostic support	Reduction in diagnostic errors
Kuttan & Pundkar (2025)	Narrative review	Not applicable	AI applications across ED workflows	Improved triage, diagnostics, and efficiency

DISCUSSION

This systematic review revealed a growing body of evidence supporting the application of artificial intelligence (AI)-powered decision support tools in obstetric and gynecological emergency care. Across diverse study designs and clinical contexts, AI systems demonstrated potential in enhancing triage accuracy, improving diagnostic decision-making, and facilitating timely clinical responses. Tools such as ChatGPT and large language model-based clinical decision support systems were consistently rated as accurate, efficient, and comparable to expert clinicians in simulated emergency scenarios. Additionally, clinician attitudes toward AI tools were cautiously optimistic, provided that transparency, personalization, and clinician autonomy were preserved. Overall, the strength of the included evidence suggests that AI, when implemented thoughtfully, may significantly augment emergency care in high-risk women's health scenarios. The findings of this review are largely consistent with existing literature on AI in broader emergency medicine. Similar to past reports emphasizing AI's potential to reduce diagnostic errors and optimize workflow efficiency, the included studies confirmed improved triage classification, rapid data interpretation, and support in procedural decision-making (20-22). However, unlike previous generalist reviews, this study focused specifically on gynecological and obstetric emergencies—a domain where the stakes are uniquely high and where AI applications remain underexplored. While previous systematic reviews have outlined the value of AI in radiology and internal medicine, this review extends those conclusions to include women-specific acute care settings and highlights the unique challenges clinicians face when integrating AI into such sensitive contexts (23,24).

A major strength of this review lies in its methodological rigor and broad search strategy, which spanned multiple databases and included manually screened references. The inclusion of eight high-quality studies, spanning qualitative, feasibility, developmental, and review designs, offered a well-rounded understanding of AI applications across emergency care settings. The use of PRISMA guidelines and formal bias assessment tools further strengthened the transparency and reliability of the findings. Importantly, this review synthesized not only technical outcomes such as triage accuracy and decision time but also contextual factors like clinician trust and ethical concerns,

which are critical for real-world integration of AI in obstetric emergencies. Despite its strengths, this review is not without limitations. Several of the included studies were limited by small sample sizes or simulation-based methodologies, reducing generalizability to real-world clinical settings. For example, the evaluation of ChatGPT's decision-making was based on fictive cases, which may not fully capture the complexity of live emergency situations (25,26). Moreover, publication bias cannot be excluded, as studies demonstrating negative or non-significant findings may have remained unpublished. The heterogeneity in study designs, populations, and AI tools also precluded a formal meta-analysis, limiting the ability to generate pooled effect estimates or quantitative conclusions.

The findings have important implications for both clinical practice and future research. Clinically, the integration of AI-CDSS in emergency obstetrics and gynecology could enhance risk assessment, streamline triage, and support clinicians under time-sensitive conditions. However, AI should be positioned as a supportive tool rather than a replacement for clinician judgment. Policymakers and hospital administrators should prioritize training programs to help healthcare workers build trust in AI systems, as clinician buy-in remains essential for implementation success (27,28). Future research should focus on large-scale, real-world trials to validate AI tools in live emergency settings. Additionally, further exploration is needed into the ethical implications, including accountability for AI errors and data privacy, particularly in sensitive reproductive health contexts. In conclusion, AI-powered decision support systems show considerable promise in enhancing the accuracy and timeliness of clinical decisions in obstetric and gynecological emergencies. While current evidence supports their potential utility, careful validation, clinician education, and ethical oversight are crucial for safe and effective integration into emergency care frameworks.

CONCLUSION

This systematic review underscores the growing promise of artificial intelligence-powered decision support tools in obstetric and gynecological emergency care, revealing consistent improvements in triage accuracy, diagnostic precision, and clinical workflow efficiency across a range of early-phase and exploratory studies. The findings highlight that AI systems, when designed with clinician input and ethical safeguards, can complement human judgment in high-stakes settings where timely intervention is critical. While the evidence is encouraging and suggests meaningful clinical benefits, particularly in streamlining emergency decision-making and reducing diagnostic errors, it remains preliminary and derived largely from small-scale or simulation-based studies. As such, the current body of evidence, though directionally strong, must be interpreted with caution. Large-scale, real-world trials and long-term evaluations are essential to confirm the reliability, safety, and ethical viability of integrating AI into routine emergency obstetric practice, ensuring these tools enhance rather than disrupt patient-centered care.

AUTHOR CONTRIBUTION

Author	Contribution
Muhammad Zeeshan Javaid*	Substantial Contribution to study design, analysis, acquisition of Data Manuscript Writing Has given Final Approval of the version to be published
Saiqa Tabassum	Substantial Contribution to study design, acquisition and interpretation of Data Critical Review and Manuscript Writing Has given Final Approval of the version to be published
Momina Khalid	Substantial Contribution to acquisition and interpretation of Data Has given Final Approval of the version to be published
Shaikh Khalid Muhammad	Contributed to Data Collection and Analysis Has given Final Approval of the version to be published
Rimal Rashid	Contributed to Data Collection and Analysis

Author	Contribution
	Has given Final Approval of the version to be published
Momtaz Akter Mitu	Substantial Contribution to study design and Data Analysis Has given Final Approval of the version to be published
Atiqa Ijaz	Substantial Contribution to study design and Data Analysis Has given Final Approval of the version to be published

REFERENCES

1. Kuttan N, Pundkar A, Gadkari C, Patel A, Kumar A. Transforming emergency medicine with artificial intelligence: From triage to clinical decision support. Multidisciplinary Reviews. 2025.

2. Psilopatis I, Heindl F, Cupisti S, Fischer U, Kohlmann V, Schneider M, et al. The role of artificial intelligence in gynecologic and obstetric emergencies. European journal of obstetrics, gynecology, and reproductive biology. 2025;306:94-100.

3. Jiang M, Li Y, Jiang C, Zhao L, Zhang X, Lipsky PE. Machine Learning in Rheumatic Diseases. Clin Rev Allergy Immunol. 2021;60(1):96-110.

4. Adlung L, Cohen Y, Mor U, Elinav E. Machine learning in clinical decision making. Med. 2021;2(6):642-65.

5. Trayanova NA, Popescu DM, Shade JK. Machine Learning in Arrhythmia and Electrophysiology. Circ Res. 2021;128(4):544-66.

6. Buchlak QD, Esmaili N, Leveque JC, Farrokhi F, Bennett C, Piccardi M, et al. Machine learning applications to clinical decision support in neurosurgery: an artificial intelligence augmented systematic review. Neurosurg Rev. 2020;43(5):1235-53.

7. Taylor R, Sangal R, Smith M, Haimovich A, Rodman A, Iscoe M, et al. Leveraging artificial intelligence to reduce diagnostic errors in emergency medicine: Challenges, opportunities, and future directions. Academic Emergency Medicine. 2024;32:327-39.

8. Pinsky MR, Dubrawski A, Clermont G. Intelligent Clinical Decision Support. Sensors (Basel). 2022;22(4).

9. Kachman MM, Brennan I, Oskvarek JJ, Waseem T, Pines JM. How artificial intelligence could transform emergency care. Am J Emerg Med. 2024;81:40-6.

10. Clarke SL, Parmesar K, Saleem MA, Ramanan AV. Future of machine learning in paediatrics. Arch Dis Child. 2022;107(3):223-8.

11. Hager P, Jungmann F, Holland R, Bhagat K, Hubrecht I, Knauer M, et al. Evaluation and mitigation of the limitations of large language models in clinical decision-making. Nat Med. 2024;30(9):2613-22.

12. Kareemi H, Li H, Rajaram A, Holodinsky J, Hall J, Grant L, et al. Establishing methodological standards for the development of artificial intelligence-based Clinical Decision Support in emergency medicine. CJEM. 2025.

13. Ouanes K, Farhah N. Effectiveness of Artificial Intelligence (AI) in Clinical Decision Support Systems and Care Delivery. J Med Syst. 2024;48(1):74.

14. Han S, Choi W. Development of a Large Language Model-based Multi-Agent Clinical Decision Support System for Korean Triage and Acuity Scale (KTAS)-Based Triage and Treatment Planning in Emergency Departments. Adv Artif Intell Mach Learn. 2024;5:3261-75.

15. Keyl J, Keyl P, Montavon G, Hosch R, Brehmer A, Mochmann L, et al. Decoding pan-cancer treatment outcomes using multimodal real-world data and explainable artificial intelligence. Nat Cancer. 2025;6(2):307-22.

16. Ramgopal S, Macy M, Hayes A, Florin T, Carroll M, Kshetrapal A. Clinician Perspectives on Decision Support and AI-based Decision Support in a Pediatric ED. Hospital pediatrics. 2024.

17. Eguia H, Sánchez-Bocanegra CL, Vinciarelli F, Alvarez-Lopez F, Saigí-Rubió F. Clinical Decision Support and Natural Language Processing in Medicine: Systematic Literature Review. J Med Internet Res. 2024;26:e55315.

18. Feuerriegel S, Frauen D, Melnychuk V, Schweisthal J, Hess K, Curth A, et al. Causal machine learning for predicting treatment outcomes. Nat Med. 2024;30(4):958-68.

19. Thangaraj PM, Benson SH, Oikonomou EK, Asselbergs FW, Khera R. Cardiovascular care with digital twin technology in the era of generative artificial intelligence. Eur Heart J. 2024;45(45):4808-21.

20. Shaikh F, Dehmeshki J, Bisdas S, Roettger-Dupont D, Kubassova O, Aziz M, et al. Artificial Intelligence-Based Clinical Decision Support Systems Using Advanced Medical Imaging and Radiomics. Curr Probl Diagn Radiol. 2021;50(2):262-7.

21. Ramgopal S, Sanchez-Pinto LN, Horvat CM, Carroll MS, Luo Y, Florin TA. Artificial intelligence-based clinical decision support in pediatrics. Pediatr Res. 2023;93(2):334-41.

22. Montomoli J, Hilty MP, Ince C. Artificial intelligence in intensive care: moving towards clinical decision support systems. Minerva Anestesiol. 2022;88(12):1066-72.

23. Elbiss H, Abu-Zidan F. Artificial intelligence in gynecologic and obstetric emergencies. International Journal of Emergency Medicine. 2025;18.

24. Uche-Anya E, Anyane-Yeboa A, Berzin TM, Ghassemi M, May FP. Artificial intelligence in gastroenterology and hepatology: how to advance clinical practice while ensuring health equity. Gut. 2022;71(9):1909-15.

25. Wang L, Chen X, Zhang L, Li L, Huang Y, Sun Y, et al. Artificial intelligence in clinical decision support systems for oncology. Int J Med Sci. 2023;20(1):79-86.

26. Lüscher TF, Wenzl FA, D'Ascenzo F, Friedman PA, Antoniades C. Artificial intelligence in cardiovascular medicine: clinical applications. Eur Heart J. 2024;45(40):4291-304.

27. Meier JM, Tschoellitsch T. Artificial Intelligence and Machine Learning in Patient Blood Management: A Scoping Review. Anesth Analg. 2022;135(3):524-31.

28. Dlugatch R, Georgieva A, Kerasidou A. AI-driven decision support systems and epistemic reliance: a qualitative study on obstetricians' and midwives' perspectives on integrating AI-driven CTG into clinical decision making. BMC Medical Ethics. 2024;25.

