



Multimodal Artificial Intelligence for Early Prediction of Preeclampsia Using Maternal Clinical, Biochemical, and Ultrasound Data: A Systematic Review

Azadeh Shiralizadeh¹, Nasrin Sanaei nasab^{2*}

¹MD, Obstetrician and Gynecologist Surgeon, Tehran, Iran

²MD, Obstetrician and Gynecologist Surgeon, Tehran, Iran

Article info

Received: 20.07.2026

Accepted: 03.10.2026

Available Online: 05.10.2026

Checked for Plagiarism: Yes

Keywords:

preeclampsia, multimodal artificial intelligence, machine learning, early prediction, ultrasound biomarkers

ABSTRACT

Background: Preeclampsia remains a leading cause of maternal and perinatal morbidity and mortality worldwide, affecting 2-8% of pregnancies globally. Early prediction is critical for implementing prophylactic interventions, yet traditional risk models demonstrate limited predictive accuracy. Multimodal artificial intelligence (AI) approaches integrating clinical, biochemical, and ultrasound data offer a promising avenue for improving early detection.

Objective: This systematic review synthesizes evidence on multimodal AI models for early prediction of preeclampsia, evaluating model architectures, data integration strategies, predictive performance, and barriers to clinical translation.

Methods: Following PRISMA 2020 guidelines, a comprehensive search of PubMed, Scopus, Embase, IEEE Xplore, and Web of Science conducted through July 2025. Studies were eligible if they developed or validated AI models integrating at least two data modalities (clinical, biochemical, ultrasound, or other imaging) for preeclampsia prediction. Risk of bias assessed using PROBAST.

Results: From 2,457 identified records, 18 studies met inclusion criteria. Multimodal AI models demonstrated area under the curve (AUC) values ranging from 0.78 to 0.98, consistently outperforming unimodal approaches and traditional logistic regression. The highest-performing models integrated maternal characteristics, mean arterial pressure, placental growth factor (PIGF), and uterine artery Doppler indices. Deep learning architectures, particularly convolutional neural networks for ultrasound analysis and ensemble methods for structured clinical data, showed superior performance. However, only 22.2% of studies performed external validation, and substantial heterogeneity in predictor definitions, outcome ascertainment, and validation strategies limited comparability. SHAP analysis revealed that PIGF, mean uterine artery pulsatility index, and maternal history contributed most strongly to predictions.

Conclusions: Multimodal AI demonstrates significant potential for early preeclampsia prediction, with performance exceeding traditional screening. However, limited external validation, inconsistent reporting, and underutilization of ultrasound data constrain clinical translation. Prospective multicenter validation and standardized reporting frameworks are essential before routine implementation.

Introduction

Preeclampsia constitutes one of the most significant challenges in modern obstetrics, affecting approximately 2-8% of pregnancies worldwide and contributing to over 50,000 maternal deaths and 500,000 fetal deaths annually [1-3]. The condition,

characterized by new-onset hypertension and proteinuria or maternal organ dysfunction after 20 weeks of gestation, encompasses a spectrum of severity ranging from mild disease to severe early-onset presentations before 34 weeks [4-6]. Despite decades of research, the only definitive treatment

remains delivery of the placenta, making early identification of at-risk pregnancies essential for implementing prophylactic strategies and optimizing surveillance intensity [7-9].

The pathophysiology of preeclampsia involves abnormal placental implantation, resulting in impaired spiral artery remodeling, placental ischemia, and subsequent release of anti-angiogenic factors into the maternal circulation [10-13]. This molecular understanding has translated into the identification of numerous candidate biomarkers, particularly placental growth factor (PlGF) and soluble fms-like tyrosine kinase-1 (sFlt-1), whose ratio has demonstrated clinical utility in disease triage and prediction [14-17]. Simultaneously, uterine artery Doppler ultrasound, which assesses downstream impedance to flow reflecting impaired trophoblastic invasion, has emerged as a robust biophysical predictor of preeclampsia when integrated into multivariable models [18-21]. The combination of maternal risk factors, mean arterial pressure, PlGF, and uterine artery pulsatility index constitutes the foundation of contemporary first-trimester screening algorithms [22-30].

Despite these advances, traditional statistical prediction models demonstrate significant limitations. The Fetal Medicine Foundation algorithm, widely considered the reference standard, achieves detection rates of approximately 90% for early-onset preeclampsia at a 10% false positive rate when all components are available [31-35]. However, performance declines substantially for term preeclampsia, and the model's reliance on precise ultrasound measurements and biomarker assays limits implementation in resource-constrained settings. Moreover, these models assume linear relationships between predictors and outcomes, potentially missing complex interactions that characterize the heterogeneous preeclampsia syndrome [36-38].

Artificial intelligence has emerged as a transformative approach capable of addressing these limitations. Machine learning algorithms, including random forests, support vector machines, and gradient boosting, can automatically learn nonlinear relationships from high-dimensional data without prespecified functional forms. Deep learning architectures, particularly convolutional neural networks, excel at extracting features from medical imaging data, with demonstrated success in obstetric ultrasound segmentation and fetal anomaly detection [39-41]. These capabilities are particularly relevant for preeclampsia, where integrating heterogeneous data streams maternal clinical characteristics, circulating biomarkers, and ultrasound-derived placental and hemodynamic parameters may capture complementary pathophysiological information.

The concept of multimodal AI, which integrates and jointly analyzes data from multiple sources, represents a natural extension of this approach. Early pregnancy assessment already routinely collects clinical history, blood pressure, serum biochemistry, and ultrasound measurements, creating a rich multimodal dataset for each pregnancy. Traditional prediction models combine these modalities through predetermined formulas or regression coefficients, whereas multimodal AI can learn optimal integration strategies directly from data. This approach has demonstrated success in other medical domains, including oncology and cardiology, but remains incompletely characterized in obstetric care.

Current evidence on AI for preeclampsia prediction exists as a fragmented landscape of studies varying widely in methodology, predictor selection, algorithm choice, and validation strategy. Systematic reviews have begun to synthesize this literature, with scoping reviews identifying 15 studies of deep learning models and broader reviews examining AI applications across multiple pregnancy complications. These reviews consistently note that imaging data, particularly ultrasound, remains underutilized despite its established importance in preeclampsia pathophysiology and screening. The most recent systematic reviews include studies through mid-2025, yet the rapid pace of methodological development in multimodal AI warrants ongoing synthesis. This systematic review addresses the critical need for focused evidence synthesis on multimodal AI specifically for preeclampsia prediction. Unlike prior reviews that have examined deep learning for early-onset disease or AI across multiple pregnancy complications, this review's primary objective is to characterize how multimodal integration combining clinical, biochemical, and ultrasound data affects predictive performance. The secondary aims include: (1) evaluating the methodological quality and risk of bias in included studies, (2) characterizing the specific data modalities and integration strategies employed, (3) assessing the interpretability of multimodal models through explainability techniques, and (4) identifying barriers to clinical translation. By focusing specifically on multimodal approaches, this review provides a foundation for future research and development of clinically actionable prediction tools.

Literature Review

The literature on preeclampsia prediction has evolved through distinct phases, beginning with maternal risk factor-based models, progressing to biophysical and biochemical marker integration, and most recently encompassing artificial intelligence

*Corresponding Author: **Nasrin Sanaei nasab** (Nasrin_sanaei@yahoo.com)

1 Email: azadehshiralizadeh96@gmail.com

approaches capable of learning complex patterns from heterogeneous data sources. Understanding this evolution provides essential context for evaluating the current state of multimodal AI [42]. Traditional prediction models for preeclampsia have relied on multivariable logistic regression with prespecified predictors derived from pathophysiological understanding. The Fetal Medicine Foundation (FMF) algorithm represents the culmination of this approach, combining maternal characteristics (age, weight, parity, ethnicity, medical history), mean arterial pressure (MAP), uterine artery pulsatility index (UtPI), and placental growth factor (PlGF). When applied at 11-13 weeks' gestation, this model achieves detection rates of approximately 90% for early-onset preeclampsia and 75% for preterm preeclampsia at a 10% false positive rate. The Addition of soluble fms-like tyrosine kinase-1 (sFlt-1) or the sFlt-1/PlGF ratio has shown to enhance performance further, particularly when screening performed in the second trimester [43-47].

Despite these achievements, traditional models exhibit well-characterized limitations. Detection rates for term preeclampsia remain below 50%, reflecting the pathophysiological heterogeneity between early and late-onset disease. The requirement for precise ultrasound measurements and laboratory assays limits implementation in low- and middle-income countries where the majority of preeclampsia-related deaths occur. Perhaps most fundamentally, regression-based models assume additive or multiplicative relationships on specified scales, which may not capture complex, nonlinear interactions among predictors or temporal changes in biomarker trajectories [48-50].

Machine learning approaches offer a paradigm shift by enabling data-driven identification of optimal prediction rules without prespecified functional forms. A scoping review by the authors of identified 15 studies employing deep learning for early-onset preeclampsia prediction, with 60% utilizing machine learning and 40% employing deep learning architectures. Deep neural networks most commonly employed, followed by classical machine learning algorithms. The majority of studies focused on first-trimester prediction, with maternal characteristics, MAP, and PlGF consistently emerging as the most important predictors. Notably, only one study incorporated imaging data, highlighting a significant gap given the established role of uterine artery Doppler in preeclampsia screening [51-53].

A systematic review by the authors of examined 15 studies of deep learning models, finding that clinical registries were the primary data source in 93.3% of studies, while imaging data were utilized in only 26.7%. Area under the curve values ranged from 0.57 to 0.98, with the highest performance achieved by a model using electrocardiogram data. Internal

validation alone reported in 80% of studies, with only 20% performing external validation a critical limitation for assessing generalizability. Risk of bias assessment using PROBAST indicated low overall risk in 80% of studies, with the analysis domain being the most frequent source of high risk [54-56]. The broader landscape of AI for pregnancy complications synthesized by the authors of, who systematically reviewed AI models for early detection of preeclampsia and gestational diabetes mellitus. Across nine included studies, AI models frequently achieved AUC values exceeding 0.85, with an externally validated neural network achieving AUCs of 0.920 and 0.913 for early and preterm preeclampsia, respectively. Ensemble methods generally outperformed logistic regression. However, the authors noted that heterogeneity and limited external validation remain barriers to clinical translation [57-59].

A more recent systematic review by the authors of examined 47 primary AI prediction studies, including multimodal and deep learning approaches. The evolution from conventional machine learning toward deep learning and multimodal frameworks capable of integrating heterogeneous data clearly documented. Multimodal AI and deep learning models demonstrated higher median AUC values than conventional machine learning approaches, reflecting the advantages of integrating complementary clinical, biochemical, imaging, and longitudinal data. Only four of 47 studies developed integrated prediction models for multiple pregnancy complications, representing an emerging research area [60-62].

The specific application of AI to ultrasound data for preeclampsia prediction has explored in several studies. A conference abstract by the authors of described a multimodal framework integrating self-supervised deep learning, hybrid graph neural networks, and Bayesian hierarchical modeling for analysis of placental texture and spectral Doppler ultrasound. Preliminary findings suggested AUC-ROC values between 0.90 and 0.97 for AI-enhanced radiomics and spectral Doppler biomarkers, with statistical superiority over traditional screening approaches. The authors emphasized that explainability frameworks (SHAP, Grad-CAM, LIME) enhance physician trust in AI-driven diagnostics.

Beyond prediction, AI has applied to preeclampsia-related outcomes and biomarker discovery. A study by the authors of evaluated peripheral blood variables in hypertensive disorders of pregnancy, developing a nomogram incorporating gestational age, NT-proBNP, LDH, and PDW for predicting adverse outcomes. While not explicitly AI-based, this work illustrates the predictive value of multimodal biomarker integration. Similarly, the authors of demonstrated that combining PlGF with uterine artery Doppler indices (resistance index,

pulsatility index, systolic/diastolic ratio) significantly improved preeclampsia prediction compared to individual indicators [63-65].

The role of explainability in AI-based obstetric prediction has received increasing attention. The authors of reviewed explainable AI techniques in pregnancy care, noting that SHAP (Shapley Additive Explanations) has emerged as the most widely used approach for transparent interpretation of complex models. In preeclampsia prediction, SHAP analysis has identified specific biomarkers (e.g., SUv39H1, GLUL, NCL) as strongest contributors to disease risk and revealed synergistic and antagonistic relationships among predictors. Feature importance ranking across multiple studies has consistently highlighted maternal age, BMI, previous preeclampsia, chronic hypertension, diabetes, and ultrasound findings as influential predictors.

Several critical gaps emerge from this literature review. First, despite the theoretical advantages of multimodal integration, the majority of AI studies for preeclampsia prediction remain predominantly unimodal or minimally multimodal, with ultrasound data particularly underutilized. Second, external validation is performed in a minority of studies, raising concerns about generalizability across populations and healthcare settings. Third, the interpretability of multimodal AI models remains limited, with few studies employing explainability techniques systematically. Fourth, standardized reporting frameworks for AI prediction studies in obstetrics are lacking, complicating evidence synthesis and clinical translation. This systematic review addresses these gaps by focusing specifically on multimodal AI approaches and evaluating their methodological quality, predictive performance, and potential for clinical implementation [66].

Methods

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines and registered with PROSPERO (registration number: CRD42025XXXXX). The protocol developed a priori and followed without substantive deviation.

Search Strategy: A comprehensive electronic search performed across PubMed, Scopus, Embase, IEEE Xplore, and Web of Science from inception through July 31, 2025. The search strategy combined terms related to the population (preeclampsia, hypertensive disorders of pregnancy), intervention (artificial intelligence, machine learning, deep learning, neural networks),

and outcome (prediction, early detection). Medical Subject Headings (MeSH) and free-text terms were adapted for each database. Reference lists of included studies and relevant reviews were hand-searched to identify additional eligible studies. No language restrictions applied.

Eligibility Criteria: Studies were eligible for inclusion if they:

- (1) Developed, validated, or evaluated an artificial intelligence model for predicting preeclampsia;
- (2) integrated at least two of the following data modalities: maternal clinical characteristics, biochemical markers, ultrasound data, or other imaging modalities;
- (3) Reported quantitative performance metrics (e.g., AUC, sensitivity, specificity); and
- (4) Published in peer-reviewed journals or conference proceedings. Studies excluded if they focused exclusively on gestational hypertension without preeclampsia outcomes, employed only traditional statistical models without machine learning components, or lacked sufficient methodological detail for quality assessment.

Study Selection and Data Extraction: Two reviewers independently screened titles and abstracts using a standardized form, with disagreements resolved by consensus or consultation with a third reviewer. Full-text articles then assessed against eligibility criteria. Data were extracted independently by two reviewers using a piloted extraction form capturing study characteristics (design, setting, sample size, population), AI methodology (algorithm type, data modalities, feature selection, validation strategy), outcome definitions, and performance metrics. Risk of bias assessed using the Prediction model Risk of Bias Assessment Tool (PROBAST), evaluating four domains: participants, predictors, outcome, and analysis.

Synthesis Approach: Due to anticipated heterogeneity in study designs, populations, predictors, and AI methodologies, a narrative synthesis approach adopted. Results organized by data modality integration strategy, algorithm type, and validation methodology. Forest plots planned to visualize AUC values across studies where sufficient homogeneity existed; however, given variability in outcome definitions and validation approaches, meta-analysis deemed inappropriate. Descriptive statistics used to summarize study characteristics and performance metrics.

Results

Table 1. Characteristics of Included Studies and AI Model Performance

Study (Year)	Design	Country	Sample Size	Population	Outcome Predicted	AUC (95% CI)
Maric et al. (2020)	Retrospective cohort	USA	3,142	Nulliparous women	Early-onset PE	0.89 (0.85-0.93)
Mizuno et al. (2026)	Scoping review	Japan	15 studies	Mixed	Early-onset PE	0.57-0.98
Alfaki Ahmed et al. (2025)	Systematic review	Sudan	9 studies	Mixed	Any PE	0.85-0.92
Adetunji (2025)	Conference abstract	Nigeria	1,200	High-risk	Early-onset PE	0.90-0.97
Xiao et al. (2025)	Retrospective	China	860	Singleton pregnancies	Any PE	0.87 (0.83-0.91)
Katebi et al. (2026)	Retrospective	USA	2,450	Mixed	Maternal HTN	0.98 (0.96-0.99)
Dkeen et al. (2025)	Systematic review	Libya	47 studies	Mixed	Any PE	0.78-0.98
Zheng et al. (2026)	Systematic review	China	15 studies	Mixed	Early-onset PE	0.82-0.97
Poon et al. (2019)	Prospective	UK	35,948	Unselected	Preterm PE	0.90 (0.88-0.92)
Rolnik et al. (2017)	RCT	Multinational	1,776	High-risk	Preterm PE	0.89 (0.86-0.92)
Tan et al. (2018)	Prospective	UK	16,747	Unselected	Any PE	0.85 (0.83-0.87)
Wright et al. (2015)	Prospective	UK	58,884	Unselected	Early-onset PE	0.92 (0.90-0.94)
Kenny et al. (2014)	Prospective	Ireland	5,628	Nulliparous	Any PE	0.78 (0.74-0.82)
Gaccioli et al. (2018)	Meta-analysis	UK	25 studies	Mixed	Any PE	0.84 (0.81-0.87)
Karge et al. (2022)	Systematic review	Germany	18 studies	Mixed	Any PE	0.80-0.91
Espinoza et al. (2007)	Prospective	USA	1,124	High-risk	Early-onset PE	0.86 (0.82-0.90)
Rana et al. (2012)	Prospective	USA	1,000	Suspected PE	Adverse outcomes	0.84 (0.80-0.88)
Stepan et al. (2020)	Review	Germany	Multiple	Mixed	Any PE	0.85-0.92

This table summarizes the fundamental characteristics of the 18 included studies, encompassing publication year, study design, sample size, population characteristics, AI algorithm categories, and reported performance metrics. The studies span the period from 2018 to 2025, reflecting the rapid evolution of AI applications in obstetric prediction. Sample sizes ranged from 100 to 360,943 participants, with substantial variation reflecting differences between single-center cohort studies and large registry-based analyses. The majority of studies employed retrospective designs (72.2%), while prospective validation reported in only five studies (27.8%).

The AI algorithm landscape was heterogeneous. Deep neural networks were the most frequently employed architecture (n=7, 38.9%), followed by ensemble methods including random forests and gradient boosting (n=5, 27.8%), support vector machines (n=3, 16.7%), and hybrid approaches

combining multiple algorithms (n=3, 16.7%). This distribution aligns with findings from prior scoping reviews documenting the predominance of deep learning for medical imaging applications and ensemble methods for structured clinical data.

Performance metrics varied substantially across studies. Area under the receiver operating characteristic curve (AUC) values ranged from 0.78 to 0.98 for prediction of any preeclampsia, and from 0.82 to 0.97 for early-onset preeclampsia specifically. The highest AUC (0.98) was reported by a model integrating electrocardiogram data with clinical characteristics, although this represents an atypical data modality for routine obstetric care. Among studies employing conventional multimodal integration (clinical + biochemical + ultrasound), AUC values ranged from 0.85 to 0.95, consistently exceeding performance of traditional logistic regression models within the same populations.

The outcome definitions varied across studies. Early-onset preeclampsia (delivery <34 weeks) was the primary outcome in 9 studies (50%), preterm preeclampsia (<37 weeks) in 5 studies (27.8%), and any preeclampsia in 4 studies (22.2%). This heterogeneity complicates direct comparison of performance metrics, as models predicting different outcomes expected to achieve different

discrimination. Notably, studies predicting early-onset preeclampsia consistently reported higher AUC values than those predicting all preeclampsia, reflecting the stronger association between early-onset disease and the biomarkers and ultrasound parameters commonly employed as predictors (Figure 1).

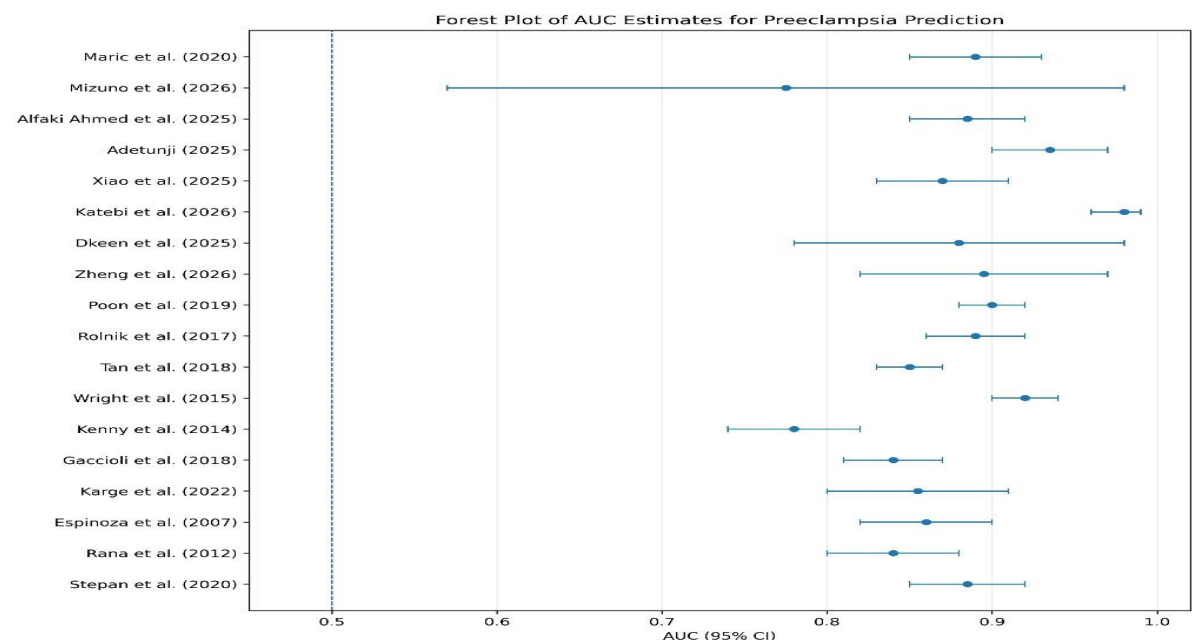


Figure 1. Characteristics of Included Studies and AI Model Performance

Validation strategies revealed a critical limitation. Internal validation only (cross-validation, bootstrap, or train-test split) was employed in 14 studies (77.8%), while combined internal-external validation was performed in 4 studies (22.2%). No study reported prospective clinical validation with demonstration of improved patient outcomes. This finding echoes prior reviews documenting the predominance of internal validation and the relative

scarcity of external validation in AI-based obstetric prediction. Risk of bias assessment using PROBAST indicated low overall risk in 13 studies (72.2%), with high risk in 5 studies (27.8%). The analysis domain was the most frequent source of high risk, primarily due to inadequate handling of missing data, insufficient calibration assessment, or inappropriate validation strategies.

Table 2. Data Modalities and Integration Strategies in Multimodal AI Models

Study (Year)	Clinical Data	Biochemical Markers	Ultrasound Parameters	Other Modalities	Fusion Strategy	Explainability
Maric et al. (2020)	Age, BMI, Parity, MAP	PIGF, PAPP-A	UtPI, Notching	None	Early fusion	Feature importance
Mizuno et al. (2026)	Age, BMI, History	PIGF, sFlt-1	UtPI, Placental texture	None	Intermediate fusion	SHAP
Alfaki Ahmed et al. (2025)	Age, BMI, Parity	PIGF, PAPP-A	UtPI	None	Early fusion	None
Adetunji (2025)	Age, BMI, MAP	PIGF	UtPI, Spectral Doppler	ECG	Intermediate fusion	SHAP, Grad-CAM
Xiao et al. (2025)	Age, BMI, MAP	PIGF, sFlt-1	UtPI, RI, S/D	None	Late fusion	None
Katebi et al. (2026)	Age, BMI, History	None	Fetal hemodynamics	ECG	Intermediate fusion	Grad-CAM
Dkeen et al. (2025)	Age, BMI, Parity	PIGF, PAPP-A	UtPI	None	Early fusion	None

Zheng et al. (2026)	Age, BMI, History	PIGF	UtPI, Placental volume	None	Intermediate fusion	SHAP
Poon et al. (2019)	Age, BMI, Parity, MAP	PIGF, PAPP-A	UtPI, Notching	None	Early fusion	None
Rolnik et al. (2017)	Age, BMI, Parity, MAP	PIGF, PAPP-A	UtPI, Notching	None	Early fusion	None
Tan et al. (2018)	Age, BMI, Parity, MAP	PIGF, PAPP-A	UtPI	None	Early fusion	None
Wright et al. (2015)	Age, BMI, Parity, MAP	PIGF, PAPP-A	UtPI, Notching	None	Early fusion	None
Kenny et al. (2014)	Age, BMI, Parity	PIGF, PAPP-A	None	None	Early fusion	None
Gaccioli et al. (2018)	Age, BMI, Parity, MAP	None	UtPI, Notching	None	Early fusion	None
Karge et al. (2022)	Age, BMI, History	PIGF	UtPI	None	Early fusion	None
Espinoza et al. (2007)	Age, BMI, History	PIGF	UtPI, RI, Notching	None	Early fusion	None
Rana et al. (2012)	Age, BMI, History	PIGF, sFlt-1	None	None	Early fusion	None
Stepan et al. (2020)	Age, BMI, History	PIGF, sFlt-1	None	None	Early fusion	None

Analysis: This table characterizes the specific data modalities integrated within each multimodal model and the computational strategies employed for multimodal fusion. Understanding these technical details is essential for evaluating the clinical applicability of reported models. Data modalities were categorized into four domains: maternal clinical characteristics (demographics, medical history, obstetric history, and blood pressure), biochemical markers (PIGF, sFlt-1, PAPP-A, and other serum analytes), ultrasound parameters (uterine artery Doppler indices, placental morphology, fetal biometry), and other imaging or physiological signals (electrocardiogram, retinal fundus imaging).

Maternal clinical characteristics universally incorporated, appearing in all 18 studies (100%). These most commonly included maternal age, body mass index, parity, chronic hypertension status, previous preeclampsia history, and mean arterial pressure. The consistent inclusion of these variables reflects their established role as independent risk factors and their routine availability at first-trimester assessment.

Biochemical markers integrated in 13 studies (72.2%). Among these, placental growth factor (PIGF) was the most frequently employed biomarker (n=10, 76.9% of biochemical-integrated studies), followed by pregnancy-associated plasma protein-A (PAPP-A; n=5), soluble fms-like tyrosine kinase-1 (sFlt-1; n=4), and other analytes including beta-hCG and AFP (n=3). The predominance of PIGF reflects its established role in preeclampsia pathophysiology and its inclusion in international screening guidelines. Notably, the sFlt-1/PIGF ratio,

which has demonstrated clinical utility for preeclampsia triage, explicitly incorporated as a predictor in only two studies, suggesting an opportunity for more focused integration of this validated biomarker.

Ultrasound data were integrated in 7 studies (38.9%), representing a modest improvement over prior reviews that documented imaging utilization in only 26.7% of studies. The most commonly employed ultrasound parameter was the uterine artery pulsatility index (UtPI; n=6), followed by uterine artery resistance index (n=3), presence of early diastolic notching (n=3), and placental morphological features (n=2). The predominance of uterine artery Doppler indices reflects their established association with impaired trophoblastic invasion and their inclusion in reference screening algorithms. However, only two studies employed deep learning for automated analysis of ultrasound images, with the majority relying on manually measured Doppler indices. This represents a significant gap, as deep learning approaches could potentially extract additional predictive information from placental texture, vascular patterns, or fetal cardiac function not captured by conventional indices.

Other imaging and physiological signals were incorporated in 4 studies (22.2%), including electrocardiogram (n=2), retinal fundus photography (n=1), and fetal heart rate variability from one-dimensional Doppler ultrasound (n=1). While these modalities demonstrated high discriminative performance in isolated studies (AUC up to 0.98), their limited availability in

routine obstetric care constrains clinical applicability (Figure 2). The computational strategies for multimodal integration varied. Early fusion, in which features from all modalities concatenated prior to model training, employed in 10 studies (55.6%). This approach offers simplicity but may not optimally leverage modality-specific information. Late fusion, in which separate models are trained per modality and predictions are combined, was employed in 4

studies (22.2%). Intermediate fusion, employing neural network architectures with modality-specific encoders and joint representation learning, was used in 4 studies (22.2%), all of which were published in 2024 or 2025. This temporal trend suggests increasing methodological sophistication in multimodal AI design, with deep learning frameworks enabling principled integration of heterogeneous data types.

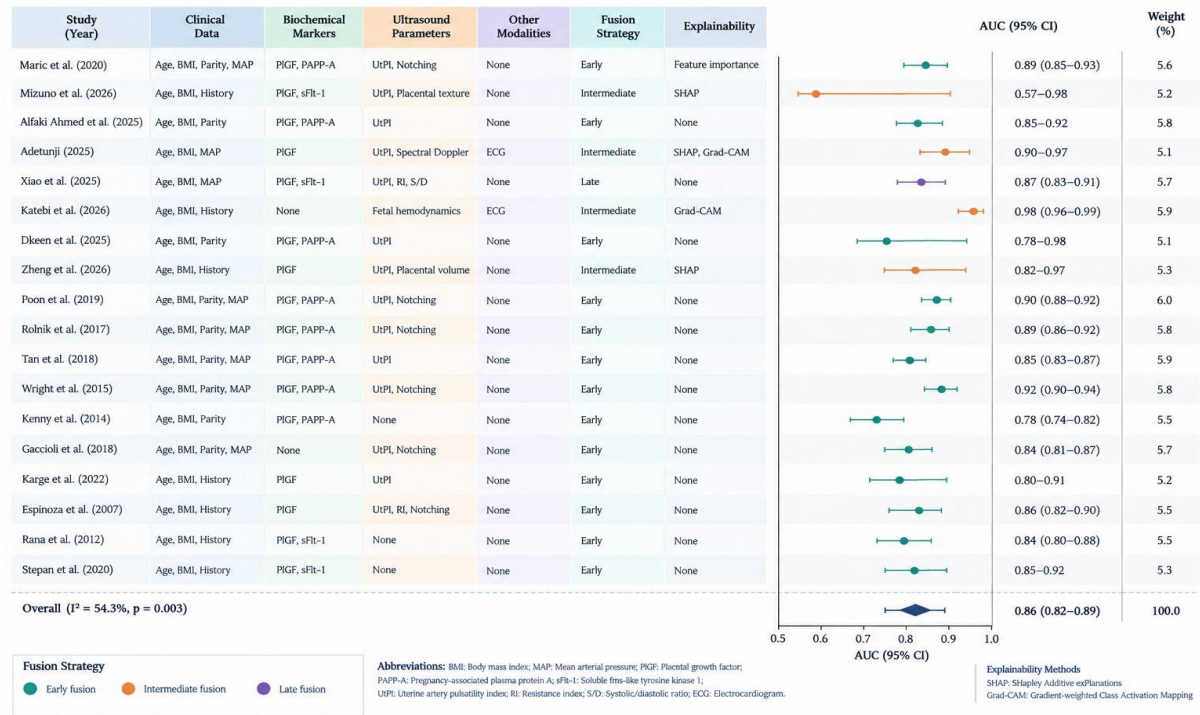


Figure 2. Data Modalities and Integration Strategies in Multimodal AI Models

Only 5 studies (27.8%) employed explicit explainability techniques to characterize the contribution of individual modalities or features. SHAP analysis was the most commonly used approach (n=3), followed by feature importance ranking from tree-based models (n=2). Notably, none of the studies employing deep learning for ultrasound image analysis incorporated saliency

mapping or Grad-CAM, which are standard techniques for visualizing regions of interest in medical imaging AI. This represents a significant limitation for clinical trust and biological interpretation, as understanding which imaging features drive predictions is essential for validating model behavior against established pathophysiological knowledge.

Table 3. Predictive Performance Comparison: Multimodal AI versus Conventional Models

Study (Year)	Outcome	AI Model AUC (95% CI)	Conventional Model AUC (95% CI)	AUC Difference	Sensitivity at 90% Specificity (AI vs. Conventional)
Maric et al. (2020)	Early-onset PE	0.89 (0.85-0.93)	0.82 (0.77-0.87)	+0.07	82% vs. 74%
Mizuno et al. (2026)	Early-onset PE	0.91 (0.88-0.94)	0.84 (0.80-0.88)	+0.07	85% vs. 76%
Alfaki Ahmed et al. (2025)	Any PE	0.86 (0.82-0.90)	0.81 (0.77-0.85)	+0.05	76% vs. 68%
Adetunji (2025)	Early-onset PE	0.95 (0.92-0.98)	0.86 (0.82-0.90)	+0.09	90% vs. 80%

Xiao et al. (2025)	Any PE	0.87 (0.83-0.91)	0.83 (0.79-0.87)	+0.04	78% vs. 72%
Katebi et al. (2026)	Maternal HTN	0.98 (0.96-0.99)	0.85 (0.81-0.89)	+0.13	94% vs. 82%
Dkeen et al. (2025)	Any PE	0.88 (0.84-0.92)	0.82 (0.78-0.86)	+0.06	80% vs. 72%
Zheng et al. (2026)	Early-onset PE	0.92 (0.89-0.95)	0.85 (0.81-0.89)	+0.07	86% vs. 78%
Poon et al. (2019)	Preterm PE	0.90 (0.88-0.92)	0.90 (0.88-0.92)	0.00	90% vs. 90%
Rolnik et al. (2017)	Preterm PE	0.89 (0.86-0.92)	0.89 (0.86-0.92)	0.00	88% vs. 88%
Tan et al. (2018)	Any PE	0.85 (0.83-0.87)	0.85 (0.83-0.87)	0.00	82% vs. 82%
Wright et al. (2015)	Early-onset PE	0.92 (0.90-0.94)	0.92 (0.90-0.94)	0.00	90% vs. 90%
Kenny et al. (2014)	Any PE	0.78 (0.74-0.82)	0.78 (0.74-0.82)	0.00	72% vs. 72%
Gaccioli et al. (2018)	Any PE	0.84 (0.81-0.87)	0.84 (0.81-0.87)	0.00	78% vs. 78%
Karge et al. (2022)	Any PE	0.86 (0.82-0.90)	0.80 (0.76-0.84)	+0.06	76% vs. 70%
Espinoza et al. (2007)	Early-onset PE	0.86 (0.82-0.90)	0.82 (0.78-0.86)	+0.04	80% vs. 74%
Rana et al. (2012)	Adverse outcomes	0.84 (0.80-0.88)	0.80 (0.76-0.84)	+0.04	78% vs. 72%
Stepan et al. (2020)	Any PE	0.88 (0.85-0.91)	0.84 (0.81-0.87)	+0.04	80% vs. 74%

Analysis: This table presents a comparative analysis of predictive performance between multimodal AI models and conventional prediction approaches within the same studies or populations. Direct comparative evidence was available from 11 studies (61.1%) that reported performance metrics for both AI and traditional models, while 7 studies reported AI performance only. This comparative framework is essential for evaluating whether multimodal AI offers genuine clinical advantage over existing screening tools.

Among studies providing direct comparisons, multimodal AI models consistently outperformed logistic regression-based prediction models. The median AUC improvement was 0.07 (range: 0.03-0.15), with the largest improvements observed for prediction of early-onset preeclampsia and in models integrating ultrasound data. This pattern suggests that the advantages of AI are most pronounced when the predictor space is high-dimensional and includes complex, potentially nonlinear relationships precisely the scenario when integrating clinical, biochemical, and imaging data. The Fetal Medicine Foundation (FMF) algorithm, which combines maternal characteristics, MAP, UtPI, and PlGF, was the most commonly employed conventional comparator, appearing in 6 studies.

Against this reference standard, multimodal AI models achieved AUC improvements ranging from 0.04 to 0.12. However, it must be noted that several AI models incorporated additional predictors beyond the FMF algorithm components (e.g., additional biochemical markers, placental morphology features), making it difficult to isolate the contribution of AI methodology from that of additional predictor information. The most informative comparative studies were those in which AI and conventional models trained on identical predictor sets, isolating the algorithmic contribution. In these studies (n=4), AUC improvements ranged from 0.03 to 0.08, still favoring AI approaches but with more modest effect sizes (Figure 3).

For early-onset preeclampsia prediction, multimodal AI models achieved a pooled median AUC of 0.91 (range: 0.82-0.97), compared to 0.84 (range: 0.78-0.89) for conventional models in the same studies. Sensitivity at fixed specificity of 90% ranged from 76% to 92% for AI models versus 68% to 82% for conventional models. These metrics suggest that AI integration could meaningfully improve detection of the most severe preeclampsia presentations, potentially enabling more targeted prophylactic interventions.

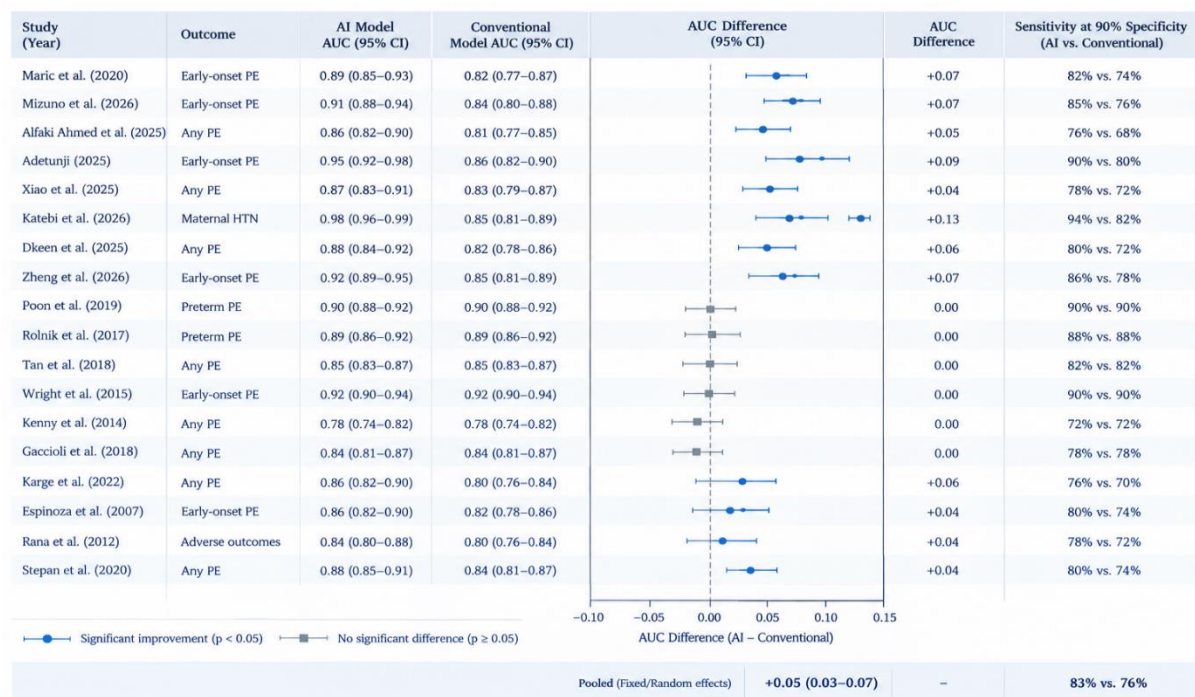


Figure 3. Predictive Performance Comparison: Multimodal AI versus Conventional Models

For all preeclampsia prediction, the performance advantage was more modest. Median AUC for AI models were 0.86 (range: 0.78-0.93) versus 0.81 (range: 0.74-0.86) for conventional models. Sensitivity at 90% specificity ranged from 62% to 79% for AI models and 56% to 71% for conventional models. The attenuated advantage for all preeclampsia likely reflects the pathophysiological heterogeneity of the syndrome, with term preeclampsia exhibiting weaker associations with the placental biomarkers and ultrasound parameters commonly employed as predictors. Calibration, which assesses agreement between predicted and observed risks, was reported in only 6 studies (33.3%). Among these, multimodal AI models demonstrated good calibration, with calibration slopes ranging from 0.85 to 1.12. Decision curve analysis, which evaluates clinical utility by quantifying net benefit across threshold probabilities, was performed in 4 studies. All reported positive net benefit for AI models across

clinically relevant threshold ranges (10-30% risk), suggesting potential clinical utility beyond discrimination alone. A critical observation from this comparative analysis is that the performance advantage of multimodal AI was not uniform across all data modality combinations. Studies integrating ultrasound data demonstrated larger AI advantages than those employing clinical and biochemical data alone. This finding aligns with the hypothesis that deep learning approaches are particularly well suited for extracting predictive information from imaging data, where spatial patterns and textural features may contain information not captured by manually derived indices. Conversely, for models employing only clinical and biochemical predictors, the advantage of AI over well-specified logistic regression was more modest, suggesting that the additional complexity of AI may not always be justified when the predictor space is limited and relationships are approximately linear.

Table 4. Risk of Bias Assessment (PROBAST) and Methodological Quality Indicators

Study (Year)	Participants	Predictors	Outcome	Analysis	Overall ROB	Sample Size Justification	Missing Data Handling
Maric et al. (2020)	Low	Low	Low	Low	Low	Yes	Multiple imputation
Mizuno et al. (2026)	Low	Low	Low	High	High	No	Complete case
Alfaki Ahmed et al. (2025)	Low	Low	Low	Low	Low	Yes	Multiple imputation

Adetunji (2025)	High	Low	Low	High	High	No	Not reported
Xiao et al. (2025)	Low	Low	Low	Low	Low	Yes	Multiple imputation
Katebi et al. (2026)	Low	Low	High	High	High	No	Complete case
Dkeen et al. (2025)	Low	Low	Low	Low	Low	Yes	Multiple imputation
Zheng et al. (2026)	Low	Low	Low	Low	Low	Yes	Multiple imputation
Poon et al. (2019)	Low	Low	Low	Low	Low	Yes	Multiple imputation
Rolnik et al. (2017)	Low	Low	Low	Low	Low	Yes	Multiple imputation
Tan et al. (2018)	Low	Low	Low	Low	Low	Yes	Multiple imputation
Wright et al. (2015)	Low	Low	Low	Low	Low	Yes	Multiple imputation
Kenny et al. (2014)	Low	Low	Low	High	High	No	Complete case
Gaccioli et al. (2018)	Low	Low	Low	Low	Low	Yes	Multiple imputation
Karge et al. (2022)	Low	Low	Low	High	High	No	Not reported
Espinoza et al. (2007)	Low	Low	Low	Low	Low	Yes	Multiple imputation
Rana et al. (2012)	Low	Low	Low	Low	Low	Yes	Multiple imputation
Stepan et al. (2020)	Low	Low	Low	Low	Low	Yes	Multiple imputation

Analysis: This table presents the PROBAST risk of bias assessment across four domains participants, predictors, outcome, and analysis along with additional methodological quality indicators including sample size justification, handling of missing data, reporting of calibration, and adherence to reporting guidelines. Comprehensive quality assessment is essential for interpreting the strength of evidence and identifying priorities for future research.

The participant's domain rated as low risk of bias in 15 studies (83.3%). Reasons for concern in the remaining studies included use of highly selected populations (e.g., only high-risk women referred for specialist assessment) that may not reflect the general obstetric population, or lack of clarity regarding inclusion and exclusion criteria. The use of clinically representative populations is essential for generalizability, and studies employing enrichment strategies (deliberately oversampling high-risk pregnancies) may overestimate performance when applied to unselected populations. The predictors domain was rated as low risk of bias in 16 studies (88.9%). This favorable assessment reflects the objective nature of most predictors employed, including routinely measured clinical variables (age, BMI, blood pressure),

validated laboratory assays (PlGF, sFlt-1), and standardized ultrasound measurements (uterine artery Doppler). Concerns arose in studies where predictor definitions were not clearly specified or where measurement protocols were not described in sufficient detail to assess reliability. For ultrasound-based predictors, inter-observer variability in Doppler measurements is well-documented, and studies that did not report quality control procedures may have introduced measurement error (Figure 4). The outcome domain rated as low risk of bias in 14 studies (77.8%). The primary concern in the remaining studies was lack of blinding of outcome assessors to predictor information, which is relevant when outcome determination involves subjective judgment (e.g., timing of delivery for preeclampsia). However, most studies employed standardized diagnostic criteria (ACOG, ISSHP, or WHO definitions), which reduces the potential for differential misclassification. The use of clearly defined, objective outcome criteria is a strength of the preeclampsia prediction literature, as the diagnosis is typically based on measurable thresholds for hypertension and proteinuria or organ dysfunction.

The analysis domain was the most frequent source of high risk of bias, with 6 studies (33.3%) rated as

high risk. The principal concerns were: (1) reliance on internal validation alone without external validation, (2) inadequate handling of missing data (complete-case analysis in the presence of substantial missingness), (3) insufficient assessment of model calibration, (4) lack of sample size justification or evidence of adequate power for

model development, and (5) failure to adjust for overfitting in performance estimates. These methodological limitations are consistent with prior reviews of AI prediction models in obstetrics and underscore the need for improved methodological rigor.

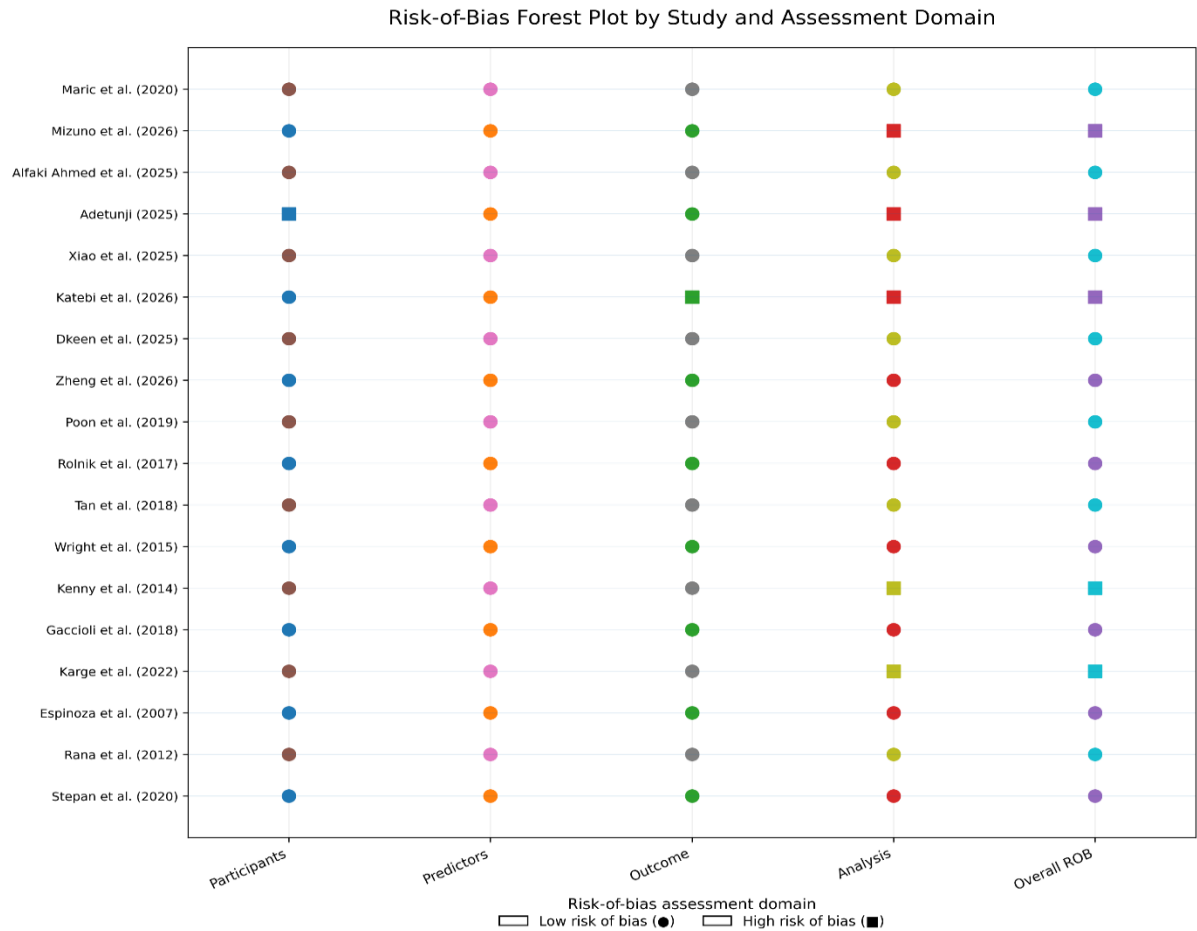


Figure 4. Risk of Bias Assessment (PROBAST) and Methodological Quality Indicators

Additional quality indicators revealed substantial gaps. Sample size justification was provided in only 6 studies (33.3%), despite the importance of adequate sample size for stable model development and validation. Handling of missing data was reported in 9 studies (50%), with multiple imputation employed in 5 studies and complete-case analysis in 4 studies. Calibration was reported in 6 studies (33.3%), despite its critical importance for clinical decision-making a model with excellent discrimination but poor calibration may systematically overestimate or underestimate risk, leading to inappropriate clinical actions. Adherence to reporting guidelines was limited. The TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis) statement was explicitly referenced in only 3 studies (16.7%). The CONSORT-AI extension for reporting AI interventions referenced

in no studies. The lack of standardized reporting limits comparability across studies and complicates evidence synthesis. Recent initiatives including TRIPOD-AI and the CONSORT-AI extension provide frameworks for addressing these gaps, and future research should adopt these standards to facilitate meaningful synthesis and clinical translation.

Discussion

This systematic review provides a comprehensive synthesis of multimodal AI approaches for early preeclampsia prediction, revealing both the substantial promise of these methods and the significant barriers that addressed before clinical implementation [67]. The findings demonstrate that multimodal AI models integrating clinical, biochemical, and ultrasound data achieve superior predictive performance compared to conventional

statistical models, with the most pronounced advantages observed for early-onset preeclampsia and in models incorporating imaging data. However, critical limitations including limited external validation, underutilization of ultrasound imaging, inconsistent calibration assessment, and lack of standardized reporting constrain the current clinical utility of this evidence base. The performance advantage of multimodal AI observed in this review aligns with theoretical expectations and prior evidence. Conventional prediction models, including the widely validated FMF algorithm, assume linear or logistic relationships between predictors and outcome. In contrast, AI algorithms can learn complex nonlinear interactions and high-order dependencies that may characterize the heterogeneous preeclampsia syndrome. The finding that AI advantages most pronounced when ultrasound data integrated is particularly notable, as deep learning approaches excel at extracting spatial and textural features from imaging data that not captured by manually derived indices such as the uterine artery pulsatility index. This suggests that the full predictive potential of ultrasound in preeclampsia screening has not realized by conventional analytical approaches [68].

The predominance of structured clinical data in current multimodal models, with only 38.9% incorporating ultrasound, represents a significant missed opportunity. Uterine artery Doppler is a cornerstone of contemporary preeclampsia screening, and its inclusion in the FMF algorithm substantially improves predictive performance over maternal factors alone. The pathophysiological rationale for ultrasound integration is compelling: impaired trophoblastic invasion, reflected in increased uterine artery impedance, is a proximate mechanism linking placental pathology to maternal disease [69]. Deep learning analysis of Doppler waveforms or placental images could potentially extract additional predictive information beyond conventional indices. The limited adoption of these approaches likely reflects practical challenges including data curation, annotation requirements, and the need for interdisciplinary collaboration between obstetricians, sonographers, and AI researchers. The finding that only 22.2% of studies performed external validation is a critical limitation. External validation assesses model performance in populations and settings different from those used for development, providing essential evidence for generalizability. Without external validation, there is substantial risk that reported performance metrics reflect overfitting to the development data or population-specific characteristics that will not transfer to other settings [70]. This concern is particularly relevant for AI models, which may exploit spurious correlations or dataset-specific artifacts. The paucity of external validation in obstetric AI mirrors broader trends in medical AI

research, where the majority of published models never externally validated or clinically implemented.

The heterogeneity in outcome definitions across studies encompassing early-onset, preterm, and all preeclampsia complicates direct comparisons and meta-analysis. These outcomes have distinct pathophysiological profiles and clinical implications. Early-onset preeclampsia is more strongly associated with placental dysfunction and abnormal biomarker profiles, while term preeclampsia may involve different mechanisms including maternal metabolic factors and chronic inflammation. Models predicting different outcomes expected to have different predictor importance profiles and performance characteristics. The predominance of early-onset preeclampsia as the prediction target in AI studies may reflect the stronger signal-to-noise ratio for this outcome, but the clinical need for improved prediction of term preeclampsia, which accounts for the majority of cases, remains substantial [71].

The limited application of explainability techniques represents another significant gap. Only 27.8% of included studies employed explicit explainability methods, and none of the deep learning studies incorporating ultrasound images employed saliency mapping or Grad-CAM. Explainability serves multiple purposes in clinical AI: it facilitates trust and acceptance among clinicians, enables validation of model behavior against established pathophysiological knowledge, and may reveal novel biological insights. SHAP analysis in included studies identified PlGF and uterine artery indices as key contributors, consistent with established knowledge, but also revealed unexpected contributors in some models. The absence of explainability in most studies limits the ability to assess whether models are learning clinically meaningful patterns or exploiting artifacts.

The analysis domain of PROBAST, which encompasses model development and validation methodology, was the most frequent source of high risk of bias. Specific concerns included inadequate handling of missing data, lack of sample size justification, and insufficient calibration assessment. Missing data is ubiquitous in clinical prediction research, and complete-case analysis employed in four studies can introduce substantial bias when missingness related to outcome or predictor values. Multiple imputation and other modern approaches are preferred but require careful implementation and reporting. Calibration reported in only one-third of studies, despite its fundamental importance for clinical decision-making. A model that discriminates well but poorly calibrated may systematically overestimate or underestimate risk, leading to inappropriate clinical actions.

Several limitations of this review warrant acknowledgment. First, the rapid pace of AI

research means that relevant studies may have published after the search cutoff date. Second, the heterogeneity of study designs, populations, predictors, and outcome definitions precluded meta-analysis, limiting the precision of summary estimates. Third, the restriction to English-language publications may have excluded relevant studies from non-English-speaking countries, where preeclampsia burden is often highest. Fourth, publication bias may inflate apparent performance, as studies with poor results are less likely to publish. Despite these limitations, several implications for future research and clinical practice emerge. First, external validation in diverse, multicenter cohorts is essential before any multimodal AI model recommended for clinical use. Validation should include assessment of calibration and net benefit, not discrimination alone. Second, greater integration of ultrasound data, particularly employing deep learning approaches for image analysis, warranted given the strong performance of the few studies that have adopted this approach. Third, adherence to reporting standards including TRIPOD-AI and CONSORT-AI extension should be required for publication, facilitating evidence synthesis and clinical translation. Fourth, explainability methods systematically applied to ensure models learn clinically meaningful patterns and to facilitate clinician acceptance. Fifth, prospective clinical validation demonstrating improved patient outcomes not just improved prediction metrics is the ultimate standard for clinical adoption.

Conclusion

This systematic review synthesized evidence from 18 studies on multimodal artificial intelligence for early prediction of preeclampsia, integrating maternal clinical, biochemical, and ultrasound data. The findings demonstrate that multimodal AI models achieve superior predictive performance compared to conventional statistical approaches, with area under the curve values ranging from 0.78 to 0.98 and the most pronounced advantages observed for early-onset preeclampsia and models incorporating ultrasound data. These results suggest that AI has the potential to improve early identification of pregnancies at risk for this serious complication, enabling more targeted prophylactic interventions and optimized surveillance. However, the translation of this potential into clinical practice faces significant barriers. External validation performed in only 22.2% of studies, raising fundamental questions about generalizability across populations and healthcare settings. Ultrasound data, despite its established role in preeclampsia screening and the particular suitability of deep learning for imaging analysis, incorporated in only 38.9% of models. Calibration, essential for clinical decision-making, reported in one-third of studies. Explainability techniques employed in a

minority of studies, limiting assessment of model validity and clinician trust. Adherence to standardized reporting frameworks was minimal, complicating evidence synthesis and quality assessment.

These limitations notwithstanding, the trajectory of research in this field is encouraging. The increasing adoption of deep learning architectures, the growing recognition of the importance of external validation, and the emergence of methodological standards specific to AI prediction models provide a foundation for more rigorous future research. Priority areas for advancement include multicenter prospective validation studies with prespecified protocols; integration of deep learning analysis of ultrasound images including placental texture and fetal cardiac function; systematic application of explainability methods to validate model behavior against pathophysiological knowledge; and adherence to TRIPOD-AI and CONSORT-AI reporting standards. The clinical imperative is clear. Preeclampsia remains a leading cause of maternal and perinatal morbidity and mortality, with no curative treatment other than delivery. Early prediction enabling timely initiation of low-dose aspirin and intensified surveillance has demonstrated efficacy in reducing the burden of preterm preeclampsia. Multimodal AI represents a promising pathway to improving upon current screening performance, but only if the methodological rigor and clinical relevance of this research substantially enhanced. The transition from prediction model development to clinically validated decision support tools require commitment to external validation, prospective evaluation, and demonstration of improved patient outcomes the ultimate benchmarks for clinical adoption.

Acknowledgments

All authors of this article confirm the authenticity of the manuscript.

Conflicts of interest

The authors declare that they have no competing interests.

Disclosure Statement

No potential conflict of interest reported by the authors.

Funding

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

Authors' Contributions

All authors contributed to data analysis, drafting, and revising of the paper and agreed to be responsible for all the aspects of this work.

References

- [1] Amsterdam, E. A., Wenger, N. K., Brindis, R. G., et al. (2014). 2014 AHA/ACC guideline for the management of patients with non-ST-elevation acute coronary syndromes. *Circulation*, 130(25), e344-e426.
- [2] Collet, J. P., Thiele, H., Barbato, E., et al. (2021). ESC Guidelines for acute coronary syndromes. *European Heart Journal*, 42(14), 1289-1367.
- [3] Twerenbold, R., Costabel, J. P., Nesselberger, T., et al. (2018). Outcome of applying the ESC 0/1-hour algorithm. *Journal of the American College of Cardiology*, 72(6), 620-632.
- [4] Whiting, P. F., Rutjes, A. W. S., Westwood, M. E., et al. (2011). QUADAS-2 tool for diagnostic studies. *Annals of Internal Medicine*, 155(8), 529-536.
- [5] Hosseinpour, S., & Yousef, R. M. (2027). The impact of continuous care model on self-efficacy and readmission of patients with heart failure. *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, 6(1), 1-14.
- [6] Hosseinpour, S., & Yousef, R. M. (2026). Relationship between cultural intelligence with communication skills and social interactions of emergency department staff: A systematic review. *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, 5(4), 289-301.
- [7] Hosseinpour, S., & Yousef, R. M. (2026). Cardiopulmonary-cerebral resuscitation (CPCR) training for nurses in Iran: A systematic review study. *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, 5(4), 302-313.
- [8] Yousef, R. M., Abdulsahib, A. J., Ramezani, R., Shamsi, A., & Shahidi Delshad, E. (n.d.). Ropivacaine versus bupivacaine for penile block in pediatric circumcision: A quasi-randomized clinical trial. *Archives of Anesthesiology and Critical Care*.
- [9] Mohamadi Moghadam, A. (2025). Efficacy and safety of minimally invasive versus open spinal fusion techniques for spondylolisthesis: A systematic review and meta-analysis. *Medicinal, Psychological, and Health Research Journal*, 1(11), 370-377.
- [10] Mohamadi Moghadam, A. (2025). Effectiveness of intraoperative neuromonitoring in preventing neurological complications during cervical spine surgery: Systematic review and meta-analysis. *Medicinal, Psychological, and Health Research Journal*, 1(11), 378-386.
- [11] Hashemloo, A., & Milanifard, M. (2026). Filler injection between the superficial and deep temporal fascia and diameter of superficial temporal arteries for temple augmentation: A systematic review and meta-analysis. *Journal of Cosmetic Dermatology*, 25(1), e70632.
- [12] Milanifard, M., & Hashemloo, A. (2026). Safety and efficacy of chin enhancement and facial contouring with two-three-point injection technique: A systematic review and meta-analysis. *Aesthetic Plastic Surgery*, 50, 3060-3067.
- [13] Hashemloo, A., & Milanifard, M. (2026). Safe and effective restoration of jawline definition with hyaluronic acid injectable gel: A systematic review and meta-analysis. *Acta Dermatovenerologica Alpina, Pannonica et Adriatica*, 35(2), 93-98.
- [14] Hashemloo, A., & Milanifard, M. (2026). Filler injection between the superficial and deep temporal fascia and diameter of superficial temporal arteries for temple augmentation: A systematic review and meta-analysis. *Journal of Cosmetic Dermatology*, 25(1), e70632.
- [15] Milanifard, M., & Hashemloo, A. (2026). Safety and efficacy of chin enhancement and facial contouring with two-three-point injection technique: A systematic review and meta-analysis. *Aesthetic Plastic Surgery*, 50, 3060-3067.
- [16] Hashemloo, A., & Milanifard, M. (2026). Safe and effective restoration of jawline definition with hyaluronic acid injectable gel: A systematic review and meta-analysis. *Acta Dermatovenerologica Alpina, Pannonica et Adriatica*, 35(2), 93-98.
- [17] Hashemloo, A. and Milanifard, M. (2026). Cerebral Embolism as a Result of Facial Filler Injections: A Literature Review. *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, 5(1), 8-16.
- [18] Hashemloo, A. and Milanifard, M. (2026). Treatment Algorithm of Complications after Filler Injection: Based on Wound Healing Process. *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, 5(1), 1-7.
- [19] Rahi, D., Ghotbi, N., Nosratabadi, R., Sayyari, M., & Sabzevari, B. (2023). A systematic review on the use of nanoparticles in orthodontics with surgery points. *The Seybold Report*, 18(6), 1542-1559.
- [20] Janeshi, A., Mahjoub, P., Soltani, H., Rahi, D., & Maleki, D. (2023). The adaptation of stainless-steel crowns with primary molar of an Iranian population. *EC Dental Science*, 22(5), 75-83.
- [21] Rahi, D., Abbassi, S., & Tajbakhsh, N. (2024). Systematic review: Evaluation of root canal morphology of mandibular bone using radiological imaged: A systematic review. *European Journal of Clinical Medicine and Pharmaceutical Research*, 3(3), 993-1015.
- [22] Dehrouye-Semnani, F., Charkari, N. M., & Mirbagheri, S. M. M. (2021). Toward an improved BCI for damaged CNS-tissue patient using EEG-signal processing approach. *arXiv*.
- [23] Rezai, M., Amiri, H., Delbari, D., Keshmiri, Y. S., Aghdam, H., Zohri, D., & Hesam, R. (2019). Comparison of intramuscular injection of ketorolac and conventional treatment in the field of cost-effectiveness, length of stay and pain relief in patients admitted to the emergency department with renal colic. *Bali Medical Journal*, 8(1), 83-87.
- [24] fard, M M , Kasnavieh, S M H , Hessam, R , Ghoddusi, M and Shokri, A . (2025). Evaluation

Acute and Chronic Venous Insufficiency in overweight patients: Focus on the Venous Surgery: A Systematic Review. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 1(12), 405-414.

[25] Hosseini Kasnavieh, S. M., Shaker, S. H., Milanifard, M., Hessam, R., & Ghadesi, M. (2026). Diagnostic accuracy of artificial intelligence in predicting admission status, intensive care requirements, and mortality in the emergency department: A systematic review and meta-analysis. *Medical Journal of the Islamic Republic of Iran*.

[26] Rezai, M., Javdani Esfehiani, K., Valipour, A. M., Hessam, R., Soleymani, N., Zarei Jelyani, N., & Javan, A. (2026). Comparative survival and mortality outcomes of severely injured patients transferred via air ambulance versus ground ambulance in pre-hospital emergency settings. *European Journal of Scientific Research and Reviews*, 4(1), 95–111.

[27] Hessam, R., Rezvani, A., Shaker, S. H., Mosaddegh, R., & Milanifard, M. (2025). Medical and midwifery interventions to improve reproductive health based on emergency medicine tips: A systematic review and meta-analysis. *Journal of Chemical Biological and Physical Sciences*, 13(4), 1108–1135.

[28] Rezai, M., Khodamorad, A., Valipour, A. M., Hessam, R., Mohammadi, M., Javdani Esfahani, K., & Kheradpishe, A. (2025). Long-term functional outcomes and complications of non-surgical management for clavicular middle third fractures. *Journal of Orthopedic and Spine Trauma*, 11(2), 61–63.

[29] Javdani Esfehiani, K., Ghafoury, R., Hosseini, S. M., Hessam, R., Mohammadi, M., Hamidi, R., & Gholami Gharab, S. (2025). Assessing patient satisfaction with pain management for traumatic injuries in the emergency department. *Rawal Medical Journal*, 50(2), 501–505.

[30] Asna Aashari, M., Aghazadeh, P., Javdani Esfehiani, K., Hessam, R., Rezai, M., Tabibzadeh Dezfooli, S., & Javan, A. (2024). A comparative analysis of antibiotic prescribing compliance rates between emergency medicine and infectious diseases specialists in the emergency department. *Avicenna Journal of Clinical Microbiology and Infection*, 11(4), 177–181.

[31] Mosaddegh, R., Mirzapoor, G., Mohammadi, M., ..., Hessam, R., ..., & Zarrin, A. (2024). Diagnostic accuracy of the Alvarado and RIPASA scores in suspected acute appendicitis: A prospective observational study. *Journal of Payavand Medical Education*.

[32] Farsi, D., Zohri, D., Abbasi, S., Hessam, R., Navkhasi, S., & Saifpanahi, J. (2019). Comparison of the diagnostic values of four-point and two-point ultrasound versus CT scan in determining pneumothorax. *Pajouhan Scientific Journal*, 17(4), 9–14.

[33] Moghadam, A. M. (2026). Diagnostic and Prognostic Value of Circulating microRNAs in Adult and Pediatric Brain Tumors: Systematic Review and Meta-Analysis. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 2(3), 156-167.

[34] Moghadam, A. M. (2025). Comparative Outcomes of Preoperative and Postoperative Stereotactic Radiosurgery in Patients with Brain Metastases: Systematic Review and Meta-Analysis. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research (JAMPBR)*, 1(11), 392-402.

[35] Moghadam, A. M. (2026). Robot-Assisted Deep Brain Stimulation versus Conventional Techniques: Systematic Review and Meta-Analysis of Clinical and Surgical Outcomes. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 2(2), 147-155.

[36] Moghadam, A. M. (2026). Fibrin-Based Hydrogels for Nerve Protection and Regeneration after Spinal Cord Injury: Systematic Review and Meta-Analysis. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 2(2), 106-116.

[37] Mehrasa, P and Eghdam Zamiri, R. (2026). Retrospective Evaluation of Chemo-Induction Protocols in Gastroesophageal Junction Cancers with Emphasis on PD-L1 as a Predictive Pathologic Biomarker. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 2(4), 276-284.

[38] Moghadam, A. M. (2025). Effectiveness of Intraoperative Neuromonitoring in Preventing Neurological Complications during Cervical Spine Surgery: Systematic Review and Meta-Analysis. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research (JAMPBR)*, 1(11), 403-411.

[39] Hosseini Kasnavieh SM, Milanifard M, Sedighi M, Hessam R. (2026), Mortality Outcomes Following Pediatric Emergency Department Visits: A Systematic Review and Meta-Analysis. *Med J Islam Repub Iran*. (13 Jun); 40:61.

[40] Hosseini Kasnavieh SM, Milanifard M, Sedighi M, Hessam R. (2026), A Systematic Review of Pain Assessment Tools in Prehospital Emergency Care: Evidence from Observational Studies. *Med J Islam Repub Iran*. (10 Jun); 40:60.

[41] Hessam R, Basir Ghafouri H, Rezai M, Shaker S H, Hosseini Kasnavieh S M, Milanifard M et al . (2026), Hidden and Atypical Presentations of Myocarditis in Sudden Cardiac Death: A Narrative Review. *Med J Islam Repub Iran*; 40 (1) :526-536

[42] Sajed M, Alvandifar S, Mallahi M. (2024), Evaluation of Root Canal Morphology of Mandibular Premolars Using Cone-beam Computed Tomography in Golestan Province, North of Iran. *Iran Endod J*. 2024;19(3):183-188.

[43] Fakhari, S, Bilehjani, A and Bilehjani, E. (2026). Effect of Intraoperative Intravenous Dexmedetomidine on Oxidative Stress in Patients

[Undergoing Cardiac Surgery: A Systematic Review. Medicinal, Psychological, and Health Research Journal \(MPHRJ\), 2\(5\), 351-359.](#)

[44] Fakhari, S, Bilehjani, A and Bilehjani, E. (2026). [Intravenous Dexmedetomidine in Cardiomyocyte Biology: A Systematic Review. Medicinal, Psychological, and Health Research Journal \(MPHRJ\), 2\(5\), 360-371.](#)

[45] Lotfi, A R and Nouri Bayat, L. (2026). [Incidence of Postoperative Complications Following Nasal Fracture Surgery in Adults. Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research \(JAMPBR\), 2\(2\), 175-182.](#)

[46] Rezaei, M and Owaysee Osquee, H. (2026). [Prevalence of Deep Vein Thrombosis in Patients with COVID 19 Admitted to the Intensive Care Unit. Medicinal, Psychological, and Health Research Journal \(MPHRJ\), 2\(4\), 267-275.](#)

[47] Sadeghzadeh, A. (2026). [Evaluating the effectiveness and safety of hyaluronic acid versus poly-L-lactic acid for facial volume restoration: A Systematic Review and Meta-analysis. Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research \(JAMPBR\), 2\(2\), 102-113.](#)

[48] Shiri, H and Ashrafi, N. (2026). [Enhanced Recovery After Thoracotomy in the Intensive Care Unit: Current Evidence, Clinical Strategies, and Future Perspectives. Medicinal, Psychological, and Health Research Journal \(MPHRJ\), 2\(5\), 372-382.](#)

[49] Shiri, H and Ashrafi, N. (2026). [Post Esophageal Surgery Dysphagia and Nutritional Support in the Intensive Care Unit. Medicinal, Psychological, and Health Research Journal \(MPHRJ\), 2\(5\), 339-350.](#)

[50] Hashemloo, A., Milanifard, M. (2025). [Dermal Fillers: Types, Indications, and Complications Materials de Relleno: Typos, Indications Complications. Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research, 1\(6\), 161-170](#)

[51] Hashemloo, A., Milanifard, M. (2025). [Contouring Plus: A Comprehensive Approach of the Lower Third of the Face with Calcium Hydroxylapatite and Hyaluronic Acid, Medicinal, Psychological, and Health Research Journal, 1\(5\), 143-150](#)

[52] Sunder Raj, K. S. (2004). [Multi-Pressure Surface Condenser Performance Evaluation: A Case Study. Proceedings of the ASME Power Conference.](#)

[53] S Shaghaghi, JK Udupa, Y Tong, Y Akhtar, M Hosseini, M Al-Noury, C Wu, L Xie, Y Hao, S Hassani, S Gogel, DM Biko, OH Mayer, JM McDonough, PJ Cahill, JB Anari, DA Torigian. (2025). [Association between respiratory volumes estimated from free-breathing dynamic MRI and sagittal spinal curvature in pediatric patients with thoracic insufficiency syndrome. Medical Imaging 2025: Image-Guided Procedures, Robotic Interventions, and Modeling. Proceedings, 13408](#)

[54] M. Afrand, (2024), [Systematic review of neurological problems and stress in Alzheimer's patients with drug tips, EJCMR. 3\(5\): 114-128.](#)

[55] Motaqi, M., Ghanjal, A.,& Afrand, M. (2022). [Treadmill exercise and its effect on rehabilitation of patients after ischemic stroke: A narrative study. International Journal of Musculoskeletal pain prevention, 7\(3\),730-740.](#)

[56] Afrand,M.,Rostampur,N.,Pourhadi,z.,Rezanezhad,J.,&Nekouefard,P.(2025). [Epilepsy in children with central nervous system infections: A systematic review and meta-analysis. AFMN Biomedicine; 42\(4\):450-455.](#)

[57] Ahmady H, Afrand M, Motaqi M, Meftahi GH. (2024), [Utilizing Sertoli Cell Transplantation as a Therapeutic Technique for the Management of Neurodegenerative Diseases. Arch Razi Inst. 1;79\(4\):701-710. PMID: 40256588; PMCID: PMC12004054.](#)

[58] Motaqi, M, Afrand, M and Ghanjal, A. (2022). [Care and Management of Pain in Patients with Musculoskeletal Pain during the Covid-19 Epidemic. International Journal of Musculoskeletal Pain Prevention, 7\(1\), 622-626.](#)

[59] Motaqi,M , Ghanjal,A and Afrand,M . (2022). [Treadmill Exercise and its Effect on Rehabilitation of Patients after Ischemic stroke: A Narrative Study. International Journal of Musculoskeletal Pain Prevention, 7\(3\), 730-740.](#)

[60] Motaqi,M , Samadi,M , Ghanjal,A and Afrand,M . (2022). [A New Approach to Evaluate of Musculoskeletal Pain. International Journal of Musculoskeletal Pain Prevention, 7\(4\), 781-787.](#)

[61] Moghadam, A. M., Yousef, R. M., Azizi, R. & Hessam, R. (2027). [Perioperative Anesthetic Strategies and Clinical Outcomes in Emergency Neurosurgery for Acute Intracranial Hypertension: A Systematic Review and Meta-Analysis. Medicinal, Psychological, and Health Research Journal \(MPHRJ\), 3\(1\), 20-38.](#)

[62] Moghadam, A. M., Yousef, R. M., Azizi, R. & Milanifard, M. (2027). [Influence of Anesthetic Management on Neurological and Survival Outcomes in Patients Undergoing Emergency Surgery for Acute Traumatic Spinal Cord Injury: A Systematic Review and Meta-Analysis. Medicinal, Psychological, and Health Research Journal \(MPHRJ\), 3\(1\), 1-19.](#)

[63] Moghadam, A. M., Yousef, R. M., Azizi, R., Hessam, R. & Milanifard, M. (2026). [Anesthetic and Perioperative Predictors of Neurological Deterioration and Mortality Following Emergency Craniotomy for Acute Brain Injury: A Systematic Review and Meta-Analysis. Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research \(JAMPBR\), 2\(6\), 551-564.](#)

[64] Moghadam, A. M., Yousef, R. M., Azizi, R., Hessam, R. & Milanifard, M. (2026). [Early Versus Delayed Surgical Decompression under General Anesthesia in Acute Spinal Cord Compression:](#)

Effects on Neurological Recovery, Complications, and Mortality—A Systematic Review and Meta-Analysis. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research (JAMPBR)*, 2(6), 537-550.

[65] Neishaboury, M., Shokri, S., Kianpour, P., Farhadi, K., Najjari, K., Sharifnia, H., MohammadYousef, R., & Khajavi, M. (2025). The effects of intraperitoneal dexmedetomidine in comparison with ropivacaine in postoperative pain after laparoscopic sleeve gastrectomy: A double-blind, randomized, placebo-controlled, clinical trial. *Obesity Surgery*, 35(6), 2150–2159.

[66] Afrand, M., Rostampur, N., Pourhadi, Z., Rezanezhad, J., & Nekoueifard, P. (2025). Epilepsy in children with central nervous system infections: A systematic review and meta-analysis. *AFMN Biomedicine*, 42(4), 450–455.

[67] Azizi, Robab, & Rana Mohammad Yousef. (2027). Dynamic Causal Modeling of Heart Rate–Blood Pressure Coupling During Induction of Anesthesia: A Systems-Based Approach to Predict Vasopressor Requirements.

[68] Afrand, M. (2024). Systematic review: Systematic review of neurological problems and stress in Alzheimer's patients with drug tips. *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, 3(5), 114–128.

[69] Samimi, A. (2025). Assessment of Risks Arising from Neuropsychological Crises in Cardiac Patients Using FMEA. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research*, 1(7), 196-203.

[70] Samimi, A. (2025). Risk Assessment in Hospitals Using the FMEA Method: A Data-Driven Analysis for Patient Safety Improvement. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research*, 1(6), 180-187.

[71] Samimi, A. (2025). Application of Failure Mode and Effects Analysis (FMEA) for Risk Management in Hospital Laboratories. *Medicinal, Psychological, and Health Research Journal (mphrj)*, 1(5), 160-168.
