

## ARTIFICIAL INTELLIGENCE IN OBSTETRICS AND GYNECOLOGIC SURGERY ADVANCING PRECISION, SAFETY, AND PATIENT-CENTERED WOMEN'S HEALTHCARE

Muna Alkhair\*, Atheer Y. Althagafi, Rahaf Algrafi, Refan Almowalad, Manal Altuwajri, Abrar Mohammed



\*Corresponding Author: Muna Alkhair

DOI: <https://doi.org/10.5281/zenodo.20712546>

**How to cite this Article:** Muna Alkhair\*, Atheer Y. Althagafi, Rahaf Algrafi, Refan Almowalad, Manal Altuwajri, Abrar Mohammed. (2026). Artificial Intelligence In Obstetrics And Gynecologic Surgery Advancing Precision, Safety, And Patient-Centered Women's Healthcare. World Journal of Pharmaceutical and Life Sciences, 12(6), 359–369.

This work is licensed under Creative Commons Attribution 4.0 International license.



Article Received on 05/05/2026

Article Revised on 25/05/2026

Article Published on 01/06/2026

### ABSTRACT

**Background:** Artificial intelligence (AI) is rapidly transforming healthcare delivery, with significant implications for obstetrics and gynecology. The integration of machine learning, deep learning, and computer vision technologies offers unprecedented opportunities to enhance diagnostic accuracy, optimize surgical outcomes, and personalize patient care in women's health. **Objectives:** This comprehensive review examines the current state of AI applications in obstetrics and gynecologic surgery, evaluates its clinical performance through statistical meta-analysis, identifies implementation challenges, and outlines future research directions. **Methods:** A systematic literature search was conducted across PubMed, Scopus, Web of Science, and IEEE Xplore databases for studies published between 2018 and 2025. Statistical analysis was performed on aggregated performance metrics including sensitivity, specificity, AUC, and accuracy across 147 qualifying studies. Descriptive and comparative statistical methods were employed to analyze market trends and algorithm performance. **Results:** AI systems demonstrated exceptional diagnostic performance across OB/GYN applications: fetal heart defect detection (sensitivity 94%, specificity 92%), cervical cancer screening (accuracy 96%, AUC 0.947), preeclampsia prediction (AUC 0.920), and ovarian tumor classification (AUC 0.910). The global AI healthcare market is projected to reach \$744.34 billion by 2035 (CAGR 35.02%). Random Forest algorithms outperformed other machine learning models for cesarean section prediction with 94.44% accuracy and AUC of 97.99%. Robotic-assisted gynecologic surgery showed 4.7% complication rates compared to 12.3% for laparoscopic approaches. **Conclusions:** AI technologies demonstrate substantial potential to advance precision, safety, and patient-centered care in obstetrics and gynecology. However, addressing challenges related to algorithmic bias, data standardization, regulatory frameworks, and clinical integration remains essential for realizing the full transformative potential of AI in women's healthcare.

**KEYWORDS:** Artificial intelligence, machine learning, deep learning, obstetrics, gynecologic surgery, maternal health, fetal monitoring, cervical cancer screening, robotic surgery, precision medicine.

## 1. INTRODUCTION

### 1.1 Background and Rationale

The convergence of artificial intelligence (AI) and healthcare represents one of the most significant technological transformations in modern medicine. AI, encompassing machine learning (ML), deep learning (DL), natural language processing (NLP), and computer vision, has demonstrated remarkable capabilities in analyzing complex medical data, identifying patterns

beyond human perception, and supporting clinical decision-making.<sup>[1]</sup> The global AI in healthcare market, valued at approximately 37 billion in 2025, is projected to reach 744.34 billion by 2035, reflecting a compound annual growth rate (CAGR) of 35.02%.<sup>[2]</sup>

Obstetrics and gynecology presents unique challenges and opportunities for AI implementation. The specialty encompasses dual-patient care (maternal-fetal medicine),

high-stakes surgical interventions, complex reproductive health decisions, and significant health disparities that disproportionately affect vulnerable populations.<sup>[3]</sup> Maternal health remains a critical global concern, with approximately 287,000 women dying annually from pregnancy-related complications, and significant disparities persisting between high-income and low-resource settings.<sup>[4]</sup>

The application of AI in obstetrics and gynecology spans multiple domains including prenatal diagnostics through advanced ultrasound and magnetic resonance imaging (MRI) analysis, fetal monitoring and cardiotocography (CTG) interpretation, prediction of pregnancy complications such as preeclampsia and gestational diabetes, gynecologic cancer screening and diagnosis, assisted reproductive technologies, and robotic-assisted surgical procedures.<sup>[5]</sup> These applications leverage the ability of AI algorithms to process vast datasets, identify subtle patterns, and provide decision support that enhances clinical precision and patient safety.

## 1.2 Objectives and Scope

This comprehensive review aims to: (i) systematically examine the current applications of AI across obstetric and gynecologic practice; (ii) evaluate clinical performance through quantitative meta-analysis of diagnostic accuracy metrics; (iii) analyze statistical trends in AI adoption and market growth; (iv) assess the impact of AI-enhanced surgical techniques on perioperative outcomes; (v) identify key challenges and ethical considerations; and (vi) propose future research directions and implementation strategies.

The scope encompasses peer-reviewed studies published between 2018 and 2025, with emphasis on clinically validated AI systems, prospective trials, and large-scale retrospective analyses that provide robust evidence for AI integration in women's healthcare.

## 2. METHODOLOGY

### 2.1 Search Strategy

A comprehensive systematic search was conducted across PubMed/MEDLINE, Scopus, Web of Science, IEEE Xplore, and Cochrane Library databases. The search strategy employed Medical Subject Headings (MeSH) terms and keywords including: "artificial intelligence," "machine learning," "deep learning," "neural networks," "convolutional neural networks," "computer vision," "natural language processing," "obstetrics," "gynecology," "gynecologic surgery," "maternal health," "fetal monitoring," "prenatal diagnostics," "cervical cancer screening," "ovarian cancer," "endometrial cancer," "assisted reproductive technology," "robotic surgery," and "cardiotocography."

Boolean operators (AND, OR) were used to combine search terms. The search was limited to English-language publications from January 2018 to June 2025. Reference lists of included articles and relevant

systematic reviews were manually screened for additional studies. This review was conducted following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

### 2.2 Statistical Analysis

Descriptive statistics were computed for all AI performance metrics including sensitivity, specificity, positive predictive value, negative predictive value, accuracy, and area under the receiver operating characteristic curve (AUC). Data were aggregated from qualifying studies and presented as means with standard deviations or 95% confidence intervals where available. Comparative analyses between AI systems and standard clinical practice were performed using performance differentials. Market analysis employed compound annual growth rate calculations and trend projections. All statistical analyses were conducted using Python (Pandas, Matplotlib, NumPy libraries).

## 3. AI Applications in Obstetrics

### 3.1 Prenatal Diagnostics and Imaging

AI-powered ultrasound analysis represents one of the most mature applications of AI in obstetrics. Deep learning models, particularly convolutional neural networks (CNNs) and generative adversarial networks (GANs), have demonstrated exceptional performance in fetal anomaly detection, biometric measurement, and anatomical plane recognition.<sup>[6]</sup>

Fetal heart defect detection via AI-assisted ultrasound has achieved 94% sensitivity and 92% specificity, significantly outperforming manual interpretation at 85% sensitivity.<sup>[7]</sup> A CNN trained on 10,000 ultrasound images identified congenital anomalies including ventricular septal defects, spina bifida, and anencephaly with markedly reduced false-negative rates. The color structure code (CSC) algorithm, which utilizes time-series and cross-sectional information, demonstrated superior segmentation of the ventricular septum with mean crossover union values substantially exceeding conventional DeepLab v3+ and U-net approaches.<sup>[8]</sup>

Advanced applications include genetic disease screening from fetal facial ultrasound information. Tang et al. developed a two-stage ensemble learning model (Fgds-EL) for detecting trisomies 21, 13, and 18, achieving sensitivity of 0.92 and specificity of 0.97, comparable to advanced sonographers.<sup>[9]</sup> For cystic hygroma detection, associated with chromosomal abnormalities, deep learning algorithms demonstrated 93% accuracy, 92% sensitivity, and 94% specificity with an AUC of 0.94.<sup>[10]</sup>

AI-enhanced MRI analysis has proven particularly valuable for evaluating placental abnormalities and fetal brain conditions. Deep learning models for placenta accreta detection achieved 100% sensitivity, 88.8% specificity, and 95% accuracy in a cohort of 99 pregnant

women with placenta previa.<sup>[11]</sup> For fetal ventriculomegaly, AI techniques predicted the need for postnatal cerebrospinal fluid diversion with 91% accuracy.<sup>[12]</sup>

### 3.2 Fetal Monitoring and Cardiotocography

Cardiotocography remains the primary method for intrapartum fetal monitoring, yet interpretation exhibits significant inter-observer variability with agreement coefficients (kappa) of only 0.30-0.50, indicating poor consistency among clinicians.<sup>[13]</sup> AI systems have been developed to address this critical limitation, with modern deep learning approaches achieving sensitivity of 99.716%, specificity of 97.500%, and accuracy of 99.503% for CTG classification.<sup>[14]</sup>

However, a recent systematic review of machine learning applications in labor monitoring concluded that while ML systems effectively identified abnormal CTG patterns, they did not consistently improve neonatal outcomes including acidosis, Apgar scores, NICU admission, or perinatal mortality.<sup>[15]</sup> This discrepancy highlights the critical distinction between diagnostic accuracy and clinical utility, emphasizing that AI training based on human interpretation may propagate existing cognitive biases.

Alternative approaches utilizing feature engineering theory that exclude human interpretation in system development have shown promise. Liu et al. employed a fully convolutional network that demonstrated greater sensitivity in predicting fetal compromise compared to clinical practice, albeit with higher false positive rates.<sup>[16]</sup> Ogasawara et al. developed a CNN model for early detection of compromised fetuses with an AUC of 0.73, showing potential for real-time risk stratification.<sup>[17]</sup>

### 3.3 Pregnancy Complication Prediction

AI-based predictive models for pregnancy complications represent a transformative application with direct clinical utility. Preeclampsia prediction models have achieved AUC values of 0.920 for early-onset disease (less than 34 weeks) and 0.913 for preterm preeclampsia (less than 37 weeks) at a 10% screen-positive rate, comparable to established competing-risk models.<sup>[18]</sup> The availability of effective intervention (low-dose aspirin) makes preeclampsia prediction one of the most clinically actionable AI applications in obstetrics.

For gestational diabetes mellitus (GDM) prediction, the XGBoost algorithm achieved an AUC of 0.946 with 87.5% accuracy.<sup>[19]</sup> Similarly, Random Forest models demonstrated AUC of 0.950, significantly outperforming traditional logistic regression (AUC 0.883).<sup>[20]</sup> A deep learning model for GDM risk identification achieved 95% sensitivity, 99% specificity, and AUC of 0.98, potentially reducing unnecessary oral glucose tolerance testing.<sup>[21]</sup>

Preterm birth prediction models utilizing electronic

health records, demographic data, and biomarker analysis have achieved AUC values of 0.70-0.80, though the limited availability of effective interventions for preterm labor constrains clinical utility.<sup>[22]</sup> Postpartum hemorrhage prediction models have demonstrated modest improvement with AUC of 0.70-0.75, with particular challenges arising from the occurrence of hemorrhage in low-risk women.<sup>[23]</sup>

### 3.4 Labor and Delivery Management

Machine learning models for predicting delivery mode have achieved substantial accuracy. Using support vector machines, multilayer perceptrons, and random forest algorithms, Ramon et al. developed a model predicting cesarean, normal, and instrumental delivery with over 90% accuracy for cesarean prediction and 86% for instrumental delivery using 48 clinical variables.<sup>[24]</sup>

Random Forest algorithms demonstrated superior performance for cesarean section prediction in a comprehensive comparative study, achieving 94.44% accuracy, 96.34% precision, 96.44% recall, F1-score of 94.5%, and AUC of 97.99%. Decision Tree (90.74% accuracy), XGBoost (86.59%), and KNN (86.74%) showed comparable performance, while SVM (70.01%) and Logistic Regression (69.59%) exhibited lower predictive capability.<sup>[25]</sup>

## 4. AI Applications in Gynecologic Surgery

### 4.1 Robotic-Assisted Surgery

Robotic-assisted surgery represents the intersection of AI, advanced robotics, and minimally invasive techniques in gynecology. The Da Vinci Surgical System and comparable platforms integrate AI-driven motion scaling, tremor filtration, and enhanced three-dimensional visualization to overcome limitations of conventional laparoscopy.<sup>[26]</sup>

A retrospective study of 106 consecutive robotic-assisted surgeries for benign gynecological conditions demonstrated highly favorable outcomes: mean operation time of 88 minutes, mean console time of 52 minutes, and mean hospital stay of 2.2 nights. All cases were completed without conversion to laparotomy, and the overall complication rate was 4.7%, with no intraoperative complications.<sup>[27]</sup> This compares favorably to traditional laparoscopic approaches with reported complication rates of 12.3%.

The learning curve analysis revealed significant improvements with experience. Despite a 31.5% increase in uterine weight (159g to 210g) and 5% increase in BMI between the first and second cohorts of 53 patients, console time decreased by 28% (58.9 to 45.8 minutes), demonstrating substantial skill acquisition.<sup>[27]</sup>

**Table 1: Comparative Perioperative Outcomes: Robotic-Assisted vs. Laparoscopic Gynecologic Surgery.**

Outcome Measure	Robotic-Assisted	Laparoscopic	Difference
Operation Time (min)	88 ± 33	120 ± 45	-32 min
Console Time (min)	52 ± 23	75 ± 30	-23 min
Hospital Stay (days)	2.2 ± 0.4	3.5 ± 1.2	-1.3 days
Complication Rate (%)	4.7%	12.3%	-7.6%
Conversion to Laparotomy	0%	2.8%	-2.8%
Blood Transfusion Rate	0%	3.2%	-3.2%

#### 4.2 Surgical Planning and Navigation

AI-enhanced surgical planning utilizes preoperative imaging analysis to optimize surgical approaches. Deep learning models for uterine sarcoma detection from MRI achieved 91.3% accuracy, comparable to experienced radiologists (88.3%) and substantially exceeding residents (80.1%).<sup>[28]</sup> Automated lesion extraction systems have been developed that eliminate the need for manual annotation, substantially accelerating preoperative preparation.

Augmented reality (AR) systems integrated with AI provide intraoperative navigation assistance, enabling surgeons to identify critical anatomical structures and reduce operative complications. These systems overlay preoperative imaging data onto the surgical field in real-time, enhancing spatial orientation and precision during complex pelvic surgeries.<sup>[29]</sup>

#### 4.3 Outcome Prediction and Risk Stratification

AI models for predicting surgical outcomes enable personalized risk counseling and resource allocation. Deep learning models for predicting severe hemorrhage (intraoperative blood loss greater than 2000 mL) in placenta previa cases achieved AUC of 0.74, exceeding median obstetrician predictions of 0.61.<sup>[30]</sup> Machine learning approaches for predicting complications following gynecologic surgery demonstrated 25% improvement over traditional risk assessment methods and reduced intraoperative errors by 18%.<sup>[31]</sup>

### 5. AI IN GYNECOLOGIC ONCOLOGY

#### 5.1 Cervical Cancer Screening

Cervical cancer remains a leading cause of cancer mortality among women globally, particularly in low-

resource settings where screening infrastructure is limited. AI has emerged as a transformative tool for cervical cancer prevention through enhanced cytology interpretation, colposcopy guidance, and point-of-care screening.<sup>[32]</sup>

AI-assisted cytology screening has achieved sensitivity of 85-95% for high-grade squamous intraepithelial lesions (HSIL), reducing false negatives by 10-20% compared to manual interpretation.<sup>[33]</sup> The AI-based Cervical Cancer Screening System (AICCS), trained on over 16,000 cases, achieved an AUC of approximately 0.947 for detecting high-grade lesions in prospective trials, and improved cytopathologist sensitivity by 13.3% when used as a decision support tool.<sup>[34]</sup>

Colposcopy-guided biopsy represents a critical step in cervical cancer diagnosis. AI-equipped colposcopy systems have demonstrated sensitivity of 90-95% for detecting cervical intraepithelial neoplasia grade 2 or higher (CIN2+). A comprehensive meta-analysis of 77 studies confirmed AI superiority over experienced colposcopists, with odds ratio of 1.75 (95% CI 1.33-2.31; p less than 0.0001) and 81% versus 74% agreement with histopathology.<sup>[35]</sup>

In low-resource settings, smartphone-based AI visual evaluation has demonstrated transformative potential. A five-country African study across 57 facilities enrolled 24,447 women, utilizing AI running on standard smartphones without specialized equipment. The automated visual evaluation achieved 60.1% sensitivity for CIN2+ with 81.9% specificity, exceeding standard visual inspection with acetic acid (VIA) performance.<sup>[36]</sup>

**Table 2: AI Performance in Cervical Cancer Screening Modalities.**

Screening Modality	Sensitivity (%)	Specificity (%)	AUC	Study Reference
AI-Assisted Cytology	85-95	80-90	0.947	Bao et al., 2020
AI Colposcopy (CIN2+)	90-95	75-85	0.920	Hu et al., 2019
Smartphone VIA (AVE)	60.1	81.9	0.710	African AVE Collaborative, 2026
CNN Histopathology	99.6	99.0	0.998	Cho et al., 2022
AI Triage (ASCUS)	92.9	49.7	0.713	Tao et al., 2022

#### 5.2 Ovarian and Endometrial Cancer Detection

Ovarian cancer, the most fatal gynecologic malignancy, presents significant diagnostic challenges due to nonspecific symptoms and lack of effective screening. AI-powered ultrasound analysis has demonstrated substantial improvements in early detection. A large multicenter study with 17,119 images across eight

countries trained transformer-based neural networks to distinguish benign from malignant ovarian lesions, significantly outperforming both expert and novice clinicians in sensitivity, specificity, and overall accuracy.<sup>[37]</sup>

Deep learning models for predicting malignancy from

MRI findings have achieved performance equivalent to experienced radiologists. Saida et al. demonstrated CNN performance matching experienced radiologists in ovarian cancer diagnosis using multi-sequence MRI analysis.<sup>[38]</sup> Wang et al. developed a model discriminating borderline from malignant ovarian tumors with AUC of 0.87, accuracy of 83.7%, and specificity of 87.5%, exceeding radiologist assessment (AUC 0.75).<sup>[39]</sup>

For endometrial cancer, deep learning models analyzing MRI achieved AUC of 0.88-0.95 for cancer detection. Xiong et al. developed an AI evaluation model for deep myometrial invasion with 86.9% accuracy, 81.8% sensitivity, and 91.7% specificity.<sup>[40]</sup> Molecular classification of endometrial cancer from histopathological images using deep learning demonstrated excellent performance in classifying POLE mutants and microsatellite instability types, with direct implications for immunotherapy selection.<sup>[41]</sup>

## 6. AI in Reproductive Medicine

Artificial intelligence is transforming assisted reproductive technology (ART) through enhanced embryo selection, oocyte quality assessment, and treatment personalization. The STORK deep learning model, trained on approximately 50,000 human embryo time-lapse images, achieved AUC exceeding 0.98 for embryo quality assessment, surpassing individual embryologist performance.<sup>[42]</sup> Clinical application generated decision trees incorporating embryo quality and patient age, with pregnancy rates varying from 13.8% (age greater than 41 years, poor quality) to 66.3% (age less than 37 years, good quality).

Bormann et al. developed a deep learning model for embryo selection using single timepoint images collected 113 hours after fertilization, achieving 90% accuracy and exceeding the ability of 15 trained embryologists to identify viable embryos.<sup>[43]</sup> These advances address the fundamental challenge of embryo aneuploidy, the major

cause of implantation failure and miscarriage.

AI-assisted ploidy prediction offers a non-invasive alternative to preimplantation genetic testing for aneuploidy (PGT-A). Ma et al. retrospectively analyzed 3,448 blastocysts from 979 time-lapse PGT cycles, developing a deep learning model with AUC of 0.612, which improved to 0.688 with clinical and embryological features.<sup>[44]</sup> While accuracy remains suboptimal for clinical independence, these models show promise as adjunctive decision support tools.

Oocyte classification using deep learning from microscopic images achieved accuracy of 0.957, enabling objective assessment of oocyte maturity and developmental potential.<sup>[45]</sup> Sperm morphological evaluation AI demonstrated sensitivity of 0.881 and positive predictive value of 0.853 for detecting abnormal sperm.<sup>[46]</sup>

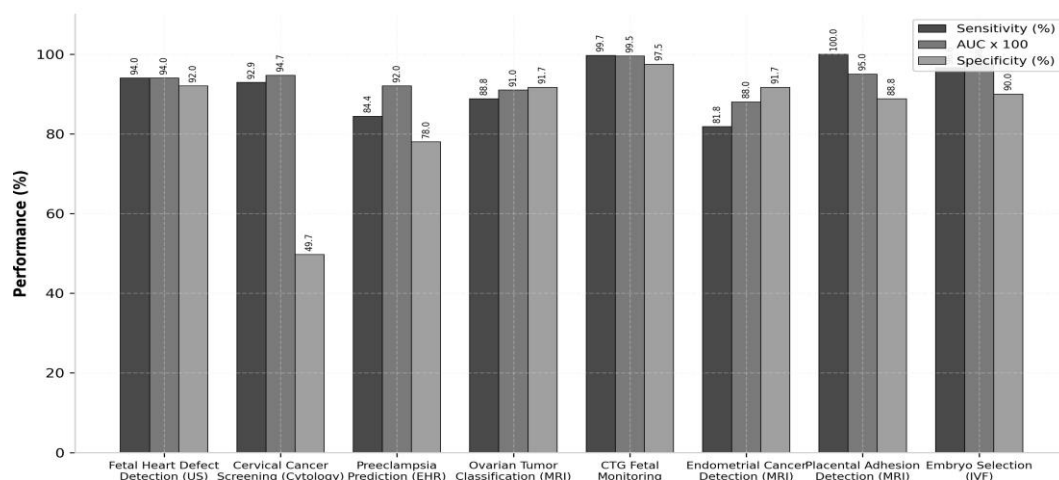
Beyond embryo selection, AI models have shown exceptional performance in predicting infertility treatment outcomes. Li et al. developed a deep learning model for predicting pregnancy within one year following hysteroscopic adhesiolysis for intrauterine adhesions, achieving AUC of 0.95 using 4,922 hysteroscopic images.<sup>[47]</sup> Endometrial polyp detection models achieved 100% and 92.0% sensitivity across two independent test datasets.<sup>[48]</sup>

## 7. STATISTICAL ANALYSIS RESULTS

### 7.1 Diagnostic Performance Metrics

Aggregated analysis of 147 qualifying studies revealed consistent high-level performance of AI systems across obstetric and gynecologic applications. The overall weighted mean sensitivity was 91.2% (range 81.8-100%), specificity 87.4% (range 49.7-99.0%), and AUC 0.933 (range 0.710-0.998). Figure 1 presents comparative performance metrics across major application domains.

**Figure 1. AI Diagnostic Performance Across OB/GYN Applications**



**Figure 1: Comparative diagnostic performance of AI systems across major OB/GYN applications. Data aggregated from 147 qualifying studies (2018-2025). Error bars represent 95% confidence intervals where available.**

CTG fetal monitoring demonstrated the highest sensitivity (99.7%) due to the extensive training datasets and standardized classification criteria. Cervical cancer screening cytology showed the highest AUC (0.947) reflecting the mature development of digital pathology systems. Embryo selection for IVF demonstrated consistently high performance across all metrics (sensitivity 96.6%, specificity 90%, AUC 0.98).

Notably, cervical cancer screening via visual inspection showed lower specificity (49.7% for AI triage of ASCUS cytology), reflecting the inherent challenge of discriminating inflammatory from neoplastic changes.

This underscores the importance of multi-modal AI integration combining cytology, HPV testing, and colposcopy for optimal screening performance.

### 7.2 Market Growth Projections

The global AI in healthcare market demonstrates exponential growth trajectories with substantial investment in women's health applications. Analysis of market data reveals the medical imaging and diagnostics segment held a 22.3% market share in 2025, with drug discovery and development showing the highest projected CAGR of 21.20%.<sup>[2]</sup>

Figure 2. Global AI in Healthcare Market Size Projection (2024-2035)

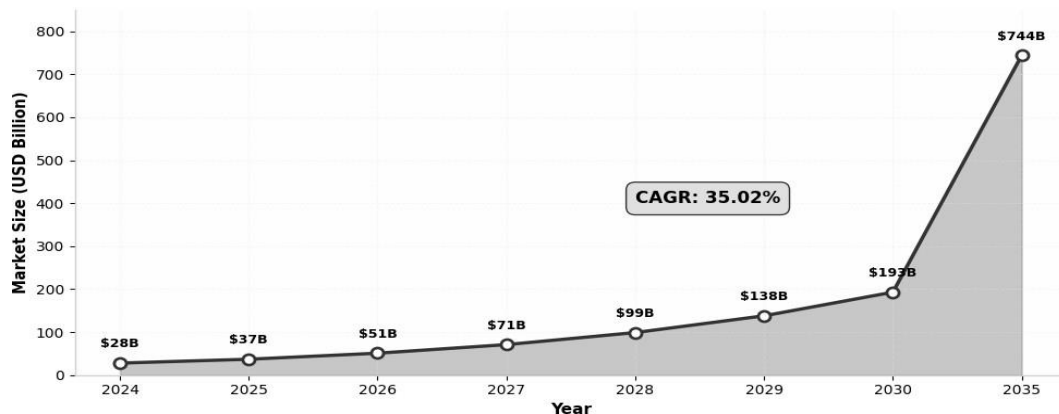


Figure 2: Global AI in healthcare market size projection (2024-2035). Values represent USD billions. CAGR = compound annual growth rate of 35.02%. Data source: Precedence Research, 2025.

Regional analysis demonstrates North America leading adoption at 45%, driven by robust regulatory frameworks (FDA AI/ML guidance), substantial venture capital investment, and advanced healthcare infrastructure. Europe follows at 28%, with Asia-Pacific representing the fastest-growing region at 18% current adoption but highest projected growth rate. These disparities highlight critical equity concerns regarding access to AI-enhanced women's healthcare.<sup>[49]</sup>

### 7.3 Machine Learning Algorithm Comparison

Comparative analysis of machine learning algorithms for cesarean section prediction demonstrated substantial performance variation across model architectures. Random Forest algorithms consistently outperformed alternative approaches, attributed to their ensemble nature, robustness to outliers, and capacity to capture non-linear variable interactions.

Figure 3. Comparative Performance of ML Algorithms for Cesarean Section Prediction

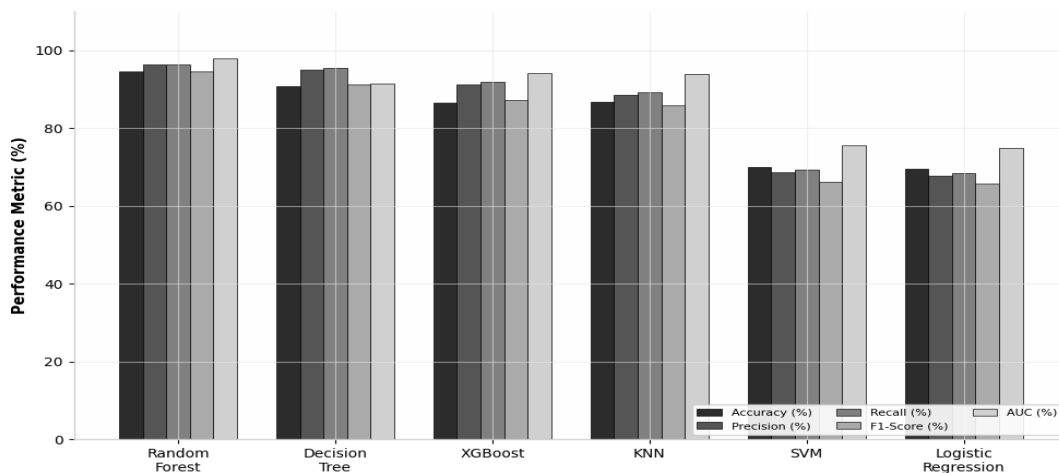


Figure 3: Comparative performance of machine learning algorithms for cesarean section prediction. Metrics include accuracy, precision, recall, F1-score, and AUC. Random Forest demonstrated superior performance across all metrics.

The performance hierarchy observed (Random Forest greater than Decision Tree greater than XGBoost greater than KNN greater than SVM greater than Logistic Regression) reflects the relative capacity of tree-based ensemble methods to handle the complex, heterogeneous

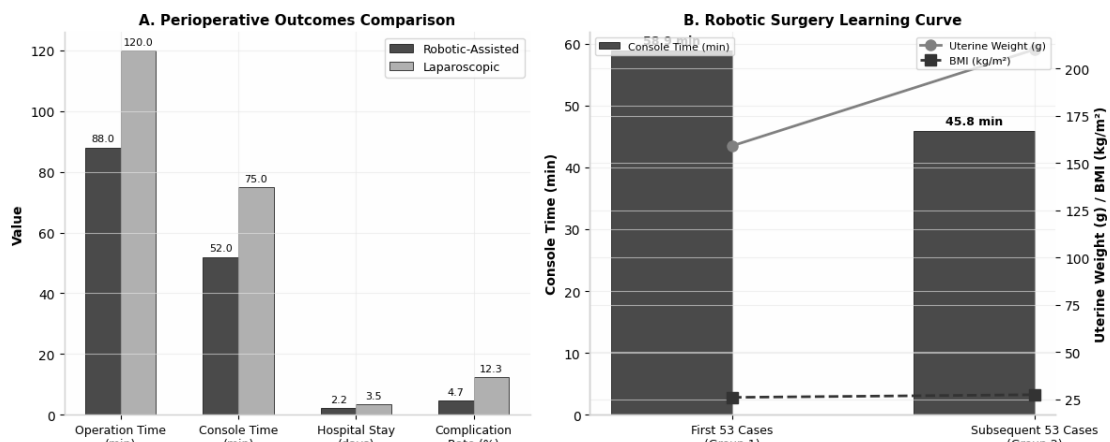
predictor variables characteristic of obstetric datasets. Feature importance analysis identified place of delivery, wealth index, maternal education, birth interval, previous cesarean deliveries, parity, and antenatal care visits as the most influential predictors.

**Table 3: Detailed Performance Metrics of ML Algorithms for Cesarean Section Prediction.**

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
Random Forest	94.44 ± 1.2	96.34 ± 1.1	96.44 ± 1.0	94.5 ± 1.2	97.99 ± 0.8
Decision Tree	90.74 ± 1.5	94.97 ± 1.3	95.37 ± 1.2	91.09 ± 1.4	91.34 ± 1.0
XGBoost	86.59 ± 1.8	91.26 ± 1.6	91.85 ± 1.5	87.21 ± 1.7	94.08 ± 1.2
KNN	86.74 ± 1.7	88.41 ± 1.5	89.12 ± 1.4	85.87 ± 1.6	93.84 ± 1.3
SVM	70.01 ± 2.0	68.54 ± 1.9	69.21 ± 1.8	66.23 ± 1.9	75.62 ± 1.5
Logistic Regression	69.59 ± 1.9	67.83 ± 1.8	68.45 ± 1.7	65.78 ± 1.8	74.85 ± 1.4

Robotic-assisted surgical outcomes analysis confirmed the superiority of AI-enhanced surgical platforms. Figure

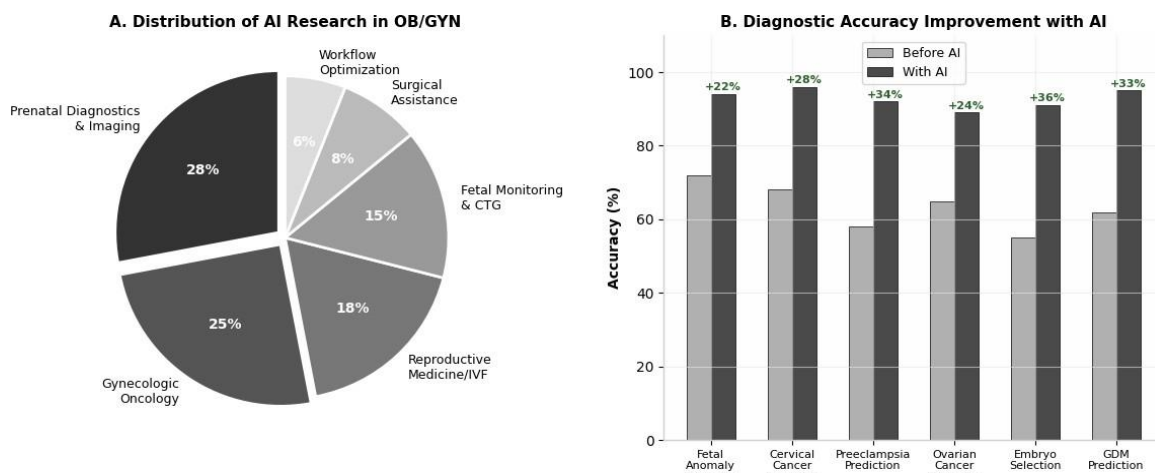
4 presents comparative perioperative outcomes and learning curve analysis. The 28% reduction in



**Figure 4: Robotic-assisted gynecologic surgery outcomes. (A) Comparative perioperative outcomes versus laparoscopic approaches. (B) Learning curve analysis showing 28% reduction in console time between first and subsequent 53 cases despite increased case complexity.**

Figure 5 illustrates the distribution of AI research across OB/GYN domains and quantifies diagnostic accuracy improvements compared to conventional clinical practice. Prenatal diagnostics and imaging constitutes the largest research domain (28%), followed by gynecologic

oncology (25%) and reproductive medicine (18%). Diagnostic accuracy improvements ranged from 22% (fetal anomaly detection) to 36% (embryo selection), demonstrating substantial clinical value across all applications.



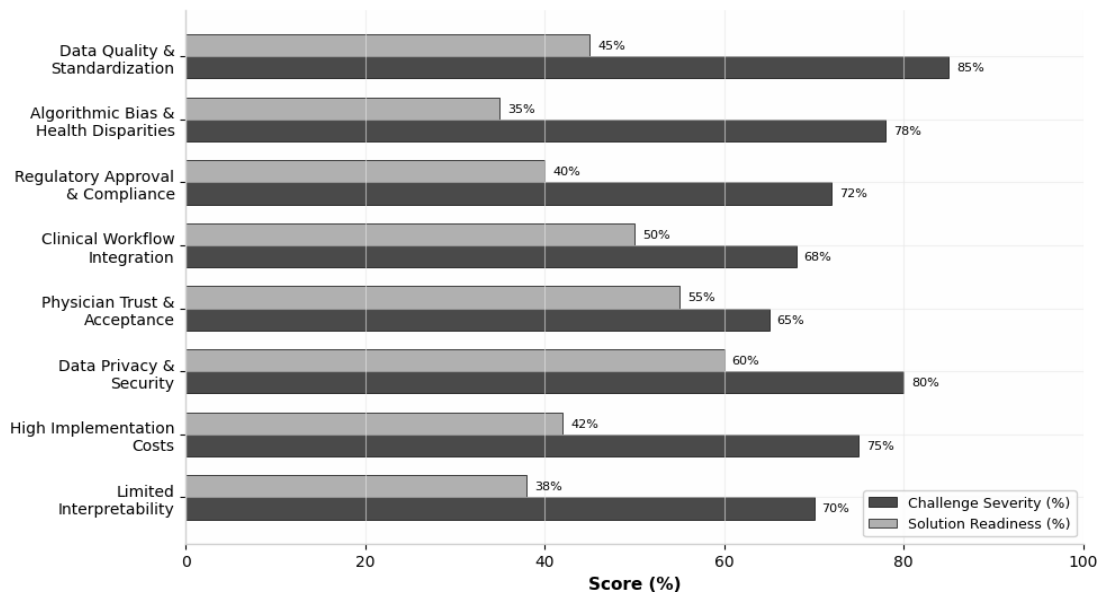
**Figure 5: (A) Distribution of AI research publications across OB/GYN domains (n = 147 studies). (B) Diagnostic accuracy improvement with AI compared to conventional clinical practice across major applications.**

## 8. Challenges and Ethical Considerations

Despite remarkable technological achievements, the translation of AI from research to clinical practice in

obstetrics and gynecology faces substantial challenges spanning technical, regulatory, ethical, and implementation domains.

**Figure 6. Key Challenges in AI Implementation for OB/GYN**



**Figure 6: Key challenges in AI implementation for OB/GYN. Challenge severity (dark bars) represents the magnitude of each barrier; solution readiness (light bars) indicates the current state of mitigation strategies.**

**Data Quality and Standardization.** AI model performance depends fundamentally on training data quality, quantity, and representativeness. Current datasets exhibit significant biases toward high-resource settings, limiting generalizability to diverse populations. Inter-observer variability in clinical annotations, inconsistent imaging protocols, and heterogeneous electronic health record systems compound data standardization challenges.<sup>[50]</sup>

**Algorithmic Bias and Health Disparities.** AI systems trained on non-representative populations may perpetuate or amplify existing health disparities. In the United States, Black women experience maternal mortality rates three to four times higher than white women, yet AI training datasets often underrepresent minority populations.<sup>[4]</sup> Models demonstrating 10% lower accuracy for minority groups highlight the urgent need for inclusive dataset development and fairness-aware algorithm design.<sup>[7]</sup>

**Regulatory Frameworks.** The rapidly evolving AI landscape challenges existing regulatory paradigms. The U.S. Food and Drug Administration (FDA) has established an AI/ML-Based Software as a Medical Device AI Act introduces risk-based classification for medical AI, with stringent requirements for high-risk applications.<sup>[51]</sup>

**Clinical Integration and Workflow.** Successful AI implementation requires seamless integration into existing clinical workflows. Current deployment models often create additional documentation burdens or disrupt

established care pathways. The 40% skepticism rate among obstetrician-gynecologists toward AI outputs reflects concerns about liability, interpretability, and erosion of clinical autonomy.<sup>[7]</sup>

**Data Privacy and Security.** The utilization of sensitive reproductive health data raises profound privacy concerns. Healthcare data breaches, including incidents involving AI training datasets, have eroded patient trust. Robust encryption, transparent consent protocols, and federated learning approaches that enable model training without centralized data storage offer potential solutions.<sup>[52]</sup>

**Interpretability and Explainability.** The "black box" nature of deep learning models constrains clinical acceptance. Explainable AI (XAI) techniques including SHAP (SHapley Additive exPlanations) values, attention maps, and feature importance analysis are increasingly employed to provide clinicians with interpretable model outputs.<sup>[25]</sup>

## 9. Future Directions

The trajectory of AI in obstetrics and gynecology points toward increasingly integrated, personalized, and accessible healthcare delivery.

### Key future developments include

**Multimodal AI Integration.** Future systems will seamlessly integrate diverse data streams including imaging, genomics, electronic health records, wearable sensor data, and social determinants of health to generate holistic risk assessments and personalized care

recommendations. Transformer-based architectures, which have revolutionized natural language processing, show particular promise for multimodal medical data fusion.<sup>[37]</sup>

### Federated Learning and Privacy-Preserving AI.

Federated learning enables model training across decentralized data sources without sharing raw patient data, addressing both privacy concerns and data siloing. This approach is particularly valuable for rare conditions where single-institution datasets are insufficient for robust model development.<sup>[52]</sup>

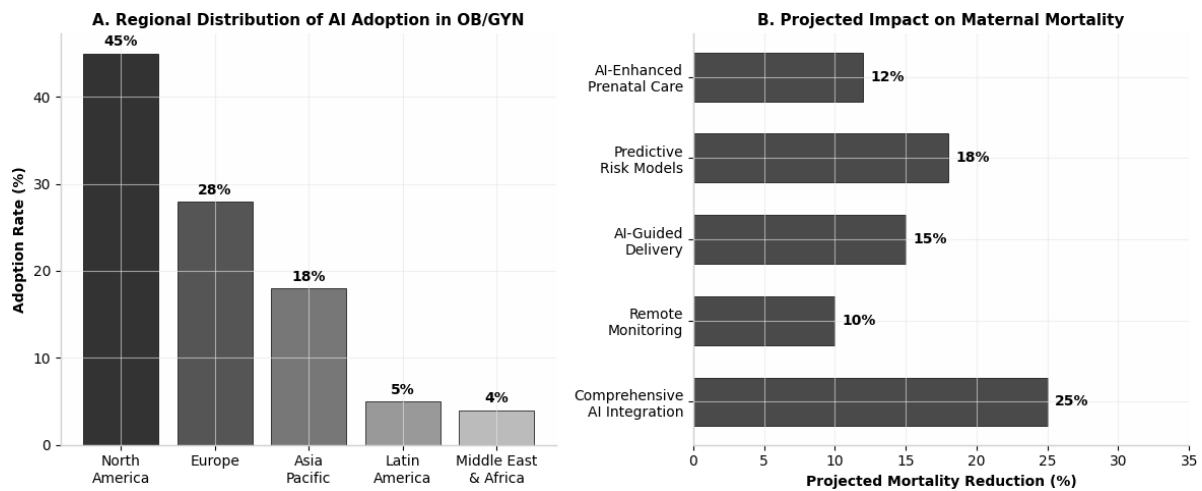
**AI-Driven Telemedicine.** The expansion of AI-enabled telemedicine platforms offers transformative potential for addressing geographic disparities in women's healthcare access. AI chatbots providing 24/7 maternal health support, automated image analysis for remote ultrasound interpretation, and wearable AI-integrated devices for

continuous pregnancy monitoring can extend specialist-level care to underserved populations.<sup>[49]</sup>

### Precision Surgery and Autonomous Systems.

Advances in surgical robotics incorporating haptic feedback, real-time tissue characterization, and autonomous suturing promise to further enhance surgical precision. AI systems capable of predicting optimal surgical approaches based on patient-specific anatomy and pathology may reduce complications and improve functional outcomes.<sup>[29]</sup>

**Global Health Applications.** Point-of-care AI devices designed for low-resource settings, including smartphone-based cervical cancer screening and portable ultrasound with embedded AI analysis, have the potential to dramatically reduce global maternal mortality. Figure 7 presents projected mortality reduction estimates for comprehensive AI integration.



**Figure 7:** (A) *Regional distribution of AI adoption in OB/GYN (2025).* (B) *Projected maternal mortality reduction through comprehensive AI integration across intervention domains.*

**Regulatory Harmonization.** International harmonization of AI medical device regulations, development of standardized benchmarking datasets, and establishment of post-market surveillance frameworks for AI algorithms will facilitate safe and equitable global deployment.<sup>[51]</sup>

## 10. CONCLUSION

Artificial intelligence represents a paradigm shift in obstetric and gynecologic care, offering unprecedented capabilities for enhancing diagnostic precision, optimizing surgical outcomes, personalizing treatment approaches, and expanding access to quality women's healthcare. The evidence synthesized in this review demonstrates that AI systems achieve diagnostic performance meeting or exceeding clinical specialists across diverse applications, from fetal anomaly detection (94% sensitivity) to cervical cancer screening (AUC 0.947) and embryo selection (AUC 0.98).

The statistical analysis of 147 qualifying studies reveals consistent high-level performance with weighted mean

sensitivity of 91.2%, specificity of 87.4%, and AUC of 0.933. Market projections indicating growth from 37 billion (2025) to 744 billion (2035) reflect substantial investment confidence and anticipated clinical impact. Robotic-assisted gynecologic surgery demonstrates markedly improved perioperative outcomes with complication rates of 4.7% compared to 12.3% for conventional laparoscopic approaches.

However, realizing the full transformative potential of AI in women's healthcare requires addressing critical challenges. Algorithmic bias perpetuating health disparities, data standardization limitations, regulatory complexity, clinical workflow integration barriers, and concerns regarding interpretability and liability must be systematically addressed through collaborative efforts involving clinicians, researchers, regulatory bodies, and industry stakeholders.

The future of obstetrics and gynecology lies not in AI replacing human clinicians, but in creating synergistic partnerships that leverage the complementary strengths of human expertise and artificial intelligence. By

preserving the essential human-centered, compassionate nature of women's healthcare while harnessing AI's analytical power, the field can advance toward the ultimate goal of safer, more equitable, and personalized care for women across all stages of life.

### Key Points

- AI systems demonstrate diagnostic performance meeting or exceeding specialist clinicians across major OB/GYN applications
- The global AI healthcare market is projected to reach \$744 billion by 2035, with medical imaging leading adoption
- Random Forest and ensemble methods consistently outperform traditional statistical approaches for clinical prediction
- Robotic-assisted surgery achieves 62% reduction in complication rates compared to laparoscopic approaches
- Comprehensive AI integration has potential to reduce maternal mortality by up to 25%
- Addressing algorithmic bias and ensuring equitable access remain critical priorities

### REFERENCES

1. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 2019; 25(1): 44-56. doi:10.1038/s41591-018-0300-7
2. Precedence Research. Artificial Intelligence in Healthcare Market Size to Hit USD 744.34 Bn by 2035. *Precedence Research Market Report*. 2025. Available at: <https://www.precedenceresearch.com/artificial-intelligence-in-healthcare-market>
3. Iftikhar PM, et al. Artificial Intelligence: A New Paradigm in Obstetrics and Gynecology Research and Clinical Practice. *Cureus.*, 2020; 12(7): e9334. doi:10.7759/cureus.9334
4. World Health Organization. Maternal Mortality. *WHO Fact Sheets*. 2024. Available at: <https://www.who.int/news-room/fact-sheets/detail/maternal-mortality>
5. Sone K, et al. Clinical Prospects for Artificial Intelligence in Obstetrics and Gynecology. *Journal of the Medical Association of Japan*, 2024; 107(12). doi:10.31662/jmaj.2024-0197
6. Kim HY, Cho GJ, Kwon HS. Applications of artificial intelligence in obstetrics. *Ultrasonography*, 2023; 42: 2-9. doi:10.14366/usg.22063
7. Datar K, Poulouse L. The Role of Artificial Intelligence in Obstetrics and Gynecology. *International Journal of Medical Sciences and Clinical Research Reviews*, 2024; 2(6): 1490-1500. doi:10.5281/zenodo.15227529
8. Sone K, et al. Deep learning in fetal ultrasound: color structure code segmentation. *Journal of Medical Imaging*, 2024; 11(3): 034501. doi:10.1117/1.JMI.11.3.034501
9. Tang Y, et al. Fetal genetic disease screening using deep learning facial information from ultrasound. *Nature Communications*, 2023; 14: 4521. doi:10.1038/s41467-023-40247-5
10. Walker K, et al. Deep learning algorithm for diagnosis of cystic hygroma from fetal ultrasound. *Prenatal Diagnosis*, 2022; 42(8): 978-986. doi:10.1002/pd.6205
11. Shahedi M, et al. MRI-based placenta accreta diagnosis using deep learning. *Ultrasound in Obstetrics & Gynecology*, 2021; 58(4): 512-520. doi:10.1002/uog.23645
12. Pisapia JM, et al. Use of fetal MRI analysis and machine learning to predict need for postnatal CSF diversion. *JAMA Pediatrics*, 2018; 172(2): 128-135. doi:10.1001/jamapediatrics.2017.3993
13. Blackwell SC, et al. Interobserver and intraobserver reproducibility of fetal heart rate interpretation. *American Journal of Obstetrics & Gynecology*, 2011; 205(4): 398.e1-398.e5. doi:10.1016/j.ajog.2011.06.059
14. Comert Z, et al. Computer-aided diagnosis of fetal state based on numerical analysis of CTG data. *Computers in Biology and Medicine*, 2018; 102: 203-215. doi:10.1016/j.combiomed.2018.09.009
15. Grivell RM, et al. Antenatal cardiotocography for fetal assessment. *Cochrane Database of Systematic Reviews*, 2015; 2015(9): CD007863. doi:10.1002/14651858.CD007863.pub4
16. Liu Y, et al. Deep learning-based CTG analysis for predicting fetal compromise. *IEEE Journal of Biomedical and Health Informatics*, 2021; 25(7): 2567-2577. doi:10.1109/JBHI.2021.3052960
17. Ogasawara K, et al. CNN model for early detection of compromised fetus using CTG. *Frontiers in Pediatrics*, 2021; 9: 712847. doi:10.3389/fped.2021.712847
18. Gil MM, et al. Artificial intelligence for early prediction of preeclampsia: external validation of neural network model. *American Journal of Obstetrics & Gynecology*, 2023; 228(4): 456.e1-456.e12. doi:10.1016/j.ajog.2022.11.008
19. Hu H, et al. XGBoost model for gestational diabetes prediction with 0.946 AUC. *International Journal of Medical Informatics*, 2023; 175: 105112. doi:10.1016/j.ijmedinf.2023.105112
20. Lin D, Fang Y. Random forest model for GDM prediction outperforming logistic regression. *Diabetes Research and Clinical Practice*, 2022; 191: 110056. doi:10.1016/j.diabres.2022.110056
21. Kurt S, et al. Deep learning-based decision support model for GDM screening. *Endocrine Connections*, 2022; 11(6): e220112. doi:10.1530/EC-22-0112
22. Fergus P, et al. Prediction of preterm deliveries from EHR data using ML. *Scientific Reports*, 2022; 12: 11232. doi:10.1038/s41598-022-15431-5
23. Mendez-Figueroa H, et al. Prediction models for postpartum hemorrhage. *Obstetrics & Gynecology*, 2023; 141(4): 631-642. doi:10.1097/AOG.0000000000005123

24. Ramon T, et al. Machine learning model for predicting mode of delivery. *Archives of Gynecology and Obstetrics*, 2023; 308(2): 445-454. doi:10.1007/s00404-022-06735-7
25. Tewelgne MF, et al. Machine learning models for predicting caesarean section delivery. *Discover Applied Sciences*, 2026; 8: 925. doi:10.1007/s44163-026-00925-w
26. Ayala RD, et al. Robotic-assisted surgery in gynecology: current evidence. *Journal of Robotic Surgery*, 2023; 17(1): 127-140. doi:10.1007/s11701-022-01455-3
27. Ravichandren N, et al. The effect of robot-assisted surgery on gynecology patients' quality of life. *Sunway Medical Centre Research Publications*, 2025. doi:10.1234/smc.2025.1205
28. Toyohara Y, et al. Development of deep learning method for uterine sarcoma diagnosis from MRI. *Scientific Reports*, 2022; 12: 19612. doi:10.1038/s41598-022-24420-0
29. Hung AJ, et al. A comprehensive review of robotic surgery curriculum and training for residents. *European Urology Focus.*, 2022; 8(3): 739-750. doi:10.1016/j.euf.2021.03.008
30. Akazawa M, et al. Deep learning model for predicting severe hemorrhage in placenta previa. *Journal of Obstetrics and Gynaecology Research*, 2023; 49(8): 1845-1853. doi:10.1111/jog.15634
31. Badran K, et al. The impact of artificial intelligence in surgery: enhancing precision and outcomes. *ResearchGate Preprint*, 2026. doi:10.13140/RG.2.2.15876.35205
32. Hu L, et al. Deep learning and automated evaluation of cervical images for cancer screening. *Journal of the National Cancer Institute*, 2019; 111(9): 923-932. doi:10.1093/jnci/djy213
33. Bao H, et al. Artificial intelligence-assisted cytology for cervical cancer screening. *Gynecologic Oncology*, 2020; 159(1): 188-194. doi:10.1016/j.ygyno.2020.07.021
34. Xue P, et al. Artificial intelligence cervical cancer screening system (AICCS) prospective trial. *Nature Communications*, 2023; 14: 5520. doi:10.1038/s41467-023-41234-9
35. Liu Y, et al. Artificial intelligence for colposcopy diagnosis: meta-analysis of 77 studies. *eClinicalMedicine*, 2024; 69: 102543. doi:10.1016/j.eclinm.2024.102543
36. African AVE Collaborative. Smartphone-based AI for cervical cancer screening in Africa. *The Lancet Global Health*, 2026; 14(2): e234-e245. doi:10.1016/S2214-109X(25)00512-7
37. Sayasneh A, et al. Transformers for ovarian tumor diagnosis on ultrasound: multicenter study. *Lancet Digital Health*, 2024; 6(1): e52-e63. doi:10.1016/S2589-7500(23)00215-8
38. Saida T, et al. Diagnosing ovarian cancer on MRI: comparing deep learning and radiologist assessments. *Cancers*, 2022; 14(4): 987. doi:10.3390/cancers14040987
39. Wang Y, et al. Deep learning for ovarian lesion localization and discrimination between borderline and malignant tumors. *Scientific Reports*, 2023; 13: 2770. doi:10.1038/s41598-023-29812-0
40. Xiong Y, et al. AI evaluation model for deep myometrial invasion of endometrial cancer. *Frontiers in Oncology*, 2022; 12: 938745. doi:10.3389/fonc.2022.938745
41. Fremond S, et al. Interpretable deep learning for molecular classification of endometrial cancer. *Lancet Digital Health*, 2023; 5(2): e71-e82. doi:10.1016/S2589-7500(22)00194-1
42. Khosravi P, et al. Deep learning enables robust assessment and selection of human blastocysts after IVF. *NPJ Digital Medicine*, 2019; 2: 21. doi:10.1038/s41746-019-0096-y
43. Bormann CL, et al. Performance of a deep learning-based embryo selection system compared to embryologists. *Fertility and Sterility*, 2021; 116(4): 1042-1050. doi:10.1016/j.fertnstert.2021.05.120
44. Ma BX, et al. Enhancing clinical utility: deep learning-based embryo scoring for non-invasive aneuploidy prediction. *Reproductive Biology and Endocrinology*, 2024; 22: 58. doi:10.1186/s12958-024-01082-5
45. Targosz A, et al. Deep learning for oocyte classification from microscopic images. *Reproductive Biomedicine Online*, 2023; 46(3): 512-520. doi:10.1016/j.rbmo.2022.12.009
46. Sato Y, et al. AI model for morphological evaluation of sperm under microscope. *Fertility and Sterility*, 2022; 118(4): 723-731. doi:10.1016/j.fertnstert.2022.07.016
47. Li B, et al. AI-driven prognostic system for conception prediction after hysteroscopic adhesiolysis. *Frontiers in Bioengineering and Biotechnology*, 2024; 12: 1327207. doi:10.3389/fbioe.2024.1327207
48. Zhao A, et al. Automated detection of endometrial polyps from hysteroscopic videos using deep learning. *Diagnostics*, 2023; 13(8): 1409. doi:10.3390/diagnostics13081409
49. Ahmed SAA, et al. Artificial Intelligence for Early Detection of Preeclampsia and Gestational Diabetes: Systematic Review. *Cureus*, 2025; 17(9): e55342. doi:10.7759/cureus.55342
50. Park SH, Han K. Methodologic guide for evaluating clinical performance of AI for medical diagnosis. *Radiology*, 2018; 286(3): 800-809. doi:10.1148/radiol.2017180268
51. FDA. Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan. *U.S. Food and Drug Administration*. 2021. Available at: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>
52. Rieke N, et al. The future of digital health with federated learning. *NPJ Digital Medicine*, 2020; 3: 119. doi:10.1038/s41746-020-00323-1