International Journal of Women's Health

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Prediction of Recurrence After Endometrial Polypectomy in Women of Childbearing Age Based on Machine Learning Algorithm

Xiangyu Chen, Guange Chen, Dandan Bao, Jianfei Liu, Xinxin Zhang

Department of Obstetrics and Gynecology, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, People's Republic of China

Correspondence: Xinxin Zhang, Department of Obstetrics and Gynecology, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, No. 306 Hualongqiao Road, Lucheng District, Wenzhou, Zhejiang, 325000, People's Republic of China, Tel +86-577-88002232; +86-150-6781-0815, Email Zhxinxin1122@163.com

Objective: Research on machine learning (ML) prediction models for recurrence after endometrial polypectomy remains unexplored. In this study, we aim to establish an ML-driven model based on multi-biomarkers to predict recurrence after endometrial polypectomy.
Methods: A retrospective cohort of 606 patients who underwent endometrial polypectomy was analyzed. Variables including age, BMI, reproductive history, prior uterine surgery, coexisting uterine disorders, polyp characteristics, and routine blood indices were extracted. Seven ML algorithms (logistic regression, SVM, MLP, random forest, KNN, Naïve Bayes, and decision tree) were trained using 10-fold cross-validation. Performance was evaluated by AUC, accuracy, sensitivity, specificity, PPV, NPV, and F1 score.
Results: Of the 606 patients, 179 (29.5%) developed recurrence within one year postoperatively. The cohort was randomly divided into a training set (n = 424) and a validation set (n = 182). In the training set, the random forest (RF) algorithm achieved the best performance (AUC = 0.838, accuracy = 79.5%, specificity = 0.930, F1 score = 0.576). In the validation set, RF remained superior (AUC = 0.760, accuracy = 75.3%, specificity = 0.875, F1 score = 0.526), underscoring its strong generalizability. SHAP analysis identified age, posterior-wall polyp location, prior uterine surgery, histopathological subtype, and hemoglobin level as the most influential predictors of recurrence.
Conclusion: The RF-based model, using demographic, clinical, and hematologic features, showed high accuracy in predicting recurrence risk after endometrial polypectomy. This interpretable ML framework can help clinicians identify high-risk patients early and personalize postoperative surveillance.
Keywords: machine learning, endometrial polypectomy, prediction model, recurrence

Introduction

Endometrial polyps (EPs) are benign proliferative lesions characterized by localized overgrowth of endometrial glands and stroma that extend beyond the surrounding endometrium, and they represent a frequent gynecological disorder.[1] Owing to the asymptomatic nature of many EPs, the true prevalence remains uncertain; population-based estimates range from 7.8% to 34.9% among women of reproductive age, perimenopause, and postmenopause.[2,3] The etiology of EPs remains poorly understood, and no effective pharmacological therapy is currently available. Existing evidence suggests that EP development is associated with endocrine dysregulation, fluctuations in estrogen levels, and inflammatory processes.[4,5] Endometrial polypectomy is the primary therapeutic approach, yet postoperative recurrence remains a major clinical challenge. Epidemiological studies report recurrence rates ranging from 2.5% to 43.6%,[6] and repeated surgical interventions can increase the risk of abnormal uterine bleeding and compromise fertility. These concerns underscore the importance of early identification of patients at elevated risk of recurrence and the timely adoption of preventive strategies, particularly in women who wish to preserve fertility.

Machine learning (ML) algorithms are increasingly recognized as powerful artificial intelligence tools capable of extracting predictive insights from large-scale clinical datasets. Their application in healthcare has expanded

https://doi.org/10.2147/IJWH.S591932
Received: 26 December 2025
Accepted: 20 March 2026
Published: 30 March 2026

substantially, demonstrating strong utility in constructing risk prediction models. Prior studies have successfully harnessed ML for disease risk stratification. For example, Deberneh HM et al developed an integrated model based on patient medical history to predict the risk of type 2 diabetes,[7] while Sun et al[8] evaluated six ML algorithms and identified the optimal approach for predicting post-chemotherapy lung infection in patients with lung cancer, achieving robust predictive performance. In contrast, research on recurrence after endometrial polypectomy has primarily focused on identifying clinical risk factors,[9,10] with limited attention given to predictive model development—particularly models based on ML methodologies. To fill this gap, the present study employs ML algorithms to construct a clinically interpretable model for predicting recurrence risk in patients following endometrial polypectomy.

Patients and Methods

Patients

This retrospective study included 606 patients aged 22–45 years who underwent endometrial polypectomy at the second affiliated hospital of Wenzhou Medical University between March 2020 and March 2024.

Inclusion criteria were: (1) women aged 18–45 years; (2) diagnosis of endometrial polyps according to standard criteria[11] and treatment with endometrial polypectomy; (3) availability of complete clinical records; (4) provision of written informed consent after full explanation of the study.

Exclusion criteria were: (1) contraindications to hysteroscopy or allergy to anesthetic agents; (2) use of hormonal contraception, hormone replacement therapy, controlled ovarian stimulation, or other medications affecting the endometrium within three months prior to enrollment; (3) previous history of endometrial polypectomy; (4) recent uterine surgery; (5) active infection of the reproductive tract; (6) cervical scarring resulting in a rigid cervix or inadequate dilation; (7) malignant tumors of the reproductive system.

Research Variables

This study evaluated 36 predictors. Demographic factors included age, height, weight, and body mass index. Surgical history comprised prior cesarean section and other uterine surgeries. Reproductive history variables included the number of full-term births, abortions, and total pregnancies. Polyp-related characteristics encompassed polyp location within the uterine cavity, histopathological classification, diameter, and number. Laboratory parameters included hemoglobin level and complete blood count indices: red blood cells, white blood cells, platelets, neutrophils, lymphocytes, monocytes, eosinophils, basophils, red cell distribution width, and mean platelet volume. A detailed summary of these variables is provided in Table 1.

Follow-Up and Recurrence Criteria

Patients underwent follow-up transvaginal Doppler ultrasonography at 3 and 12 months after the initial endometrial polypectomy to assess for recurrence. All examinations were performed by the same physician using a standardized protocol. Suspicious ultrasonographic findings were defined as endometrial thickening, heterogeneous echogenicity, or mildly increased intrauterine echoes. Patients could be either asymptomatic or present with symptoms such as abdominal pain, abnormal vaginal bleeding, or increased vaginal discharge. In the absence of contraindications, patients with suspicious findings underwent repeat hysteroscopic surgery. Recurrence was defined as histopathological confirmation of endometrial polyps following repeat polypectomy.

Machine Learning Algorithms

Patients were randomly allocated to training and test sets in a 7:3 ratio. Model construction was conducted using Python 3.11. Within the training set, 10-fold cross-validation was applied to develop models using seven machine learning (ML) algorithms: logistic regression (LR), support vector machine (SVM), random forest (RF), Naïve Bayes (BAY), multi-layer perceptron (MLP), K-nearest neighbor (KNN), and decision tree (DT). Model performance was assessed by accuracy, area under the receiver operating characteristic curve (AUC), F1 score, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The algorithm with the best overall performance was identified as optimal.

Table 1 Description of the Study Variables

SN	Predictors	Description	Types	Values
1	Age	Age	Continuous	22 - 45 years
2	Height	Height	Continuous	140 - 175 cm
3	Weight	Weight	Continuous	40 - 95 kg
4	BMI	Body mass index	Continuous	15.63–37.88 kg/m^2
5	CSH	Cesarean section history	Categorical	0 = no, 1 = yes
6	USH	Uterine surgery history	Categorical	0 = no, 1 = yes
7	FF	Full-term fetus	Continuous	0 - 3
9	Abortion	Abortion	Continuous	0 - 10
10	NP	Number of pregnancies	Continuous	0 - 11
11	PLU	Polyp location in uterine cavity	Categorical	0 = Posterior wall, 1 = Anterior wall, 2 = Lateral wall, 3 = Multiple
12	PC	Pathological classification	Categorical	1 = Functional polyps, 2 = No-functional polyps
13	COM	Other uterine diseases	Categorical	0 = no, 1 = yes
14	DP	Diameter of polyp	Continuous	0.2–5.0 cm
15	NPP	Number of polyps	Categorical	1 = 1, 2 = 2–5, 3 = ≥6
16	HAE	Haemoglobin	Continuous	52 - 161 g/L
17	RBC	Red blood cells	Continuous	2.16–9.56 × 10^{12}/L
18	WBC	White blood cell	Continuous	2.41–19.52 ×10^9/L
19	PLA	Platelet	Continuous	88 - 544×10^9/L
20	NEUT	Neutrophil	Continuous	0.69–15.89×10^9/L
21	LYMP	Lymphocyte	Continuous	0.25–4.09×10^9/L
22	MONO	Monocytes	Continuous	0.043–1.04×10^9/L
23	EOSI	Eosinophils	Continuous	0 – 0.77×10^9/L
24	BASO	Basophils	Continuous	0 – 0.13×10^9/L
25	MPV	Mean platelet volume	Continuous	7.4–13.9 fL
26	RDW	Red blood cell distribution width	Continuous	9.1–26.9 fL

Statistical Analysis

Data in the present study was analyzed by Statistic Package for Social Science (SPSS) software (version 20.0, IBM, Armonk, NY, USA). The group F-test, Wilcoxon rank-sum test or $\chi 2$ test were used to compare demographic data and other baseline value indicators to measure the equilibrium between groups. The differences between the measured data were compared using the *t*-test and the Wilcoxon rank-sum test. And for the count data, the $\chi 2$ test or Fisher's exact probability method was used for comparison. The ROC curve was used to analyze the prediction model and calculate AUC, sensitivity and specificity. $P<0.05$ indicated that the difference was statistically significant.

Results

Follow-Up Recurrence Outcomes

A total of 1014 patients who underwent endometrial polypectomy were initially screened. Of these, 69 did not meet the inclusion criteria, 164 met exclusion criteria, and 32 were lost to follow-up. Ultimately, 749 patients were eligible for the final analysis (Figure 1). Follow-up revealed recurrence in 237 patients, while 512 patients remained free of recurrence.

Patient Characteristics

To develop the prediction model, 606 patients were randomly assigned to a training set (n = 424) and a test set (n = 186) in a 7:3 ratio. Baseline characteristics did not differ significantly between the two groups, confirming comparability (Tables 2 and 3).

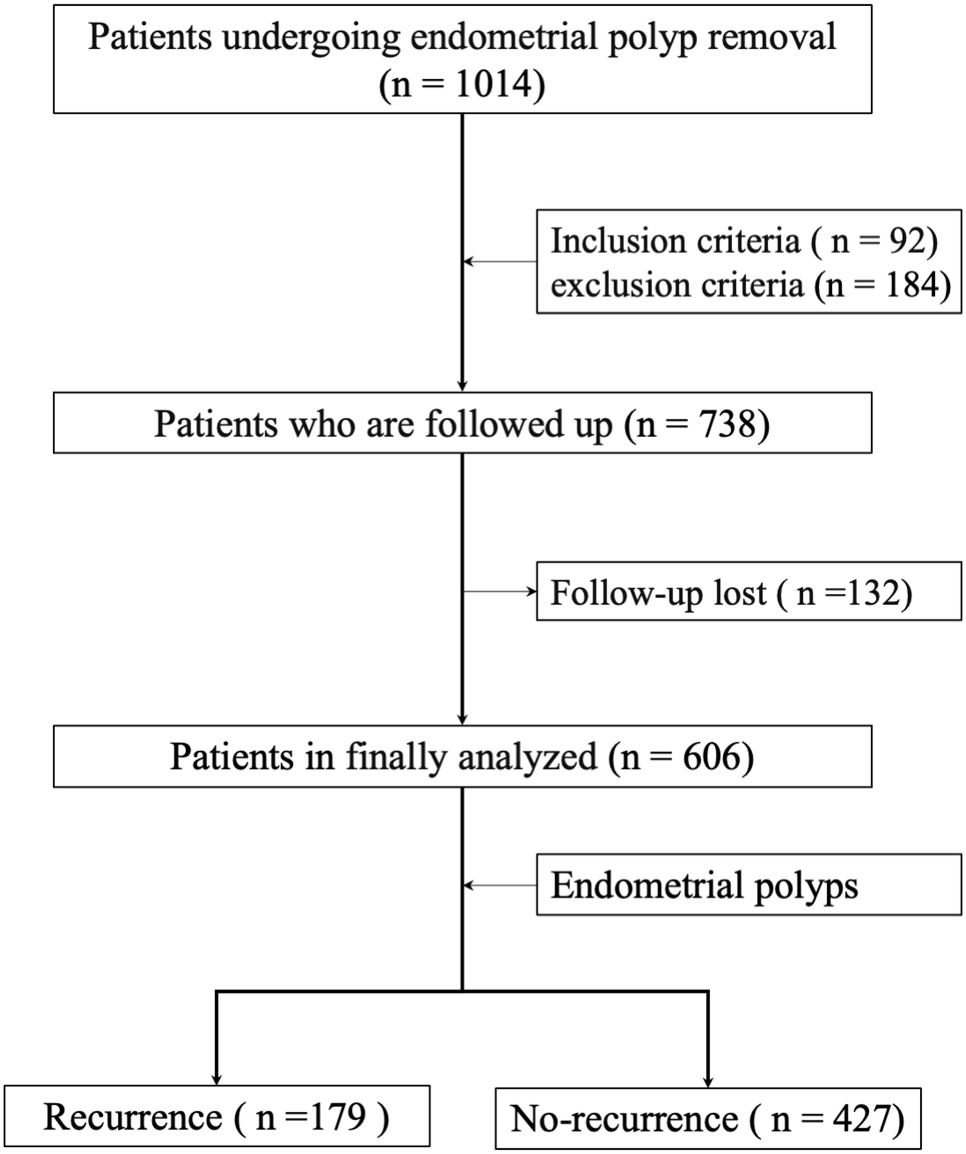

Figure 1 Flowchart of patient selection and grouping.

Clinical Characteristics of Patients with versus without Recurrence in the Training Cohort

Within the training cohort, significant differences were observed between patients with and without postoperative recurrence (Table 3). Patients in the recurrence group were older (40.02 ± 3.81 vs 34.68 ± 5.35 years, $P < 0.001$) and more likely to have a history of cesarean section (41.6% vs 21.7%, $P < 0.001$) or other uterine surgery (40.8% vs 20.7%, $P < 0.001$). They also had higher numbers of full-term deliveries, pregnancies, and abortions. Polyps were more frequently located on the posterior uterine wall ($P < 0.001$), and functional polyps were more prevalent ($P = 0.010$). Hemoglobin levels were significantly lower in the recurrence group compared with the non-recurrence group (120.28 ± 14.84 vs 125.33 ± 17.11 g/L, $P = 0.004$). No significant group differences were noted in height, weight, BMI, or most hematologic parameters.

Predictors of Postoperative Recurrence in Endometrial Polyp Patients

Multivariate logistic regression analysis in the training cohort (Table 4) identified five independent predictors of post-operative EP recurrence: older age (OR = 1.237, 95% CI: 1.158–1.321, $P < 0.001$), history of uterine surgery (OR = 2.913, 95% CI: 1.628–5.211, $P < 0.001$), posterior-wall polyp location (OR = 7.540, 95% CI: 3.440–16.529, $P < 0.001$),

Table 2 Comparison of Baseline Data in Recurrence and No-Recurrence Group

Variables	Total (n=606)	Training Set (n = 424)	Test set (n = 182)	t/χ^2	P
Age (years)	35.98±5.39	36.25±5.51	35.36±5.08	1.86	0.063
Height (cm)	159.81±4.59	159.77±4.46	159.89±4.89	0.289	0.773
Weight (kg)	56.10±8.42	55.89±8.36	56.59±8.54	0.936	0.35
BMI (kg/m^2)	21.97±3.23	21.89±3.18	22.15±3.34	0.893	0.372
CSH, n (%)	175 (28.88)	117 (27.59)	58 (31.87)	1.132	0.287
USH, n (%)	175 (28.88)	113 (26.65)	62 (34.07)	3.409	0.065
FF (n)	1.17±0.80	1.19±0.79	1.12±0.81	0.959	0.338
Abortion (n)	0.91±1.15	0.92±1.17	0.90±1.11	0.258	0.796
NP (n)	2.08±1.53	2.11±1.52	2.02±1.54	0.714	0.476
PLU, n (%)				2.807	0.422
Posterior wall	95 (15.68)	61 (14.39)	34 (18.68)		
Anterior wall	126 (20.79)	92 (21.70)	34 (18.68)		
Lateral wall	60 (9.90)	45 (10.61)	15 (8.24)		
Multiple	325 (53.63)	226 (53.30)	99 (54.40)		
PC, n (%)				2.827	0.093
Functional	441 (72.77)	317 (74.76)	124 (68.13)		
Nonfunctional	165 (27.23)	107 (25.24)	58 (31.87)		
COM, n (%)	246 (40.59)	179 (42.22)	67 (36.81)	1.542	0.214
DP (cm)	1.26±0.56	1.24±0.53	1.32±0.61	1.716	0.087
NPP (n)				0.254	0.881
1, n (%)	149 (24.59)	105 (24.76)	44 (24.18)		
2–5, n (%)	289 (47.69)	204 (48.11)	85 (46.70)		
≥6, n (%)	168 (27.72)	115 (27.12)	53 (29.12)		
HAE (g/L)	123.60±16.53	123.84±16.61	123.04±16.38	0.548	0.584
RBC (10^{12}/L)	4.36±0.45	4.37±0.48	4.34±0.35	0.707	0.48
WBC (10^9/L)	6.13±1.93	6.11±1.92	6.19±1.95	0.489	0.625
PLA (10^9/L)	265.64±68.68	263.23±66.49	271.25±73.40	1.318	0.188
NEUT (10^9/L)	3.89±1.73	3.87±1.70	3.96±1.80	0.605	0.545
LYMP (10^9/L)	1.80±0.54	1.79±0.55	1.80±0.50	0.187	0.852
MONO (10^8/L)	3.32±1.20	3.34±1.24	3.29±1.11	0.459	0.647
EOSI (10^7/L)	9.54±8.94	9.40±9.19	9.87±8.35	0.591	0.554
BASO (10^7/L)	1.95±1.44	1.93±1.37	1.97±1.60	0.321	0.748
MPV (fL)	9.98±1.06	10.02±1.04	9.90±1.12	1.22	0.223
RDW (fL)	13.65±1.96	13.69±2.00	13.58±1.86	0.604	0.546

Table 3 Baseline Characteristics in Recurrence and Non-Recurrence Group in Training Set

Variables	No-Recurrence Group (n = 299)	Recurrence Group (n = 125)	t/χ^2	P
Age (years)	34.68±5.35	40.02±3.81	11.613	<0.001
Height (cm)	159.94±4.30	159.37±4.81	1.208	0.228
Weight (kg)	55.95±8.42	55.76±8.27	0.213	0.832
BMI (kg/m^2)	21.88±3.27	21.94±2.98	0.185	0.853
CSH, n (%)	65 (21.74)	52 (41.60)	17.403	<0.001
USH, n (%)	62 (20.74)	51 (40.80)	18.153	<0.001
FF (n)	1.04±0.83	1.54±0.55	7.337	<0.001
Abortion (n)	0.76±1.05	1.30±1.36	4.434	<0.001
NP (n)	1.81±1.47	2.85±1.40	6.756	<0.001

(*Continued*)

Table 3 (Continued).

Variables	No-Recurrence Group (n = 299)	Recurrence Group (n = 125)	t/χ^2	P
PLU, n (%)			44.914	<0.001
Posterior wall	22 (7.36)	39 (31.20)		
Anterior wall	77 (25.75)	15 (12.00)		
Lateral wall	36 (12.04)	9 (7.20)		
Multiple	164 (54.85)	62 (49.60)		
PC, n (%)			6.686	0.010
Functional	213 (71.24)	104 (83.20)		
Nonfunctional	86 (28.76)	21 (16.80)		
COM, n (%)	115 (38.46)	64 (51.20)	5.864	0.015
DP (cm)	1.22±0.53	1.28±0.54	1.054	0.292
NPP (n)			1.772	0.412
1, n (%)	70 (23.41)	35 (28.00)		
2–5, n (%)	143 (47.83)	61 (48.80)		
≥6, n (%)	86 (28.76)	29 (23.20)		
HAE (g/L)	125.33±17.11	120.28±14.84	2.879	0.004
RBC (10^{12}/L)	4.36±0.41	4.37±0.64	0.167	0.868
WBC (10^9/L)	6.18±2.03	5.92±1.62	1.298	0.195
PLA (10^9/L)	267.07±62.81	254.06±74.05	1.841	0.066
NEUT (10^9/L)	3.92±1.76	3.74±1.52	0.973	0.331
LYMP (10^9/L)	1.81±0.54	1.76±0.57	0.962	0.337
MONO (10^8/L)	3.36±1.26	3.29±1.20	0.488	0.626
EOSI (10^7/L)	9.86±10.13	8.30±6.32	1.918	0.056
BASO (10^7/L)	1.92±1.28	1.97±1.58	0.393	0.695
MPV (fL)	9.97±1.06	10.13±0.97	1.445	0.149
RDW (fL)	13.66±1.85	13.74±2.34	0.357	0.721

Table 4 Results of Logistic Regression Analysis of the Training Set

Variable	*β*	*S.E.*	Walds	OR (95% CI)	*P*
Age	0.212	0.034	40.026	1.237 (1.158–1.321)	<0.001
USH	1.069	0.297	12.977	2.913 (1.628–5.211)	<0.001
PLU (Posterior wall)	2.020	0.400	25.453	7.540 (3.440–16.529)	<0.001
PC (Functional)	0.956	0.341	7.864	2.601 (1.333–5.072)	0.005
HAE	−0.019	0.008	5.454	0.981 (0.966–0.997)	0.020
Const	−8.524	1.591	28.716	–	–

functional histopathological type (OR = 2.601, 95% CI: 1.333–5.072, P = 0.005), and lower hemoglobin concentration (OR = 0.981, 95% CI: 0.966–0.997, P = 0.020). These variables were incorporated as core predictors in the subsequent machine learning models.

Machine Learning Models for Predicting Postoperative Recurrence in Endometrial Polyp Patients

In the training cohort, seven machine learning algorithms—LR, SVM, RF, NB, MLP, KNN, and DT—were used to develop predictive models with 10-fold cross-validation (Figure 2A and B). AUC values ranged from 0.752 to 0.838. Both RF and DT achieved the highest AUC (0.838), followed by MLP (0.812) and LR (0.807). RF provided the most balanced performance, with high specificity (0.930) and a moderate F1 score (0.576). By contrast, NB achieved perfect

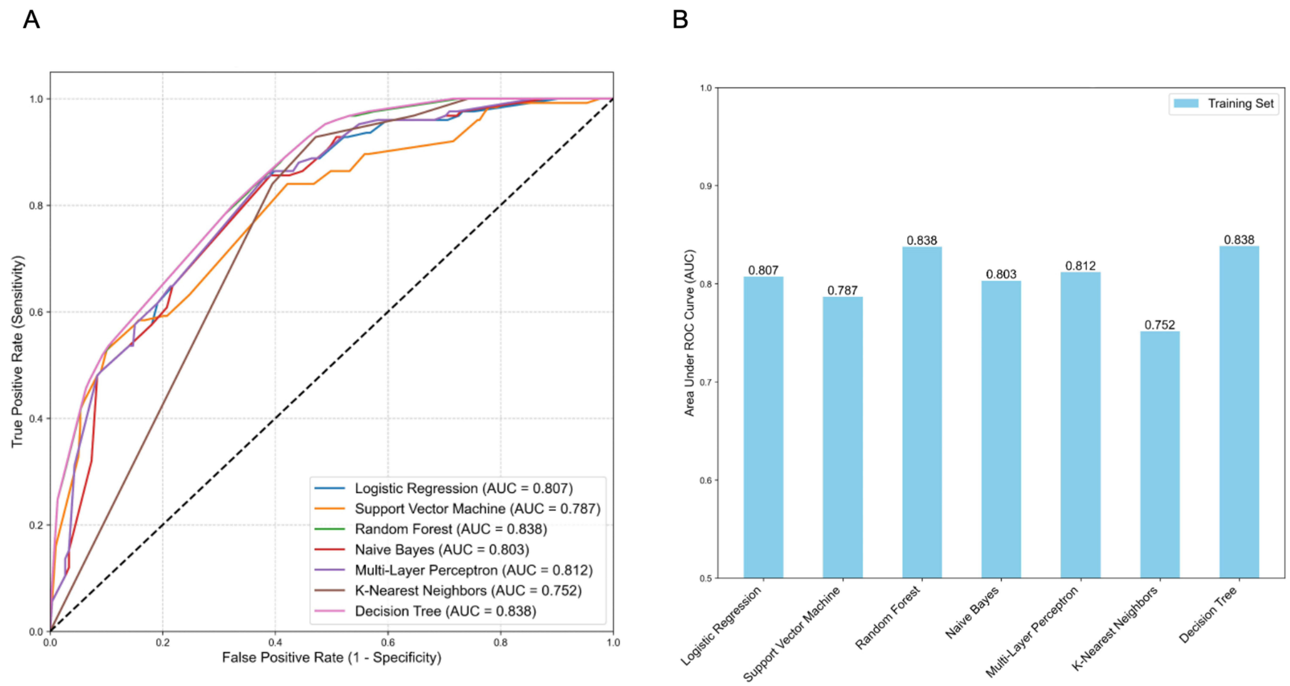

Figure 2 Performance comparison of seven machine learning models in the training set. (**A**) ROC curves of the seven machine learning models in the training set. (**B**) Bar chart showing the AUC values for each model in the training set.

recall (1.000) but extremely low specificity (0.030), reflecting strong sensitivity but very limited discriminative ability. KNN exhibited the lowest AUC (0.752), confirming its weak discriminative capacity in this dataset (Table 5).

Validation of Machine Learning Models for Postoperative Recurrence in Endometrial Polyp Patients

In the independent validation cohort, all models showed reduced performance compared with the training cohort (Figure 3A and B). RF achieved the highest AUC (0.760), followed by MLP (0.759) and NB (0.743). RF also demonstrated superior accuracy (0.753), specificity (0.875), and F1 score (0.526). By contrast, KNN again performed the worst (AUC = 0.671). These results indicate that RF provided the most consistent generalization across datasets and was therefore selected as the optimal model for subsequent interpretability analyses (Table 6).

SHAP-Based Interpretability of the Optimal Machine Learning Model for Postoperative Recurrence in Endometrial Polyp Patients

The RF algorithm was selected as the optimal predictive model. To interpret model outputs, SHapley Additive exPlanations (SHAP) analysis was applied to the RF model (Figure 4A–C). Age was identified as the most influential predictor, with

Table 5 Performance Metrics of Seven Machine Learning Models in the Training Set

Models	*AUC*	*AUC 95% CI*	Accuracy	Precision	Recall	F1 Score	Specificity
LR	0.807	0.764–0.851	0.785	0.693	0.488	0.573	0.910
SVM	0.787	0.739–0.832	0.785	0.693	0.488	0.573	0.910
RF	0.838	0.797–0.872	0.795	0.738	0.472	0.576	0.930
BAY	0.803	0.758–0.843	0.316	0.301	1.000	0.463	0.030
MLP	0.812	0.772–0.854	0.764	0.615	0.536	0.573	0.860
KNN	0.752	0.714–0.789	0.646	0.451	0.928	0.607	0.528
DT	0.838	0.796–0.875	0.795	0.750	0.456	0.567	0.936

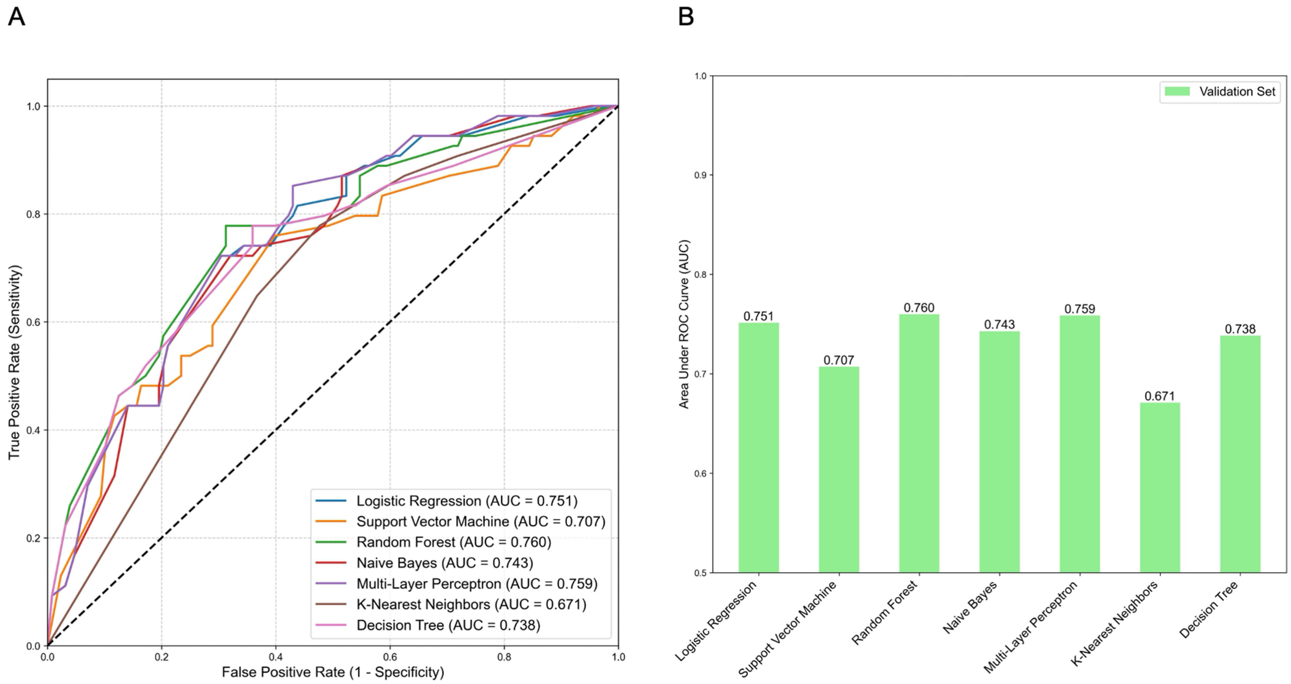

Figure 3 Performance comparison of seven machine learning models in the validation set. (**A**) ROC curves of the seven machine learning models in the validation set. (**B**) Bar chart showing the AUC values for each model in the validation set.

older age positively associated with a higher recurrence risk. Polyp location (PLU), particularly posterior-wall polyps, was the second most important predictor. A history of uterine surgery (USH) and functional histopathological type (PC) also significantly increased recurrence risk. By contrast, hemoglobin level (HAE) was negatively associated with recurrence, suggesting that lower hemoglobin (anemia) may predispose to recurrence. SHAP dependence plots further revealed that combinations of multiple high-risk factors acted synergistically to markedly increase the predicted probability of recurrence, providing clinically interpretable insights to support individualized clinical decision-making.

Discussion

Endometrial polyps (EPs) are a common gynecological condition, with onset ranging from reproductive age to postmenopause, and their clinical manifestations and impact differ by age group.[12,13] In women of reproductive age, active ovarian function and elevated estrogen secretion often present as abnormal uterine bleeding, infertility, or menstrual irregularities, and are characterized by rapid polyp growth and an increased risk of postoperative recurrence.[14,15] In this study, 179 patients experienced recurrence within one year after polypectomy, yielding a recurrence rate of 29.53%, which was higher than that observed in other age groups. A retrospective study from Spain reported a recurrence rate of 24.35% among adults, with more than 70% of recurrences occurring within one year

Table 6 Performance Metrics of Seven Machine Learning Models in the Validation Set

Models	*AUC*	*AUC 95% CI*	Accuracy	Precision	Recall	F1 Score	Specificity
LR	0.751	0.671–0.826	0.725	0.545	0.444	0.490	0.844
SVM	0.707	0.617–0.793	0.725	0.545	0.444	0.490	0.844
RF	0.760	0.677–0.832	0.753	0.610	0.463	0.526	0.875
BAY	0.743	0.664–0.817	0.330	0.307	1.000	0.470	0.047
MLP	0.759	0.682–0.824	0.698	0.490	0.444	0.466	0.805
KNN	0.671	0.597–0.748	0.599	0.408	0.778	0.535	0.523
DT	0.738	0.646–0.820	0.747	0.611	0.407	0.489	0.891

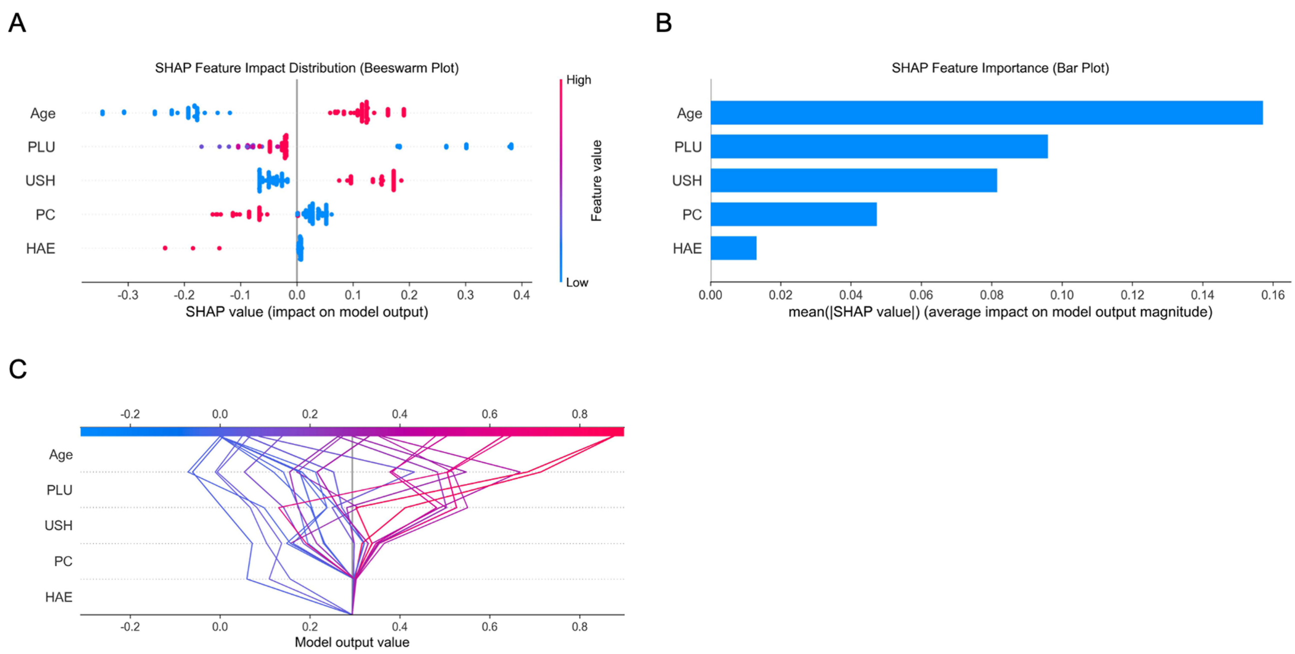

Figure 4 SHAP-based interpretability analysis of the random forest model. (**A**) SHAP beeswarm plot illustrating the impact of each feature on the model output. (**B**) Bar chart of absolute SHAP values showing the relative importance of each feature. (**C**) SHAP dependence plots demonstrating the interaction effects among the major predictive factors.

postoperatively.[16] Similarly, data from a Chinese cohort showed a recurrence rate of 15.96% in premenopausal women during up to three years of follow-up.[17] Collectively, these findings indicate that reproductive-age women face an elevated risk of recurrence after surgery, underscoring the importance of early screening to identify high-risk patients and implement timely preventive interventions.

In this study, seven ML algorithms were evaluated to develop prediction models based exclusively on readily available preoperative clinical data, including demographic variables, polyp characteristics, and routine hematologic indices. By restricting the model to core variables, we avoided complex feature selection and adopted a pragmatic approach suitable for clinical implementation. Among the algorithms tested, the RF model demonstrated the most robust performance in both the training set (AUC = 0.838) and the validation set (AUC = 0.760), outperforming logistic regression and other ML methods. The RF model identified age, history of uterine surgery, polyp location, pathological subtype, and hemoglobin concentration as the key predictors of recurrence, consistent with logistic regression analysis. Importantly, the model was deliberately constructed using only fundamental clinical data—demographic features, polyp characteristics, and preoperative blood indices—thereby involving fewer predictors and obviating the need for prior feature screening.

Advancing age is associated with progressive endocrine dysregulation, characterized by abnormal secretion of multiple hormones that disrupt endometrial homeostasis and increase the risk of polyp recurrence.[18,19] A history of uterine surgery was also identified as a significant risk factor, consistent with prior studies that linked such history to both endocrine disturbances[20] and sustained intrauterine inflammation.[21] In some patients, inadequate postoperative management may exacerbate chronic inflammatory responses within the uterine cavity, leading to alterations in endometrial tissue architecture and the local microenvironment that favor recurrence.[22,23]

Polyp subtype was also an important determinant of recurrence after polypectomy, reflecting differences in underlying pathogenesis and histological features.[6,24] Endometrial polyps are typically categorized as benign, hyperplastic, malignant, or special types. In our cohort, all cases were benign, including 464 functional and 142 non-functional polyps. We found that non-functional polyps carried a higher risk of recurrence, whereas functional polyps acted as a protective factor in the prediction model. Functional polyps generally synchronize with the endometrial cycle and are estrogen-dependent, which reduces the likelihood of recurrence following complete resection.[25] In contrast, non-functional polyps, characterized by cystic dilation, irregular glandular arrangement, and prominent stromal fibrosis, are associated with chronic endometrial inflammation and are therefore more prone to recurrence, particularly when postoperative management is suboptimal.[26]

We further found that hemoglobin level served as a protective factor against recurrence following endometrial polypectomy. Hemoglobin, a key diagnostic indicator of anemia, is the major protein in red blood cells, responsible for transporting oxygen and facilitating carbon dioxide exchange in the lungs.[27–29] Among EP patients, abnormal uterine bleeding is a common clinical symptom, primarily due to the fragile, highly vascularized surface of polyps that is prone to rupture and hemorrhage.[2,30] Previous studies have demonstrated that polyp number and volume are positively associated with bleeding risk; patients with more numerous or larger polyps are more likely to develop severe anemia, as chronic blood loss over time results in decreased hemoglobin levels.[31,32]

In conclusion, we developed a machine learning–based model to predict recurrence following endometrial polypectomy in women of reproductive age. By integrating age, history of uterine surgery, polyp histopathological subtype, and hemoglobin concentration, the model demonstrated high predictive accuracy for assessing recurrence risk in patients with EP.

Data Sharing Statement

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Ethics Approval

All study protocols received approval from Research Ethics Committee of the Second Affiliated Hospital of Wenzhou Medical University and Yuying Children's Hospital of Wenzhou Medical University (Number: 2024-K-091-01). The responsibilities, composition, procedures, and documentation of this Medical Ethics Committee are conducted in accordance with the Measures for the Ethical Review of Biomedical Research Involving Humans, the International Ethical Guidelines for Health-related Research Involving Humans, the Declaration of Helsinki, Good Clinical Practice (GCP), ICH-GCP, and other international ethical standards, together with the relevant national laws and regulations.

Acknowledgments

The authors would like to thank the clinical staff of the Department of Obstetrics and Gynecology, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, for their support in data collection and patient follow-up. We are also grateful to all the patients who participated in this study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

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

Disclosure

The authors declare that they do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

1. Vitale SG, Haimovich S, Laganà AS, et al. Endometrial polyps. An evidence-based diagnosis and management guide. *Eur J Obstet Gynecol Reprod Biol*. 2021;260:70–77. doi:10.1016/j.ejogrb.2021.03.017
2. Clark TJ, Stevenson H. Endometrial polyps and abnormal uterine bleeding (AUB-P): what is the relationship, how are they diagnosed and how are they treated? *Best Pract Res Clin Obstet Gynaecol*. 2017;40:89–104. doi:10.1016/j.bpobgyn.2016.09.005

3. Weigel GM, Baison GN, Mihalov L, Mupombwa T. Prevalence of and risk factors for endometrial polyps among asymptomatic postmenopausal women with uterovaginal prolapse. *Am J Obstet Gynecol*. 2025;232(2):191–196. doi:10.1016/j.ajog.2024.08.001
4. Raz N, Feinmesser L, Moore O, Haimovich S. Endometrial polyps: diagnosis and treatment options - a review of literature. *Minimally Invasive Ther Allied Technol*. 2021;30(5):278–287. doi:10.1080/13645706.2021.1948867
5. Li Y, Yang S, Wang P. Expected treatment for endometrial polyp. *Taiwan J Obstet Gynecol*. 2024;63(1):1–3. doi:10.1016/j.tjog.2023.11.001
6. Ciscato A, Zare SY, Fadare O. The significance of recurrence in endometrial polyps: a clinicopathologic analysis. *Hum Pathol*. 2020;100:38–44. doi:10.1016/j.humpath.2020.03.005
7. Deberneh HM, Kim I. Prediction of type 2 diabetes based on machine learning algorithm. *Int J Environ Res Public Health*. 2021;18(6):3317. doi:10.3390/ijerph18063317
8. Sun T, Liu J, Yuan H, Li X, Yan H. Construction of a risk prediction model for lung infection after chemotherapy in lung cancer patients based on the machine learning algorithm. *Front Oncol*. 2024;14:1403392. doi:10.3389/fonc.2024.1403392
9. Liu J, Liang Y, Ouyang J, Yang S. Analysis of risk factors and model establishment of recurrence after endometrial polypectomy. *Ann Palliat Med*. 2021;10(11):11628–11634. doi:10.21037/apm-21-2747
10. Yang J, Chen C, Chen S, Yang Y, Chen M. Factors influencing the recurrence potential of benign endometrial polyps after hysteroscopic polypectomy. *PLoS One*. 2015;10(12):e0144857. doi:10.1371/journal.pone.0144857
11. American AOGL. AAGL practice report: practice guidelines for the diagnosis and management of endometrial polyps. *J Minim Invasive Gynecol*. 2012;19(1):3–10. doi:10.1016/j.jmig.2011.09.003
12. Hu Y, Xiang Y, Huang W. Endometrial polyps effect on pregnancy outcomes in infertile women with minimal/mild endometriosis: a retrospective study. *J Obstetrics Gynaecol Res*. 2023;49(12):2946–2951. doi:10.1111/jog.15794
13. Hiraoka T, Osuga Y, Hirota Y. Current perspectives on endometrial receptivity: a comprehensive overview of etiology and treatment. *J Obstetrics Gynaecol Res*. 2023;49(10):2397–2409. doi:10.1111/jog.15759
14. Bittencourt CA, Dos Santos Simões R, Bernardo WM, et al. Accuracy of saline contrast sonohysterography in detection of endometrial polyps and submucosal leiomyomas in women of reproductive age with abnormal uterine bleeding: systematic review and meta-analysis. *Ultrasound Obstet Gynecol*. 2017;50(1):32–39. doi:10.1002/uog.17352
15. Jang S, Hwang SO. The risk factors for premalignant and malignant endometrial polyps in premenopausal and postmenopausal women and trends over the past decade: a retrospective study in a single center, South Korea. *Eur J Obstet Gynecol Reprod Biol*. 2024;295:118–123. doi:10.1016/j.ejogrb.2024.01.033
16. Cea García J, Jiménez Caraballo A, Ríos Vallejo MDM, Zapardiel I. Retrospective cohort study on the symptomatic recurrence pattern after hysteroscopic polypectomy. *Gynecol Minim Invasive Ther*. 2020;9(4):209–214. doi:10.4103/GMIT.GMIT_102_19
17. Wang Y, Yang M, Huang X, Li X, Lin E, Feng Y. Prevention of benign endometrial polyp recurrence using a levonorgestrel-releasing intrauterine system in premenopausal patients: a retrospective cohort study. *J Minim Invasive Gynecol*. 2020;27(6):1281–1286. doi:10.1016/j.jmig.2019.11.023
18. Gu F, Zhang H, Ruan S, et al. High number of endometrial polyps is a strong predictor of recurrence: findings of a prospective cohort study in reproductive-age women. *Fertil Steril*. 2018;109(3):493–500. doi:10.1016/j.fertnstert.2017.11.029
19. Jiménez-Lopez JS, Miguel AG, Tejerizo-Garcia A, Muñoz-Gonzalez JL, Lopez-Gonzalez G. Effectiveness of transcervical hysteroscopic endometrial resection based on the prevention of the recurrence of endometrial polyps in post-menopausal women. *BMC Womens Health*. 2015;15(1):20. doi:10.1186/s12905-015-0179-0
20. Bíró J, Ritzén EM, Hall K, Eneroth P. Effects of hysterectomy and uterine extracts on growth hormone, somatomedin, prolactin, thyrotrophin and thyroid hormones in adult rats. *Acta Endocrinologica*. 1983;103(2):172–179. doi:10.1530/acta.0.1030172
21. Qu D, Liu Y, Zhou H, Wang Z. Chronic endometritis increases the recurrence of endometrial polyps in premenopausal women after hysteroscopic polypectomy. *BMC Womens Health*. 2023;23(1):88. doi:10.1186/s12905-023-02232-3
22. Vitagliano A, Cialdella M, Cicinelli R, et al. Association between endometrial polyps and chronic endometritis: is it time for a paradigm shift in the pathophysiology of endometrial polyps in pre-menopausal women? Results of a systematic review and meta-analysis. *Diagnostics*. 2021;11(12):2182. doi:10.3390/diagnostics11122182
23. Cicinelli E, Bettocchi S, de Ziegler D, et al. Chronic endometritis, a common disease hidden behind endometrial polyps in premenopausal women: first evidence from a case-control study. *J Minim Invasive Gynecol*. 2019;26(7):1346–1350. doi:10.1016/j.jmig.2019.01.012
24. Huang J, Li Y, Chen P, Zhi Z. The risk factors and prognostic impact of different benign pathologic types of background endometrium surrounding endometrial polyps. *Int J Gynaecol Obstet*. 2025;169(1):247–257. doi:10.1002/ijgo.16067
25. Carvalho FM, Aguiar FN, Tomioka R, De oliveira RM, Frantz N, Ueno J. Functional endometrial polyps in infertile asymptomatic patients: a possible evolution of vascular changes secondary to endometritis. *Eur J Obstet Gynecol Reprod Biol*. 2013;170(1):152–156. doi:10.1016/j.ejogrb.2013.05.012
26. Nijkang NP, Anderson L, Markham R, Manconi F. Endometrial polyps: pathogenesis, sequelae and treatment. *SAGE Open Med*. 2019;7:2050312119848247. doi:10.1177/2050312119848247
27. Hess JR. Hemoglobin-based oxygen carriers. *Transfusion*. 2024;64(6):959–962. doi:10.1111/trf.17856
28. Zhang W, Zhang Y, Shi X, Wang S, Bao Y. Hemoglobin wonders: a fascinating gas transporter dive into molluscs. *Crit Rev Biochem Mol Biol*. 2023;58(2–6):132–157. doi:10.1080/10409238.2023.2299381
29. Musallam KM, Barella S, Origa R, et al. Pretransfusion hemoglobin level and mortality in adults with transfusion-dependent β-thalassemia. *Blood*. 2024;143(10):930–932. doi:10.1182/blood.2023022460
30. Munro MG, Critchley HOD, Fraser IS; FIGO MDC. The two FIGO systems for normal and abnormal uterine bleeding symptoms and classification of causes of abnormal uterine bleeding in the reproductive years: 2018 revisions. *Int J Gynaecol Obstet*. 2018;143(3):393–408. doi:10.1002/ijgo.12666
31. Donnez J, Carmona F, Maitrot-Mantelet L, Dolmans M, Chapron C. Uterine disorders and iron deficiency anemia. *Fertil Steril*. 2022;118(4):615–624. doi:10.1016/j.fertnstert.2022.08.011
32. Jobda K, Szeszko Ł, Wróbel G, et al. Hemoglobin levels and platelet counts after hysteroscopy due to abnormal uterine bleeding. *Diagnostics*. 2022;12(3):594. doi:10.3390/diagnostics12030594