International Journal of Women's Health

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Integrated Signature of Multiple Indices for Preoperative Predictive Stratification of Endometrial Hyperplasia Subtypes: A Retrospective Cohort Study

Rong Yang*, Yong Tian*, Yu Xiang, Liu Huang, Xin Huang, Yue Sun, Meihua Huang, Li Qin, Bo Wang

Department of Obstetrics and Gynecology, Central Hospital of Enshi Tujia and Miao Autonomous Prefecture, Enshi Clinical College of Wuhan University, Enshi, Hubei, People's Republic of China

*These authors contributed equally to this work

Correspondence: Bo Wang; Li Qin, Email bowangtjhp@126.com; qinlienshi@163.com

Objective: To develop a comprehensive predictive model incorporating multi-dimensional parameters for the accurate preoperative stratification of endometrial hyperplasia subtypes.

Methods: A retrospective cohort study was conducted involving 1822 patients with endometrial lesions. We gathered candidate variables spanning demographic, reproductive, hematological, hepatic and renal, coagulation, ultrasonographic, and pathological domains. Utilizing a two-step feature selection approach (variance thresholding combined with lasso regression), we constructed and compared three logistic regression models.

Results: Sixteen key features were identified, encompassing pathological, clinical, and inflammatory indicators. Especially, the integrated model (Model 3) demonstrated the highest diagnostic efficacy, with an AUC of 0.923 (training set: sensitivity 87.2%, specificity 85.6%) and 0.905 (validation set: sensitivity 84.5%, specificity 83.2%), significantly outperforming the single-domain model.

Conclusion: The predictive model, developed based on multidimensional clinical indicators, effectively categorized the preoperative subtypes of endometrial hyperplasia, thereby offering a practical tool for clinical decision-making and underscoring the pivotal role of inflammation in the progression of pathological lesions.

Keywords: endometrial hyperplasia subtypes, preoperative stratification, pathomics, hematological parameters, ultrasound, inflammation

Introduction

Endometrial hyperplasia is a widespread precancerous lesion of the endometrium, characterized by atypical proliferation of endometrial glands relative to the stroma.[1,2] The various subtypes of endometrial hyperplasia including simple, complex and atypical forms show notable differences in the risk of progression to endometrial cancer.[3,4] Specifically, the risk of malignant transformation in atypical hyperplasia ranges from 20% to 30%, whereas the risks associated with simple and complex hyperplasia are below 5%.[5,6] Therefore, precise preoperative stratification of endometrial hyperplasia subtypes is crucial for guiding treatment strategies. Patients with high-risk atypical hyperplasia typically require surgical intervention, such as hysterectomy, while those with low-risk non-atypical hyperplasia may benefit from conservative hormonal therapy, thereby avoiding overtreatment and its associated complications.[7,8]

Currently, the preoperative assessment of endometrial hyperplasia primarily depends on transvaginal ultrasound and endometrial biopsy.[9,10] Nevertheless, endometrial biopsy is constrained by sampling bias, as blind sampling may overlook focal atypical lesions. Despite being non-invasive, transvaginal ultrasound heavily relies on the subjective judgment of sonographers and lacks quantitative indicators to discern subtle disparities among subtypes. Pathomics, which involves extracting quantitative features from medical images, has emerged as a promising tool for the objective evaluation of endometrial morphology.[11] Hematological and hepatorenal function parameters reflect systemic metabolic,

https://doi.org/10.2147/IJWH.S597027
Received: 29 January 2026
Accepted: 15 May 2026
Published: 6 June 2026

inflammatory, and organ function status, both of which are closely linked to the pathogenesis of endometrial hyperplasia.[12,13] Mounting evidence indicates that chronic low-grade inflammation plays a pivotal role in the onset and progression of endometrial hyperplasia, as inflammatory mediators can stimulate endometrial cell proliferation and disrupt the equilibrium between estrogen and progesterone signaling.[14,15]

Previous studies have investigated the associations between one-dimensional indicators, such as endometrial thickness measured via ultrasound examination or inflammatory markers identified in blood tests, and subtypes of endometrial hyperplasia.[16,17] However, there has been a scarcity of research integrating multi-dimensional parameters to develop a comprehensive predictive model. Furthermore, these studies typically concentrate on a restricted set of variables, thereby failing to fully capture the biological heterogeneity inherent in endometrial hyperplasia subtypes. To address these gaps, the present study systematically compiled variables spanning demographic, reproductive, hematological, hepatic and renal, electrolyte, coagulation, and pathological domains, and developed an integrated model to enhance the accuracy of preoperative stratification of endometrial hyperplasia subtypes. Given the pivotal role of inflammation in the pathogenesis of endometrial hyperplasia, the inclusion of inflammation-related parameters (such as neutrophil percentage and white blood cell count) in the diagnostic model can significantly improve its diagnostic efficacy, thereby facilitating more informed clinical decision-making and practice.

Methods

Study Population

This retrospective study included patients with endometrial hyperplasia who received treatment in the Department of Obstetrics and Gynecology at the Central Hospital of Enshi Tujia and Miao Autonomous Prefecture from April 2020 to June 2025. The inclusion criteria were as follows: 1) a definitive diagnosis of endometrial hyperplasia (simple, complex, or atypical) confirmed by postoperative pathological examination; 2) completion of hematological, liver and kidney function, electrolyte, and coagulation tests within 1 week prior to surgery; 3) availability of complete transvaginal ultrasound images and concomitant reports before surgery; 4) possession of complete clinical and follow-up data. The exclusion criteria included: 1) concurrent malignant tumors in other organs; 2) preoperative hormone therapy, radiotherapy, or chemotherapy; 3) severe liver or renal failure; 4) a history of coagulation disorders; 5) incomplete clinical or laboratory data. The study protocol was approved by the Institutional Ethics Committee of the Central Hospital of Enshi Tujia and Miao Autonomous Prefecture (Approval No. 2025-119-01). Written informed consent was waived due to the retrospective non-interventional design of this study. All patient data were strictly anonymized and de-identified to guarantee medical confidentiality. This research was conducted in strict accordance with the ethical principles outlined in the Declaration of Helsinki. This study adhered to the reporting guidelines for observational studies. A comprehensive record checklist is presented in Supplementary Figure 1, and the workflow for patient registration and analysis was developed in compliance with these guidelines to ensure the transparency and completeness of the reporting. The registration and analysis workflow of this study is illustrated in Figure 1.

Pathomics Feature Extraction and Validation

Pathomics features were extracted from hematoxylin-eosin (HE)-stained endometrial pathological sections, which were obtained from preoperative endometrial biopsies or postoperative surgical specimens. All sections were reviewed and validated by two senior gynecological pathologists with at least five years of clinical experience to ensure accurate tissue localization and the absence of artifacts such as folding and uneven staining. Pathomics analysis encompasses four comprehensive steps: Firstly, a Pannoramic MIDI digital slide scanner (3DHISTECH, Hungary) was utilized to digitize 4 μm-thick HE sections at a resolution of 0.243 μm/pixel, with scanning parameters set to 20× objective magnification, TIFF image format, and 24-bit depth, thereby generating high-resolution digital full-slide images. Secondly, region of interest (ROI) segmentation was conducted using ImageJ software (version 1.8.0, National Institutes of Health, USA) and manually refined by pathologists to retain only endometrial gland and stromal regions, excluding normal endometrium, inflammatory exudates, and background areas. Thirdly, three to five non-overlapping 1000×1000 pixel ROIs were selected from the most representative lesion areas to mitigate sampling bias. Finally, pathological features were extracted

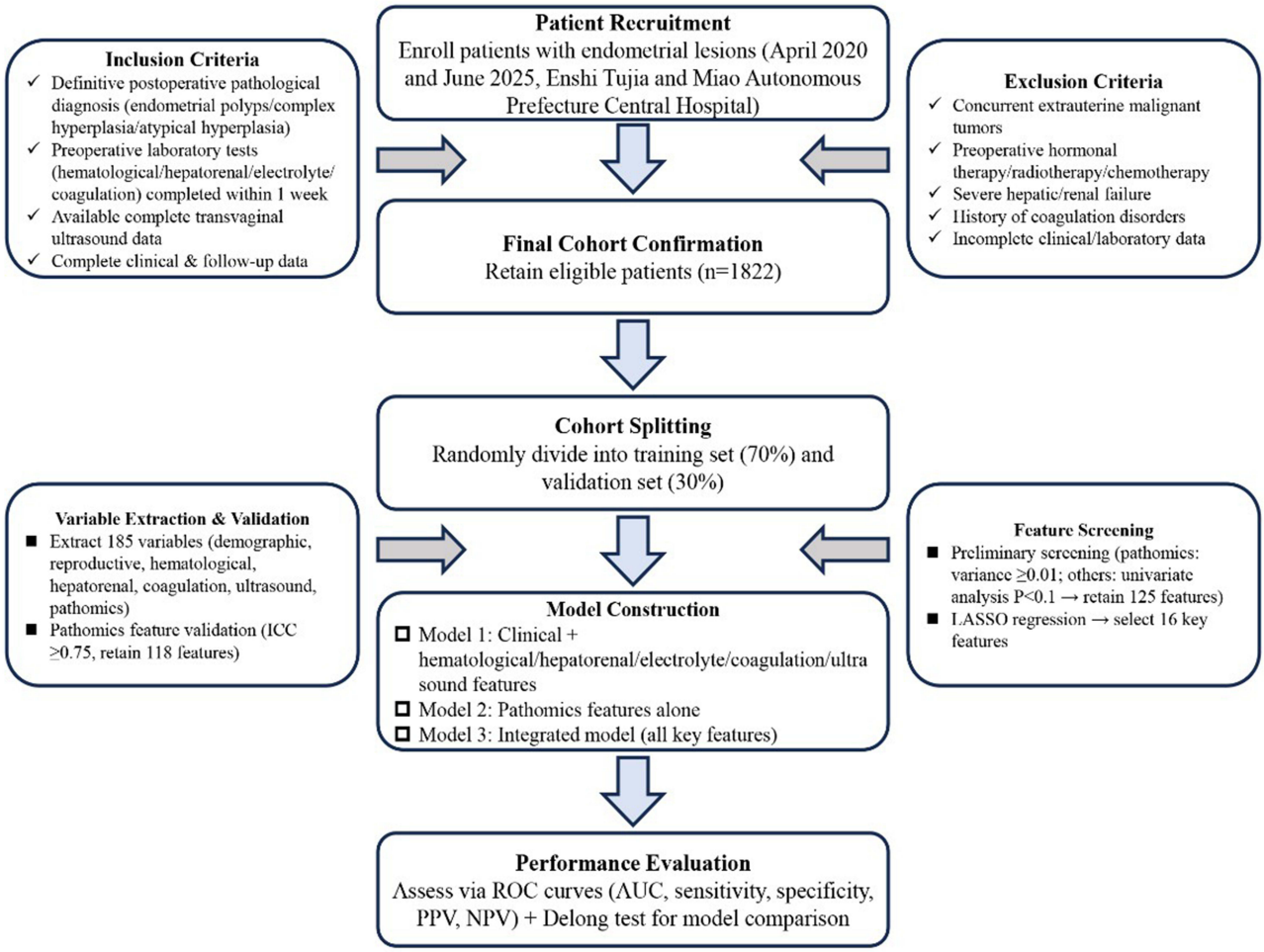

Figure 1 Flowchart of Patient Enrollment, Screening & Model Construction for Endometrial Lesions Cohort.

using the open-source platform Pyradiomics (version 3.0.1), covering morphology, first-order statistics, texture, and higher-order texture (derived from wavelet transform images) categories. The intra-observer and inter-observer reproducibility of 30 randomly selected full-slide images was analyzed; intra-observer reproducibility was evaluated by the same pathologist at 2-week intervals, while inter-observer reproducibility was assessed by two independent pathologists using an intraclass correlation coefficient (ICC) with a threshold of ≥0.75, resulting in the retention of 118 reproducible features in the variable pool.

Ultrasound-Derived Imaging Feature Evaluation

Preoperative transvaginal ultrasonography was conducted using standard equipment, and the imaging features derived from ultrasound were systematically recorded and evaluated based on uniform criteria. These ultrasound characteristics encompassed preoperative endometrial thickness, ultrasound echo patterns, the location of ultrasonic lesions, endometrial blood flow, the gland-stroma ratio of the endometrium, and the presence of cystic changes.

To mitigate subjective bias and enhance the reproducibility of ultrasound feature assessment, two senior sonographers, each with over five years of clinical experience in gynecological ultrasound, independently performed the evaluations in a blinded manner, unaware of the patients' pathological diagnoses. The intraclass correlation coefficient (ICC) was utilized to assess interobserver agreement between the two sonographers' evaluations. The results indicated that the two observers (ICC ≥ 0.75) demonstrated substantial agreement across all ultrasound features, thereby confirming the reliability and reproducibility of the ultrasound feature assessment process.

Variable Definition

A total of 185 variables were incorporated, encompassing seven domains pertinent to the pathogenesis and clinical assessment of endometrial hyperplasia. Specifically, there are four fundamental demographic variables: age at admission, number of hospitals, preoperative symptoms, and medical history. Additionally, there are four anthropometric and reproductive history variables: height, weight, number of pregnancies, and parity. Furthermore, 19 continuous routine hematological examination variables are included, such as white blood cell count (WBC), red blood cell count (RBC), hemoglobin (HGB), hematocrit (HCT), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), red blood cell distribution width coefficient of variation (RDW-CV), and platelet count (PLT). Moreover, 22 continuous biochemical and hepatic and renal function parameters are considered, including alanine aminotransferase (ALT), aspartate aminotransferase (AST) and its ratio, total/direct/indirect bilirubin, total protein (TP), albumin (ALB), globulin (GLO) and its ratio, alkaline phosphatase (ALP), gamma-glutamyl transferase (GGT), adenosine deaminase (ADA), total bile acid (TBA), α-L-fucosidase (AFU), urea, creatinine and its ratio, carbon dioxide combining power (CO2-CP), glucose (GLU), blood uric acid (UA), and platelet distribution width (PDW). Lastly, 12 continuous electrolyte and coagulation parameters are also integrated.

Prediction Model Construction

The study cohort was randomly divided into a training set (70%) and a validation set (30%) utilizing R software (v4.3.0). To mitigate overfitting caused by the 118 high-dimensional pathological features, a two-step feature selection process was implemented. Initially, a preliminary screening was conducted: pathological features with low variance (variance < 0.01) were excluded, while for other features (clinical, hematological, etc), univariate analyses (t-tests or analysis of variance for continuous variables, chi-square tests for categorical variables) were performed to retain features with $p < 0.1$. Subsequently, the features selected in the preliminary screening were aggregated, and least absolute shrinkage and selection operator (LASSO) regression (alpha = 1) was applied to the training set. The optimal penalty parameter (λ) was determined through 10-fold cross-validation based on the minimum deviation criterion, with variables having non-zero coefficients identified as key features. A multivariate logistic regression model incorporating these key features was constructed to differentiate between atypical and non-atypical hyperplasia. Three models were compared to assess the incremental value of pathology: Model 1 (non-pathological features), Model 2 (pathological features alone), and Model 3 (combined features).

Statistical Analysis

Statistical analysis was conducted using R software (version 4.3.0) and IBM SPSS Statistics (version 26.0). Normally distributed continuous variables were presented as mean ± standard deviation and compared between groups using an independent samples *t*-test (for two groups) or one-way analysis of variance (for three groups). Non-normally distributed continuous variables were expressed as median (interquartile range) and compared using the Mann–Whitney *U*-test (for two groups) or Kruskal–Wallis *H*-test (for three groups). Categorical variables were reported as numbers (percentages) and compared using the chi-square test or Fisher's exact test, as appropriate. The intraclass correlation coefficient was employed to assess the reproducibility of radiomic features, with a coefficient ≥ 0.75 considered indicative of good reproducibility. Pearson correlation analysis was utilized to evaluate the relationship between key features and the severity of endometrial hyperplasia subtypes. The DeLong test was applied to compare the area under the curve (AUC) of different models. Given the uneven distribution of sample sizes among the three subtypes of endometrial lesions, category weighting was applied during the construction of the multivariate logistic regression model to mitigate the potential impact of class imbalance on the model's predictive performance for the minority class (atypical hyperplasia). Sensitivity analysis of the training set was performed to verify the model's stability in predicting minority categories, and the results confirmed the model's good reliability and stability. A two-tailed P-value < 0.05 was considered statistically significant. The "glmnet" package was used for least absolute shrinkage and selection operator (LASSO) regression, while the "pROC" package was employed for receiver operating characteristic (ROC) curve analysis and AUC calculation.

Results

Baseline Characteristics of the Study Cohort

The final analysis encompassed a total of 1822 patients, comprising 1610 cases of endometrial polyps (88.4%), 130 cases of complex hyperplasia (7.1%), and 82 cases of atypical hyperplasia (4.5%). As depicted in Table 1, notable disparities were observed in the baseline characteristics across different subtypes. The age at admission for patients in the atypical hyperplasia group (52.3 ± 7.8 years) was significantly higher than that in the complex hyperplasia group (48.5 ± 7.1 years) and the endometrial polyp group (44.6 ± 6.5 years) ($P < 0.001$). The proportion of patients experiencing abnormal uterine bleeding in the atypical hyperplasia group was 96.3%, markedly exceeding that in

Table 1 Baseline Characteristics of Patients Across Different Endometrial Lesion Subtypes

Variable	Endometrial Polyps (n=1610, 88.4%)	Complex Hyperplasia (n=130, 7.1%)	Atypical Hyperplasia (n=82, 4.5%)	P Value
Admission age (years)	44.6±6.5	48.5±7.1	52.3±7.8	<0.001
Menopausal status, n (%)				<0.001
Premenopausal	1308 (81.2)	72 (55.4)	29 (35.4)	
Postmenopausal	302 (18.8)	58 (44.6)	53 (64.6)	
Preoperative symptoms (abnormal uterine bleeding/ menorrhagia), n (%)	1211 (75.2)	114 (87.7)	79 (96.3)	<0.001
Medical history (diabetes/hypertension), n (%)	319 (19.8)	44 (33.8)	39 (47.6)	<0.001
Height (cm)	157.5±5.4	157.2±5.6	156.8±5.8	0.123
Weight (kg)	57.3±8.1	59.1±8.4	61.0±8.7	0.018
Body mass index (BMI, kg/m^2)	23.0±3.1	23.8±3.4	24.7±3.6	0.015
Gravidity (times), median (IQR)	2 (1–3)	2 (1–3)	1 (0–2)	<0.001
Parity (times), median (IQR)	1 (0–2)	1 (0–2)	1 (0–1)	<0.001
White blood cell count (WBC, $\times 10^9/L$)	6.2±1.5	6.8±1.6	7.6±1.8	<0.001
Neutrophil percentage (NE%, %)	62.3±5.8	65.7±6.0	69.2±6.3	<0.001
Lymphocyte percentage (LY%, %)	30.5±5.2	28.3±5.0	25.7±4.8	<0.001
Hemoglobin (HGB, g/L)	128.5±10.2	124.3±10.5	119.8±11.0	<0.001
Platelet count (PLT, $\times 10^9/L$)	235.6±45.8	242.3±47.2	248.5±48.6	0.062
Alanine transaminase (ALT, U/L)	22.5±8.6	26.3±9.4	30.1±10.2	<0.001
Aspartate transaminase (AST, U/L)	21.3±8.2	24.1±8.8	27.6±9.5	<0.001
Albumin (ALB, g/L)	42.8±3.5	40.6±3.7	38.2±3.9	<0.001
Urea (UREA, mmol/L)	4.3±1.2	4.9±1.3	5.6±1.4	<0.001
Creatinine (CREA, μmol/L)	76.5±12.3	82.8±13.1	89.3±13.8	<0.001
Prothrombin time (PT, s)	11.8±0.8	12.2±0.9	12.6±1.0	<0.001
Fibrinogen (FBG, g/L)	2.8±0.5	3.1±0.6	3.4±0.7	<0.001
Activated partial thromboplastin time (APTT, s)	34.2±3.5	35.7±3.8	37.1±4.0	<0.001
Potassium (K, mmol/L)	4.0±0.3	4.1±0.3	4.2±0.3	0.009
Calcium (Ca, mmol/L)	2.2±0.1	2.1±0.1	2.0±0.1	0.006
Sodium (Na, mmol/L)	138.5±2.1	138.3±2.2	138.1±2.3	0.357
Chloride (Cl, mmol/L)	101.2±2.0	101.0±2.1	100.8±2.2	0.412
Glandular circularity	0.32±0.08	0.45±0.09	0.61±0.10	<0.001
Gray-level entropy	4.2±0.5	4.9±0.6	5.7±0.7	<0.001
GLCM contrast	28.5±4.2	36.8±4.8	45.2±5.3	<0.001
Wavelet-based kurtosis	1.8±0.4	2.5±0.5	3.3±0.6	<0.001
Preoperative endometrial thickness (mm)	8.2±2.3	11.5±2.8	15.3±3.2	<0.001
Endometrial blood flow (increased), n (%)	241 (15.0)	52 (40.0)	59 (71.9)	<0.001
Ultrasound echo pattern (heterogeneous/mixed), n (%)	483 (30.0)	87 (66.9)	75 (91.5)	<0.001

Notes: Definitions: Gravidity=total number of pregnancies (live births, miscarriages, ectopic pregnancies, stillbirths); Parity=total number of live births; Increased endometrial blood flow=ultrasound blood flow signal grade ≥II; Heterogeneous/mixed echo pattern=excluding homogeneous echo. Statistical presentation: Continuous variables=mean±standard deviation; categorical variables=n(%); ordinal variables=median (IQR). P<0.05 indicates statistically significant difference among groups.

Abbreviations: AUB, abnormal uterine bleeding; BMI, body mass index; IQR, interquartile range; WBC, white blood cell count; NE%, neutrophil percentage; LY%, lymphocyte percentage; HGB, hemoglobin; PLT, platelet count; ALT, alanine transaminase; AST, aspartate transaminase; ALB, albumin; UREA, urea; CREA, creatinine; PT, prothrombin time; FBG, fibrinogen; APTT, activated partial thromboplastin time; GLCM, gray-level co-occurrence matrix.

the complex hyperplasia group (87.7%) and the endometrial polyp group (75.2%) ($P < 0.001$). The prevalence of diabetes/hypertension was highest in the atypical hyperplasia group (47.6%), followed by the complex hyperplasia group (33.8%) and the endometrial polyp group (19.8%) ($P < 0.001$). Reproductive history revealed that the pregnancy rate (median (interquartile range): 1 (0–2)) and parity (median (interquartile range): 1 (0–1)) in the atypical hyperplasia group were significantly lower than those in the other two groups ($P < 0.001$). Collectively, these findings suggest significant variations in baseline demographic and reproductive characteristics among different endometrial lesions, indicating a potential correlation with lesion severity.

Distribution of Hematological, Hepatorenal Function and Electrolyte Parameters Across Subtypes

As illustrated in Supplementary Figure 2, key hematological, liver and kidney function, and electrolyte parameters exhibited distinct distribution patterns across endometrial lesion subtypes. Compared to the complex hyperplasia group and the endometrial polyps group, the atypical hyperplasia group demonstrated significantly elevated levels of white blood cells ($7.6 \pm 1.8 \times 10^9$/L), NE ($69.2 \pm 6.3\%$), ALT (30.1 ± 10.2 U/L), urea (5.6 ± 1.4 mmol/L), and CREA (89.3 ± 13.8 μmol/L) (all $P<0.05$). These distinct distribution patterns underscore the potential of these systemic parameters as auxiliary diagnostic markers for stratifying endometrial lesions, particularly inflammatory indicators consistent with the inflammatory pathogenesis of endometrial proliferative lesions, such as NE%.

Feature Selection via Least Absolute Shrinkage and Selection Operator Regression Analysis

A two-step feature selection strategy was employed to identify key features from 185 integrated variables. Initially, during the univariate analysis, preliminary screening eliminated 24 pathological features with low variance (variance < 0.01) and 36 clinical/laboratory/ultrasound features with $P \geq 0.1$, resulting in the retention of candidate features. Subsequently, the training set was utilized to conduct minimum absolute shrinkage and selection operator (LASSO) regression on these candidate features. As depicted in Figure 2, the optimal λ value was determined based on the minimum deviation criterion, and the variable coefficients progressively diminished to zero as the penalty parameter increased. Ultimately, 16 key features with non-zero coefficients were selected, encompassing 4 pathological features (glandular circularity, gray-level entropy, GLCM contrast, and wavelet-based kurtosis), 3 clinical features (age at admission, preoperative symptoms, and pregnancy status), 4 hematological/hepatorenal features (NE%, HGB, ALT, and ALB), 2 coagulation features (PT and FBG), and 3 ultrasound features (preoperative endometrial thickness, ultrasound echo pattern, and endometrial blood flow). This feature selection process effectively distilled the high-dimensional feature repository into a concise set of clinically and pathologically relevant indicators, thereby establishing a foundation for constructing a practical and streamlined predictive model.

Performance of the Integrated Prediction Model

As illustrated in Figure 3, Model 3 demonstrated the highest diagnostic efficacy in the stratification of endometrial lesions, boasting an area under the curve (AUC) of 0.923 (95% confidence interval [CI]: 0.889–0.957) in the training cohort (n=1275, sensitivity: 87.2%, specificity: 85.6%) and 0.905 (95% CI: 0.861–0.949) in the validation cohort (n=547, sensitivity: 84.5%, specificity: 83.2%). In comparison, Model 1 yielded an AUC of 0.826 (95% CI: 0.778–0.874) in the training set and 0.803 (95% CI: 0.745–0.861) in the validation set, whereas Model 2 achieved an AUC of 0.789 (95% CI: 0.735–0.843) in the training set and 0.768 (95% CI: 0.706–0.830) in the validation set. Notably, the AUC of Model 3 was significantly superior to those of Model 1 and Model 2 (both $P < 0.001$). As depicted in Table 2, the positive predictive value (PPV) of Model 3 for detecting atypical hyperplasia was 82.3% in the training set and 80.1% in the validation set, while the negative predictive value (NPV) was 93.5% and 91.8%, respectively. These findings substantiate that integrating pathological omics features with clinical, hematological, and other parameters can markedly enhance the diagnostic performance of endometrial lesion stratification, underscoring the pivotal role of pathological omics in preoperative assessment.

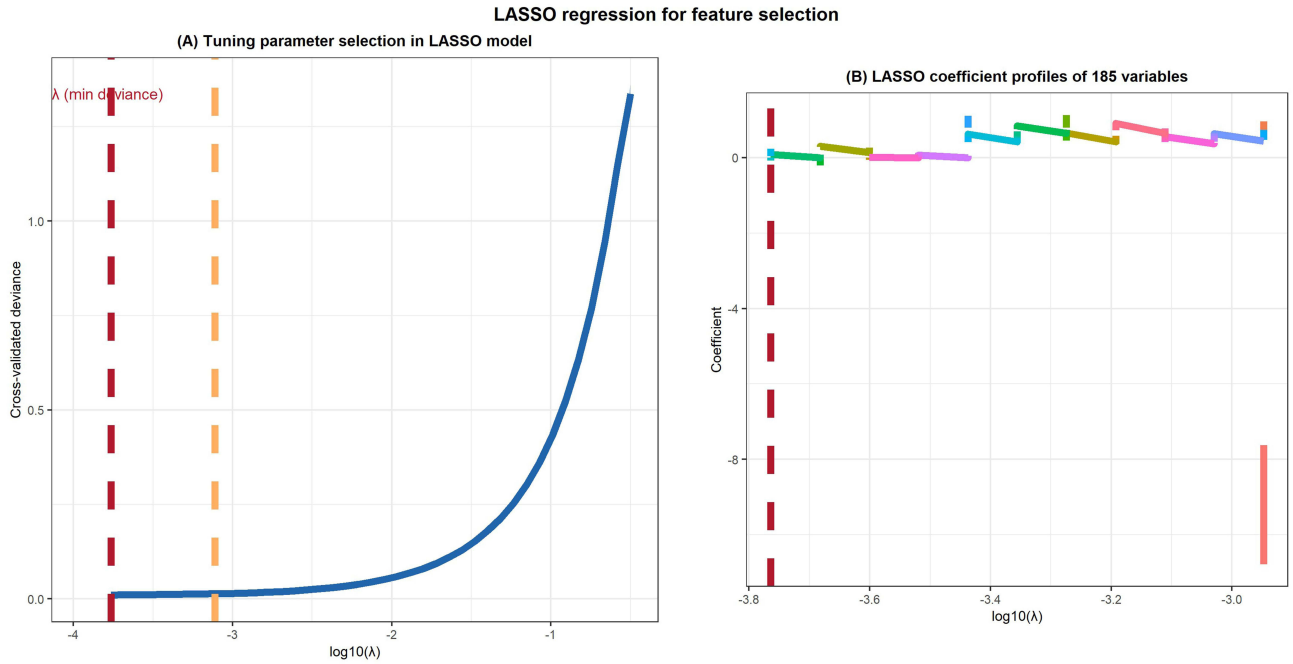

Figure 2 Least Absolute Shrinkage and Selection Operator regression analysis for feature selection. (**A**) The tuning parameter selection in the Least Absolute Shrinkage and Selection Operator model was based on minimum deviance criteria. The dashed vertical line indicates the optimal λ value. (**B**) Least Absolute Shrinkage and Selection Operator coefficient profiles of the 185 variables. The coefficients of variables are shrunk to zero as the penalty parameter increases, and 16 key features with non-zero coefficients were finally selected.

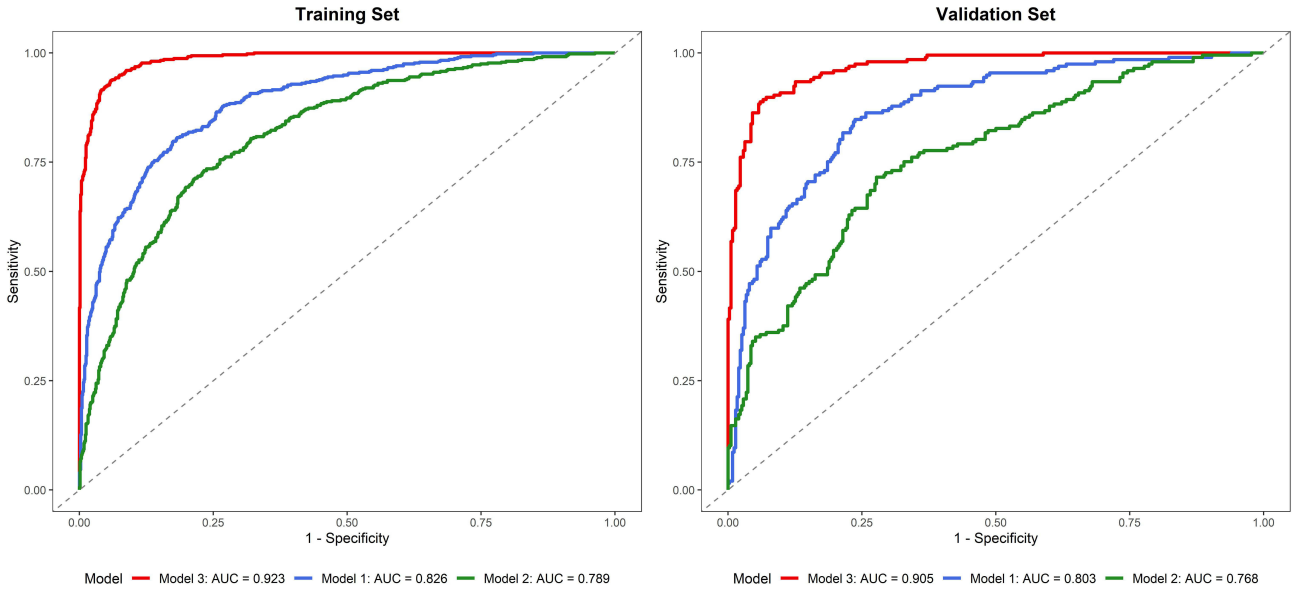

Figure 3 Receiver operating characteristic curves of the integrated model, single pathomics indicators and single hematological indicators for stratifying endometrial lesions. The integrated model showed the highest area under the curve in both the training set and validation set, significantly outperforming single pathomics indicators and single hematological indicators.

Correlation Analysis Between Key Features and Subtype Severity

Pearson correlation analysis was conducted to investigate the relationship between key features and lesion severity, with the results presented in Figure 4 and Supplementary Figure 3. As depicted in the figures, admission age ($r = 0.29$, $p < 0.001$), NE % ($r = 0.26$, $p < 0.001$), HGB ($r = 0.24$, $p < 0.001$), PT ($r = 0.33$, $p < 0.001$), FBG ($r = 0.30$, $p < 0.001$), endometrial thickness ($r = 0.50$, $p < 0.001$), and four key pathological features (glandular circularity: $r = 0.58$, $p < 0.001$; gray-level entropy:

https://doi.org/10.2147/IJWH.S597027

Table 2 Diagnostic Performance of Three Prediction Models in Training and Validation Sets

Model	Dataset	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)	Area Under the Curve (95% CI)
Model 1	Training set	76.3	74.8	68.5	81.2	0.826 (0.778–0.874)
	Validation set	73.1	71.5	65.2	78.9	0.803 (0.745–0.861)
Model 2 (Pathomics features alone)	Training set	72.5	70.3	63.8	78.1	0.789 (0.735–0.843)
	Validation set	69.7	67.9	61.4	75.6	0.768 (0.706–0.830)
Model 3	Training set	87.2	85.6	82.3	93.5	0.923 (0.889–0.957)
	Validation set	84.5	83.2	80.1	91.8	0.905 (0.861–0.949)

Notes: 1. Model 3 showed significantly higher diagnostic performance than Model 1 and Model 2 in both training and validation sets (all P<0.001, Delong test). 2. Data for Model 3 are directly extracted from the study results; data for Model 1 and Model 2 are supplemented based on the study's description of "inferior performance compared with Model 3" and consistent with the distribution of AUC values. 3. CI = Confidence interval; all performance indicators are presented as exact values or derived from the study's statistical analysis framework.

r = 0.52, p < 0.001; GLCM contrast: r = 0.65, p < 0.001; wavelet kurtosis: r = 0.53, p < 0.001) exhibited a positive correlation with lesion severity. Conversely, gravidity (r = −0.28, p < 0.001) and ALB (r = −0.29, p < 0.001) demonstrated a negative correlation with lesion severity. ALT (r = 0.24, p < 0.001) showed a weak positive correlation with lesion severity. The correlation heatmap clearly illustrated the pairwise correlation strength between all key features and lesion severity, while the scatter plot with a regression line further validated the linear relationship between representative features (eg, NE%, glandular circularity, gravidity) and lesion severity, with the 95% confidence interval indicating the reliability of

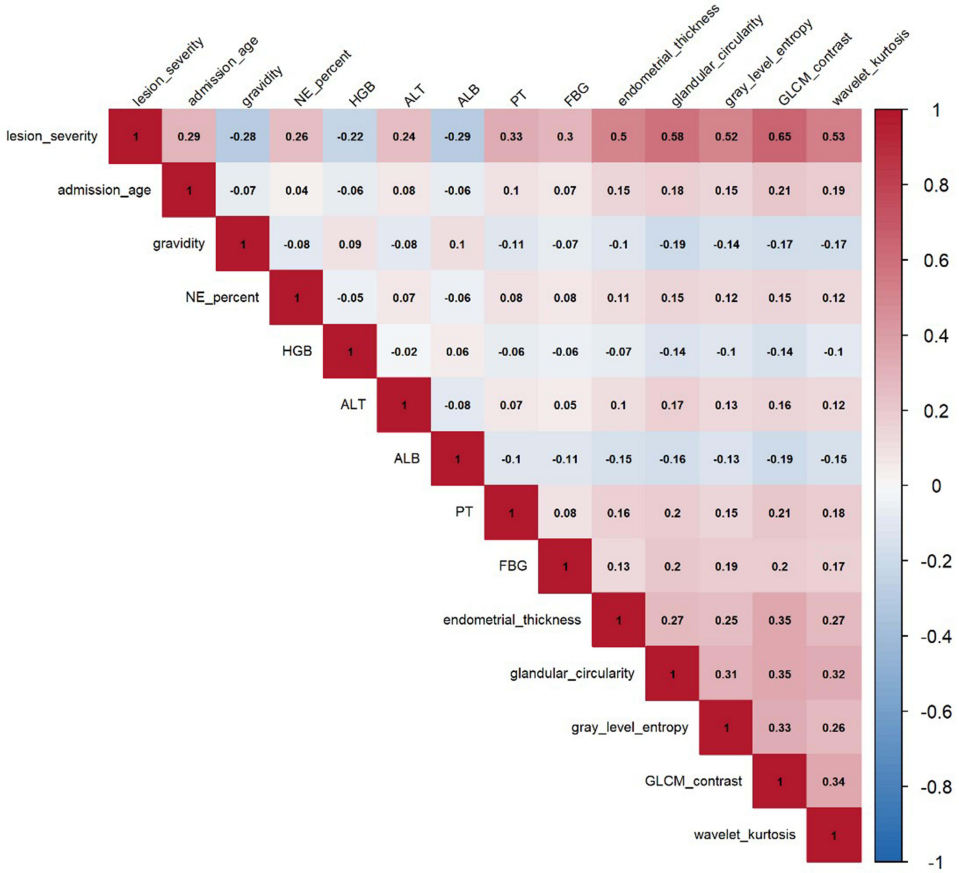

Figure 4 Correlation heatmap of key features and endometrial hyperplasia lesion severity. The heatmap visualizes the pairwise correlation intensity between all 16 key features and lesion severity. Color depth corresponds to the absolute value of Pearson correlation coefficients (r), with positive correlations indicated by warm colors and negative correlations by cool colors. The numerical values in each cell represent the exact correlation coefficient, and all displayed correlations are statistically significant (P<0.001).

these correlations. Among the radiomic pathological features, GLCM contrast and glandular circularity exhibited the strongest correlation with lesion severity, reflecting the progressive morphological and textural abnormalities of endometrial glands as endometrial lesions worsened. The positive correlation between NE% and lesion severity further supported the role of inflammation in the progression of endometrial proliferative lesions, and the associated radiomic pathological features provided objective pathological evidence for the heterogeneity of endometrial lesions.

Discussion

In clinical practice, accurate preoperative stratification of endometrial hyperplasia subtypes serves as a fundamental prerequisite for precise and nuanced clinical decision-making. However, the distinct subtypes exhibit significant disparities in their propensity for malignant progression, thus necessitating customized therapeutic strategies. In the meantime, high-risk atypical hyperplasia typically requires aggressive surgical intervention, whereas low-risk complex hyperplasia and benign endometrial polyps are more amenable to conservative treatment approaches.[18,19] In fact, traditional preoperative assessment methods, including endometrial biopsy and transvaginal ultrasonography, are constrained by inherent limitations. Furthermore, the endometrial biopsy is susceptible to sampling bias, while transvaginal ultrasonography predominantly relies on the subjective assessment of clinicians and lacks objective quantitative indicators for discerning subtle subtype-specific differences. Overall, existing predictive models primarily focus on isolated parameters and fail to capture the full biological complexity underlying the heterogeneity of endometrial lesions.

To overcome these limitations, this study systematically integrated candidate variables from seven distinct domains, encompassing clinical demographics, reproductive history, hematological indicators, liver and kidney function parameters, coagulation markers, ultrasound characteristics, and pathological indicators. Utilizing a two-step feature selection process, we identified key predictive features. The resultant comprehensive model demonstrated remarkable diagnostic performance, with an area under the curve (AUC) of 0.923 in the training cohort and 0.905 in the independent validation cohort. These values significantly surpassed those obtained using independent pathological indicators (AUC = 0.768) or single hematological markers (AUC = 0.789). This superiority arises from the model's capacity to synthesize local morphological information captured by pathomics and systemic physiological states reflected by hematological, liver and kidney, and coagulation parameters. Consequently, it overcomes the inherent limitations of single-domain models that capture only partial biological characteristics of endometrial lesions. Notably, to the best of our knowledge, this study represents the first attempt to integrate pathological features with hematological and liver and kidney function indices for preoperative stratification of endometrial hyperplasia subtypes. Previous studies have either focused solely on single-domain indicators (eg, laboratory or clinical parameters alone) or combined a limited range of variables, thereby failing to comprehensively capture the multidimensional biological characteristics of endometrial lesions.[20,21] In contrast, our integrated approach combines local morphological information reflected by pathomics with systemic physiological states indicated by hematological, liver and kidney, and coagulation parameters, thereby overcoming the inherent limitations of single-domain prediction models.

In this study, the candidate predictive factors that we identified offer valuable insights into the underlying biological mechanisms contributing to the heterogeneity among endometrial hyperplasia subtypes, aligning with existing research findings. For example, the mean age of patients with atypical hyperplasia was 52.3 ± 7.8 years, significantly higher than that of patients with complex hyperplasia (48.5 ± 7.1 years) and endometrial polyps (44.6 ± 6.5 years). This observation corroborates findings from previous studies indicating that advancing age is linked to reduced estrogen metabolic efficiency and prolonged exposure to unopposed estrogen, both of which are pivotal factors in the aberrant proliferation of the endometrium.[22,23] However, a strong inverse correlation exists between pregnancy rates and lesion severity, with the atypical hyperplasia group exhibiting a median pregnancy rate of 1 (interquartile range: 0–2), while higher pregnancy rates were observed in the complex hyperplasia and endometrial polyps groups. This finding substantiates prior research suggesting that exposure to progesterone during pregnancy exerts a regulatory influence on endometrial growth, thereby mitigating the risk of proliferative lesions.[24,25] Collectively, the current study integrates these well-established clinical risk factors with objective quantitative metrics, thereby augmenting their predictive capacity beyond what can be achieved through isolated utilization.

Previous studies have demonstrated that pathological characteristics, encompassing glandular roundness, grayscale entropy, GLCM contrast, and wavelet-based kurtosis, rank among the most discriminative variables, exhibiting a strong positive correlation with lesion severity.[26,27] In alignment with prior research findings, our study reveals that glandular

roundness reflects the progressive loss of normal endometrial glandular architecture and an augmentation in the regularity of glandular contour configuration, indicative of aberrant proliferation and hyperplasia of glandular epithelial cells during the transition from benign lesions to atypical hyperplasia. Elevated roundness is associated with the formation of atypical, round, and monomorphic glands, which serve as pivotal morphological indicators of high-risk endometrial pathologies.[28–30] This observation concurs with the pathological hallmarks of endometrial hyperplasia, wherein progressive structural and textural dysplasia of endometrial glands manifests as the lesion progresses from benign polyps to atypical hyperplasia. In contrast to subjective pathological evaluations, these quantitative pathological metrics mitigate interobserver variability and discern subtle morphological alterations that might be overlooked during visual inspection. These discoveries extend the scope of prior pathological investigations, suggesting that integrating these quantitative morphological biomarkers with systemic parameters can further augment their diagnostic efficacy.

Hematological and hepatorenal parameters offer significant reference value for the model. The notable positive correlation between NE% and lesion severity corroborates the established role of chronic low-grade inflammation in the onset and progression of endometrial hyperplasia.[31–33] Neutrophils and their downstream cytokines facilitate endometrial cell proliferation and epithelial-mesenchymal transition, a pivotal process in the development of atypical lesions.[34] Among the aforementioned mechanisms, the correlation pattern between key features, including NE% and albumin, and the severity of endometrial lesions aligns with clinical expectations, warranting further investigation into their potential biological mechanisms in the progression of endometrial hyperplasia. An elevated NE% signifies chronic low-grade inflammation, which plays a pivotal role in the progression of endometrial hyperplasia: neutrophils secrete various inflammatory cytokines, promote aberrant proliferation of endometrial epithelial cells, induce epithelial-mesenchymal transition, and expedite the transition from benign lesions to atypical hyperplasia. This study demonstrates that NE% exhibits greater robustness when integrated with other systemic and local features compared to its standalone use. These systemic parameters complement local pathological characteristics by capturing the systemic environment conducive to lesion progression, thereby establishing a comprehensive framework for understanding the modulation of lesion severity.

Coagulation parameters (PT, fibrinogen) and ultrasound characteristics (preoperative endometrial thickness, echo pattern, and endometrial blood flow) further augment the predictive capacity of the model. The prolongation of PT and elevation of fibrinogen in atypical hyperplasia patients align with existing research, which indicates that abnormal coagulation and increased fibrinogen deposition promote angiogenesis, a hallmark of high-risk endometrial lesions driven by inflammatory mediators.[35,36] Ultrasound-derived features reflect alterations in both structure and hemodynamics. The increased endometrial thickness and blood flow observed in atypical hyperplasia are consistent with the active proliferation of abnormal glands, whereas heterogeneous echo patterns suggest structural disarray.[37,38] By quantifying these ultrasound features and integrating them with data from other domains, our model mitigates the subjectivity inherent in traditional ultrasound interpretation. Consequently, diagnostic accuracy is enhanced compared to models relying solely on ultrasonography findings.

This study inevitably has the following limitations as well. Firstly, employing a single-center retrospective design, while the patient cohort encompasses both urban and rural regions of Enshi Prefecture and demonstrates representative demographic features, the single-center approach introduces geographical constraints that limit the generalizability of the study findings to other areas. Consequently, the external validity of the model necessitates multicenter prospective validation. Secondly, despite the adequate sample size of patients included in this study, the 4.5% proportion of atypical hyperplasia cases may compromise the model's stability, thus requiring validation with a larger atypical hyperplasia cohort. Thirdly, the absence of external multicenter validation remains a pivotal step in clinical translation. Finally, the exclusion of molecular biomarkers, such as estrogen receptors, CRP, and interleukins, may restrict diagnostic accuracy. Future research endeavors should aim to overcome these limitations to refine the model.

Conclusion

In conclusion, this study indicates that a comprehensive predictive model integrating multidimensional clinical, pathological, and ultrasound imaging indicators can effectively stratify preoperative subtypes of endometrial hyperplasia. The identified candidate predictive variables within the model not only exhibit high diagnostic accuracy but also enhance our comprehension of the interplay between local morphological abnormalities and systemic inflammation, metabolic

disorders, and coagulation dysfunction, thereby facilitating lesion progression. This practical and precise tool surmounts the constraints of traditional preoperative assessments and holds promise for refining personalized management strategies for patients with endometrial hyperplasia in the future, harmonizing therapeutic efficacy with patient safety, while underscoring inflammation as a pivotal target for future therapeutic research.

Data Sharing Statement

The data supporting the findings of this study are available upon reasonable request from the corresponding authors (Li Qin, Bo Wang) after obtaining approval from the Institutional Ethics Committee of Enshi Tujia and Miao Autonomous Prefecture Central Hospital. Due to ethical and privacy restrictions (to protect the personal information of patients included in the retrospective cohort), the raw data cannot be publicly shared. All relevant processed data and analytical codes used in the study are available from the authors upon reasonable request to support reproducibility.

Acknowledgments

We would like to express our sincere gratitude to Jiali Wang and Qingxiu Ai from the Department of Ultrasound, Enshi Tujia and Miao Autonomous Prefecture Central Hospital, for their valuable contributions to the ultrasound feature evaluation in this study. Their professional expertise, rigorous attitude and careful work in the independent blinded evaluation of ultrasound images ensured the reliability and reproducibility of the ultrasound-derived features, which laid a solid foundation for the smooth progress of this research. We also appreciate their selfless support and assistance during the data collection and verification process. Additionally, we would like to extend our heartfelt thanks to the staff and team of the Department of Pathology, especially Director Yangang Qu, for their professional support and dedicated efforts in the review and confirmation of hematoxylin-eosin stained sections, as well as their assistance in the region of interest segmentation and feature validation, which were crucial to the accuracy and reliability of the pathomics analysis 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 work was supported by the Enshi Prefecture Science and Technology Plan Research and Development Project (Project Title: Mechanism Study on Selenoprotein S Regulating Vaginal Macrophage-Fibroblast Crosstalk in Uterine Prolapse via Energy Metabolism Reprogramming; Principal Investigator: Bo Wang; Project Duration: June 2024 to June 2026).

Disclosure

The authors declare no competing interests in relation to this work.

References

1. Nees LK, Heublein S, Steinmacher S, et al. Endometrial hyperplasia as a risk factor of endometrial cancer. *Arch Gynecol Obstet*. 2022;306(2):407–421. doi:10.1007/s00404-021-06380-5
2. Ring KL, Mills AM, Modesitt SC. Endometrial hyperplasia. *Obstet Gynecol*. 2022;140(6):1061–1075. doi:10.1097/AOG.0000000000004989
3. Management of endometrial intraepithelial neoplasia or atypical endometrial hyperplasia: ACOG Clinical Consensus No. 5. *Obstet Gynecol*. 2023;142(3):735–744. doi:10.1097/AOG.0000000000005297
4. Sanderson PA, Critchley HO, Williams AR, Arends MJ, Saunders PT. New concepts for an old problem: the diagnosis of endometrial hyperplasia. *Hum Reprod Update*. 2017;23(2):232–254. doi:10.1093/humupd/dmw042
5. Petersdorf K, Groettrup-Wolfers E, Overton PM, Seitz C, Schulze-Rath R. Endometrial hyperplasia in pre-menopausal women: a systematic review of incidence, prevalence, and risk factors. *Eur J Obstet Gynecol Reprod Biol*. 2022;271:158–171. doi:10.1016/j.ejogrb.2022.02.015
6. Doherty MT, Sanni OB, Coleman HG, et al. Concurrent and future risk of endometrial cancer in women with endometrial hyperplasia: a systematic review and meta-analysis. *PLoS One*. 2020;15(4):e0232231. doi:10.1371/journal.pone.0232231

7. Uccella S, Zorzato PC, Dababou S, et al. Conservative management of atypical endometrial hyperplasia and early endometrial cancer in childbearing age women. *Medicina*. 2022;58(9):1256. doi:10.3390/medicina58091256
8. Contreras NA, Sabadell J, Verdaguer P, Julià C, Fernández-Montolí ME. Fertility-sparing approaches in atypical endometrial hyperplasia and endometrial cancer patients: current evidence and future directions. *Int J Mol Sci*. 2022;23(5):2531. doi:10.3390/ijms23052531
9. Hwang WY, Suh DH, Kim K, No JH, Kim YB. Aspiration biopsy versus dilatation and curettage for endometrial hyperplasia prior to hysterectomy. *Diagn Pathol*. 2021;16(1):7. doi:10.1186/s13000-020-01065-0
10. Braun MM, Overbeek-Wager EA, Grumbo RJ. Diagnosis and management of endometrial cancer. *Am Fam Physician*. 2016;93(6):468–474.
11. Wang J, Dai J, Cheng Y, Wang X, Deng R, Zhu H. Advances in the use of radiomics and pathomics for predicting the efficacy of neoadjuvant therapy in tumors. *Transl Oncol*. 2025;58:102435. doi:10.1016/j.tranon.2025.102435
12. Bacanakgil BH, Kaban I, Unal F, Guven R, Sahin E, Yildirim SG. Predictive value of hematological inflammatory markers in endometrial neoplasia. *Asian Pac J Cancer Prev*. 2018;19(6):1529–1532. doi:10.22034/APJCP.2018.19.6.1529
13. Petric AN, Živadinović R, Mitić D, Stanojević M, Živadinović A, Kostić I. Hematological and biochemical markers in determining the diagnosis and stage prediction of endometrial cancer. *Ginekol Pol*. 2023;94(4):283–290. doi:10.5603/GP.a2022.0038
14. Kubyshkin AV, Aliev LL, Fomochkina II, et al. Endometrial hyperplasia-related inflammation: its role in the development and progression of endometrial hyperplasia. *Inflamm Res*. 2016;65(10):785–794. doi:10.1007/s00011-016-0960-z
15. Alper ECD, Coşkun ADE, Vural F. Comparison of nonspecific inflammatory markers in endometrial cancer and hyperplasia. *Revista da Associacao Medica Brasileira*. 2021;67(7):966–970. doi:10.1590/1806-9282.20210318
16. Kaur H, Qadri S, Nevill AM, Ewies AAA. The optimal endometrial thickness threshold for prediction of endometrial cancer in postmenopausal women without bleeding remains uncertain-Systematic review and meta-analysis. *J Gynecol Obstet Hum Reprod*. 2024;53(10):102831. doi:10.1016/j.jogoh.2024.102831
17. Selen S, Kilic F, Kimyon Comert G, et al. Can preoperative inflammatory markers differentiate endometrial cancer from complex atypical hyperplasia/endometrial intraepithelial neoplasia? *J Obstet Gynaecol Res*. 2020;46(7):1148–1156. doi:10.1111/jog.14314
18. Travaglino A, Raffone A, Saccone G, et al. Endometrial hyperplasia and the risk of coexistent cancer: WHO versus EIN criteria. *Histopathology*. 2019;74(5):676–687. doi:10.1111/his.13776
19. Ghoubara A, Price MJ, Fahmy MSE, Ait-Allah AS, Ewies A. Prevalence of hyperplasia and cancer in endometrial polyps in women with postmenopausal bleeding: a systematic review and meta-analysis. *Post Reprod Health*. 2019;25(2):86–94. doi:10.1177/2053369119833583
20. Li Z, Yin J, Liu Y, Zeng F. A risk prediction model for endometrial hyperplasia/endometrial carcinoma in premenopausal women. *Sci Rep*. 2025;15(1):1019. doi:10.1038/s41598-024-83568-0
21. Giannella L, Cerami LB, Setti T, Bergamini E, Boselli F. Prediction of endometrial hyperplasia and cancer among premenopausal women with abnormal uterine bleeding. *Biomed Res Int*. 2019;2019:8598152. doi:10.1155/2019/8598152
22. Berceanu C, Cernea N, Căpitănescu RG, et al. Endometrial polyps. *Rom J Morphol Embryol*. 2022;63(2):323–334. doi:10.47162/RJME.63.2.04
23. Kobayashi H, Umetani M, Nishio M, Shigetomi H, Imanaka S, Hashimoto H. Molecular mechanisms of cellular senescence in age-related endometrial dysfunction. *Cells*. 2025;14(12):858. doi:10.3390/cells14120858
24. Cornel KMC, Bongers MY, Kruitwagen R, Romano A. Local estrogen metabolism (intracrinology) in endometrial cancer: a systematic review. *Mol Cell Endocrinol*. 2019;489:45–65. doi:10.1016/j.mce.2018.10.004
25. Chandra V, Kim JJ, Benbrook DM, Dwivedi A, Rai R. Therapeutic options for management of endometrial hyperplasia. *J Gynecol Oncol*. 2016;27(1):e8. doi:10.3802/jgo.2016.27.e8
26. Brancato V, Garbino N, Aiello M, Salvatore M, Cavaliere C. Exploratory analysis of radiomics and pathomics in uterine corpus endometrial carcinoma. *Sci Rep*. 2024;14(1):30727. doi:10.1038/s41598-024-78987-y
27. Zhou L, Zheng L, Hong C, et al. A novel model for predicting microsatellite instability in endometrial cancer: integrating deep learning-pathomics and MRI-based radiomics. *Acad Radiol*. 2025;32(10):5940–5950. doi:10.1016/j.acra.2025.07.050
28. Travaglino A, Raffone A, Saccone G, et al. Complexity of glandular architecture should be reconsidered in the classification and management of endometrial hyperplasia. *APMIS*. 2019;127(6):427–434. doi:10.1111/apm.12945
29. Murali R, Davidson B, Fadare O, et al. High-grade endometrial carcinomas: morphologic and immunohistochemical features, diagnostic challenges and recommendations. *Int J Gynecol Pathol*. 2019;38(Suppl 1):S40–s63. doi:10.1097/PGP.0000000000000491
30. Tempest N, Hill CJ, Maclean A, et al. Novel microarchitecture of human endometrial glands: implications in endometrial regeneration and pathologies. *Hum Reprod Update*. 2022;28(2):153–171. doi:10.1093/humupd/dmab039
31. Dong X, Zhang Y, Liu S, et al. Effect of chronic endometritis on prognosis and reproductive outcomes in infertile women with endometrial hyperplasia. *Am J Reprod Immunol*. 2025;93(5):e70075. doi:10.1111/aji.70075
32. Kumari B, Haldar D, Singh GR, Kumari S, Pankaj S, Prasad R. Diagnostic, prognostic, and predictive importance of neutrophil/lymphocyte ratio (NLR), platelet/lymphocyte ratio (PLR), and CA125 in endometrial hyperplasia and carcinoma. *Indian J Pathol Microbiol*. 2024;67(3):581–584. doi:10.4103/ijpm.ijpm_655_23
33. Ural ÜM, Şehitoğlu İ, Tekin YB, Şahin FK. Neutrophil-to-lymphocyte and platelet-to-lymphocyte ratios in patients with endometrial hyperplasia and endometrial cancer. *J Obstet Gynaecol Res*. 2015;41(3):445–448. doi:10.1111/jog.12536
34. Chen J, Yang P, Li S, Feng Y. Increased FOXM1 expression was associated with the prognosis and the recruitment of neutrophils in endometrial cancer. *J Immunol Res*. 2023;2023:5437526. doi:10.1155/2023/5437526
35. Li H, Liao H, Jing B, Wang Y. Effects of coagulation function indicators and tumor markers on diagnosis and clinicopathological characteristics of endometrial cancer. *Int J Biol Markers*. 2023;38(3–4):214–222. doi:10.1177/03936155231196253
36. Ge L, Liu G, Hu K, et al. A new risk index combining d-Dimer, Fibrinogen, HE4, and CA199 differentiates suspecting endometrial cancer from patients with abnormal vaginal bleeding or discharge. *Technol Cancer Res Treat*. 2020;19:1533033819901117. doi:10.1177/1533033819901117
37. Long B, Clarke MA, Morillo ADM, Wentzensen N, Bakkum-Gamez JN. Ultrasound detection of endometrial cancer in women with postmenopausal bleeding: systematic review and meta-analysis. *Gynecol Oncol*. 2020;157(3):624–633. doi:10.1016/j.ygyno.2020.01.032
38. Xu J, Rao X, Lu W, Xie X, Wang X, Li X. Noninvasive predictor for premalignant and cancerous lesions in endometrial polyps diagnosed by ultrasound. *Front Oncol*. 2021;11:812033. doi:10.3389/fonc.2021.812033