Observational Study

Medicine

OPEN

Precision identification of endometrial malignancy and precancerous lesions

Development of a machine learning model incorporating multidimensional clinical and imaging parameters

Jinpeng Jiang, MD[a], Xunya Xiong, MD[b], Meiling Li, MD[a], Dan Ma, MD[b,*]

Abstract

To develop and validate a machine learning (ML) model integrating multidimensional clinical, pathomic, and ultrasound radiomic parameters for precise identification of endometrial malignancy and precancerous lesions, with a focus on addressing the diagnostic needs of younger patients pursuing fertility preservation. This retrospective study enrolled patients with suspected endometrial lesions from a single institution. Clinical baseline data (e.g., age, body mass index, menopausal status), pathomic features (e.g., cellular atypia, gland density), and ultrasound radiomic parameters (e.g., endometrial thickness, resistance index) were collected. Key predictors were selected using Pearson correlation, SHapley Additive exPlanations analysis, and least absolute shrinkage and selection operator regression. Seven ML models were constructed and optimized via 5-fold cross-validation. Model performance was evaluated using metrics such as ROC-AUC, sensitivity, specificity, and precision–recall performance in training (70%) and testing (30%) sets. A total of 1221 patients (854 in training, 367 in testing) were included. Seven variables including age, body mass index, menopausal status, cellular atypia, gland density, endometrial thickness, and resistance index emerged as robust predictors. Among the 7 ML models, the random forest model showed superior performance, with receiver operating characteristic area under curve of 0.98 in the training set and 0.96 in the testing set (95% confidence interval [CI]: 0.93–0.98). It had balanced sensitivity (0.89, 95% CI: 0.75–0.96) and specificity (0.86, 95% CI: 0.82–0.90) in the testing set. It maintained stability across varying risk thresholds and cost-benefit ratios, outperforming other models in precision–recall balance. Integration of multidimensional data via ML, particularly the random forest model, enhances the precision of endometrial malignancy detection. This approach enables personalized risk stratification, supporting targeted management for younger patients and advancing patient-centered care.

Abbreviations: BMI = body mass index, CI = confidence interval, DT = decision tree, EC = endometrial cancer, Enet = elastic net, LASSO = least absolute shrinkage and selection operator, ML = machine learning, MLP = multilayer perceptron, PRC = precision–recall curve, RF = random forest, ROC-AUC = receiver operating characteristic area under curve, SHAP = SHapley Additive exPlanations, XGBoost = extreme gradient boosting.

Keywords: diagnostic model, endometrial malignancy, machine learning, multidimensional parameters, random forest

1. Introduction

The global burden of endometrial cancer (EC) has escalated markedly, driven by rising metabolic syndrome prevalence and evolving reproductive behaviors. This upward trend is pronounced in certain regions, where shifting lifestyles and delayed childbearing have created a growing cohort grappling with concurrent cancer management and fertility preservation needs. Globally, EC has become the most prevalent gynecological malignancy in high-income countries, with increasing incidence and mortality across developing nations.[1,2] These trends underscore an urgent need for enhanced clinical strategies.

Complicating the management of EC is the expanding cohort of younger patients, with approximately 25% of cases affecting premenopausal women and 3% to 10% diagnosed prior to

ML and DM contributed to this article equally.

JJ and XX contributed to this article equally.

The authors have no funding and conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are publicly available.

Supplemental Digital Content is available for this article.

[a] Department of Gynecology, Xianning Central Hospital, The First Affiliated Hospital of Hubei University of Science and Technology, Xianning, Hubei, China, [b] Department of Oncology, Xianning Central Hospital, The First Affiliated Hospital of Hubei University of Science and Technology, Xianning, Hubei, China.

** Correspondence: Dan Ma, Department of Oncology, Xianning Central Hospital, The First Affiliated Hospital of Hubei University of Science and Technology, 228 Jingui Road, Xian'an District, Xianning City 437100, Hubei Province, China (e-mail: 18371540758@163.com).*

How to cite this article: Jiang J, Xiong X, Li M, Ma D. Precision identification of endometrial malignancy and precancerous lesions: Development of a machine learning model incorporating multidimensional clinical and imaging parameters. Medicine 2026;105:2(e47257).

Received: 17 September 2025 / Received in final form: 21 November 2025 / Accepted: 3 December 2025

http://dx.doi.org/10.1097/MD.0000000000047257

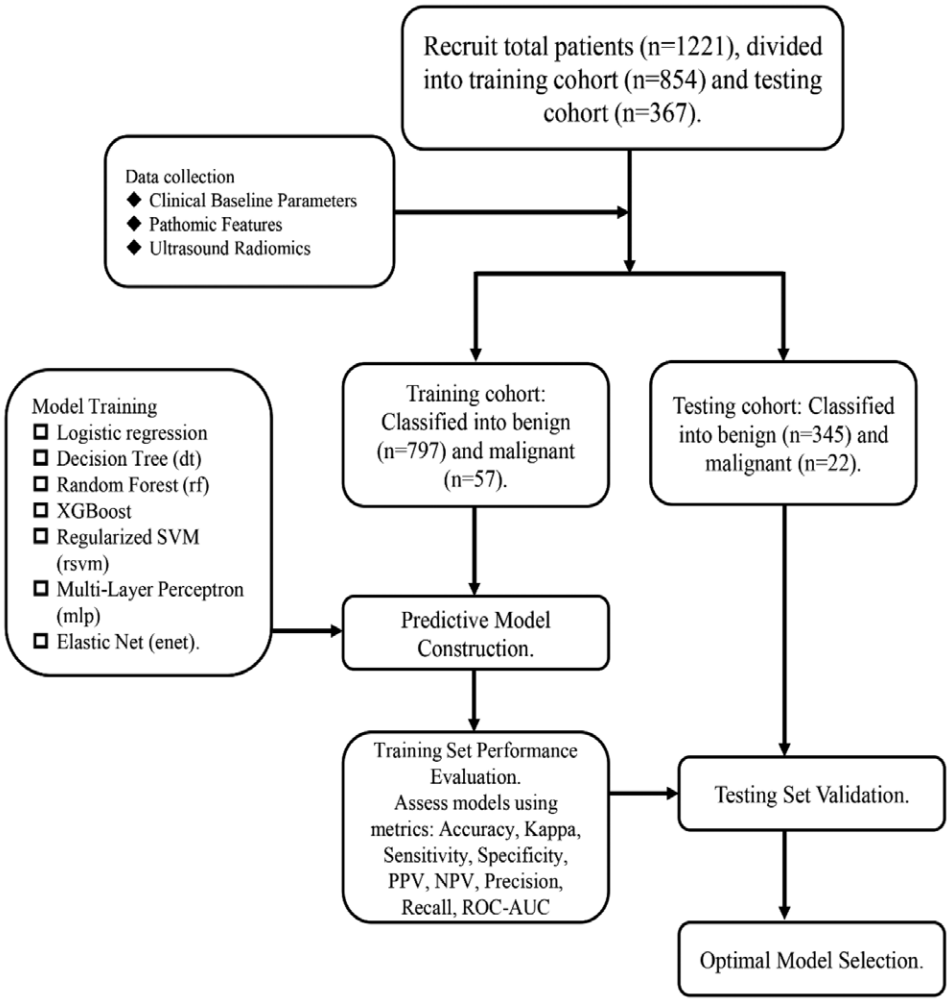

Figure 1. Flowchart for patient inclusion and predictive model construction.

40 years of age.[3–5] Socioeconomic factors delaying childbearing, coupled with evolving family planning policies in various regions, mean a significant proportion of these younger patients desire fertility preservation, conflicting with traditional radical surgical approaches. This disparity highlights the critical need for nuanced diagnostic tools to balance cancer control with reproductive aspirations, particularly given atypical endometrial hyperplasia is a precancerous lesion with 8.2% to 27.5% progression risk to EC and represents a key intervention point.[6,7]

While progestin-based fertility-sparing therapies show promise with remission rates of 75.3% to 88.7%, their success depends on early, accurate diagnosis.[8,9] Current modalities like endometrial biopsy and hysteroscopy suffer from invasiveness, variable accuracy, and inconsistency, limiting their clinical utility.[10–12] This diagnostic gap underscores the need for more reliable approaches to identify early endometrial neoplasia and precancerous changes.

Advanced computational methodologies present a viable resolution, with machine learning (ML) algorithms uniquely equipped to synthesize multidimensional datasets.[13,14] By incorporating clinical baseline metrics, pathomic characteristics, and ultrasound radiomic profiles, a predictive framework can leverage complementary perspectives that encompass patient-specific risk indicators, quantitative pathological indices, and imaging-derived patterns. This research endeavors to develop and validate such a model, aiming to enhance early detection, refine risk stratification, and enable personalized therapeutic strategies. Ultimately, this endeavor seeks to alleviate the burden of EC while addressing the fertility requirements of younger patients.

2. Materials and methods

2.1. Study population

This retrospective study enrolled patients who underwent evaluation for suspected endometrial lesions at Department of Gynecology, Xianning Central Hospital between January 2018 and December 2024. Eligibility criteria included: presence of clinical indicators for endometrial assessment (e.g., abnormal uterine bleeding, postmenopausal bleeding); availability of complete clinical records, endometrial curettage pathology reports, and three-dimensional ultrasound images; no prior history of endometrial malignancy or receipt of definitive treatment (e.g., radiotherapy, chemotherapy) before enrollment. Patients were excluded if they had incomplete data, poor-quality ultrasound images, or concurrent gynecological malignancies (e.g., ovarian cancer). The study protocol was approved by the Institutional Review Board of Xianning Central Hospital, and all procedures adhered to the Declaration of Helsinki. Written informed consent was obtained from all participants prior to data collection. The patient inclusion and construction of the prediction model are elaborated in Figure 1.

2.2. Data collection

2.2.1. Clinical baseline parameters. Demographic and clinical data were extracted from electronic medical records, including age, body mass index (BMI), menstrual status (premenopausal/postmenopausal), parity, history of diabetes

Table 1

Comparison of variables between benign and malignant cases in training and testing sets.

	Training set				Testing set			
Variables	**Overall (n = 854)**	**Benign (n = 797)**	**Malignant (n = 57)**	***P*-value**	**Overall (n = 367)**	**Benign (n = 345)**	**Malignant (n = 22)**	***P*-value**
Age (median [IQR]) (yr)	46.00 [41.00, 51.00]	46.00 [41.00, 50.00]	59.00 [54.00, 64.00]	<.001	47.00 [41.00, 51.00]	46.00 [41.00, 50.00]	60.50 [55.00, 67.75]	<.001
BMI (median [IQR]) (kg/m^2)	24.02 [22.10, 25.98]	23.85 [21.99, 25.67]	26.31 [24.80, 28.87]	<.001	24.02 [22.16, 25.77]	23.99 [22.05, 25.70]	25.46 [23.49, 28.74]	.006
Parity (median [IQR])	2.00 [1.00, 3.00]	2.00 [1.00, 3.00]	2.00 [1.00, 3.00]	.03	2.00 [1.00, 3.00]	2.00 [1.00, 3.00]	2.00 [1.25, 3.00]	.616
Menopause (%)								
Yes	208 (24.4)	166 (20.8)	42 (73.7)		91 (24.8)	74 (21.4)	17 (77.3)	
No	646 (75.6)	631 (79.2)	15 (26.3)	<.001	276 (75.2)	271 (78.6)	5 (22.7)	<.001
Hypertension (%)								
Yes	199 (23.3)	182 (22.8)	17 (29.8)		84 (22.9)	75 (21.7)	9 (40.9)	
No	655 (76.7)	615 (77.2)	40 (70.2)	.297	283 (77.1)	270 (78.3)	13 (59.1)	.07
Diabetes (%)								
Yes	112 (13.1)	97 (12.2)	15 (26.3)		53 (14.4)	47 (13.6)	6 (27.3)	
No	742 (86.9)	700 (87.8)	42 (73.7)	.004	314 (85.6)	298 (86.4)	16 (72.7)	.146
Smoking (%)								
Yes	146 (17.1)	136 (17.1)	10 (17.5)		63 (17.2)	60 (17.4)	3 (13.6)	
No	708 (82.9)	661 (82.9)	47 (82.5)	1	304 (82.8)	285 (82.6)	19 (86.4)	.872
Oral contraceptive (%)								
Yes	191 (22.4)	185 (23.2)	6 (10.5)		78 (21.3)	74 (21.4)	4 (18.2)	
No	663 (77.6)	612 (76.8)	51 (89.5)	.04	289 (78.7)	271 (78.6)	18 (81.8)	.925
Hormone therapy (%)								
Yes	106 (12.4)	91 (11.4)	15 (26.3)		58 (15.8)	51 (14.8)	7 (31.8)	
No	748 (87.6)	706 (88.6)	42 (73.7)	.002	309 (84.2)	294 (85.2)	15 (68.2)	.068
PCOS (%)								
Yes	87 (10.2)	82 (10.3)	5 (8.8)		39 (10.6)	35 (10.1)	4 (18.2)	
No	767 (89.8)	715 (89.7)	52 (91.2)	.889	328 (89.4)	310 (89.9)	18 (81.8)	.407
Polyp history (%)								
Yes	71 (8.3)	67 (8.4)	4 (7.0)		32 (8.7)	29 (8.4)	3 (13.6)	
No	783 (91.7)	730 (91.6)	53 (93.0)	.906	335 (91.3)	316 (91.6)	19 (86.4)	.65
Abnormal bleeding (%)								
Yes	498 (58.3)	454 (57.0)	44 (77.2)		221 (60.2)	202 (58.6)	19 (86.4)	
No	356 (41.7)	343 (43.0)	13 (22.8)	.004	146 (39.8)	143 (41.4)	3 (13.6)	.018
Cycle regularity (median [IQR])	2.13 [1.70, 2.63]	2.18 [1.75, 2.66]	1.28 [0.96, 1.70]	<.001	2.22 [1.71, 2.70]	2.25 [1.74, 2.73]	1.50 [0.90, 1.87]	<.001
Bleeding duration (median [IQR]), (days)	5.22 [3.71, 6.48]	5.22 [3.72, 6.48]	5.10 [3.57, 6.44]	.949	5.22 [3.79, 6.62]	5.18 [3.80, 6.54]	5.63 [3.54, 7.44]	.514
Endometrial thickness (median [IQR]) (mm)	7.90 [6.57, 9.16]	7.74 [6.45, 8.88]	10.80 [10.18, 12.47]	<.001	7.94 [6.58, 9.09]	7.80 [6.51, 8.87]	10.59 [10.19, 11.49]	<.001
Polyp count (median [IQR])	1.28 [0.66, 1.86]	1.28 [0.66, 1.82]	1.47 [0.84, 2.14]	.144	1.26 [0.69, 1.81]	1.22 [0.67, 1.78]	1.94 [1.22, 2.15]	.006
Cellular atypia (median [IQR])	1.02 [0.65, 1.38]	0.96 [0.63, 1.32]	2.36 [1.77, 2.82]	<.001	1.00 [0.65, 1.35]	0.97 [0.62, 1.30]	2.03 [1.61, 2.78]	<.001
Gland density (median [IQR])	2.26 [1.72, 2.81]	2.18 [1.66, 2.68]	3.38 [2.91, 3.94]	<.001	2.22 [1.65, 2.82]	2.18 [1.58, 2.69]	3.50 [2.98, 3.80]	<.001
Stromal density (median [IQR])	3.08 [2.40, 3.72]	3.08 [2.39, 3.70]	3.35 [2.62, 3.99]	.053	3.05 [2.48, 3.76]	3.05 [2.48, 3.72]	3.24 [2.52, 4.17]	.351
Inflammation degree (median [IQR])	1.54 [0.99, 2.11]	1.54 [1.02, 2.12]	1.29 [0.73, 1.79]	.063	1.45 [0.88, 2.07]	1.43 [0.86, 2.07]	1.70 [1.08, 2.01]	.405
Polyp length (median [IQR]) (mm)	11.89 [9.43, 14.76]	11.64 [9.29, 14.33]	19.01 [15.87, 22.00]	<.001	12.05 [9.62, 14.87]	11.76 [9.38, 14.46]	18.98 [15.95, 22.02]	<.001
Polyp width (median [IQR]) (mm)	8.61 [6.53, 10.67]	8.63 [6.54, 10.70]	8.30 [6.23, 10.31]	.709	8.42 [6.39, 10.59]	8.51 [6.50, 10.66]	6.57 [5.68, 9.49]	.057
Polyp volume (median [IQR]) (cm^3)	1.81 [1.14, 2.40]	1.80 [1.15, 2.40]	1.92 [1.13, 2.47]	.264	1.81 [1.13, 2.43]	1.82 [1.17, 2.43]	1.79 [0.92, 2.34]	.591
Resistance index (median [IQR])	0.67 [0.59, 0.75]	0.68 [0.61, 0.76]	0.52 [0.43, 0.60]	<.001	0.67 [0.59, 0.75]	0.67 [0.60, 0.76]	0.49 [0.44, 0.53]	<.001
Pulsatility index (median [IQR])	1.23 [1.00, 1.42]	1.23 [1.00, 1.42]	1.19 [0.99, 1.36]	.552	1.22 [1.00, 1.39]	1.22 [1.00, 1.39]	1.19 [1.04, 1.40]	.917
Echo homogeneity (median [IQR])	2.21 [1.70, 2.70]	2.27 [1.79, 2.73]	1.25 [0.95, 1.98]	<.001	2.26 [1.75, 2.70]	2.28 [1.81, 2.73]	1.33 [0.48, 2.01]	<.001
Location (median [IQR])	2.58 [1.81, 3.32]	2.59 [1.81, 3.33]	2.49 [1.92, 3.13]	.645	2.46 [1.71, 3.30]	2.47 [1.76, 3.28]	2.28 [1.66, 3.42]	.757
Border regularity (median [IQR])	2.26 [1.66, 2.85]	2.26 [1.66, 2.85]	2.31 [1.53, 2.99]	.982	2.32 [1.62, 2.82]	2.32 [1.62, 2.81]	2.53 [1.17, 3.07]	.739

BMI = body mass index, IQR = interquartile range, PCOS = polycystic ovary syndrome.

mellitus, hypertension, polycystic ovary syndrome, use of hormonal medications (e.g., oral contraceptives), and presence of symptoms (e.g., abnormal uterine bleeding duration).

2.2.2. Pathomic features. Endometrial tissue samples were obtained via endometrial curettage, and pathological examinations were performed by 2 experienced pathologists (with ≥5 years of expertise in gynecological pathology) who were blinded to clinical and imaging data. Discrepancies in diagnosis were resolved through consensus. Extracted pathomic parameters included: pathological diagnosis (e.g., normal endometrium, benign hyperplasia, atypical endometrial

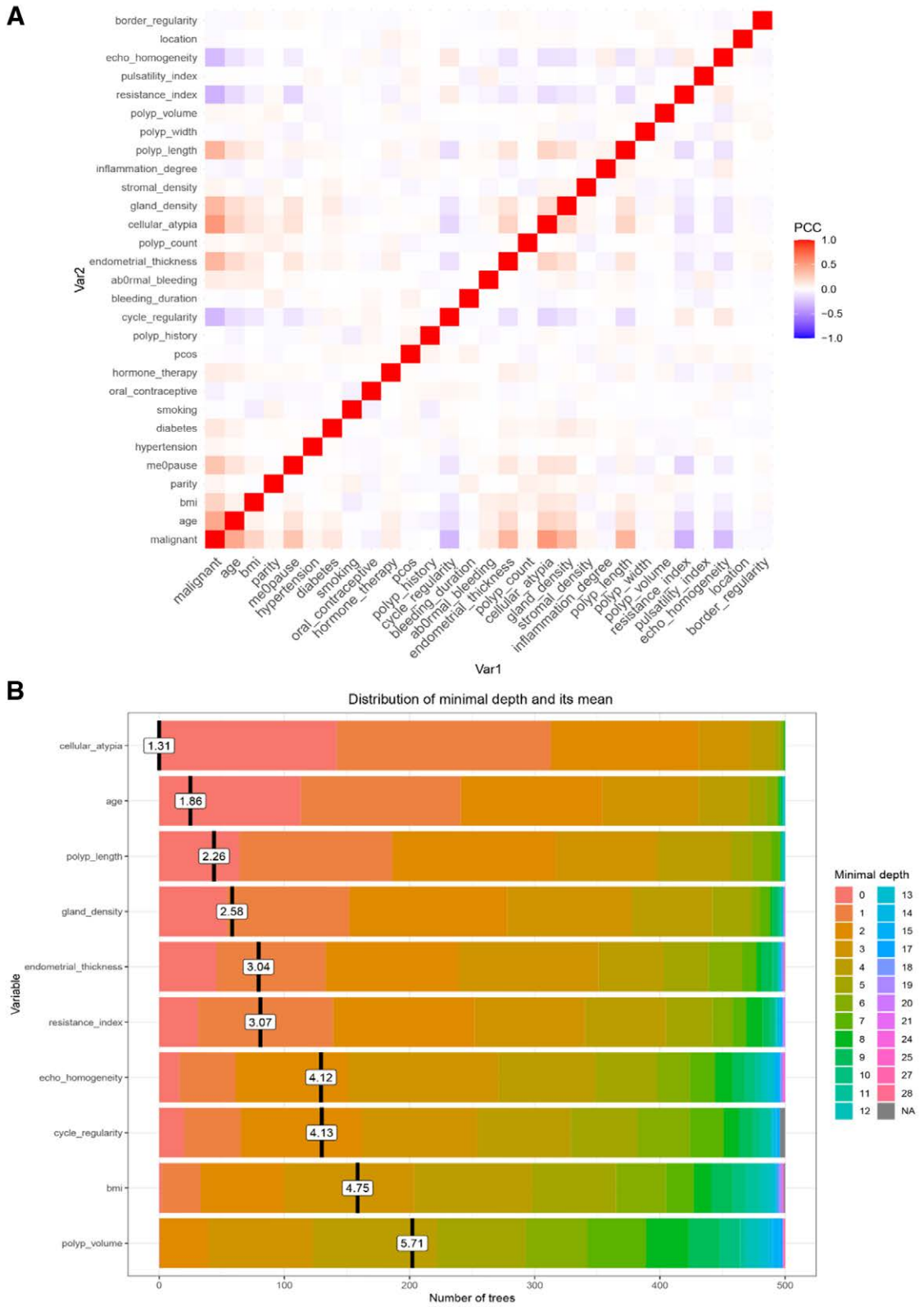

Figure 2. Associations between all candidate variables and outcome variable. (A) Pearson correlation analysis of endometrial malignancy; (B) SHapley Additive exPlanations (SHAP) analysis based on candidate variables.

hyperplasia, EC); presence of glandular crowding or atypia; nuclear pleomorphism; mitotic index.

Cellular atypia: Scored on a 0 to 3 scale (0 = no atypia; 1 = mild atypia [slight nuclear pleomorphism, uniform chromatin]; 2 = moderate atypia [moderate nuclear pleomorphism, irregular chromatin distribution]; 3 = severe atypia [marked nuclear pleomorphism, clumped chromatin, abnormal mitoses]).

Gland density: Scored on a 0 to 3 scale (0 = sparse [glands occupy < 25% of the tissue area]; 1 = mild crowding [25–50%]; 2 = moderate crowding [51–75%]; 3 = severe crowding [>75%], with glandular fusion).

2.2.3. Ultrasound radiomics. All three-dimensional ultrasound examinations were performed using either the GE Voluson E10 or Philips Affiniti 70 system, with a 5 to 9 MHz transvaginal probe or a 3.5 to 5 MHz transabdominal probe, based on patient suitability. Images were acquired by 2 radiologists (with ≥3 years of experience in gynecological imaging) and stored in DICOM format. For radiomic feature extraction: image preprocessing: DICOM images were converted to standardized formats, and regions of interest were manually segmented by the radiologists to outline the endometrial cavity, with interobserver consistency verified using intraclass correlation coefficients. Feature extraction: A total of 44 radiomic features were extracted using Pyradiomics software, including first-order statistics (e.g., mean, variance), shape features (e.g., volume, surface area), and texture features (e.g., entropy, correlation via gray-level co-occurrence matrices). Feature selection: redundant features were removed using variance inflation factor analysis (<10), and key features were selected via least absolute shrinkage and selection operator (LASSO) regression to avoid overfitting.

2.2.4. Model construction and evaluation. The dataset was randomly divided into a training set (70%) and an internal validation set (30%) using stratified sampling to maintain consistent distribution of pathological diagnoses. Missing values were imputed using multiple imputation.Several ML algorithms were tested, including decision tree (DT), elastic net (Enet), logistic regression, multilayer perceptron (MLP), random forest (RF), regularized support vector machine, and extreme gradient boosting (XGBoost). Hyperparameters for each model were optimized using 5-fold cross-validation on the training set. The final model was selected based on the highest performance in cross-validation.

The performance of the predictive model was evaluated using the following metrics: area under the receiver operating characteristic curve (AUC-ROC) to assess discriminative ability; sensitivity, specificity, accuracy, positive predictive value, and negative predictive value at the optimal cutoff (determined by maximizing the Youden index); calibration curve with Hosmer–Lemeshow test to evaluate agreement between predicted and observed probabilities; decision curve analysis and precision–recall curves (PRC) were used to assess clinical utility by quantifying net benefit across different threshold probabilities.

2.3. Statistical analysis

Statistical analyses were performed using R software (Version 4.3.1) and SPSS (Version 26.0). Continuous variables were presented as mean ± standard deviation or median (interquartile range) based on normality, and compared using Student *t* test, analysis of variance, or Kruskal–Wallis test. Categorical variables were presented as counts (percentages) and compared using the Chi-square test or Fisher exact test. A two-tailed *P*-value <.05 was considered statistically significant.

3. Results

3.1. Comparison of variables between benign and malignant cases in training and testing sets

In both the training set (n = 854) and testing set (n = 367), several variables differed significantly between benign and malignant cases (Table 1). Malignant cases had a higher median age (training set: 59.00 vs 46.00 years; testing set: 60.50 vs 46.00 years; both $P < .001$) and median BMI (training set: 26.31 vs 23.85 kg/m², $P < .001$; testing set: 25.46 vs 23.99 kg/m², $P = .006$). A larger proportion of malignant cases were postmenopausal (training set: 73.7% vs 20.8%; testing set: 77.3% vs 21.4%; both $P < .001$). In the training set, malignant cases showed higher rates of diabetes (26.3% vs 12.2%, $P = .004$) and hormone therapy use (26.3% vs 11.4%, $P = .002$), along with lower oral contraceptive use (10.5% vs 23.2%, $P = .04$), though these were not significant in the testing set. Abnormal bleeding was more common in malignant cases (training set: 77.2% vs 57.0%, $P = .004$; testing set: 86.4% vs 58.6%, $P = .018$), while cycle regularity was lower (training set: 1.28 vs 2.18; testing set: 1.50 vs 2.25; both $P < .001$).

Pathologically, malignant cases had higher cellular atypia (training set: 2.36 vs 0.96; testing set: 2.03 vs 0.97; both $P < .001$) and gland density (training set: 3.38 vs 2.18; testing set: 3.50 vs 2.18; both $P < .001$). On three-dimensional ultrasound, malignant cases had greater endometrial thickness (training set: 10.80 vs 7.74 mm; testing set: 10.59 vs 7.80 mm; both $P < .001$) and longer polyp length (training set: 19.01 vs 11.64 mm; testing set: 18.98 vs 11.76 mm; both $P < .001$). They also had a lower resistance index (training set: 0.52 vs 0.68; testing set: 0.49 vs 0.67; both $P < .001$) and lower echo homogeneity (training set: 1.25 vs 2.27; testing set: 1.33 vs 2.28; both $P < .001$). Variables such as hypertension, smoking, polycystic ovary syndrome, polyp history, bleeding duration, polyp width, volume, pulsatility index, location, and border regularity showed no significant differences between groups in either set (all $P > .05$). Overall, age, BMI, menopausal status, abnormal bleeding, cycle regularity, cellular atypia, gland density, endometrial thickness, polyp length, resistance index, and echo homogeneity emerged as key variables distinguishing benign from malignant cases, with consistent significance across both training and testing sets.

Table 2

Core performance metrics of machine learning models on the test set.

Model	Accuracy	Sensitivity	Specificity	ROC-AUC (95% CI)
Logistic regression	0.93	0.93	0.86	0.92 (0.89–0.95)
Decision tree (DT)	0.94	0.98	0.50	0.72 (0.66–0.78)
Random forest (RF)	0.89	0.89	0.86	0.96 (0.93–0.98)
Extreme gradient boosting (XGBoost)	0.83	0.83	0.82	0.90 (0.86–0.94)
Regularized support vector machine (RSVM)	0.87	0.87	0.82	0.90 (0.86–0.94)
Multilayer perceptron (MLP)	0.82	0.82	0.86	0.92 (0.89–0.95)
Elastic net (Enet)	0.89	0.89	0.86	0.93 (0.90–0.96)

CI = confidence interval, ROC-AUC = receiver operating characteristic area under curve.

3.2. Selection of predictive variables

To identify the most relevant predictors for the outcome, multiple analytical approaches were employed, including Pearson correlation analysis, SHapley Additive exPlanations (SHAP) analysis, and LASSO regression. Pearson correlation analysis (Fig. 2A) quantified the strength of associations between candidate variables and the outcome, while SHAP analysis (Fig. 2B) evaluated their relative importance in predictive performance. These methods converged to highlight 7 variables with the strongest and most consistent associations: BMI, age, menopausal status, cellular atypia, gland density, endometrial thickness, and resistance index. Specifically, these variables demonstrated notably higher correlation coefficients in the Pearson analysis and ranked at the top in the SHAP analysis, underscoring their robust predictive value.

Further validation via LASSO regression (Fig. S1, Supplemental Digital Content, https://links.lww.com/MD/R181) reinforced the stability of these 7 variables. The coefficient paths showed they retained nonzero coefficients across varying penalization intensities (log(λ) values), and their coefficient magnitudes remained substantial during the shrinkage process, confirming their reliability as core predictors. Collectively, the consistent results from Pearson correlation, SHAP, and LASSO analyses indicate that BMI, age, menopausal status, cellular atypia, gland density, endometrial thickness, and resistance index are the most robust predictive variables, warranting inclusion in subsequent model construction.

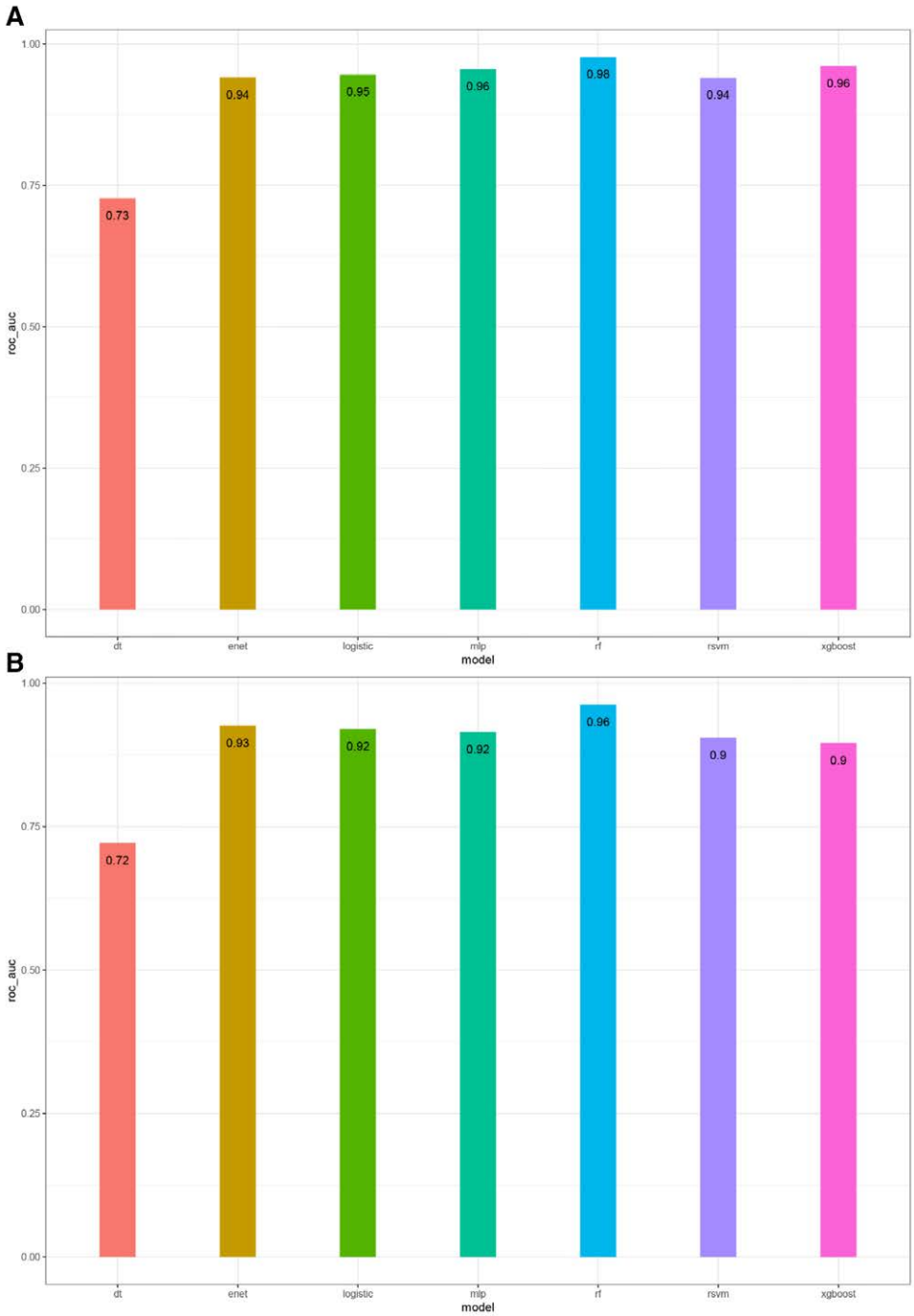

Figure 3. Performance comparison of different models on training set (A) and test set (B).

3.3. Construction and performance evaluation of predictive models

Seven predictive models were developed, including DT, Enet, logistic regression, MLP, RF, regularized support vector machine, and XGBoost, with hyperparameters optimized as visualized in Figures S2 to S8, Supplemental Digital Content, https://links.lww.com/MD/R181. Performance metrics (Table 2 and Table S1, Supplemental Digital Content, https://links.lww.com/MD/R181) revealed distinct variations across models. In the training set, DT achieved the highest accuracy (0.96) but showed poor specificity (0.46) and discriminative ability (receiver operating characteristic area under curve [ROC-AUC] = 0.73). RF exhibited the strongest discriminative power (ROC-AUC = 0.98) with moderate accuracy (0.91), while logistic regression balanced accuracy (0.93) and ROC-AUC (0.95). In the testing set, RF retained the highest ROC-AUC (0.96), followed by Enet (0.93) and logistic regression (0.92); XGBoost and MLP performed relatively lower (ROC-AUC = 0.90 and 0.92, respectively).

Consistent with these metrics, Figure 3 confirmed stable performance of RF and logistic regression across training and testing sets, whereas DT showed a notable drop in discriminative ability in the testing set. Figure 4 further highlighted RF's balanced sensitivity (0.89) and specificity (0.86) in the testing set, with logistic regression excelling in F-measure (0.96). Overall, RF demonstrated the best comprehensive performance with the highest discriminative ability in both datasets, while logistic regression showed robust stability, making them the most promising models for practical application.

3.4. Robustness and predictive efficacy of prediction models

To further evaluate the robustness and predictive efficacy of the developed models, PRC and performance under varying high-risk thresholds and cost-benefit ratios (Fig. 5) were analyzed. PRC (Fig. 6) revealed that in both the training and testing sets, RF maintained the highest precision across most recall values, followed by logistic regression. This indicates that RF and logistic regression consistently balanced the ability to correctly identify positive cases (recall) and the proportion of true positives among predicted positives (precision). DT showed lower precision at intermediate recall levels in both sets, reflecting its relatively weaker performance in balancing these metrics. Figure 5 further demonstrated model robustness under different high-risk thresholds and cost-benefit ratios. In both training and testing sets, RF exhibited stable performance across all thresholds and cost-benefit ratios, with minimal fluctuation in net benefit. Logistic regression also showed good stability, particularly at moderate risk thresholds (0.4–0.6) and balanced cost-benefit ratios (2:3–3:2). In contrast, DT and MLP displayed larger variations in performance under extreme cost-benefit ratios (e.g., 100:1 or 1:100), indicating lower robustness in such scenarios. Overall, RF emerged as the most robust model with superior predictive efficacy, characterized by high and stable precision–recall performance and consistent net benefit across diverse thresholds and cost-benefit scenarios. Logistic regression also showed strong robustness, making it a reliable alternative for practical applications where computational complexity is a consideration.

4. Discussion

The escalating global burden of EC, compounded by the rising number of younger patients seeking fertility preservation, underscores the urgency of advancing diagnostic precision in gynecological oncology.[15,16] Traditional diagnostic methods, such as endometrial biopsy and hysteroscopy, are hampered by invasiveness, variable accuracy, and inconsistency, leaving a

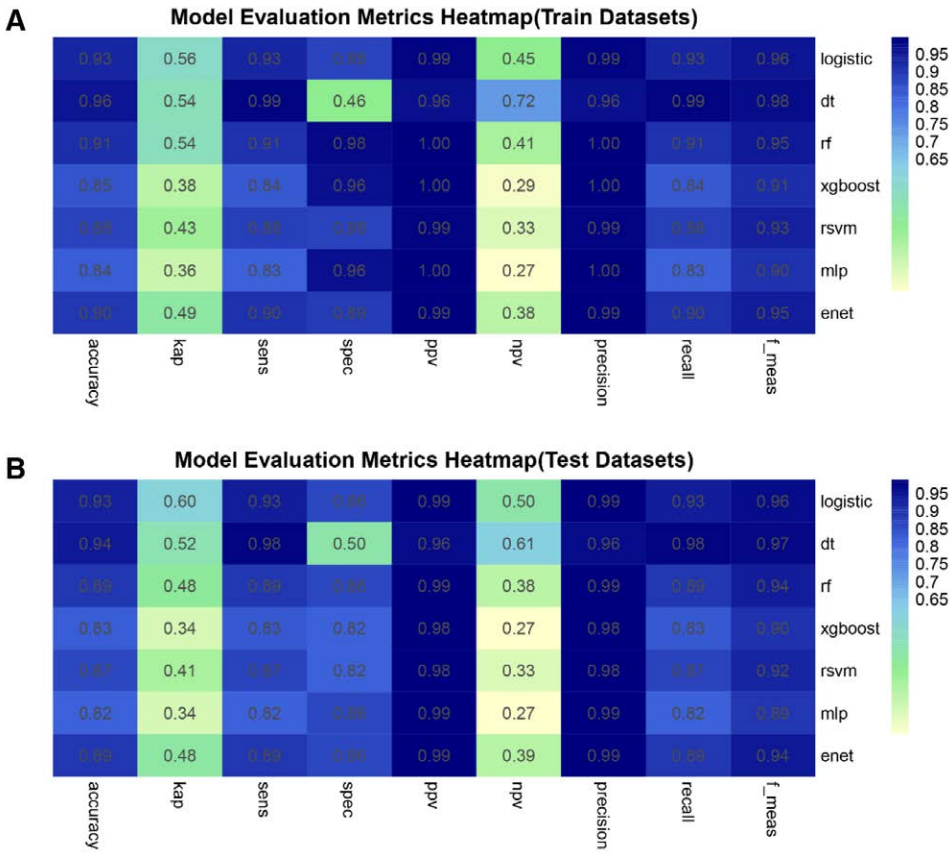

Figure 4. Performance efficacy of predictive models in training set (A) and validation set (B).

critical gap in personalized risk assessment. This study addresses this gap by developing a ML model that integrates clinical baseline data, pathomic features, and ultrasound radiomics, aiming to enhance the early detection of endometrial malignancy and precancerous lesions. The findings not only validate the potential of multidimensional data integration but also provide a framework for refining clinical decision-making in complex scenarios, particularly for patients balancing cancer management with reproductive aspirations.

The identification of key predictive variables, including age, BMI, menopausal status, cellular atypia, gland density, endometrial thickness, and resistance index, aligns with established understanding of EC pathogenesis while introducing novel insights through multimodal integration. Among these, demographic and clinical variables such as age, BMI, and menopausal status are rooted in well-characterized pathophysiological mechanisms. Age is a recognized independent risk factor for EC.[17,18] Cumulative exposure to unopposed estrogen over time drives endometrial proliferation and genomic instability.[19–21] This aligns with our findings that older patients in both training and testing sets showed a significantly higher likelihood of malignant lesions. BMI, a marker of metabolic dysregulation, contributes to EC risk through peripheral estrogen synthesis in adipose tissue, insulin resistance-mediated hyperinsulinemia, and altered adipokine signaling.[22,23] These factors promote endometrial cell

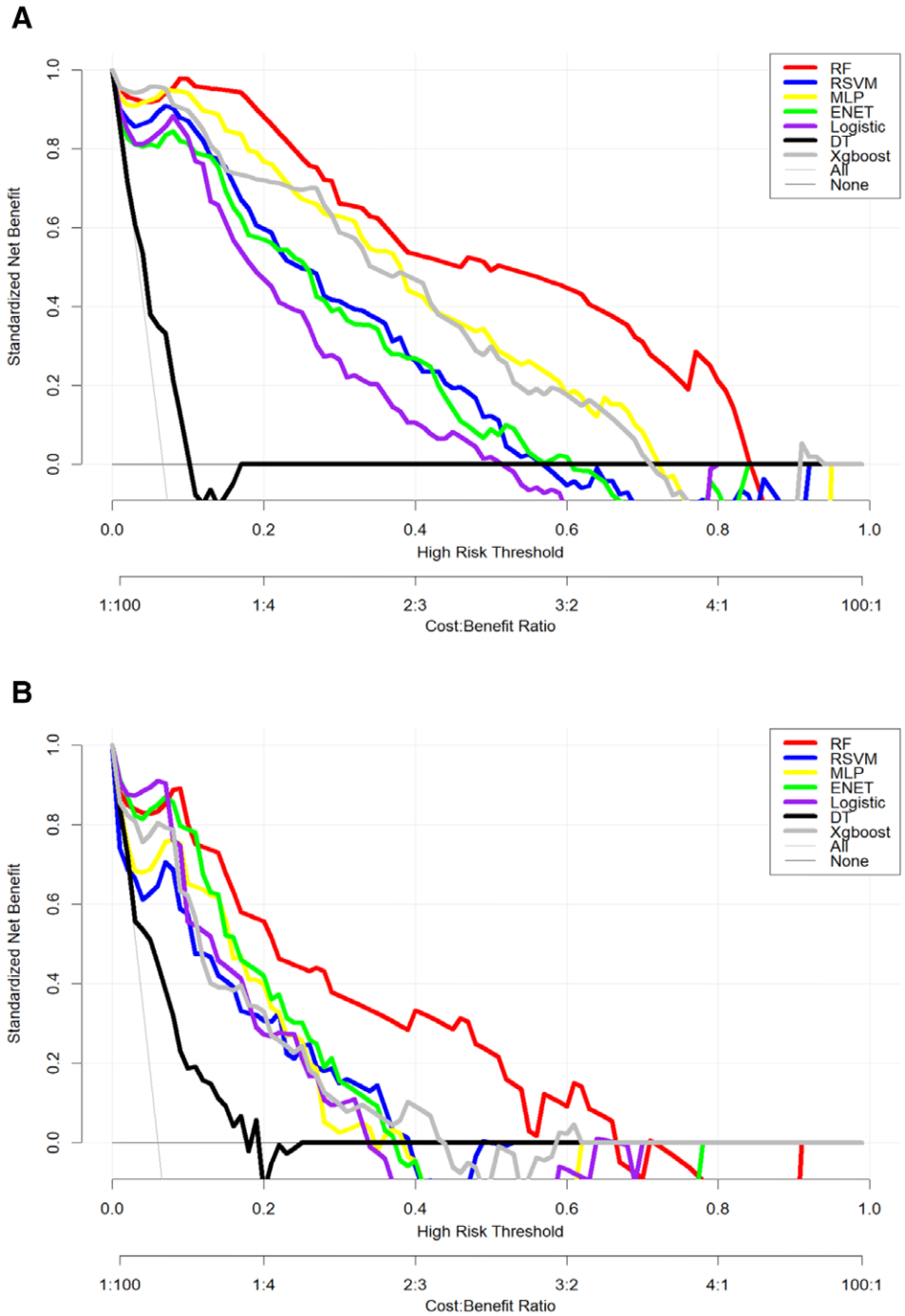

Figure 5. Performance of various models under different high risk thresholds and cost-benefit ratios: (A) training set; (B) testing set.

proliferation and inhibit apoptosis. Menopausal status further amplifies these risks. Postmenopausal women, who constituted a larger proportion of malignant cases in our cohort, often experience a hormonal milieu dominated by estrogen (without cyclic progesterone antagonism) and this creates a pro-carcinogenic environment.[24,25] Collectively, these variables capture the systemic and hormonal foundations of endometrial neoplasia, making them robust predictors in clinical practice.

Pathomic and imaging variables, including cellular atypia, gland density, endometrial thickness, and resistance index, complement these clinical factors by reflecting localized biological changes and structural abnormalities.[26–29] Cellular atypia and gland density are histopathological hallmarks of neoplastic transformation. Atypia indicates dysregulated cellular differentiation and nuclear instability, while increased gland density reflects excessive proliferation of endometrial glands.[30,31] Both are key features of atypical hyperplasia and early EC, as confirmed in our analysis showing significantly higher values in malignant cases. Endometrial thickness, a widely used ultrasound marker, correlates with the extent of proliferative activity.[32] Thicker endometria in malignant cases, as observed here, likely reflect unregulated growth driven by hormonal or genetic aberrations. Resistance index, a vascular parameter, offers unique insights into tumor angiogenesis.[33,34] Malignant lesions require increased blood supply to support rapid growth, leading to the formation of abnormal, low-resistance vessels. This phenomenon is captured by the lower resistance index in our malignant cohort. This vascular remodeling, often an early event in carcinogenesis, underscores the value of integrating functional imaging metrics to detect preclinical changes that may precede overt histological abnormalities.

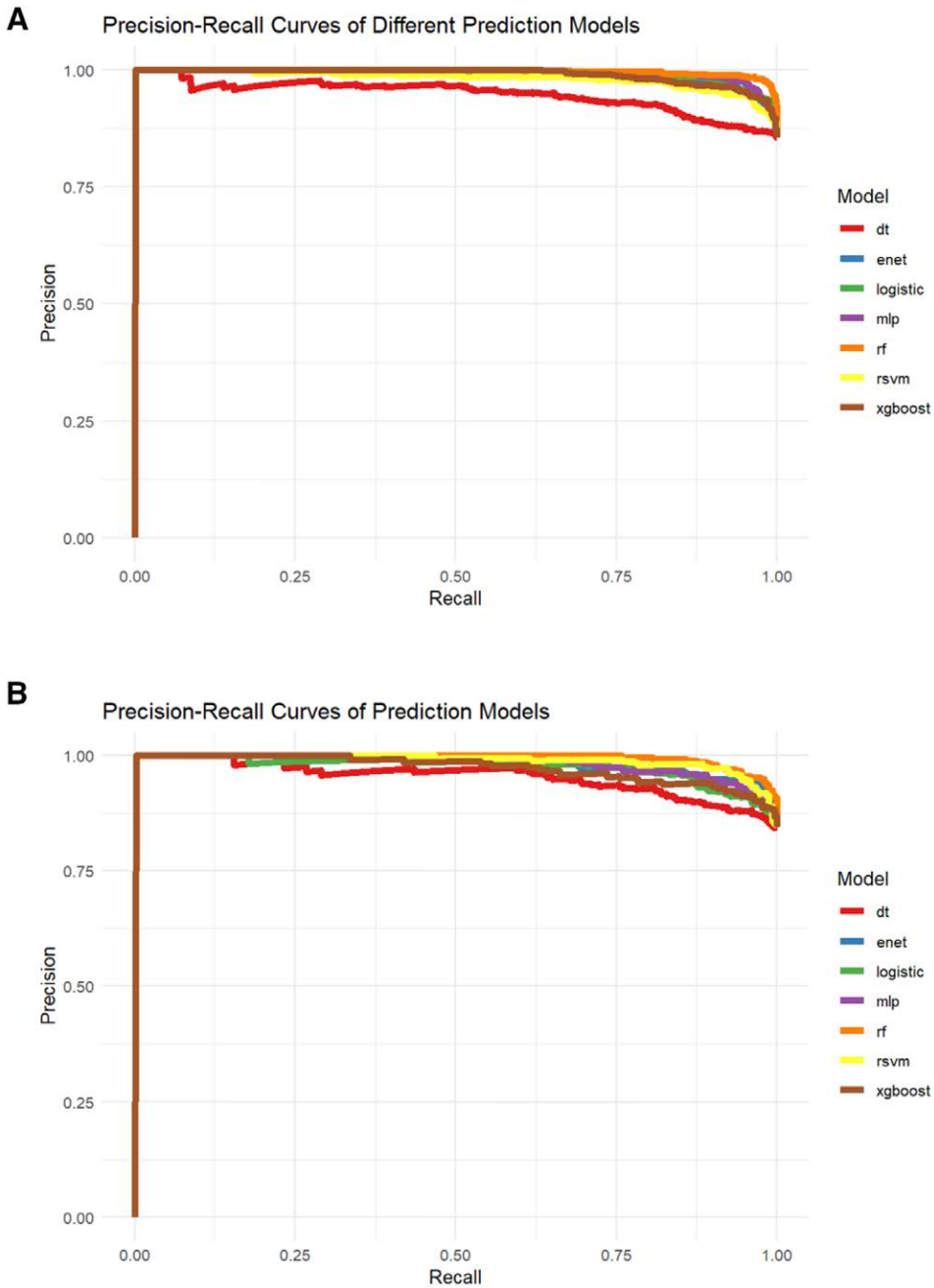

Figure 6. Precision–recall curves of different prediction models. (A) Training set and (B) testing set.

Among the models evaluated, RF demonstrated superior performance, with high ROC-AUC, balanced precision–recall metrics, and stability across varying risk thresholds and cost-benefit scenarios. Its superiority over logistic regression and other algorithms likely stems from its ability to model nonlinear relationships and variable interactions, which is critical for capturing the complex interplay of clinical, pathological, and imaging features in endometrial carcinogenesis. The consistent performance of RF in both training and testing sets, coupled with its high sensitivity and specificity, positions it as a valuable tool for risk stratification. Its ability to maintain precision across diverse recall values supports its potential to identify high-risk individuals for targeted interventions while avoiding unnecessary invasive procedures in low-risk cases, a balance particularly vital for premenopausal patients pursuing fertility preservation.

This study has several limitations. Firstly, its retrospective design and single-institution cohort may introduce selection bias and overfit to local demographics and practices, limiting global generalizability. Secondly, the dataset has class imbalance (79 total malignant cases, 22 in testing), weakening performance stability. Thirdly, subjective manual segmentation for ultrasound radiomics, even with consistency checks, may reduce feature reproducibility. Fourthly, multicenter external validation is needed to confirm performance in diverse settings. Addressing these (e.g., expanding malignant samples, automated segmentation) will be critical to translating findings into clinical tools.

5. Conclusion

In conclusion, integrating clinical, pathomic, and ultrasound radiomic data via ML, particularly the RF model, markedly improves the precision of endometrial malignancy detection. The key identified variables, rooted in hormonal, metabolic, histopathological, and vascular mechanisms, provide a biological basis for personalized risk stratification. This approach balances diagnostic accuracy and clinical utility, addressing the unique needs of younger patients and advancing EC management toward targeted, patient-centered care. Future efforts should focus on prospective validation, multi-center testing, and integration with emerging technologies to further refine its utility.

Author contributions

Conceptualization: Xunya Xiong, Dan Ma.
Data curation: Xunya Xiong, Dan Ma.
Formal analysis: Dan Ma.
Investigation: Jinpeng Jiang, Meiling Li.
Methodology: Jinpeng Jiang, Meiling Li.
Resources: Jinpeng Jiang.
Software: Jinpeng Jiang, Dan Ma.
Supervision: Dan Ma.
Visualization: Xunya Xiong.
Writing – original draft: Jinpeng Jiang, Meiling Li, Dan Ma.
Writing – review & editing: Jinpeng Jiang, Dan Ma.

References

[1] Gao S, Wang J, Li Z, Wang T, Wang J. Global trends in incidence and mortality rates of endometrial cancer among individuals aged 55 years and above from 1990 to 2021: an analysis of the global burden of disease. Int J Women's Health. 2025;17:651–62.

[2] Mazidimoradi A, Momenimovahed Z, Khalajinia Z, Allahqoli L, Salehiniya H, Alkatout I. The global incidence, mortality, and burden of uterine cancer in 2019 and correlation with SDI, tobacco, dietary risks, and metabolic risk factors: an ecological study. Health Sci Rep. 2024;7:e1835.

[3] Wise MR, Jordan V, Lagas A, et al. Obesity and endometrial hyperplasia and cancer in premenopausal women: a systematic review. Am J Obstet Gynecol. 2016;214:689.e1–689.e17.

[4] 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.

[5] Bourou MZ, Matsas A, Vrekoussis T, Mastorakos G, Valsamakis G, Panoskaltsis T. Conservative treatment of endometrial cancer in women of reproductive age (Review). Mol Clin Oncol. 2023;19:55.

[6] Barakat A, Ismail A, Chattopadhyay S, Davies Q. Endometrial cancer incidence in patients with atypical endometrial hyperplasia according to mode of management. Cancer Diagnos Prognos. 2022;2:564–8.

[7] Giannella L, Grelloni C, Bernardi M, et al. Atypical endometrial hyperplasia and concurrent cancer: a comprehensive overview on a challenging clinical condition. Cancers. 2024;16:914.

[8] Wang R, Wang J, Wu Y, Zhu A, Li X, Wang J. Prediction of a radiomics-based nomogram for complete remission in atypical endometrial hyperplasia and early endometrial cancer patients seeking fertility preservation. Comput Biol Med. 2025;187:109764.

[9] Li X, Wang Y, Wang J, Fan Y, Wang J. Prediction of complete regression in fertility-sparing patients with endometrial cancer and apical hyperplasia: the GLOBAL model in a large Chinese cohort. J Transl Med. 2024;22:127.

[10] Vitale SG, Riemma G, Carugno J, et al. Hysteroscopy in the management of endometrial hyperplasia and cancer in reproductive aged women: new developments and current perspectives. Transl Cancer Res. 2020;9:7767–77.

[11] Pace L, Actis S, Mancarella M, et al. Clinical, sonographic, and hysteroscopic features of endometrial carcinoma diagnosed after hysterectomy in patients with a preoperative diagnosis of atypical hyperplasia: a single-center retrospective study. Diagnostics (Basel, Switzerland). 2022;12:3029.

[12] Henderson I, Black N, Khattak H, Gupta JK, Rimmer MP; UKARCOG Working Group Authors. Diagnosis and management of endometrial hyperplasia: a UK national audit of adherence to national guidance 2012–2020. PLoS Med. 2024;21:e1004346.

[13] MacEachern SJ, Forkert ND. Machine learning for precision medicine. Genome. 2021;64:416–25.

[14] Deo RC. Machine learning in medicine. Circulation. 2015;132:1920–30.

[15] Corr BR, Erickson BK, Barber EL, Fisher CM, Slomovitz B. Advances in the management of endometrial cancer. BMJ (Clinical research ed). 2025;388:e080978.

[16] Passarello K, Kurian S, Villanueva V. Endometrial cancer: an overview of pathophysiology, management, and care. Semin Oncol Nurs. 2019;35:157–65.

[17] 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;23:2531.

[18] Wu Y, Sun W, Liu H, Zhang D. Age at menopause and risk of developing endometrial cancer: a meta-analysis. Biomed Res Int. 2019;2019:8584130.

[19] Gong TT, Wang YL, Ma XX. Age at menarche and endometrial cancer risk: a dose-response meta-analysis of prospective studies. Sci Rep. 2015;5:14051.

[20] Forte M, Cecere SC, Napoli M D, et al. Endometrial cancer in the elderly: characteristics, prognostic and risk factors, and treatment options. Crit Rev Oncol Hematol. 2024;204:104533.

[21] Abdol Manap N, Ng BK, Phon SE, Abdul Karim AK, Lim PS, Fadhil M. Endometrial cancer in pre-menopausal women and younger: risk factors and outcome. Int J Environ Res Public Health. 2022;19:9059.

[22] Laskov I, Zilberman A, Maltz-Yacobi L, et al. Effect of BMI change on recurrence risk in patients with endometrial cancer. Int J Gynecol Cancer. 2023;33:713–8.

[23] Kawachi A, Shimazu T, Budhathoki S, et al. Association of BMI and height with the risk of endometrial cancer, overall and by histological subtype: a population-based prospective cohort study in Japan. Eur J Cancer Prevent. 2019;28:196–202.

[24] Clarke MA, Long BJ, Del Mar Morillo A, Arbyn M, Bakkum-Gamez JN, Wentzensen N. Association of endometrial cancer risk with postmenopausal bleeding in women: a systematic review and meta-analysis. JAMA Internal Med. 2018;178:1210–22.

[25] Braun MM, Overbeek-Wager EA, Grumbo RJ. Diagnosis and management of endometrial cancer. Am Fam Physician. 2016;93:468–74.

[26] Vitale SG, Riemma G, Haimovich S, et al. Risk of endometrial cancer in asymptomatic postmenopausal women in relation to ultrasonographic endometrial thickness: systematic review and diagnostic test accuracy meta-analysis. Am J Obstet Gynecol. 2023;228:22–35.e2.

[27] Vikram SR, Robinson J, Thanawala T, et al. Ultrasound and blind endometrial sampling for detection of endometrial cancer in women

with postmenopausal bleeding. Cochrane Database Syst Rev. 2024;6:CD014568.
[28] Ribeiro-Santos P, Martins Vieira C, Viana Veloso GG, et al. Tailoring endometrial cancer treatment based on molecular pathology: current status and possible impacts on systemic and local treatment. Int J Mol Sci . 2024;25:7742.
[29] Galant N, Krawczyk P, Monist M, et al. Molecular classification of endometrial cancer and its impact on therapy selection. Int J Mol Sci . 2024;25:5893.
[30] Chou AJ, Bing RS, Ding DC. Endometrial atypical hyperplasia and risk of endometrial cancer. Diagnostics (Basel, Switzerland). 2024;14:2471.
[31] Leino A, Nostolahti A, Ahtikoski A, Huvila J. Molecular subtype and mutational profile of endometrial atypical hyperplasia/endometrioid intraepithelial neoplasia. Int J Gynecol Pathol. 2025;11:1.
[32] Saccardi C, Spagnol G, Bonaldo G, Marchetti M, Tozzi R, Noventa M. New light on endometrial thickness as a risk factor of cancer: what do clinicians need to know? Cancer Manage Res. 2022;14:1331–40.
[33] Ahmadzade A, Gharibvand MM, Azhine S. Correlation of color Doppler ultrasound and pathological grading in endometrial carcinoma. J Fam Med Primary Care. 2020;9:5188–92.
[34] Huang W, Cen S, Kang XL, Wang WF, Wang Y, Chen X. Doppler ultrasound measurement of resistance index in the diagnosis of prostate cancer. Tumori. 2015;101:644–9.