Journal of Inflammation Research

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

A Machine Learning Model Integrating Preoperative Blood-Based Indices for Early and Noninvasive Detection of Endometrial Cancer

Jia Wang[1,*], HaoTian Wu[2,*], Fei Wang[3], YingXiang Wang[1], Jian Gu[1]

[1]Department of Gynecology, The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, 510000, People's Republic of China; [2]Department of Neurology, The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, 510000, People's Republic of China; [3]Department of Spacial AI, Li Auto Inc., Beijing, 100020, People's Republic of China

*These authors contributed equally to this work

Correspondence: Jian Gu, Email gujian@mail.sysu.edu.cn

Background: Endometrial cancer (EC) incidence is rising globally, yet early diagnosis remains challenging. Our objective is to develop a non-invasive, preoperative tool to predict EC risk using machine learning (ML) techniques.

Methods: This retrospective analysis included clinical data from patients with endometrial lesions at the Third Affiliated Hospital of Sun Yat-sen University between January 2014 to August 2024. Six machine learning techniques including Random Forest (RF), Extreme Gradient Boosting (XGBoost), Support Vector Mac (SVM), Gradient Boosting Machine Model (GBDT), Logistic Regression (LR) and Multi-Layer Perceptron (MLP) were used to construct the prediction model of endometrial cancer. Receiver operating characteristic curve (ROC) was used to evaluate the model. S Hapley Additive ExPlanation (SHAP) analysis was applied to determine the predictive role of each feature in the model with the highest predictive performance.

Results: A total of 857 patients were included in the study. Eight baseline characteristics (Age, BMI, Gravidity, Parity, Family history, Menopause status, Diabetes, Hypertension), one imaging feature (Endometrial thickness) and eight peripheral blood-based markers (WBC, NLR, MLR, PLR, SII, SIRI, CA-125, HE4) were selected for develop and validate the machine learning model, these features were obtained noninvasively. Data from 686 patients were randomly assigned to the training group, and data from 171 patients were used for internal validation. Among the six-machine learning model, GBDT had the highest prediction, the model achieved an AUC of 0.95 (95% CI: 0.93–0.97), accuracy of 90.0% and a Brier score of 0.06. The SHAP analysis showed that HE4, CA-125 and SIRI were the most influential contributors to the prediction.

Conclusion: We developed and validated a GBDT prediction model, which showed the best performance in predicting endometrial cancer. This model can be applied in clinical practice to effectively predict the risk of EC for patients.

Keywords: endometrial cancer, serum inflammatory markers, machine learning, prediction model

Introduction

Endometrial cancer (EC) is a common gynecologic malignancy affecting women's health, its incidence rising in many countries over recent decades. According to the 2022 cancer statistics from the National Cancer Center of China, the incidence rate of EC stands at 6.84 per 100,000, with a mortality rate of 1.05 per 100,000.[1] This increase is hypothesized to be associated with the increasing prevalence of obesity and changes in female reproductive patterns.[2,3] EC develops through a multistep progression originating from normal or hyperplastic endometrium. Endometrial hyperplasia (EH), is histologically characterized by an abnormal increase in the gland-to-stroma ratio accompanied by architectural irregularities in glandular morphology, including variations in both shape and size. This pathological transformation is predominantly driven by prolonged exposure to unopposed estrogen stimulation.[4] Clinical evidence indicates that untreated EH carries a significant risk of malignant transformation.[5] The prognosis of EC is critically dependent on the disease stage at diagnosis. Although approximately 70% of

Received: 11 April 2025
Accepted: 28 July 2025
Published: 12 August 2025

early-stage EC cases are detected due to abnormal vaginal bleeding, nearly 30% of patients present with advanced-stage disease due to the asymptomatic, resulting in significantly poorer clinical outcomes.[6,7] Consequently, early detection of EC is important for improving patient prognosis and survival rates.

Recent studies have shown that both in the treatment and screening of EC, it is recommended to minimize damage for patients and non-invasive,[8] but the absence of a simple, non-invasive screening protocol for EC represents a significant clinical challenge. While hysteroscopy and diagnostic curettage remain the most frequently utilized methods for evaluating endometrial lesions, these approaches are associated with several substantial limitations: invasive, procedural complexity and substantial healthcare costs. Furthermore, repeated applications of these techniques may increase the risk of lesion metastasis and induce intrauterine adhesions –that are particularly relevant for young, nulliparous women. Although fine-needle aspiration offers a minimally invasive method, its diagnostic reliability is compromised, which typically evaluates less than 50% of the uterine cavity. Transvaginal ultrasound (TVUS) remains the first-line imaging modality for evaluating endometrial abnormalities, offering high sensitivity for detecting hyperplasia or polyps. However, its specificity for differentiating benign lesions from early-stage malignancies remains suboptimal. Magnetic resonance imaging (MRI), while superior in assessing myometrial invasion and tumor staging, is cost-prohibitive for routine screening. These limitations underscore the need for complementary non-invasive tools to refine preoperative risk stratification.[9] Previous studies has identified several biomarkers associated with EC clinical features and prognosis;[7,10,11] however, their diagnostic performance remains suboptimal when used in isolation. These diagnostic limitations present substantial challenges in differentiating between endometrial hyperplasia (EH) and early-stage EC. In densely populated nations such as China, there is an urgent need to develop robust, quantitative and cost-efficient predictive models for EC. Such advancements could facilitate timely intervention, and ultimately improve patient outcomes through personalized risk stratification.

Uncontrolled inflammation plays a pivotal role in both the initiation and progression of tumors, with the inflammatory state often reflected in alterations of serum inflammatory markers.[12] Beyond traditional markers such as white blood cells, lymphocytes, and platelets, emerging indices including the neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), monocyte-to-lymphocyte ratio (MLR), as well as the more recently developed systemic immune-inflammation index (SII)[13] and systemic inflammatory response index (SIRI)[14] have been recognized as valuable indicators of systemic inflammatory response.[15] These peripheral blood-based markers have demonstrated significant roles in systemic inflammation and cancer biology, encompassing cancer prediction, progression, and survival prognosis.[16–20] While these biomarkers have been studied as prognostic indicators in EC,[17] their potential utility in the diagnosis and prediction of EC remains underexplored.

Machine Learning (ML) algorithms have increasingly been integrated into the medical field for disease prediction, offering significant advantages over traditional statistical methods. ML algorithm is capable of processing large and complex datasets, identifying implicit relationships among various relevant features, and thereby enabling more accurate disease risk prediction.[21] The current landscape of EC risk assessment reveals a paucity of robust predictive models that base on real-world data. Our objective is to develop a non-invasive preoperative tool utilizing peripheral blood indices and ultrasound to predict EC risk, reduce the need for the invasive diagnostic interventions.

Materials and Methods

Study Participants

The study included women treated at the Third Affiliated Hospital of SYSU between January 2014 to August 2024, who were diagnosed by histopathology.

Inclusion Criteria:

a. Patients diagnosed with EH or EC confirmed by diagnostic curettage or surgical pathology.
b. Patients with complete clinical information and data.

Exclusion Criteria:

a. Patients with severe dysfunction of the heart, liver, kidney, or other major organs.
b. Patients with other malignant tumors or conditions affecting serum tumor marker and inflammatory marker levels.
c. Patients without complete blood cell count data available one week prior to surgery.
d. Patients with a history of fertility-preserving treatment for EC who were receiving hormone therapy.

Data Collection

Feature selection was guided by evidence-based approach: (1) established clinical relevance (eg, age, BMI, and menopausal status as known EC risk factors); (2) systematic review of biomarkers implicated in EC (eg, HE4, CA-125); and (3) relevant frontier guidelines research and literature (eg, NLR).
The clinical pathological data of patients were obtained through the hospital electronic medical record database:

1. Basic information: including age, height, weight, comorbidities, etc.; Body mass index (BMI) is calculated as the patient's weight (kg) divided by the square of the height (m).
2. Preoperative serum examinations including WBC, neutrophil count, lymphocyte count, monocyte count, platelet count;
3. Preoperative tumor markers including Serum carbohydrate antigen 125 (CA-125) and human epididymis protein 4(HE4).

Neutrophil-to-lymphocyte ratio (NLR) is calculated as: neutrophil count/lymphocyte count; Monocyte-to-lymphocyte ratio (MLR) is calculated as: monocyte count/lymphocyte count; Platelet-to-lymphocyte ratio (PLR) is calculated as: platelet count/lymphocyte count; Systemic immune-inflammation index (SII) is calculated as: platelet count × neutrophil count/lymphocyte count; Systemic inflammatory response index (SIRI) is calculated as: neutrophil count × monocyte count/lymphocyte count.

Pre-Processing and Model Development, Evaluation

To mitigate the issue of class imbalance between two groups, we implemented a comprehensive data preprocessing strategy combining the Synthetic Minority Oversampling Technique (SMOTE) with random under sampling. Subsequently, all features were standardized using Standard Scaler to prevent potential bias from features with larger numerical ranges.

The dataset was strategically partitioned through random stratified sampling into a training set (80%, n = 686) and a validation set (20%, n = 171). The training set was exclusively used for model development, while the validation set served as an independent cohort for performance evaluation. Six machine learning algorithms including Random Forest (RF), Extreme Gradient Boosting (XGBoost), Support Vector Mac (SVM), Gradient Boosting Machine Model (GBDT), Logistic Regression (LR) and Multilayer Perceptron (MLP) were used to construct the prediction model of EC. Among them, the RF classifier is a popular machine learning algorithm implemented in the Python package RF. The RF classification algorithm can be run without tuning the parameters and can give an approximate estimate of the importance of the features. Boosting refers to the use of a series of linear combinations of models to complete model tasks. It includes gradient boosting, there is a technique called GBDT. MLP is one of the simplest artificial neural networks, which consists of three layers—an input layer, an output layer, and a hidden layer.[22] LR is a member of the general linear model family.[23] Model performance was comprehensively evaluated using multiple metrics, with particular emphasis on the area under the receiver operating characteristic curve (AUC)as the primary indicator of discriminative ability. Brier score is a measure of the degree of deviation between the predicted and actual results, with lower values indicating better alignment between predicted probabilities and actual outcomes. Sensitivity and specificity were analyzed as complementary performance measures.

To elucidate feature contributions, SHapley Additive exPlanations (SHAP) values were employed to quantify and interpret feature importance in the best predictive performance model. The algorithm provides a measure of feature importance across the model.

Statistics

The Shapiro–Wilk normality test was performed to assess the data normality. Continuous variables are reported as mean (SD) or medians with interquartile ranges (IQRs) for skewed distributed variables and were compared using an unpaired, Mann–Whitney *U*-test. Categorical variables are reported as whole numbers and proportions (n [%]) and were compared using the $\chi 2$ test. Statistical significance was defined as a p-value <0.05. The strength of associations among modeling variables was assessed using Spearman correlation analysis.

All statistical analyses were performed using IBM SPSS Statistics 22 (SPSS Inc., Chicago, IL, USA). The predictive model construction and graphical representations were implemented using Python V3.7 (Python Software Foundation) and Prism 10.0 (GraphPad Software, San Diego, CA, USA), respectively.

Ethics

The study reporting adheres to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines and has obtained informed consent from all participants. This study was approved by the ethics committee of the Third Affiliated Hospital of Sun Yat-sen University (No. II2023-008-02). Our research strictly adheres to the principles of the Declaration of Helsinki.

Results

A retrospective cohort of 948 patients diagnosed with endometrial lesions was identified from the electronic medical records of the Third Affiliated Hospital of Sun Yat-sen University between January 1, 2014 and August 31, 2024. According to the inclusion and exclusion criteria,857 patients were included in the final analysis (Figure 1).

Characteristics of the Participants

The study cohort included 857 patients, stratified into two groups based on histopathological diagnosis: 208 patients in EH group and 649 patients in EC group (Table 1). Demographic analysis revealed significant between-group differences

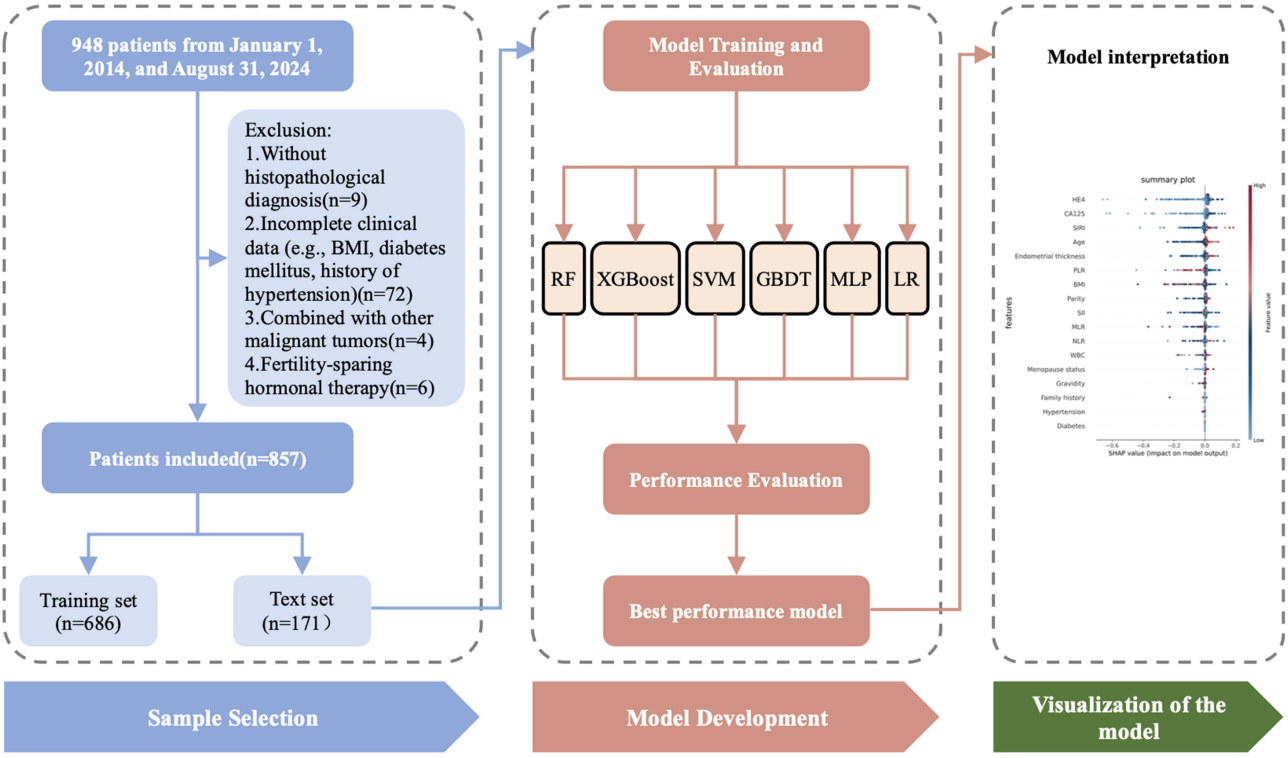

Figure 1 Flowchart of the study population.

Table 1 Baseline Characteristics and Serum Inflammatory Markers of the Participants

	EH (n=208)	EC (n=649)	p-value
Age (years)	46(41.3–50)	53(47–59)	<0.0001
BMI (kg/m2)	23.8 (21.7–25.6)	23.8 (21.5–26.5)	0.88
Gravidity	2.5 (1–4)	3 (2–4)	0.43
Parity	2 (1–2)	2 (1–3)	0.07
Family history (n, %)			0.54
Yes	5(2.4)	21(3.2)	
No	203(97.6)	628(96.8)	
Menopause status (n, %)			<0.0001
Yes	26(12.5)	345(53.2)	
No	182(87.5)	304(46.8)	
Diabetes (n, %)			< 0.0001
Yes	8(3.8)	117(18.0)	
No	200(96.2)	532(82.0)	
Hypertension (n, %)			< 0.0001
Yes	16(7.7)	173(26.7)	
No	192(92.3)	476(73.3)	
Endometrial thickness (mm)	14 (10–18)	15 (10–23)	0.004
Stage (n, %)			
1	–	524 (80.7%)	
2	–	44 (6.8%)	
3	–	68 (10.5%)	
4	–	13 (2.0%)	
Serum inflammatory markers			
WBC	6.24 (5.30–7.55)	6.63 (5.41–7.99)	0.049
NLR	2.22 (1.63–2.97)	2.21 (1.64–2.97)	0.94
MLR	0.23 (0.17–0.30)	0.22 (0.17–0.29)	0.55
PLR	160.9 (134.1–232.7)	149.3 (116.4–202.1)	0.0003
SII	687.3 (479.9–914.9)	619 (430.6–868.1)	0.046
SIRI	0.82 (0.61–1.23)	0.88 (0.58–1.41)	0.49
CA-125	15.2 (10.65–20.15)	19.4 (13.8–35.55)	<0.0001
HE4	34.5 (29.1–43.2)	48.25 (34.98–76.6)	<0.0001

in median age (EH group: 46 years [IQR 41.3–50] vs EC group: 53 years [IQR 47–59]; p<0.001). Furthermore, statistically significant differences were observed in menopausal status, hypertension, diabetes mellitus and endometrial thickness between the two groups. Among EC patients: Stage I (n=524, 80.7%), Stage II (n=44, 6.8%), Stage III (n=68, 10.5%), and Stage IV (n=13, 2.0%).

Performed Spearman's rank correlation analysis to quantify the strength of associations among these differential variables, with the results visualized in a heatmap (Figure 2). These variables may play important roles in cancer pathogenesis and progression.

Construction and Evaluation of Prediction Model

The predictive performance of these selected features was evaluated using six ML model: Random Forest (RF), Extreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), Gradient Boosting Decision Tree (GBDT), Logistic Regression (LR), and Multilayer Perceptron (MLP). As detailed in Table 2, the GBDT model demonstrated superior discriminative performance, achieving an AUC of 0.95 (95% CI: 0.93–0.97), with specificity of 0.90, and F1-score of 0.90. Subsequent validation in an independent cohort confirmed the model's performance. With the lowest integrated Brier score (0.06), the GBDT model demonstrated significant advantages in predicting EC compared to other models. The ROC curves for all six models are presented in Figure 3A and B, providing a comprehensive comparison of their predictive capabilities across different models.

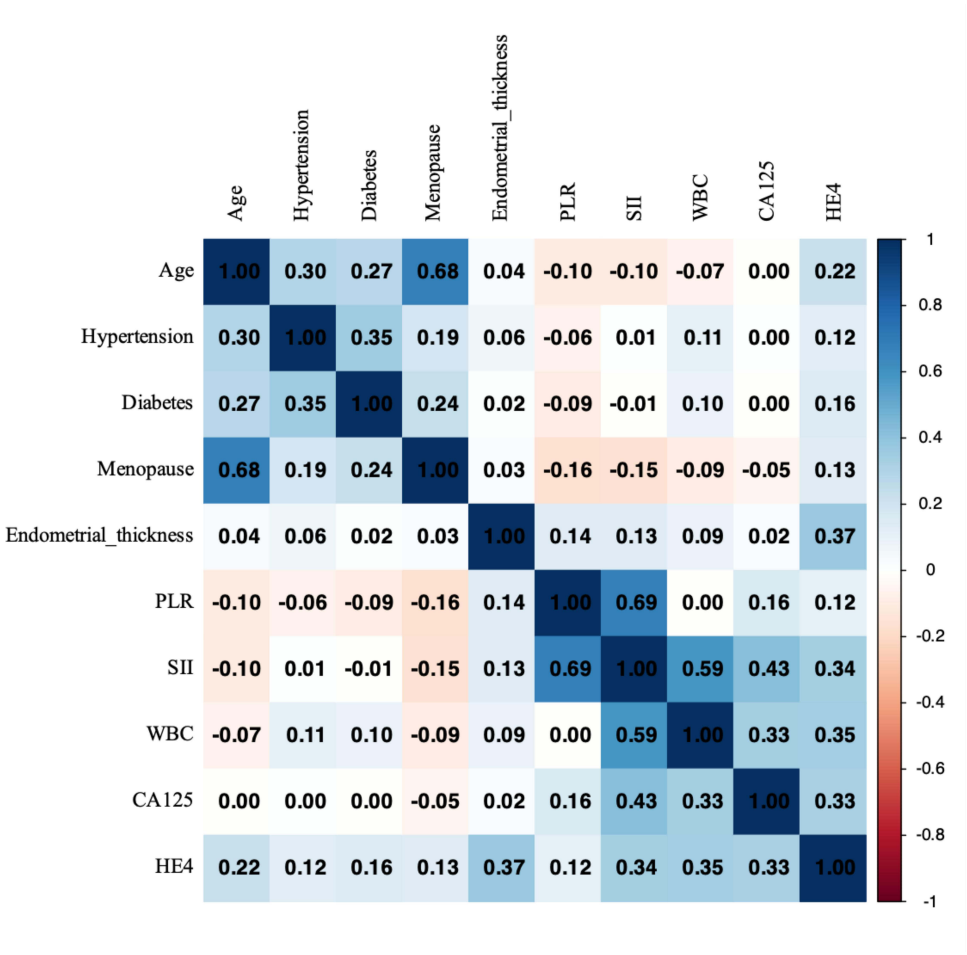

Figure 2 The overall correlation between parameters in EC patients.

Importance of Features Interpreted by SHAP Value

To elucidate feature contributions and interpret model predictions, we implemented SHAP analysis, a robust game-theoretic approach that quantifies the relative importance of each predictive feature (Figure 4A and B). The higher the SHAP value of a feature, the higher your log odds of risk. Red to blue represents the eigenvalue from large to small. The thickness of the line represents the sample distribution. In the optimal performing GBDT model, the top three predictive features for EC identification were HE4 (0.03), CA-125 (0.02) and SIRI (0.02).

Table 2 Predictive Performances of the Six ML Models for EC

Model	AUC (95% CI)	Accuracy	F1 score	Specificity	Recall	Brier
RF	00.94 (0.92–0.96)	00.93	00.93	00.93	00.94	00.09
XGBoost	00.94 (0.93–0.96)	00.90	00.90	00.90	00.90	00.07
SVM	00.84 (0.81–0.86)	00.80	00.82	00.87	00.80	00.18
GBDT	00.95 (0.93–0.97)	00.90	00.90	00.90	00.90	00.06
MLP	00.87 (0.84–0.89)	00.81	00.82	00.85	00.81	00.15
LR	00.94 (0.92–0.95)	00.81	00.83	00.86	00.81	00.11

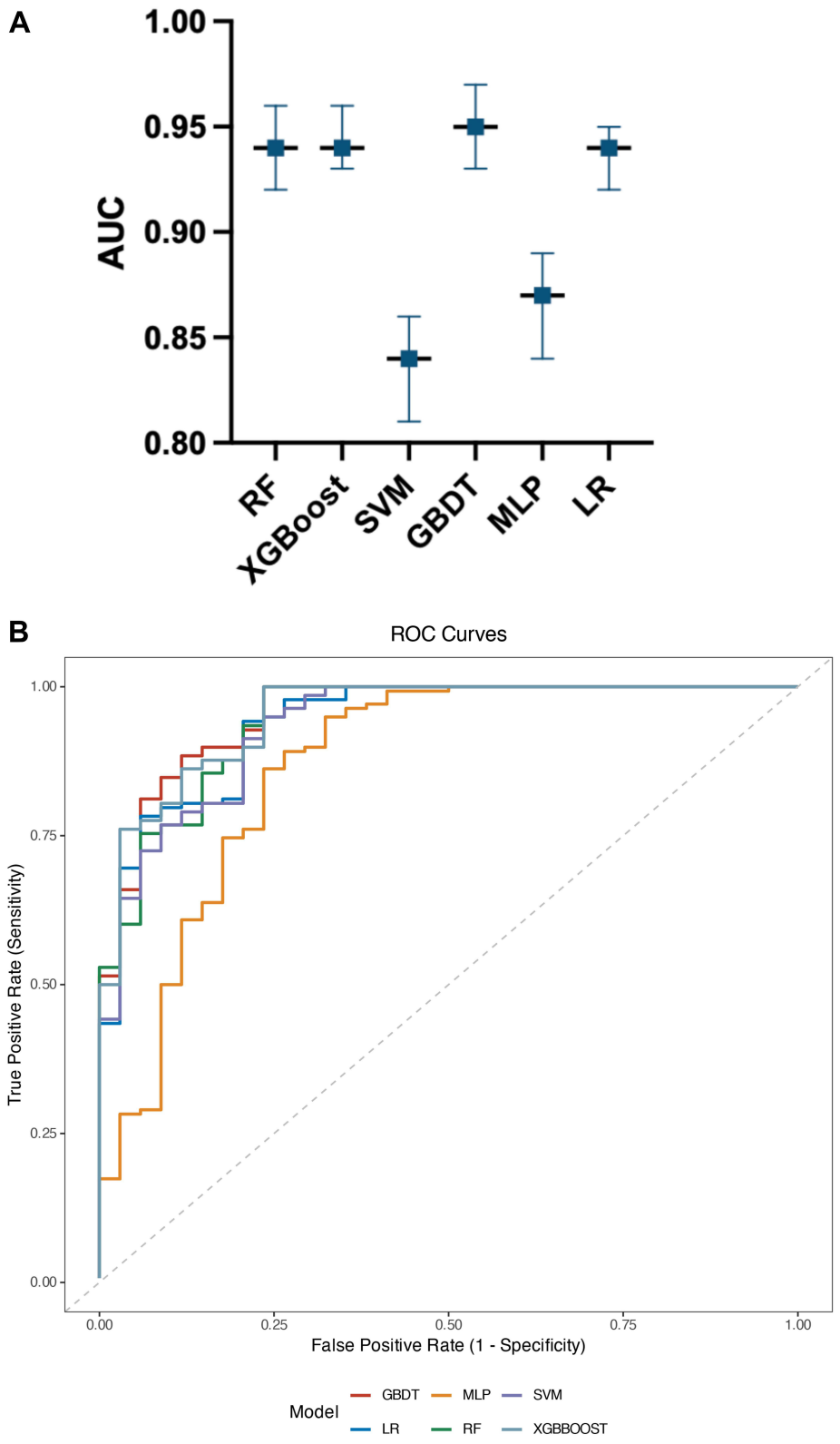

Figure 3 Receiver operating characteristic curves for the 6 machine learning models. (**A**). Comparison of area under the curve. (**B**). Receiver operating characteristic curves.

https://doi.org/10.2147/JIR.S530974

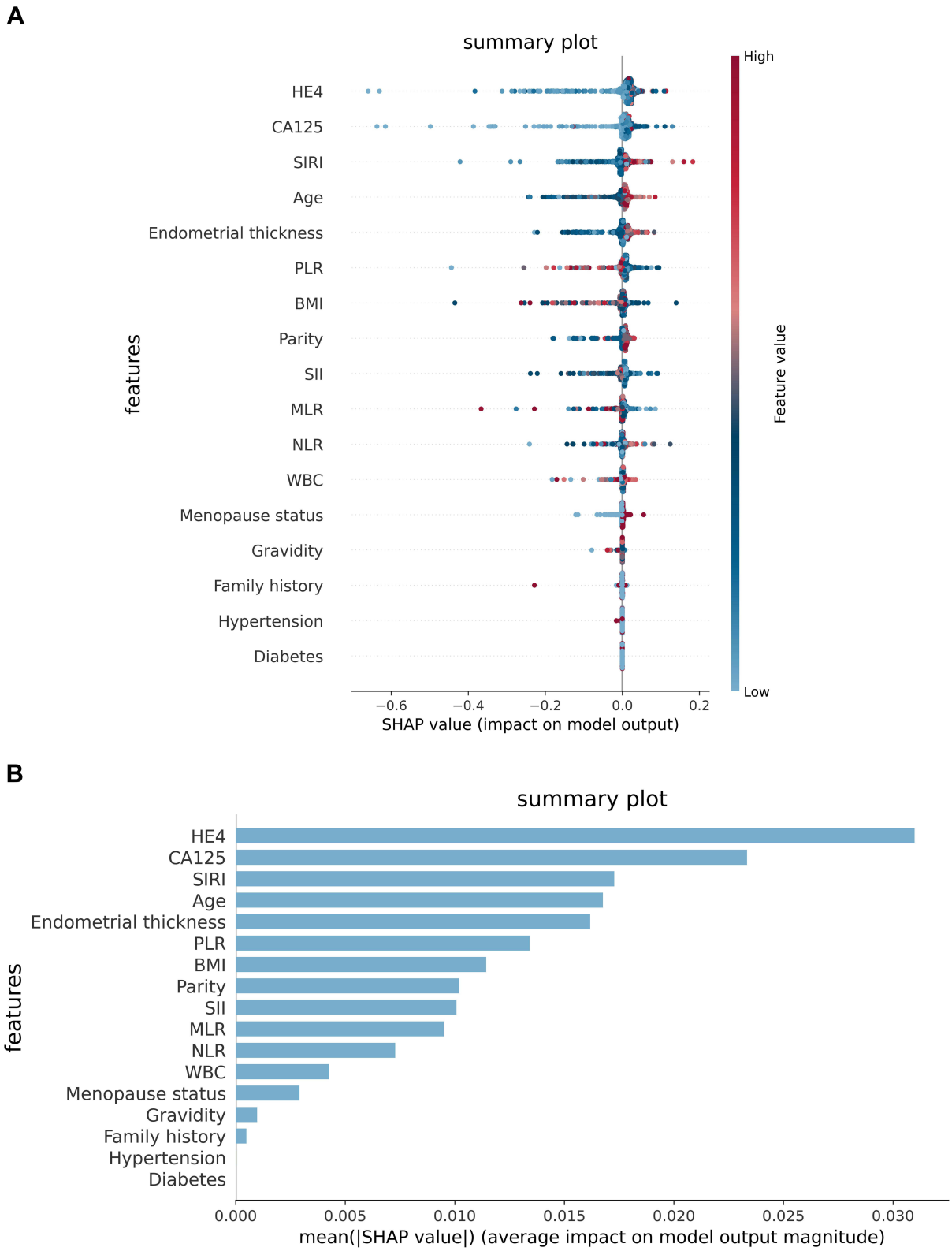

Figure 4 SHAP interpretation of the GBDT model. (**A**). Importance score ranking of the model prediction characteristics. (**B**). Every feature's impact on the model's output.

Discussion

This study represents an advancement in EC prediction by developing and validating a machine learning (ML) model that integrates baseline characteristics with non-invasive biomarkers. Among six ML models, the GBDT model demonstrated superior predictive performance, achieving an AUC of 0.95, Brier score of 0.06. SHAP interpretability analysis identified HE4, CA-125 and SIRI as key contributors to the model's predictions. These findings provide a novel technical pathway for EC risk prediction.

EC has emerged as the most prevalent gynecological malignancy globally, surpassing cervical cancer in disease burden. Late-stage diagnosis is associated with poor clinical outcomes. Among gynecologic tumors, cervical and ovarian cancers can be screened early and non-invasively. The significant reduction in cervical cancer incidence and mortality rates has been largely attributed to the implementation of population-based screening programs and the development of robust risk-prediction algorithms.,[24,25] EC lacks effective early detection tools. Early identification and management of high-risk precancerous lesions remain the most cost-effective strategy for reducing cancer-related morbidity and mortality.

The management of EH, particularly atypical hyperplasia, presents significant clinical challenges. That may progressively evolve into EC if left undetected or untreated. While current clinical guidelines recommend periodic endometrial surveillance via diagnostic curettage or hysteroscopic sampling,[26,27] these invasive procedures carry inherent risks of iatrogenic endometrial damage, including irreversible basal layer injury and intrauterine adhesions—complications particularly detrimental to young patients with fertility preservation requirements. Therefore, developing cost-effective, non-invasive methods for EC prediction is crucial for improving risk stratification and guiding conservative management strategies.

In the non-invasive screening of tumors, tumor markers have emerged as pivotal tools for the early detection of malignancies. While tumor markers like HE4 and CA125 have been evaluated for EC detection,[28] our study revealed significant limitations: 72% of EC patients showed CA125 levels below the diagnostic threshold, and 67% had subthreshold HE4 levels, despite significant. These results align with multicenter studies,[29,30] emphasizing the insufficiency of single-marker strategies. Using ultrasound alone for prediction also has the problem of low sensitivity.[31]

The intricate relationship between inflammation and cancer, initially posited by Virchow in 1863,[32] extensive research has elucidated the role of inflammatory cells and cytokines in tumorigenesis and progression. These inflammatory cells are implicated in tumor growth, progression, and metastasis.[33] Among inflammatory cells, leukocytes constitute the largest group, with neutrophils contributing to tumor progression through the release of tumor necrosis factor, interleukin-1, and interleukin-6.[34] Lymphocytes and Monocytes play a crucial role in tumor-specific immune responses by inducing cytotoxic cell death and inhibiting tumor cell proliferation and migration.[35] Platelets influence the metastatic potential of cancer cells via multiple biological pathways.[36] A single blood parameter as a marker may not adequately reflect the inflammatory state, composite markers such as NLR, PLR, MLR, SII and SIRI can sometimes provide more information. Markers derived from peripheral blood serum can provide predictive information when evaluated preoperatively, and their analysis is cost-effective and readily accessible.

Machine learning models have gained significant traction in disease prediction due to their ability to handle complex datasets and uncover intricate patterns. The field of EC detection lacks validated machine learning-based prediction models utilizing real-world clinical data, which is essential to improve the screening and diagnostic precision for EC. Previous studies, such as those by Li, Vetter and Su et al,[37–39] utilized traditional statistical logistic regression methods to construct prediction models in postmenopausal populations. Qiu et al[40] employed genetic data for predictive modeling, which is less feasible for widespread clinical application. While Erdemoglu et al[41] used the artificial intelligence in EC prediction, their model's performance was suboptimal (F1 score: 0.59), potentially due to only demographic data and ultrasonic endometrial thickness were used for modeling. Our investigation addresses these critical limitations through a comprehensive approach that: (1) including pre- and post-menopausal populations; (2) employs advanced machine learning algorithms to identify complex, nonlinear interactions among multidimensional clinical features; and (3) using clinical data and blood markers thereby enhancing model generalizability. Unlike previous studies that predominantly focused on single-type indicators, our research combines demographic characteristics (age, menopausal status, hypertension, etc), imaging metrics (endometrial thickness),

tumor markers (CA-125, HE4), and inflammatory markers (PLR, SIRI, etc) to construct a highly discriminative prediction model. Among the six ML models developed, the GBDT demonstrated the highest predictive performance, with an AUC of 0.95, outperforming the other five models.

To improve the interpretability and intuitiveness of the ML approach, we applied SHAP values to the model, facilitating a better understanding of the impact of key features. SHAP values are widely recognized in ML, particularly in medical applications, for their ability to quantify the contribution of each feature to the model's output. SHAP decision plots provide clinicians with an intuitive grasp of the results. Our analysis revealed that HE4, CA-125 and SIRI are the primary influencing factors of EC.

During the clinical application process, the data characteristics of patients are collected and input into the model for risk prediction, when patients are identified as high-risk for EC, timely invasive procedures such as hysteroscopy and curettage can be performed to confirm the diagnosis and facilitate referral to gynecologic oncologists. Conversely, for patients deemed low-risk, non-invasive screening and predictive methods can be employed for regular monitoring.

Strengths and Limitations

Our study has demonstrated a satisfactory predictive capability of the model, indicating that the GBDT model could be utilized in the future to assess the risk of EC, offering a non-invasive approach particularly suitable for the long-term follow-up of younger patients. Secondly, the findings of this study can be applied in clinical settings, assisting physicians in managing patients with endometrial lesions more effectively, especially in resource-limited environments.

However, our study has several limitations. Firstly, the research was conducted in China, with participant selection primarily based on the local population. Consequently, extrapolating these results to a global population may introduce potential biases.[42] Secondly, the retrospective single-center design may inherently introduce selection bias,[43] and healthy patients were not included in the development of the current model, which limits the generalization of the model to asymptomatic women. Fortunately, compared to previously published studies, our sample size is relatively large.[38,39] Future research should involve multicenter, large-sample, prospective studies to further optimize the model.

Conclusion

This study establishes a GBDT model integrating preoperative blood-based indices and endometrial thickness achieves high accuracy in predicting endometrial cancer risk. The SHAP- analysis identified three principal determinants: HE4, CA-125, and SIRI, aligning with their established roles in oncogenesis and inflammation. This non-invasive tool holds promise for preoperative risk stratification, particularly in reducing unnecessary invasive procedures. Future prospective studies are warranted to confirm its generalizability in asymptomatic populations and diverse clinical settings.

Data Sharing Statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Xia C, Dong X, Li H, et al. Cancer statistics in China and United States, 2022: profiles, trends, and determinants. *Chin Med J*. 2022;135:584–590.
2. Ninimiya SY, Ansu-Mensah M, Bawontuo V, Kuupiel D. Mapping evidence on ovarian, endometrial, vaginal, and vulva cancer research in Africa: a scoping review protocol. *Syst Rev*. 2021;10:108. doi:10.1186/s13643-021-01654-0
3. Raglan O, Kalliala I, Markozannes G, et al. Risk factors for endometrial cancer: an umbrella review of the literature. *Int, J, Cancer*. 2019;145:1719–1730. doi:10.1002/ijc.31961
4. Barakat A, Ismail A, Chattopadhyay S, Davies Q. Endometrial cancer incidence in patients with atypical endometrial hyperplasia according to mode of management. *Cancer Diagn Progn*. 2022;2(5):564–568. doi:10.21873/cdp.10143
5. Ring KL, Mills AM, Modesitt SC. Endometrial hyperplasia. *Obstet Gynecol*. 2022;140:1061–1075. doi:10.1097/AOG.0000000000004989
6. Bell DW, Ellenson LH. Molecular genetics of endometrial carcinoma. *Annu Rev Pathol*. 2019;14:339–367. doi:10.1146/annurev-pathol-020117-043609

7. Brooks RA, Fleming GF, Lastra RR, et al. Current recommendations and recent progress in endometrial cancer. *CA Cancer J Clin*. 2019;69:258–279. doi:10.3322/caac.21561
8. Bogani G, Di Donato V, Papadia A, et al. Evaluating long-term outcomes of three approaches to retroperitoneal staging in endometrial cancer. *Gynecol Oncol*. 2022;166:277–283. doi:10.1016/j.ygyno.2022.06.007
9. Braun MM. Diagnosis and management of endometrial cancer. *Endometrial Cancer*. 2016;93:468–474.
10. Espiau Romera A, Cuesta Guardiola T, Benito Vielba M, De Bonrostro Torralba C, Coronado Martín PJ, Baquedano Mainar L HE4 tumor marker as a predictive factor for lymphatic metastasis in endometrial cancer. *Int J Gynaecol Obstet*. 2020;149(3):265–268. doi:10.1002/ijgo.13140
11. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71:209–249. doi:10.3322/caac.21660
12. Arneth B. Tumor microenvironment. *Med Kaunas Lith*. 2019;56:15.
13. Zhang Y, Chen B, Wang L, Wang R, Yang X. Systemic immune-inflammation index is a promising noninvasive marker to predict survival of lung cancer: a meta-analysis. *Medicine*. 2019;98:e13788. doi:10.1097/MD.0000000000013788
14. Wang L, Zhou Y, Xia S, et al. Prognostic value of the systemic inflammation response index (SIRI) before and after surgery in operable breast cancer patients. *Cancer Biomark Sect Dis Markers*. 2020;28:537–547. doi:10.3233/CBM-201682
15. Zhong J-H, Huang D-H, Chen Z-Y. Prognostic role of systemic immune-inflammation index in solid tumors: a systematic review and meta-analysis. *Oncotarget*. 2017;8:75381–75388. doi:10.18632/oncotarget.18856
16. Zheng L, Zou K, Yang C, et al. Inflammation-based indexes and clinicopathologic features are strong predictive values of preoperative circulating tumor cell detection in gastric cancer patients. *Clin Transl Oncol*. 2017;19:1125–1132. doi:10.1007/s12094-017-1649-7
17. Song H, Jeong MJ, Cha J, et al. Preoperative neutrophil-to-lymphocyte, platelet-to-lymphocyte and monocyte-to-lymphocyte ratio as a prognostic factor in non-endometrioid endometrial cancer. *Int J Med Sci*. 2021;18:3712–3717. doi:10.7150/ijms.64658
18. Takahashi R, Mabuchi S, Kawano M, et al. Prognostic significance of systemic neutrophil and leukocyte alterations in surgically treated endometrial cancer patients: a monoinstitutional study. *Gynecol Oncol*. 2015;137:112–118. doi:10.1016/j.ygyno.2015.02.006
19. Cong R, Kong F, Ma J, et al. Combination of preoperative neutrophil-lymphocyte ratio, platelet-lymphocyte ratio and monocyte-lymphocyte ratio: a superior prognostic factor of endometrial cancer. *BMC Cancer*. 2020;20(1):464. doi:10.1186/s12885-020-06953-8
20. Xu M, Wu Q, Cai L, et al. Systemic inflammatory score predicts overall survival in patients with cervical cancer. *J Cancer*. 2021;12:3671–3677. doi:10.7150/jca.56170
21. Rajula HSR, Verlato G, Manchia M, Antonucci N, Fanos V. Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. *Med Kaunas Lith*. 2020;56:455.
22. Etemad L, Wu T-H, Heiner P, et al. Machine learning from clinical data sets of a contemporary decision for orthodontic tooth extraction. *Orthod Craniofac Res*. 2021;24:193–200. doi:10.1111/ocr.12502
23. Montolío A, Martín-Gallego A, Cegoñino J, et al. Machine learning in diagnosis and disability prediction of multiple sclerosis using optical coherence tomography. *Comput Biol Med*. 2021;133:104416. doi:10.1016/j.compbiomed.2021.104416
24. Bumrungthai S, Ekalaksananan T, Kleebkaow P, et al. Mathematical modelling of cervical precancerous lesion grade risk scores: linear regression analysis of cellular protein biomarkers and human papillomavirus E6/E7 RNA staining patterns. *Diagnostics*. 2023;13:1084. doi:10.3390/diagnostics13061084
25. Tisler A, Võrk A, Tammemägi M, et al. Nationwide study on development and validation of a risk prediction model for CIN3+ and cervical cancer in Estonia. *Sci Rep*. 2024;14:24589. doi:10.1038/s41598-024-75697-3
26. Concin N, Matias-Guiu X, Vergote I, et al. ESGO/ESTRO/ESP guidelines for the management of patients with endometrial carcinoma. *Int J Gynecol Cancer*. 2021;31:12–39. doi:10.1136/ijgc-2020-002230
27. Yang B, Xu Y, Zhu Q, et al. Treatment efficiency of comprehensive hysteroscopic evaluation and lesion resection combined with progestin therapy in young women with endometrial atypical hyperplasia and endometrial cancer. *Gynecol Oncol*. 2019;153:55–62. doi:10.1016/j.ygyno.2019.01.014
28. Liu J, Han L, Sun Q, Li Y, Niyazi M. Meta-analysis of the diagnostic accuracy of HE4 for endometrial carcinoma. *Eur J Obstet Gynecol Reprod Biol*. 2020;252:404–411. doi:10.1016/j.ejogrb.2020.07.015
29. Ginath S, Menczer J, Fintsi Y, et al. Tissue and serum CA125 expression in endometrial cancer. *Int J Gynecol Cancer off J Int Gynecol Cancer Soc*. 2002;12:372–375. doi:10.1136/ijgc-00009577-200207000-00007
30. Dotters DJ. Preoperative CA 125 in endometrial cancer: is it useful? *Am J Obstet Gynecol*. 2000;182:1328–1334. doi:10.1067/mob.2000.106251
31. Opolskiene G, Sladkevicius P, Valentin L. Prediction of endometrial malignancy in women with postmenopausal bleeding and sonographic endometrial thickness ≥ 4.5 mm. *Ultrasound Obstet Gynecol*. 2011;37(2):232–240. doi:10.1002/uog.8871
32. Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow?. *Lancet*. 2001;357(9255):539–45 doi:10.1016/S0140-6736(00)04046-0
33. Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. *Cell*. 2010;140:883–899. doi:10.1016/j.cell.2010.01.025
34. Cupp MA, Cariolou M, Tzoulaki I, Aune D, Evangelou E, Berlanga-Taylor AJ. Neutrophil to lymphocyte ratio and cancer prognosis: an umbrella review of systematic reviews and meta-analyses of observational studies. *BMC Med*. 2020;18(1):360. doi:10.1186/s12916-020-01817-1
35. Patel A, Ravaud A, Motzer RJ, et al. Neutrophil-to-lymphocyte ratio as a prognostic factor of disease-free survival in postnephrectomy high-risk locoregional renal cell carcinoma: analysis of the S-TRAC trial. *Clin Cancer Res*. 2020;26:4863–4868. doi:10.1158/1078-0432.CCR-20-0704
36. Song M, Graubard BI, Rabkin CS, Engels EA. Neutrophil-to-lymphocyte ratio and mortality in the United States general population. *Sci Rep*. 2021;11:464. doi:10.1038/s41598-020-79431-7
37. Vetter MH, Smith B, Benedict J, et al. Preoperative predictors of endometrial cancer at time of hysterectomy for endometrial intraepithelial neoplasia or complex atypical hyperplasia. *Am J Obstet Gynecol*. 2020;222(1):e60.1–e60.7. doi:10.1016/j.ajog.2019.08.002
38. Su J, She L, Fan Y, et al. Model predicting the risk of endometrial hyperplasia developing into endometrial cancer. *J Inflamm Res*. 2024:6159–6171—.
39. Li X, Wang H, Wang T, et al. Combining demographic data and transvaginal ultrasonography: a predictive model for endometrial carcinoma in postmenopausal patients. *BMC Women's Health*. 2024;24(1):539. doi:10.1186/s12905-024-03374-8
40. Sang Q, Yang L, Zhao H, et al. Risk prediction model of uterine corpus endometrial carcinoma based on immune-related genes. *BMC Women's Health*. 2024;24(1):429. doi:10.1186/s12905-024-03237-2
41. Erdemoglu E. Artificial intelligence for prediction of endometrial intraepithelial neoplasia and endometrial cancer risks in pre- and postmenopausal women. *AJOG Glob Rep*. 2023;3:100154.doi: 10.1016/j.xagr.2022.100154

42. Shi Y, Zhang G, Ma C, et al. Machine learning algorithms to predict intraoperative hemorrhage in surgical patients: a modeling study of real-world data in shanghai, China. *BMC Med Inf Decis Mak*. 2023;23:156. doi:10.1186/s12911-023-02253-w
43. de Ron J, Fried EI, Epskamp S. Psychological networks in clinical populations: investigating the consequences of berkson's bias. *Psychol Med*. 2021;51:168–176. doi:10.1017/S0033291719003209