scientific reports

Check for updates

OPEN

A risk prediction model for endometrial hyperplasia/ endometrial carcinoma in premenopausal women

Zhen Li[1], Juan Yin[1], Yu Liu[2] & Fanqing Zeng[1]✉

This study investigated the risk factors for endometrial hyperplasia (EH) and endometrial carcinoma (EC) in premenopausal women. The goal was to establish a nomogram model to predict the risk of EH/EC and quantitative standards in clinical practice, which improved the clinical prognosis of EH/EC patients. Data were collected from premenopausal women with suspected EH/EC who underwent hysteroscopic endometrial biopsy. Patients ($n = 1541$) were divided into training and validation groups at a 3:1 ratio. Univariable and multivariable logistic regression analyses were conducted to identify risk factors for EH/EC and establish a predictive model. The model's discrimination was evaluated using the area under the receiver operating characteristic curve (AUC), its calibration was assessed using calibration plots, and its clinical effectiveness was evaluated using decision curve analysis (DCA). The optimal score and probability cutoff values were determined to differentiate between low and high-risk populations, guiding clinical medical practice. BMI, age at menarche, intrauterine device (IUD), diabetes, polycystic ovary syndrome (PCOS), endometrial thickness (ET), and uterine cavity fluid were identified as independent risk factors for EH/EC and were incorporated into the predictive nomogram model. The model demonstrated good discrimination with AUCs of 0.845 and 0.905 in the training and validation sets, respectively. The calibration plots and DCA showed excellent model calibration and clinical effectiveness. EH/EC is significantly associated with BMI, age at menarche, IUD use, diabetes, PCOS, ET, and uterine cavity fluid. The nomogram model can be used to predict the risk of EH/EC in premenopausal women and facilitate rapid screening.

Keywords Nomogram, Endometrial hyperplasia, Premenopausal women, Risk factors

Due to factors such as declining fertility rates and increasing obesity rates, the incidence of EC has been progressively rising, ranking as the second most common gynecologic malignancy worldwide[1]. EH represents a critical pathogenic process leading to EC[2], particularly endometrial atypical hyperplasia (AH)/endometrioid intraepithelial neoplasia (EIN), which carries a cancer transformation rate as high as 25–33%[3,4]. Analysis and classification of EH is not without challenge, as endometrium is a dynamic, multicellular tissue structure, especially for premenopausal and peri-menopausal women[5]. Despite the provision of relevant risk factor recommendations in clinical guidelines or studies[5–7], there is a lack of specific quantitative standards in clinical practice, with physicians often relying on their experience-based judgment, which results in taking a more proactive approach in the diagnosis and treatment of patients presenting with relevant symptoms, including abnormal uterine bleeding (AUB), vaginal fluid, uterine cavity fluid, and hypogastric pain. However, this aggressive management has led to unnecessary invasive diagnostic procedures, such as curettage or hysteroscopy, which inflict physical and psychological harm upon women, resulting in intrauterine adhesions, infertility, diminished quality of sexual life, and even uterine arteriovenous malformations[8–10]. Given the higher demands for fertility and quality of sexual life among premenopausal women, utmost caution should be exercised when considering invasive procedures involving the uterus.

Hence, the establishment of a predictive model based on pertinent risk factors would better guide clinical practice. Previous research has employed clinical variables and ultrasound indicators to build predictive models[11–17]. Nonetheless, these models have predominantly targeted postmenopausal women or revolved

[1]Department of Gynecology, Chongqing Ninth People's Hospital, 69, Jialing Village, Beibei District, Chongqing 400700, China. [2]Hainan Hospital of PLA General Hospital, Sanya, Hainan Province, China. ✉email: zfq123456zfq@sina.com

around AH/EIN and EC, overlooking premenopausal women, and endometrial hyperplasia without atypia. Moreover, the majority of these models incorporate a restricted set of variables. Thus, a comprehensive system to assess the probability of pathological changes in the endometrium is warranted, which would facilitate precise quantification to identify individuals who require invasive procedures. This approach would effectively minimize the harms resulting from aggressive diagnosis and treatment in premenopausal women, empowering outpatient physicians to manage premenopausal patients presenting with relevant symptoms promptly and accurately. The nomogram currently represents a widely applied and effective predictive model[18]. Accordingly, we have collected comprehensive clinical data to construct a nomogram-based risk prediction model, addressing the question of "Who needs to undergo hysteroscopy with diagnostic curettage?"

Results

Patient characteristics

Based on the inclusion and exclusion criteria, a total of 1541 patients who were treated at the Department of Gynecology in Chongqing Ninth People's Hospital from January 2016 to June 2023 were enrolled in this study. In addition, these eligible patients were divided into the training cohort ($n=1156$) and the validation cohort ($n=385$) at a ratio of 3:1. In the training cohort, the age of the patients was 42.65 ± 7.34 years and the mean BMI was 23.23 (15.24–39.06) kg/m^2. Among these patients, the symptoms at consultation were AUB (86.25%), vaginal fluid (7.61%), uterine cavity fluid (5.71%) and hypogastric pain (13.32%). A total of 624(53.98%) patients' histopathologies showed EH/EC, and other patients had proliferative/secretory endometrium. In the validation cohort, the mean age of the patients was 42.91 ± 7.10 years and the mean BMI was 22.98 (18.03–38.27) kg/m^2. Among these patients, the symptoms at consultation were AUB (84.16%), vaginal fluid (5.19%), uterine cavity fluid (5.97%) and hypogastric pain (10.65%). A total of 209(54.29%) patients' histopathologies showed EH/EC, and other patients had proliferative/secretory endometrium. The results for the medical history, comorbidities and other data of the patients in the training and validation cohorts are shown in Table 1. There were no significant differences regarding any variables between the two cohorts (all $p>0.05$).

Factors associated with EH/EC

To identify potential risk factors for EH/EC, univariable logistic regression analysis was performed for EH/EC patients (Table 2). Univariable logistic regression indicated that nine candidate factors namely, BMI, age at menarche, IUD use, diabetes, hypertension, hyperuricemia, PCOS, ET, and uterine cavity fluid were positively associated with EH/EC development ($p<0.05$). Subsequently, the multivariable logistic regression analysis was conducted to identify the independent risk factors and evaluate their effect on the development of EH/EC for patients in the training cohort. Significant independent predictors of EH/EC included BMI (OR = 1.401, 95% CI: 1.328–1.478, $p<0.001$), age at menarche (OR = 0.745, 95% CI: 0.677–0.819, $p<0.001$), IUD use (OR = 3.012, 95% CI: 2.102–4.317, $p<0.001$), diabetes (OR = 2.542, 95% CI: 1.137–5.685, $p=0.023$), PCOS (OR = 3.784, 95% CI: 1.940–7.379, $p<0.001$), ET (OR = 5.769, 95% CI: 3.894–8.546, $p<0.001$), and uterine cavity fluid (OR = 3.784, 95% CI: 1.940–7.379, $p<0.001$). The detailed results of the multivariate logistic regression analysis are presented in Table 2.

Nomogram development and validation

A nomogram was established to predict a patient's probability of developing EH/EC based on the seven independent risk factors: BMI, age at menarche, IUD use, diabetes, PCOS, ET, and uterine cavity fluid. Among these risk factors, ET and uterine cavity fluid were evaluated using ultrasound. From the nomogram, BMI and ET had the greatest influence on EH/EC risk, followed by age at menarche, PCOS, uterine cavity fluid, IUD use diabetes and diabetes. A line is drawn straight upward to the points axis to assign a weighted score to each of the independent risk factors. The number of points received for each variable value on the point scale are added together. The total points reflect the sum of the score of each factor, which is then converted to a probability of EH/EC for a given patient by drawing a line straight down from the total points axis to the EH/EC risk axis. The highest total score is 180 points, and the scale of the EH/EC probability ranges from 0.1 to 0.99. Therefore, a larger total point score indicates a greater possibility of developing EH/EC. The nomogram is shown in Fig. 1. To ensure that model application is simple in clinical practice, we transformed the nomogram into a web-based calculator. This calculator can be installed on a doctor's computer for real-time calculation. (https://li123.shinyapps.io/DynNomapp/)

The performance of the nomogram was evaluated by discrimination, calibration, and clinical usefulness. First, we examined the discrimination of the nomogram; the nomogram demonstrated high discrimination and good prediction accuracy as indicated by the AUC value of 0.845 (95% CI 0.823–0.868) with a significant p value ($p<0.001$) in the training cohort as shown in Fig. 2A. In the validation cohort, the AUC value was 0.905 (95% CI 0.874–0.935) for the nomogram with a significant p value ($p<0.001$) as shown in Fig. 2B. Subsequently, the calibration plot showed a favorable consistency between the predicted and actual probabilities in both the training and the validation groups, as shown in Fig. 3A and B, which indicated the adequate fit of the nomogram for predicting EH/EC. To further evaluate the clinical benefit of the nomogram, we conducted DCA, which showed the great benefit obtained from the application of our nomogram, as shown in Fig. 4A and B.

Optimal threshold of the nomogram

Individualized scores for each patient were accurately calculated according the nomogram. Therefore, we used an optimal cutoff value that maximized "(sensitivity + specificity)-1" in the receiver operating characteristic (ROC) curve. According to the optimal cutoff value of the nomogram, two groups were identified: a low-risk EH/EC group (score<76.411 points, EH/EC probability < 0.548) and a high-risk EH/EC group (score ≥ 76.411 points, EH/EC probability ≥ 0.548). An EH/EC probability of 0.548 corresponded to the optimal threshold of

Characteristic	Training cohort ($n=1156$)	Validation cohort ($n=385$)	*P* value
Age, year	42.65 ± 7.34	42.91 ± 7.10	0.542
BMI, kg/m^2	23.23 (15.24–39.06)	22.98 (18.03–38.27)	0.851
Symptom			
AUB	997 (86.25%)	324 (84.16%)	0.310
Vaginal fluid	88 (7.61%)	20 (5.19%)	0.108
Uterine cavity fluid	66 (5.71%)	23 (5.97%)	0.847
Hypogastric pain	154 (13.32%)	41 (10.65%)	0.172
Histopathology			0.917
Proliferative/secretory endometrium	532 (46.02%)	176 (45.71%)	
Simple/complex/atypia EH and EC	624 (53.98%)	209 (54.29%)	
Perimenopause			0.738
Yes	125 (((10.81%)	44 (11.43%)	
No	1031 (89.19%)	341 (88.57%)	
Menarche	13 (10–19)	13 ((10–19)	0.383
Parity	1 (0–6)	1 (0–5)	0.757
IUD			0.390
Yes	249 (21.54%)	75 (19.48%)	
No	907 (78.46%)	310 (80.52%)	
Iatrogenic factors			0.437
Yes	29 (2.51%)	7 (1.82%)	
No	1127 (97.49%)	378 (98.18%)	
Smoking			0.299
Yes	17 (1.47%)	3 (0.78%)	
No	1139 (98.53%)	382 (99.22%)	
Comorbidity			
Diabetes	59 (5.10%)	26 (6.75%)	0.219
Hypertension	77 (6.66%)	35 (9.09%)	0.112
HLP	76 (6.57%)	19 (4.94%)	0.247
Hyperuricemia	16 (1.38%)	5 (1.30%)	0.9
Hypercholesteremia	25 (2.16%)	7 (1.82%)	0.681
Hyperthyroidism	12 (1.04%)	5 (1.30%)	0.672
Hypothyroidism	15 (1.30%)	4 (1.04%)	0.69
PCOS	34 (2.94%)	14 (3.64%)	0.496
Hyperprolactinemia	14 (1.21%)	3 (0.78%)	0.482
ET	0.87 ± 0.43	0.85 ± 0.44	0.76

Table 1. Clinical characteristics of patients. *BMI* body mass index, *AUB* abnormal uterine bleeding, *EH* endometrial hyperplasia, *EC* endometrial cancer, *IUD* intrauterine device, *HLP* hyperlipidaemia, *PCOS* polycystic ovary syndrome, *ET* endometrial thickness.

the nomogram in terms of clinical utility with an overall diagnostic accuracy of 83.40%. At this threshold, in the training cohort, the sensitivity and specificity of the nomogram were 0.758(95% CI: 0.724, 0.792) and 0.801(95% CI: 0.767, 0.835), respectively, while in the validation cohort, the sensitivity and specificity of the nomogram were 0.837(95% CI: 0.787, 0.887) and 0.847(95% CI: 0.793, 0.900), respectively.

Discussion

AUB, vaginal fluid, uterine cavity fluid, and hypogastric pain represent the key symptoms of EH[19]. Nevertheless, EH is observed in only approximately 10% of symptomatic premenopausal women[20]. Blindly subjecting all symptomatic patients to invasive procedures is unwarranted and may lead to unnecessary harm, including infertility, increased miscarriage risk[8–10], Therefore, we have developed an effective, quantitative, and intuitive disease prediction model to help gynecologists make accurate clinical decisions in symptomatic premenopausal women.

Initially, we conducted an analysis of clinical practice guidelines and diverse risk models related to diagnosis and treatment. Our study evaluated 20 clinical indicators to analyze the risk factors for EH/EC. The results identified seven independent risk factors: BMI, age at menarche, IUD use, diabetes, PCOS, ET, and uterine cavity fluid.

Obesity, is an independent risk factor for EH/EC, and EH/EC risk increases with increasing BMI levels. Michelle et al.[21] investigated the risk factors for AH and EC in premenopausal women, revealing obesity as a significant independent risk factor, particularly women with a BMI > 30 kg/m^2, have four times higher likelihood

Variables	Endometrial hyperplasia			
	Univariate		Multivariate	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Age	1.014 (0.998–1.030)	0.085		
BMI	1.436 (1.366–1.509)	***<0.001***	1.401 (1.328–1.478)	***<0.001***
AUB				
No	Ref			
Yes	0.976 (0.698–1.365)	0.887		
Menarche	0.721 (0.666–0.780)	***<0.001***	0.745 (0.677–0.819)	***<0.001***
Perimenopause				
No	Ref			
Yes	1.179 (0.810–1.716)	0.39		
Parity	1.084 (0.911–1.291)	0.363		
IUD				
No	Ref			
Yes	2.147 (1.596–2.889)	***<0.001***	3.012 (2.102–4.317)	***<0.001***
Iatrogenic factors				
No	Ref			
Yes	0.512 (0.240–1.095)	0.084		
Smoking				
No	Ref			
Yes	2.067 (0.723–5.904)	0.175		
Diabetes				
No	Ref			
Yes	3.530 (1.852–6.727)	***<0.001***	2.542 (1.137–5.685)	***0.023***
Hypertension				
No	Ref			
Yes	1.965 (1.195–3.229)	***0.008***	1.252 (0.657–2.386)	0.494
HLP				
No	Ref			
Yes	1.594 (0.982–2.588)	0.059		
Hyperuricemia				
No	Ref			
Yes	3.752 (1.063–13.237)	***0.04***	1.716 (0.325–9.053)	0.524
Hypercholesteremia				
No	Ref			
Yes	2.228 (0.923–5.375)	0.075		
Hyperthyroidism				
No	Ref			
Yes	1.714 (0.513–5.725)	0.381		
Hypothyroidism				
No	Ref			
Yes	0.742 (0.268–2.063)	0.569		
PCOS				
No	Ref			
Yes	2.582 (1.468–4.539)	***0.001***	3.784 (1.940–7.379)	***<0.001***
Hyperprolactinemia				
No	Ref			
Yes	0.851 (0.297–2.442)	0.764		
ET	5.434 (3.897–7.577)	***<0.001***	5.769 (3.894–8.546)	***<0.001***
Uterine cavity fluid				
No	Ref			
Yes	2.582 (1.468–4.539)	***0.001***	3.784 (1.940–7.379)	***<0.001***

Table 2. Analysis of risk factors for endometrial hyperplasia. *OR* odd ritio, *CI* confidence interval, *BMI* body mass index, *AUB* abnormal uterine bleeding, *IUD* intrauterine device, *HLP* hyperlipidaemia, *PCOS* polycystic ovary syndrome, *ET* endometrial thickness. Significant values are in bold with italics.

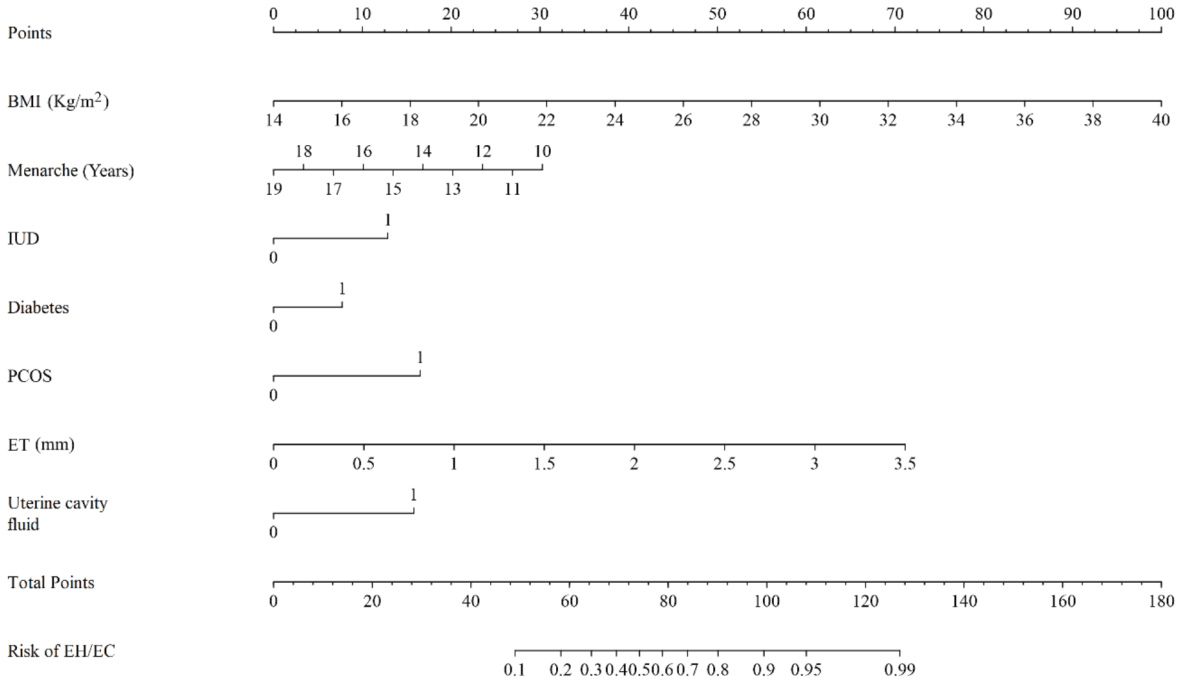

Fig. 1. Nomogram prediction model for the risk of developing endometrial hyperplasia and endometrial cancer. Endometrial hyperplasia nomogram prediction models were developed in this retrospective analysis, including BMI, menarche, IUD use, diabetes, PCOS, ET and hydrocele. BMI: body mass index, PCOS: polycystic ovary syndrome, IUD: intrauterine device, EH: endometrial hyperplasia, EC: endometrial cancer; 0 and 1 in IUD, Diabetes, PCOS and Uterine cavity fluid is no and yes.

of developing EH/EC compared to normal-weight women. A study[22] proposed that the increased estrogen aromatization in the fat cells of individuals with a higher BMI might contribute to this phenomenon. Additionally, there might be a decrease in serum progesterone and sex hormone-binding globulin levels, possibly involving the cytokines associated with obesity. Both early menarche and high BMI are considered risk factors, likely due to estrogen's pivotal role in EH/EC development. Approximately 7% of women of childbearing age suffer from PCOS, the most common endocrine disorder in this population[23]. Multiple studies[24,25] have indicated that PCOS is a risk factor for EC, with PCOS patients having a three to four times higher risk than women without PCOS. EH is considered a precursor to EC, and PCOS is also closely associated with EH. In PCOS, chronic anovulation leads to prolonged estrogen exposure without the counterbalance of progesterone, stimulating the endometrium and increasing the risk of EH and possibly EC[26]. Moreover, PCOS is closely linked to an imbalance in local endometrial oxidative stress, resulting in poor endometrial receptivity and an elevated risk of EH/EC[27]. Long-term studies have established a connection between diabetes and various cancers, including EC[28]. Diabetes increases the risk of EH by up to five times[29], and may be associated with insulin resistance and hyperinsulinemia stimulating cell proliferation, although the exact mechanisms remain unclear[28]. The IUDs utilized in this study were metallic contraceptive rings, and the principle is based on foreign objects acting to prevent fertilized eggs from implanting, concurrently, they induce persistent inflammatory stimulation to the endometrium. Research findings substantiate that IUD use represents a risk factor for chronic endometritis and may additionally contribute to EH[30]. In light of its accessibility and cost-effectiveness, transvaginal ultrasound (TVUS) is recommended as the primary examination for patients with AUB, following clinical guidelines[5–7]. For postmenopausal women with AUB, an ET greater than 4 mm is a well-established high-risk factor for EH and EC[31]. However, a consensus on the ET threshold for premenopausal women has not been reached. Kumari et al.[32] proposed that women in the perimenopausal phase with an ET $\geq$ 10.5 mm should undergo endometrial biopsy due to an increased risk of endometrial lesions. Giannella et al.[33] identified an ET > 11 mm as a risk factor for endometrial thickening in premenopausal women. Cong et al.[34], through the pathology-based analysis of patients with sonographic endometrial echo abnormalities, indicated that an ET $\geq$ 7 mm is a risk factor for endometrial lesions. Nevertheless, the measurement of ET is influenced by multiple factors and offers limited added value in these cases. For these women, other features such as grayscale ultrasound morphology and Doppler patterns may be more discriminative[35]. However, ET measurement is still the best choice[16,36]. Uterine cavity fluid had not been hitherto considered a contributing risk factor. Our center's observations have disclosed a correlative relationship between uterine cavity fluid and EH, consequently warranting its inclusion as a variable in this investigation. Subsequent analysis has unequivocally established uterine cavity fluid as a risk factor, corroborating our initial clinical suppositions. A study showed a significant increase in the proportion of

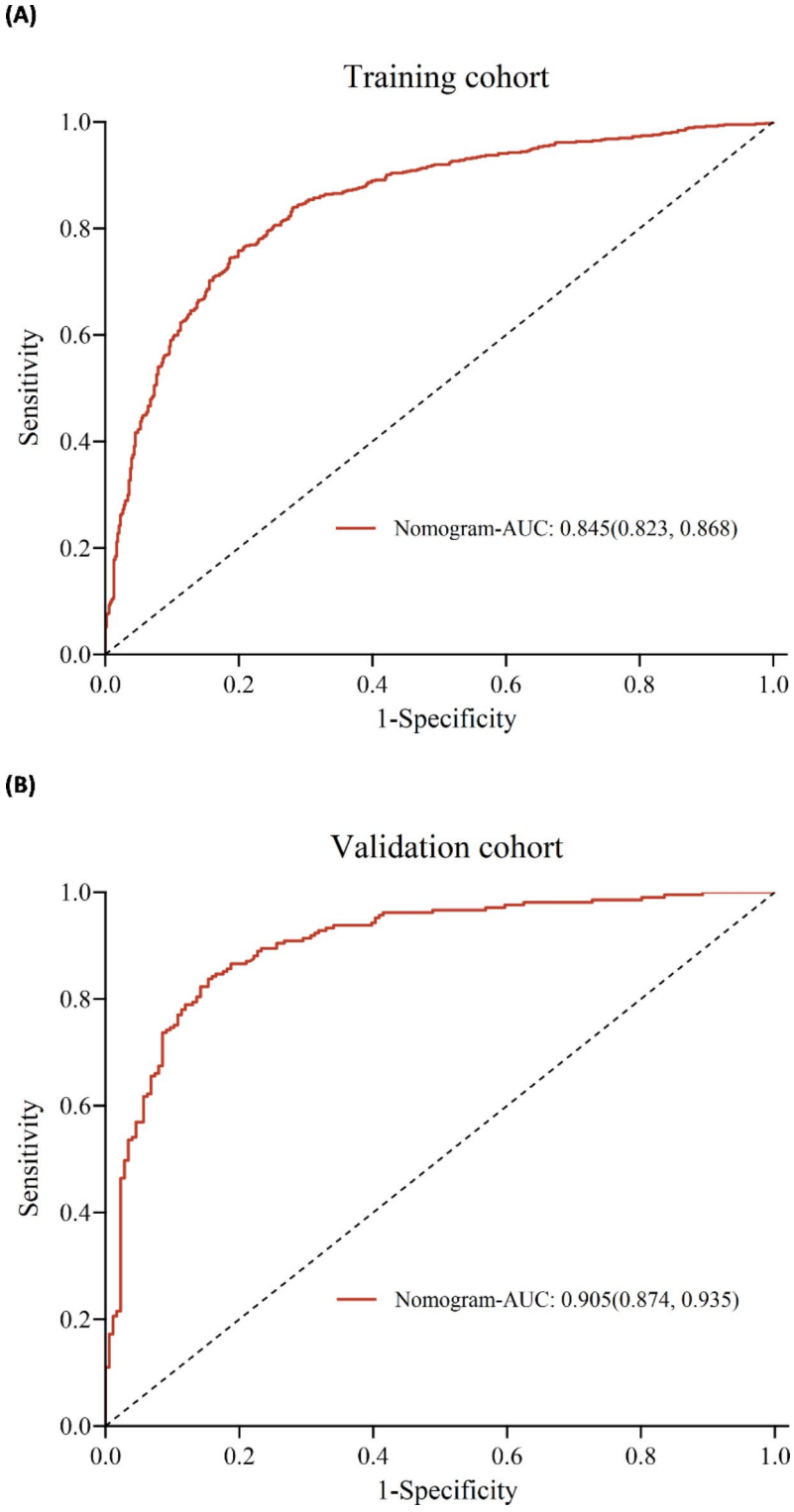

Fig. 2. (**A**) ROC curves of the nomogram for predicting EH/EC in the training cohort. (**B**) ROC curves of the nomogram for predicting EH/EC in the validation cohort.

macrophages in tissue samples from women with EH[37], which may be related to the increase and accumulation of fluid in the uterine cavity. Interestingly, a study found that uterine fluid cleared significantly more slowly after mating in dogs with EH than in healthy dogs[38]. The specific mechanism is not yet clear.

In this pertinent investigation, age, smoking, alcohol consumption, hypertension, and certain metabolic disorders were examined as potential risk factors. Our study, which focused on premenopausal patients, revealed a relatively low proportion of elderly women. Notably, menopausal status emerged as a more suitable

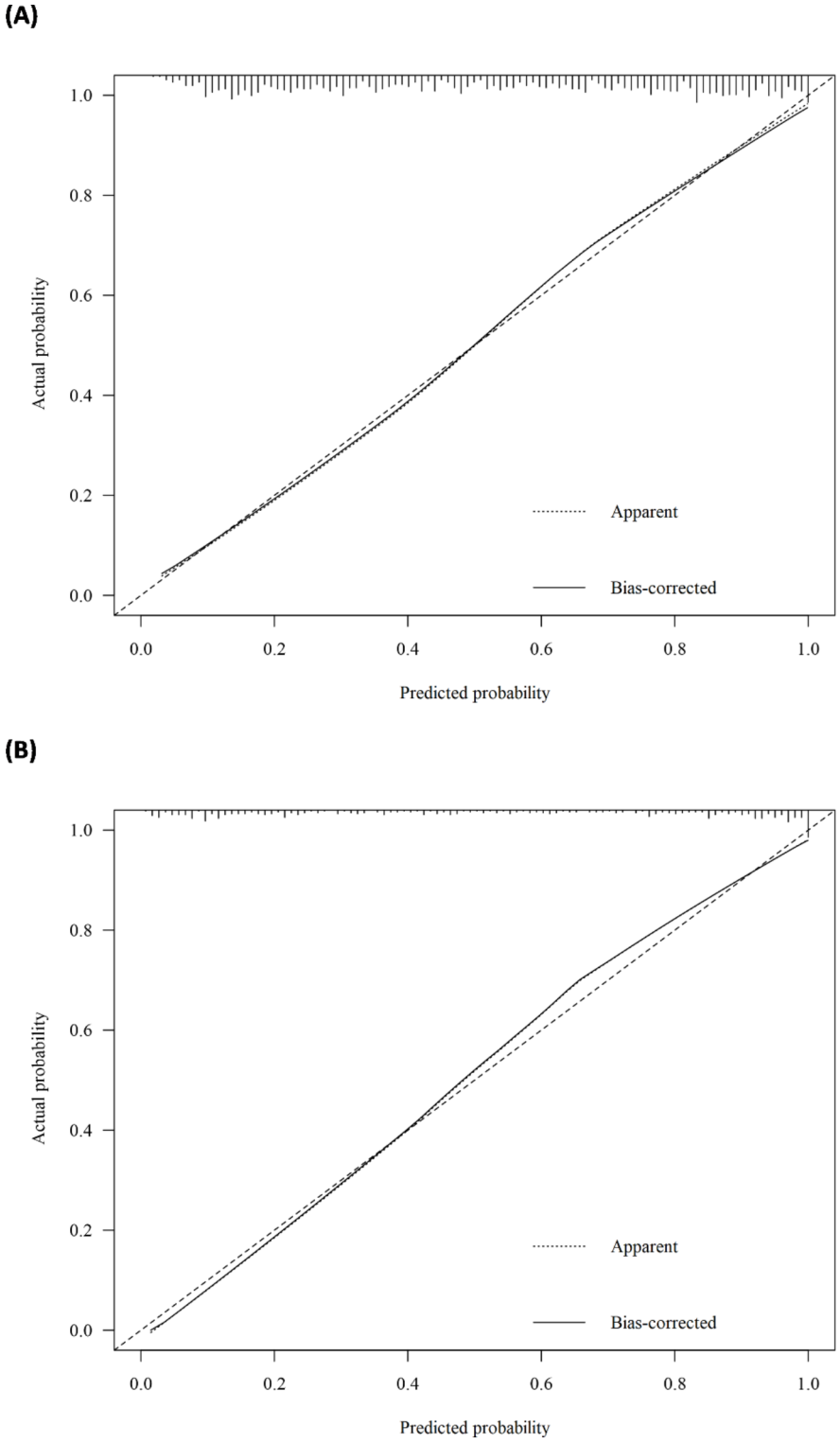

Fig. 3. (**A**) The calibration curves of the nomogram prediction model in the training cohort. (**B**) The calibration curves of the nomogram prediction model in the validation cohort.

risk indicator than age[39]. Consequently, age alone does not represent an absolute risk factor. Some studies have elucidated a mechanistic association between the antiestrogenic effects of smoking[40] and reduced risk of EC among smokers[41–43], Alcohol has been shown to have a J-shaped association with EC risk[44]. However, cultural disparities and other factors contribute to a comparatively low prevalence of smoking and alcohol consumption among Chinese women. The study's statistical analysis was limited by the small sample size, resulting in no statistically significant differences. While hypertension and certain metabolic disorders have been identified as

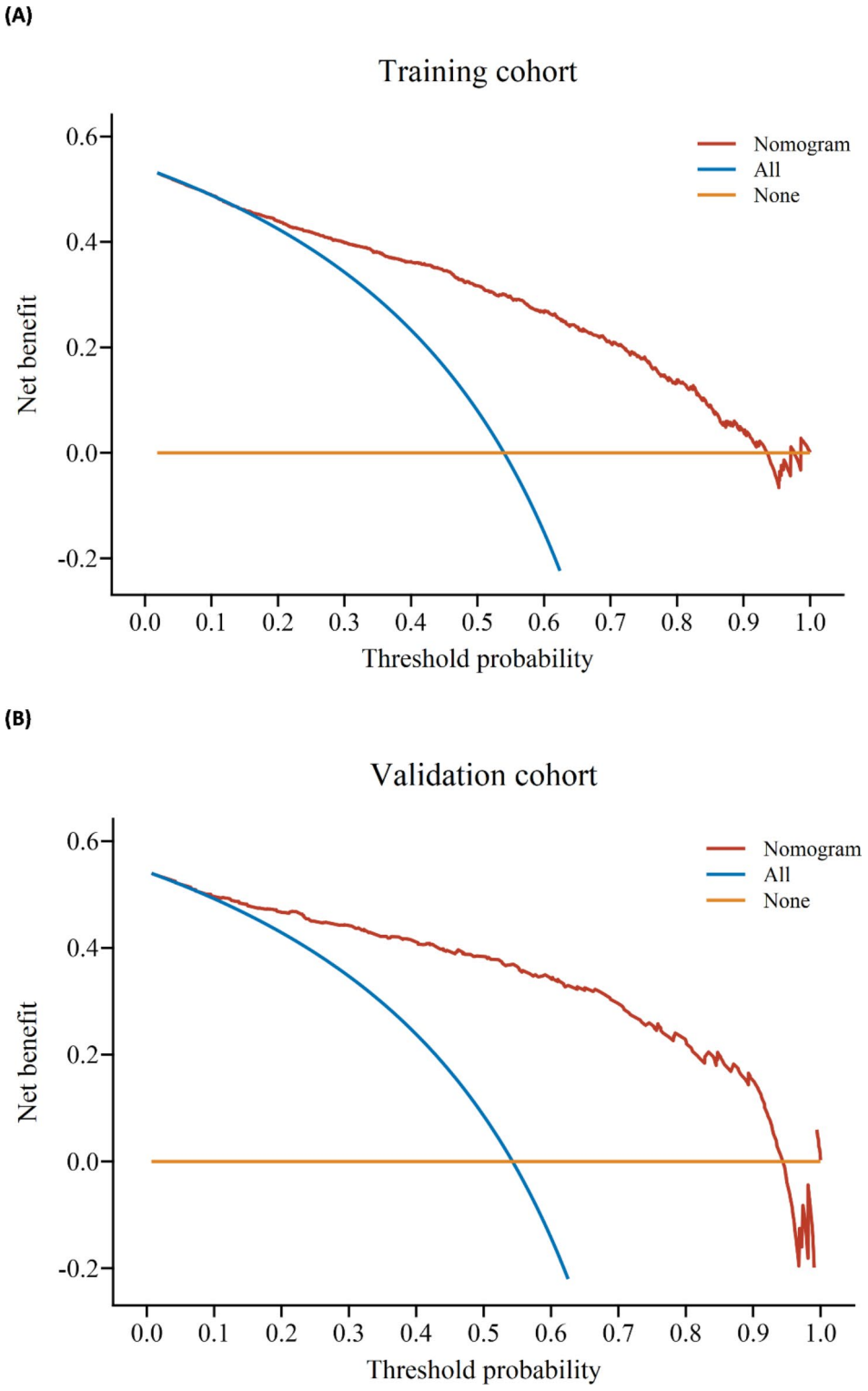

Fig. 4. (**A**) Decision curve analysis of the nomogram prediction model in the training cohort. (**B**) Decision curve analysis of the nomogram prediction model in the validation cohort.

risk factors in numerous studies[16,45], their impact may be confounded by factors such as obesity and diabetes, thereby restricting their individual risk contributions. Moreover, our study's findings indicate that hypertension does not operate as an independent risk factor.

Clinical prediction models are statistical models based on disease characteristics that predict the likelihood of specific events occurring in relevant populations. Currently, most clinical prediction models focus on predicting EC in postmenopausal women[34,46,47], and there is a lack of an authoritative risk assessment model for EH in

premenopausal women, and no consensus has been reached. Srinivas et al.[17] developed a predictive scoring model for premenopausal women with AH/EC. The model includes four variables: an age ≥ 45 years, anovulatory bleeding, a BMI ≥ 30 kg/m^2, and diabetes, achieving an AUC of 0.848. Giannella et al.[33] established a clinical prediction model for EH/EC in symptomatic premenopausal women, incorporating three independent risk factors: a BMI ≥ 30 kg/m^2, diabetes, and an ET > 11 mm, with an AUC of 0.854. Jha et al.[48] developed a model for predicting EH/EC in premenopausal women (aged < 55 years) presenting with AUB. Their model incorporated variables such as an age > 40 years, a BMI > 25 kg/m^2, an ET > 13 mm, hypothyroidism, and menstrual pattern (intermenstrual bleeding), achieving an AUC of 0.971. Although these studies provide reference points for EH/EC prediction models for premenopausal women, their limitations include only focusing on AUB patients and limited including of factors. In our preliminary research, we combined risk factors mentioned in guidelines and various predictive models to identify 20 variables and indices. We integrated seven independent risk factors (BMI, age at menarche, IUD use, diabetes, PCOS, ET, and uterine cavity fluid) into the model using a nomogram, a predictive model widely utilized for the rapid calculation of a patient's disease risk and prognosis in clinical practice. Kuai et al.[49] developed a clinical prediction model for EH/EC in young women (age ≤ 40 years), incorporating variables such as BMI, PCOS, infertility, anemia, menstruation, bleeding description, and ET, achieving AUCs of 0.899, 0.867, and 0.956 for EH/EC, EH, and AH/EC, respectively. Our results are similar, but we expanded the model to include premenopausal women over 40 years of age, making it applicable to a broader population. After constructing the model, we evaluated its discrimination, calibration, and clinical usefulness, obtaining favorable results. Additionally, we validated the model using an independent validation set, yielding satisfactory outcomes. These findings demonstrated that the model can accurately predict EH/EC risk in clinical practice, providing gynecologists with decision-making support.

In clinical settings, patients with scores below 76.411 can be considered low-risk individuals, receiving primarily noninvasive diagnostic and therapeutic interventions to address their symptoms, meet their fertility demands, and avoid declining sexual health due to complications. For high-risk patients with scores above 76.411, invasive procedures are necessary to diagnose EH/EC promptly, reducing misdiagnosis and treatment delays, safeguarding the reproductive health of young women.

There are some limitations to our study. This was a retrospective single-center study. The nomogram requires validation in other medical centers. We plan to perform external validation using data from other institutions and design prospective studies for further verification. During participant selection, the inclusion or exclusion of individuals with missing medical record data may introduce bias. In this study, we opted to exclude such individuals, leading to potential biases. Moreover, cultural and racial disparities between the East and West resulted in factors such as smoking and alcohol consumption not being included as criteria in the nomogram. Hence, certain limitations exist when applying the model. Last, due to cost considerations, this study did not incorporate laboratory or imaging examinations. In the future, if costs decrease or guidelines recommend it, the inclusion of such data could enhance the nomogram's accuracy.

Conclusion

EH/EC is significantly associated with BMI, age at menarche, IUD use, diabetes, PCOS, ET, and uterine cavity fluid. The nomogram model can be used to predict the risk of EH/EC in premenopausal women and facilitate rapid screening.

Methods

Study population

We conducted a retrospective study at the Department of Gynecology in Chongqing Ninth People's Hospital, selecting patients who underwent hysteroscopy with diagnostic curettage from January 2016 to June 2023. The study was approved by the ethics committee board of Chongqing Ninth People's Hospital [2023 K(IRB)014] and exempted patients from informed consent, and adhered to the principles of the Declaration of Helsinki. The inclusion criteria were as follows: (a) patients who subsequently underwent hysteroscopy with diagnostic curettage in premenopausal period; and (b) patients with complete electronic medical records (including data on the following variables). The exclusion criteria were as follows: (a) postmenopausal patients; (b) patients who underwent hysteroscopy with diagnostic curettage due to high bleeding volume or emergency situations; (c) patients taking medications affecting coagulation; (d) patients with Mirena intrauterine devices; and (e) patients with severe liver and kidney function damage, autoimmune diseases, and thrombotic and hemorrhagic diseases.

Cohort definition and recoded variable

Patients were divided into the training cohort ($n = 1156$) and the validation cohort ($n = 385$) at a ratio of 3:1. The training cohort, consisting of 1156 patients, was utilized to screen variables and construct the nomogram. The validation cohort, comprising 385 patients, served as an independent dataset for internal validation.

Based on previous articles and expert discussions, the following variables were extracted from each patient's electronic medical records:

(1) Basic characteristics: age, BMI, and symptoms (AUB, vaginal fluid, uterine cavity fluid, and hypogastric pain).
(2) Medical history: menopause status, age at menarche, parity, IUD use, iatrogenic factors (such as long-term use of estrogen without progesterone antagonism or tamoxifen), and smoking status.
(3) Comorbidities: diabetes, hypertension, hyperlipidemia (HLP), hyperuricemia, hyperthyroidism, hypothyroidism, PCOS, and hyperprolactinemia.
(4) Other data: Histopathology results and ultrasonic examination findings (ET and the presence of uterine cavity fluid).

*The IUDs utilized in this study were metallic contraceptive rings.

Statistical analysis

Categorical variables are presented as numbers and percentages (%), while continuous variables are expressed as the mean ± standard deviation (SD) if they followed a normal distribution; otherwise, they are presented as the median (range). To compare differences between the training and validation cohorts, two-sample t tests were used for normally distributed continuous variables, and Wilcoxon rank-sum tests were used for nonnormally distributed continuous variables. Chi-square or Fisher's exact tests were applied for categorical variables.

Univariable logistic regression analysis was performed to assess the association of each factor with EH/EC within all clinicopathological parameters. Variables showing significant differences ($p < 0.05$) in the univariate analysis were included in a multivariable logistic analysis. Independent risk factors ($P < 0.05$), identified from the multivariate analysis, were incorporated into the nomogram.

The performance of the nomogram was evaluated in terms of discrimination, calibration, and clinical usefulness. Discrimination was quantified using the AUC, where values of 0.5 and 1.0 indicate no discrimination and complete discrimination, respectively. Calibration was assessed through calibration plots, which visualize the agreement between the nomogram-based predictions and the actual observed EH/EC probabilities in both the training and validation cohorts. DCA was used to evaluate clinical effectiveness by measuring standardized net benefits at different threshold probabilities. These analyses assess the model's ability to predict outcomes based on a set of risk parameters.

Statistical analyses were conducted using SPSS statistical software, version 23.0 (SPSS Inc.), and R Studio, with the rms, pROC, ggplot2, and ggDCA packages. All tests were two-sided, and a p value < 0.05 was considered statistically significant, unless otherwise indicated.

Data availability

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

Received: 5 August 2024; Accepted: 16 December 2024
Published online: 06 January 2025

References

1. Sung, H. et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**(3), 209–249. https://doi.org/10.3322/caac.21660 (2021).
2. Reed, S. D. et al. Incidence of endometrial hyperplasia. *Am. J. Obstet. Gynecol.* **200**(6), 678.e1–678.e6786. https://doi.org/10.1016/j.ajog.2009.02.032 (2009).
3. Li, L. et al. Chinese guideline on the management of endometrial hyperplasia. *Chin. J. Obstet. Gynecol.* **57**(8), 566–574. https://doi.org/10.3760/cma.j.cn112141-20220628-00418 (2022). (in Chinese).
4. WHO Classification of Tumours Editorial Board. *WHO Classification of Tumours: Female Genital Tumours [M]* 5th edn 248–251 (IARC, 2020).
5. Sanderson, P. A., Critchley, H. O., Williams, A. R., Arends, M. J. & Saunders, P. T. New concepts for an old problem: The diagnosis of endometrial hyperplasia. *Hum. Reprod. Update.* **23**(2), 232–254. https://doi.org/10.1093/humupd/dmw042 (2017).
6. Armstrong, A. J., Hurd, W. W., Elguero, S., Barker, N. M. & Zanotti, K. M. Diagnosis and management of endometrial hyperplasia. *J. Minim. Invasive Gynecol.* **19**(5), 562–571. https://doi.org/10.1016/j.jmig.2012.05.009 (2012).
7. Chandra, V., Kim, J. J., Benbrook, D. M., Dwivedi, A. & Rai, R. Therapeutic options for management of endometrial hyperplasia. *J. Gynecol. Oncol.* **27**(1), e8. https://doi.org/10.3802/jgo.2016.27.e8 (2016).
8. Hooker, A. B. et al. Systematic review and meta-analysis of intrauterine adhesions after miscarriage: Prevalence, risk factors and long-term reproductive outcome. *Hum. Reprod. Update.* **20**(2), 262–278. https://doi.org/10.1093/humupd/dmt045 (2014).
9. Singh, P. Abnormal uterine bleeding- evaluation by endometrial aspiration. *J. Midlife Health.* **9**(1), 32–35. https://doi.org/10.4103/jmh.JMH_109_17 (2018).
10. Peitsidis, P., Manolakos, E., Tsekoura, V., Kreienberg, R. & Schwentner, L. Uterine arteriovenous malformations induced after diagnostic curettage: A systematic review. *Arch. Gynecol. Obstet.* **284**(5), 1137–1151. https://doi.org/10.1007/s00404-011-2067-7 (2011).
11. Hüsing, A. et al. An epidemiological model for prediction of endometrial cancer risk in Europe. *Eur. J. Epidemiol.* **31**(1), 51–60. https://doi.org/10.1007/s10654-015-0030-9 (2016).
12. Hutt, S. et al. Statistical meta-analysis of risk factors for endometrial cancer and development of a risk prediction model using an artificial neural network algorithm. *Cancers (Basel).* **13**(15), 3689. https://doi.org/10.3390/cancers13153689 (2021).
13. Burbos, N. et al. Estimating the risk of endometrial cancer in symptomatic postmenopausal women: A novel clinical prediction model based on patients' characteristics. *Int. J. Gynecol. Cancer.* **21**(3), 500–506. https://doi.org/10.1097/IGC.0b013e31820c4cd6 (2011).
14. Bhardwaj, V. et al. Machine learning for endometrial cancer prediction and prognostication. *Front. Oncol.* **12**, 852746. https://doi.org/10.3389/fonc.2022.852746 (2022).
15. Zhang, H. et al. Correlation of metabolic factors with endometrial atypical hyperplasia and endometrial cancer: Development and assessment of a new predictive nomogram. *Cancer Manag. Res.* **13**, 7937–7949. https://doi.org/10.2147/CMAR.S335924 (2021).
16. Ruan, H. et al. Development and validation of a nomogram prediction model for endometrial malignancy in patients with abnormal uterine bleeding. *Yonsei Med. J.* **64**(3), 197–203. https://doi.org/10.3349/ymj.2022.0239 (2023).
17. Bagepalli Srinivas, S. et al. A novel risk-scoring model for prediction of premalignant and malignant lesions of uterine endometrium among symptomatic premenopausal women. *Int. J. Womens Health.* **12**, 883–891. https://doi.org/10.2147/IJWH.S268169 (2020).
18. Balachandran, V. P., Gonen, M., Smith, J. J. & DeMatteo, R. P. Nomograms in oncology: More than meets the eye. *Lancet Oncol.* **16**(4), e173–e180. https://doi.org/10.1016/S1470-2045(14)71116-7 (2015).
19. Wang Gang, C. et al. Expert advice on long-term management of endometrial proliferative diseases. *Chin. J. Family Plann. Gynecotokology* (7), 7–11 (2022) (**in Chinese**).
20. Göl, K., Saraçoğlu, F., Ekici, A. & Sahin, I. Endometrial patterns and endocrinologic characteristics of asymptomatic menopausal women. *Gynecol. Endocrinol.* **15**(1), 63–67 (2001).
21. Wise, M. R., Gill, P., Lensen, S., Thompson, J. M. & Farquhar, C. M. Body mass index trumps age in decision for endometrial biopsy: Cohort study of symptomatic premenopausal women. *Am. J. Obstet. Gynecol.* **215**(5), 598. .e1-598.e8 (2016).

22. Aune, D. et al. Anthropometric factors and endometrial cancer risk: A systematic review and dose-response meta-analysis of prospective studies. *Ann. Oncol.* **26**(8), 1635–1648. https://doi.org/10.1093/annonc/mdv142 (2015).
23. Azziz, R. et al. The prevalence and features of the polycystic ovary syndrome in an unselected population. *J. Clin. Endocrinol. Metab.* **89**(6), 2745–2749. https://doi.org/10.1210/jc.2003-032046 (2004).
24. Haoula, Z., Salman, M. & Atiomo, W. Evaluating the association between endometrial cancer and polycystic ovary syndrome. *Hum. Reprod.* **27**(5), 1327–1331. https://doi.org/10.1093/humrep/des042 (2012).
25. Barry, J. A., Azizia, M. M. & Hardiman, P. J. Risk of endometrial, ovarian and breast cancer in women with polycystic ovary syndrome: A systematic review and meta-analysis. *Hum. Reprod. Update.* **20**(5), 748–758. https://doi.org/10.1093/humupd/dmu012 (2014).
26. Xue, Z. et al. Research progress on the mechanism between polycystic ovary syndrome and abnormal endometrium. *Front. Physiol.* **12**, 788772. https://doi.org/10.3389/fphys.2021.788772 (2021).
27. Shan, H. et al. Abnormal endometrial receptivity and oxidative stress in polycystic ovary syndrome. *Front. Pharmacol.* **13**, 904942. https://doi.org/10.3389/fphar.2022.904942 (2022).
28. Gallagher, E. J. & LeRoith, D. Diabetes, cancer, and metformin: Connections of metabolism and cell proliferation. *Ann. N. Y. Acad. Sci.* **1243**, 54–68. https://doi.org/10.1111/j.1749-6632.2011.06285.x (2011).
29. Kaya, S., Kaya, B., Keskin, H. L., Kayhan Tetik, B. & Yavuz, F. A. Is there any relationship between benign endometrial pathologies and metabolic status? *J. Obstet. Gynaecol.* **39**(2), 176–183. https://doi.org/10.1080/01443615.2018.1469606 (2019).
30. Ozalp, S., Kabukcuoglu, S. & Tanir, H. M. Should endometrial hyperplasia be regarded as a reason for abnormal uterine bleeding in users of the intrauterine contraceptive device? *Eur. J. Contracept. Reprod. Health Care.* **8**(1), 17–20 (2003).
31. Timmermans, A. et al. Endometrial thickness measurement for detecting endometrial cancer in women with postmenopausal bleeding: A systematic review and meta-analysis. *Obstet. Gynecol.* **116**(1), 160–167. https://doi.org/10.1097/AOG.0b013e3181e3e7e8 (2010).
32. Kumari, P., Gaikwad, H. S. & Nath, B. Endometrial cut off thickness as predictor of endometrial pathology in perimenopausal women with abnormal uterine bleeding: A cross-sectional study. *Obstet. Gynecol. Int.* **2022**, 5073944. https://doi.org/10.1155/2022/5073944 (2022).
33. Giannella, L., Cerami, L. B., Setti, T., Bergamini, E. & Boselli, F. Prediction of endometrial hyperplasia and cancer among premenopausal women with abnormal uterine bleeding. *Biomed. Res. Int.* 2019:8598152 (2019). https://doi.org/10.1155/2019/8598152. Erratum in: Biomed. Res. Int. 2020;2020:3653414.
34. Cong, Q. et al. Histopathology of women with non-uniform endometrial echogenicity and risk factors for atypical endometrial hyperplasia and carcinoma. *Am. J. Transl. Res.* **13**(5), 4500–4509 (2021).
35. Van Den Bosch, T. et al. Typical ultrasound features of various endometrial pathologies described using International Endometrial Tumor Analysis (IETA) terminology in women with abnormal uterine bleeding. *Ultrasound Obstet. Gynecol.* **57**(1), 164–172. https://doi.org/10.1002/uog.22109 (2021).
36. Nasheeha, N. & Gk, P. Diagnostic accuracy of uterine artery and spiral artery Doppler for evaluation of endometrial pathology in postmenopausal bleeding. *J. Gynecol. Obstet. Hum. Reprod.* **50**(10), 102209. https://doi.org/10.1016/j.jogoh.2021.102209 (2021).
37. Lai, Z. Z. et al. Changes in subsets of immunocytes in endometrial hyperplasia. *Am. J. Reprod. Immunol.* **84**(4), e13295. https://doi.org/10.1111/aji.13295 (2020).
38. England, G. C., Moxon, R. & Freeman, S. L. Delayed uterine fluid clearance and reduced uterine perfusion in bitches with endometrial hyperplasia and clinical management with postmating antibiotic. *Theriogenology* **78**(7), 1611–1617. https://doi.org/10.1016/j.theriogenology.2012.07.009 (2012).
39. Verbakel, J. Y. et al. Risk assessment for endometrial cancer in women with abnormal vaginal bleeding: Results from the prospective IETA-1 cohort study. *Int. J. Gynaecol. Obstet.* **159**(1), 103–110. https://doi.org/10.1002/ijgo.14097 (2022).
40. Michnovicz, J. J., Hershcopf, R. J., Naganuma, H., Bradlow, H. L. & Fishman, J. Increased 2-hydroxylation of estradiol as a possible mechanism for the anti-estrogenic effect of cigarette smoking. *N. Engl. J. Med.* **315**(21), 1305–1309. https://doi.org/10.1056/NEJM198611203152101 (1986).
41. Lacey, J. V. Jr et al. Endometrial cancer and menopausal hormone therapy in the National Institutes of Health-AARP Diet and Health Study cohort [published correction appears in Cancer. 2007;110(4):937]. *Cancer* **109**(7), 1303–1311. https://doi.org/10.1002/cncr.22525 (2007).
42. Al-Zoughool, M. et al. Risk of endometrial cancer in relationship to cigarette smoking: Results from the EPIC study. *Int. J. Cancer.* **121**(12), 2741–2747. https://doi.org/10.1002/ijc.22990 (2007).
43. Lindemann, K., Vatten, L. J., Ellstrøm-Engh, M. & Eskild, A. Body mass, diabetes and smoking, and endometrial cancer risk: A follow-up study. *Br. J. Cancer.* **98**(9), 1582–1585. https://doi.org/10.1038/sj.bjc.6604313 (2008).
44. Friberg, E., Orsini, N., Mantzoros, C. S. & Wolk, A. Alcohol intake and endometrial cancer risk: A meta-analysis of prospective studies. *Br. J. Cancer.* **103**(1), 127–131. https://doi.org/10.1038/sj.bjc.6605698 (2010).
45. Yang, X. & Wang, J. The role of metabolic syndrome in Endometrial Cancer: A review. *Front. Oncol.* **9**, 744. https://doi.org/10.3389/fonc.2019.00744 (2019).
46. Vetter, M. H. et al. Preoperative predictors of endometrial cancer at time of hysterectomy for endometrial intraepithelial neoplasia or complex atypical hyperplasia. *Am. J. Obstet. Gynecol.* **222**(1), 60e.1–60.e7 (2020).
47. Rajadurai, V. A. et al. Predictors of endometrial carcinoma in patients with atypical endometrial hyperplasia at a tertiary gynaecological cancer centre in Western Australia. *Aust. N. Z. J. Obstet. Gynaecol.* **61**(2), 275–283. https://doi.org/10.1111/ajo.13300 (2021).
48. Jha, S. et al. Rate of premalignant and malignant endometrial lesion in low-risk premenopausal women with abnormal uterine bleeding undergoing endometrial biopsy. *Obstet. Gynecol. Sci.* **64**(6), 517–523. https://doi.org/10.5468/ogs.21150 (2021).
49. Kuai, D., Tang, Q., Tian, W. & Zhang, H. Rapid identification of endometrial hyperplasia and endometrial endometrioid cancer in young women. *Discov. Oncol.* **14**(1), 121. https://doi.org/10.1007/s12672-023-00736-w (2023).

Author contributions

Zhen Li: project development, data collection, data analysis, manuscript writing.Juan Yin: Data collection or management.Yu Liu: Data analysis.Fanqing Zeng: project development, data analysis, manuscript editing.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to F.Z.

Reprints and permissions information is available at www.nature.com/reprints.