Artificial intelligence for prediction of endometrial intraepithelial neoplasia and endometrial cancer risks in pre- and postmenopausal women

Evrim Erdemoglu, MD; Tekin Ahmet Serel, MD; Erdener Karacan, MD; Oguz Kaan Köksal, MD; İlyas Turan, MD; Volkan Öztürk, MD; Kemal Kürşat Bozkurt, MD

BACKGROUND: The current approach to endometrial cancer screening requires that all patients be able to recognize symptoms, report them, and carry out appropriate interventions. The current approach to endometrial cancer screening could become a problem in the future, especially for Black women and women from minority groups, and could lead to disparities in receiving proper care. Moreover, there is a lack of literature on artificial intelligence in the prediction and diagnosis of endometrial intraepithelial neoplasia and endometrial cancer.

OBJECTIVE: This study analyzed different artificial intelligence methods to help in clinical decision-making and the prediction of endometrial intraepithelial neoplasia and endometrial cancer risks in pre- and postmenopausal women. This study aimed to investigate whether artificial intelligence may help to overcome the challenges that statistical and diagnostic tests could not.

STUDY DESIGN: This study included 564 patients. The features that were collected included age, menopause status, premenopausal abnormal bleeding and postmenopausal bleeding, obesity, hypertension, diabetes mellitus, smoking, endometrial thickness, and history of breast cancer. Endometrial sampling was performed on all women with postmenopausal bleeding and asymptomatic postmenopausal women with an endometrial thickness of at least 3 mm. Endometrial biopsy was performed on premenopausal women with abnormal uterine bleeding and asymptomatic premenopausal women with suspected endometrial lesions. Python was used to model machine learning algorithms. Random forest, logistic regression, multilayer perceptron, Catboost, Xgboost, and Naive Bayes methods were used for classification. The synthetic minority oversampling technique was used to correct the class imbalance in the training sets. In addition, tuning and boosting were used to increase the performance of the models with a 5-fold cross-validation approach using a training set. Accuracy, sensitivity, specificity, positive predictive value, and F1 score were calculated.

RESULTS: The prevalence of endometrial or preuterine cancer was 7.9%. Data from 451 patients were randomly assigned to the training group, and data from another 113 patients were used for internal validation. Of note, 3 of 9 features were selected by the Boruta algorithm for use in the final modeling. Age, body mass index, and endometrial thickness were all associated with a high risk of developing precancerous and cancerous diseases, after fine-tuning for the multilayer computer to have the highest area below the receiver operating characteristic curve (area under the curve, 0.938) to predict a precancerous disease. The accuracy was 0.94 for predicting a precancerous disease. Precision, recall, and F1 scores for the test group were 0.71, 0.50, and 0.59, respectively.

CONCLUSION: Our study found that artificial intelligence can be used to identify women at risk of endometrial intraepithelial neoplasia and endometrial cancer. The model is not contingent on menopausal status or symptoms. This may be an advantage over the traditional methodology because many women, especially Black women and women from minority groups, could not recognize them. We have proposed to include patients to provide age and body mass index, and measurement of endometrial thickness by either sonography or artificial intelligence may help improve healthcare for women in rural or minority communities.

Key words: artificial intelligence, endometrial cancer, endometrial intraepithelial neoplasia, machine learning, minority, prediction

From the Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, Suleyman Demirel University, Isparta, Turkey (Drs Erdemoglu, Turan, and Öztürk); Departments of Urology (Dr Serel); Obstetrics and Gynecology (Drs Karacan and Köksal); Pathology (Dr Bozkurt), Suleyman Demirel University, Isparta, Turkey

Patient consent is not required in this study because no personal information or detail is included.

The authors report no conflict of interests.

Individual deidentified participant data, including data dictionaries, will be shared if the request is approved by the university.

Cite this article as: Erdemoglu E, Serel TA, Karacan E, et al. Artificial intelligence for prediction of endometrial intraepithelial neoplasia/endometrial cancer risk in pre- and postmenopausal women. Am J Obstet Gynecol Glob Rep 2023;XX:x.ex−x.ex.

Corresponding author: Evrim Erdemoglu, MD. evrimmd@yahoo.com

2666-5778/$36.00

http://dx.doi.org/10.1016/j.xagr.2022.100154

AJOG Global Reports at a Glance

Why was this study conducted?

Endometrial cancer risk management requires patients to be able to recognize and report symptoms. This might be problematic in underserved women. Artificial intelligence (AI) is a new popular subject in cancer diagnosis and risk prediction, but it is not well studied in endometrial cancer diagnosis and risk prediction. This study aimed to investigate how AI may accurately predict endometrial cancer risk.

Key findings

Age, body mass index, and endometrial thickness were associated with the risk of precancerous and cancerous diseases. AI modeling did not require symptoms or menopause status. Precision, recall, and F1 scores were 0.71, 0.50, and 0.59, respectively.

What does this add to what is known?

AI can be used to predict endometrial intraepithelial neoplasia and endometrial cancer risks and is not contingent on menopausal status or symptoms. Moreover, AI could help to overcome some barriers and inequalities to women's health.

Introduction

Artificial Intelligence (AI) is becoming increasingly popular in cancer diagnosis and risk prediction; however, it is not well studied in endometrial cancer diagnosis and risk prediction. Endometrial cancer is particularly common and more deadly in women who have less access to adequate healthcare. If these women can become aware of their risk with the help of AI, they can more likely seek help early in the disease progression.

The most common gynecologic malignancy is endometrial cancer, which is estimated to occur in approximately 25.7 women per 100,000 women each year.[1] In the past 2 decades, the mortality rate for endometrial cancer has increased by 21%,[1] despite the availability of more effective diagnostic and treatment options. These trends in the United States are indicative of a global problem. The incidence of endometrial cancer is higher in industrialized countries, but the incidence of the disease has also been reported to be high even in Asian countries and countries with lower-middle income.[2,3] In 2007, endometrial cancer morbidity exceeded that of cervical cancer and affected 13,606 women in 2012, which was the highest number of gynecologic malignant tumor cases in Japan.[2]

There is no screening test for endometrial cancer. The current approach for the detection and treatment of endometrial cancer requires that all patients are able to recognize symptoms and report them and that an adequate intervention is undertaken.[4] The current approach to endometrial cancer screening could be a problem in the future and could lead to disparities in receiving appropriate care.[4] The recognition of postmenopausal bleeding (PMB) was less common in Black women than in White women. Such a scenario was associated with a higher risk of dying from endometrial cancer within 5 years in Black women than in White women.[5,6] The risk of endometrial intraepithelial neoplasia (EIN) and endometrial cancer is high in pre- and perimenopausal women who have abnormal uterine bleeding (AUB),[7] but women from minority groups and women living in rural areas are also exposed to the risk of being unaware of their medical risk, obtaining inadequate medical care, receiving an incomplete diagnosis, and dying from endometrium cancer.[8] As women approach menopause, the difference among physiological bleeding, premenopausal bleeding, and PMB becomes increasingly difficult to determine.[9] Similarly, doctors might find it difficult to distinguish between physiological and abnormal symptoms in perimenopausal women. Furthermore, clinicians and various guidelines have imprecise cutoff values for endometrial thickness to trigger endometrial biopsy and risk thresholds for further investigation.[10]

General programming algorithms use the input data and the given rules to produce outputs, whereas AI can use the input data and the output data to produce rules and patterns. AI can reliably predict results from new input.[9] AI uses complex algorithms to learn the potential relationships among different biological data. This information is used to reason and perform cognitive functions, including problem-solving and decision-making to assist clinical activities.[11,12] AI has the potential to improve the accuracy of diagnosis and treatment in human clinical practice, making predictions on health risks in real time.[11]

Digitalization, smartphone applications (apps), and Internet usage are widely adopted in various socioeconomic classes in many countries. AI and easy-to-use smartphone apps may be able to help assess the risk and diagnose EIN and endometrial cancer across different populations.[13]

AI in predicting and diagnosing EIN and endometrial cancer is understudied. Of note, 1 of 13 studies on AI and endometrial cancer have analyzed demographic data for the prediction of endometrial cancer.[14] However, this study only included postmenopausal women with abnormal bleeding and endometrial thickness of >5 mm.[15] Other studies have used image-based data or clinical parameters to predict myometrial invasion, lymph node metastasis, hysteroscopic diagnosis of endometrial cancer, and response to treatment. To fill the gap in existing models, we analyzed different AI methodologies that may help in clinical decision-making and the prediction of EIN and endometrial cancer risks in pre- and postmenopausal women.

Material and Methods

This study was conducted in the Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, Suleyman Demirel University, Isparta, Turkey. This study was registered and approved by the institutional review board (approval number 166). Data from consecutive patients aged 35 years or older were collected between January 2015 and May 2022. The inclusion criteria for this study were women who had undergone transvaginal ultrasound, endometrial biopsy, dilation and curettage, or hysterectomy. Patients who had Lynch syndrome or were followed up for a diagnosis of endometrial pathology, who had a history of fertility-preserving treatment for endometrial cancer, or who were receiving hormone replacement therapy or selective estrogen receptor modulators were excluded.

The features that were collected included age, menopause, premenopausal abnormal bleeding and PMB, obesity, hypertension, diabetes mellitus, smoking, endometrial thickness, and history of breast cancer. Endometrial sampling was performed on all women with PMB and asymptomatic postmenopausal women with an endometrial thickness of at least 3 mm. Endometrial biopsy was performed on premenopausal women with AUB and asymptomatic premenopausal women with suspected endometrial lesions.

The histopathological diagnosis from endometrial biopsy or hysterectomy specimens was classified according to the 2014 World Health Organization guidelines. Benign lesions and hyperplasia without atypia were coded as benign diseases, and atypical endometrial hyperplasia, EIN, and carcinoma were coded as precancerous diseases.[16] The diagnostic target was the highest histopathological diagnosis (main outcome), and the management target was hysterectomy (secondary outcome).

Data handling and machine learning analysis

Patient data were randomly sampled and divided into 2 groups, with a 0.80:0.20 ratio, based on the 80/20 rule. Feature selection was performed using the Boruta algorithm on 2 groups. The Boruta algorithm is a wrapper around the random forest (RF) classifier, which is a popular machine learning algorithm implemented in the Python package RF. The RF classification algorithm can be run without tuning of parameters and can give an approximate estimation of the feature importance. Python was used to model machine learning algorithms. RF, logistic regression (LR), multilayer perceptron (MLP), Catboost, Xgboost, and Naive Bayes methods were used for classification in both groups. The synthetic minority oversampling technique was used to correct the class imbalance in the training sets. Tuning and boosting were also used to increase the performance of the models with a 5-fold cross-validation approach using a training set. Accuracy, sensitivity, specificity, positive predictive value, and F1 score were calculated. A heatmap of precision and recall was generated from the following website: https://mlu-explain.github.io/precision-recall/.

Results

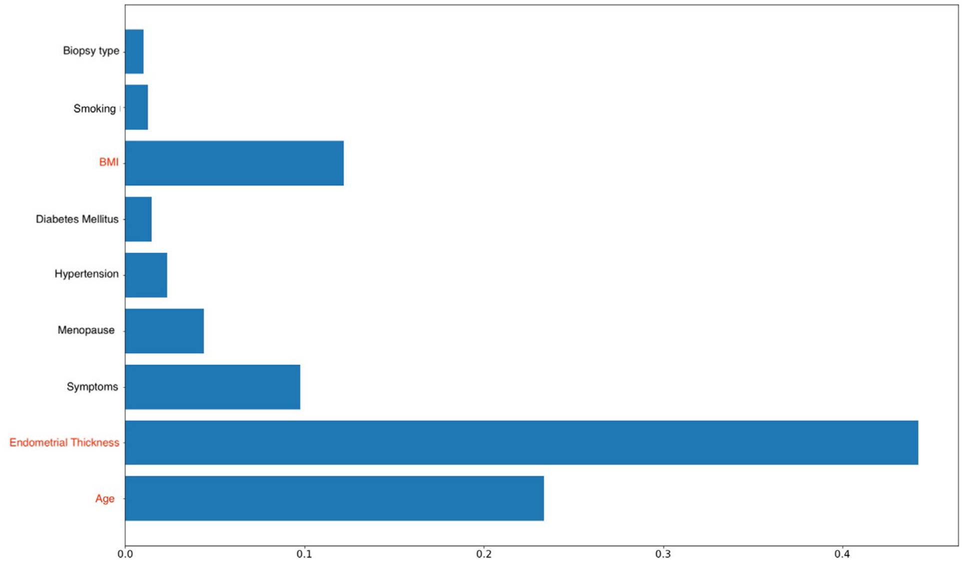
The study included 564 patients. The prevalence of precancer or endometrial cancer was 7.9% (45 of 564 cases). The mean age of the patients was 50.9±10.6 years. The mean body mass index (BMI) of the patients was 29.5±5.3 kg/m^2. The mean endometrial thickness was 8.8±6.5 cm. The data from 451 patients were used for the training cohort, and the data from the other 113 patients were used for internal validation. Of note, 3 of 9 features were selected by the Boruta algorithm for use in the final modeling. Age, BMI, and endometrial thickness were all associated with a high risk of developing precancerous and cancerous diseases (EIN and endometrial cancer). The coefficients of the features are shown in Figure 1. Of note, 6 machine learning models were used to analyze the training cohort: LR, support vector machine, K-nearest neighbors, RF, gradient-boosted decision tree, and neural network (hangileri ahmet abi sor). Cross-validation was performed 5 times to adjust the parameters of these models. After fine-tuning for the MLP to have the highest area below the receiver operating characteristic (ROC) curve (area under the curve [AUC]) to predict precancerous disease, RF has the highest AUC to predict hysterectomy. The AUC for precancer and cancer prediction was 0.94 in the internal validation cohort, showing near-perfect discriminative ability, but for the prediction of hysterectomy as a treatment, the AUC was 0.53. Excluding endometrial thickness from the model yielded very poor precision, recall, and F1 scores. The Table shows the accuracy, precision, recall, and F1 scores of the test group and studied models. The diagnostic prediction of MLP was good with an F1 score of 0.59. Figure 2 shows a heatmap of the precision, recall, and F1 scores.

Comment

Principal findings

To predict endometrial precancer and cancer diseases, age and BMI must be taken into account. These are 2 basic pieces of information that are typically obtained from patients. This information can be conveyed by women from different educational backgrounds and social classes. In contrast, recognizing symptoms, menopause status, and the transition to menopause can be occasionally difficult to discern. Our analysis included symptoms and menopausal status, but the AI dropped these features and instead built a model with more simple information, such as age, BMI, and endometrial thickness. The parameter precision is valuable in MLP studies, as it is not a standard measure in diagnostic studies. The model had a precision of 0.71 for diagnosing endometrial cancer. The recall and sensitivity were 0.50 and 0.58, respectively, showing a trade-off between false positives and false negatives. The F1 score was 0.59. The machine learning model that was developed may provide an opportunity for women to be able to recognize their risk of developing endometrial cancer. Predicting treatment when a patient requires a hysterectomy has been unsuccessful.

FIGURE 1
Coefficients of the features



The age, BMI, and endometrial thickness were all associated with a high risk of developing precancerous and cancerous diseases (endometrial intraepithelial neoplasia and endometrial cancer). In contrast to standard diagnostic tests, artificial intelligence did not identify menopause and abnormal bleeding as factors to be considered in its models. This type of modeling can be beneficial for some women.

BMI, body mass index.

Erdemoglu. Artificial intelligence and risk of endometrial intraepithelial neoplasia. Am J Obstet Gynecol Glob Rep 2023.

Results in the context of what is known

Our study provided new and better information because this study used a new research method, AI, to analyze a well-known previously studied subject. The study's methodology was more reliable, and its data were larger than most others. Our study did not specifically aim to identify risk factors for endometrial cancer, such as in a multivariate or logistic regression analysis.[17–19] The risk factors analyzed in our study were based on previous reports and known risk factors.[20,21] Thus, in medical records, there is information about BMI, hypertension, smoking, and other characteristics. Our main goal was to develop a simple and useful AI model to diagnose endometrial precancer and endometrial cancer. We have included pre- and perimenopausal women and postmenopausal women in the study. According to our clinical protocol, all postmenopausal women had endometrial biopsy regardless of endometrial thickness. In previous studies about AI and endometrial cancer, postmenopausal women with bleeding and a thickness of at least 4 mm have had biopsies.

There is 1 study that has been conducted to predict which patients will

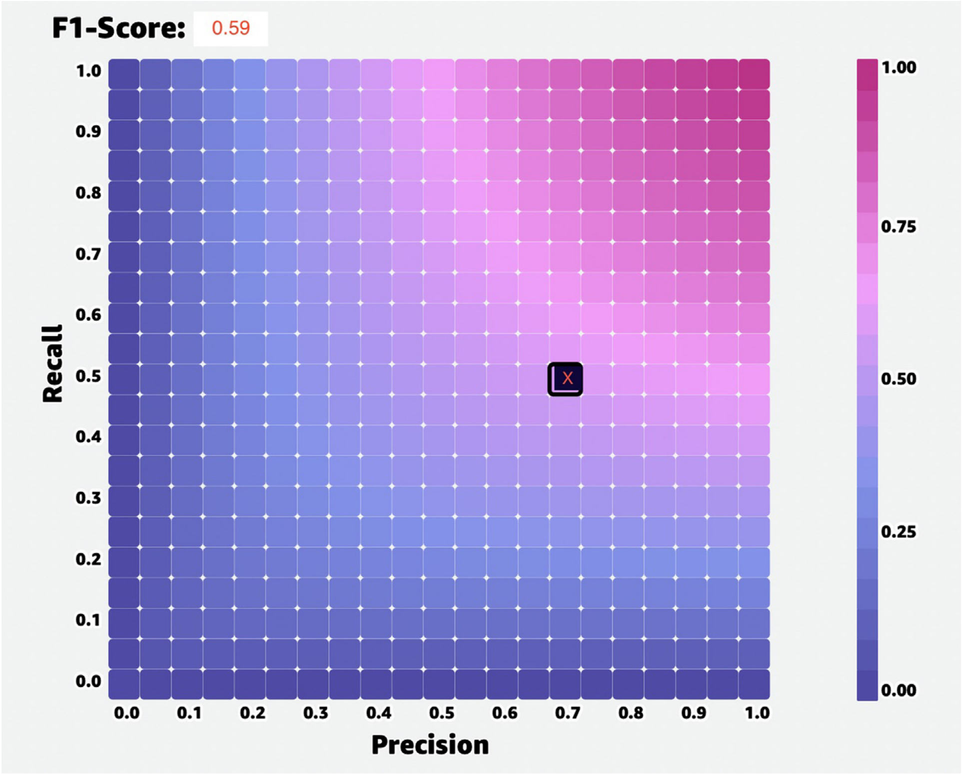
TABLE
Accuracy, precision, recall, and F1 scores

Variable	Recall	Precision	F1 score	Accuracy
Prediction of a precancerous disease with the model using age, BMI, and endometrial thickness				
Presence of a precancerous or cancerous disease	0.50	0.71	0.59	0.94
Prediction of a precancerous disease with the model using age and BMI				
Presence of a precancerous or cancerous disease	0.25	0.14	0.18	0.92
Prediction of hysterectomy as the treatment modality				
Presence of a precancerous or cancerous disease	0.62	0.14	0.22	0.69

BMI, body mass index.

Erdemoglu. Artificial intelligence and risk of endometrial intraepithelial neoplasia. Am J Obstet Gynecol Glob Rep 2023.

FIGURE 2
Heatmap for F1 score



The feedforward artificial neural network accurately diagnosed the precancerous or cancerous disease with an F1 score of 0.59. In the model, the *magenta* area indicates a good F1 score, whereas the *purple navy* area indicates having no predictive power.

Erdemoglu. Artificial intelligence and risk of endometrial intraepithelial neoplasia. Am J Obstet Gynecol Glob Rep 2023.

develop a particular pathology based on their clinical features.[15] This study included only women who have PMB and an endometrial thickness >5 mm. There were 178 women in this study, and among the women who participated in the study, 106 had endometrial cancer. There was an unusually high incidence of endometrial cancer in women who experienced PMB, which may affect the results of the study. Previous reports suggested that the incidence of endometrial cancer in women who experienced PMB is approximately 8% to 10%. Moreover, these findings were consistent with our findings.

In addition, the biopsy procedure may affect the accuracy of test results. Pipelle biopsy has been reported to have a 10% false-negative rate (FNR). Until recently, dilatation of the cervical canal and diagnostic curettage of the endometrium (D&C) has been the gold standard procedure, and currently, hysteroscopy and D&C were regarded as the best options to investigate women with abnormal bleeding. In our study, all women were referred for a D&C, and if it was not possible to perform a D&C, a high-pressure vacuum biopsy was performed. Some previous studies investigating abnormal bleeding using AI have used a pipelle method, whereas other studies have not mentioned their method.

This study did not evaluate the clinical management of abnormal bleeding. We applied a risk-based AI approach to evaluate clinical and diagnostic tools to inform the decision-making of women with abnormal bleeding. In clinical studies, symptoms, such as PMB, and endometrial thickness were found to be important factors to identify women at risk of EIN and endometrial cancer. However, in our study using AI machine learning methods, symptoms were not found to be a significant predictor. It was found that age, BMI, and endometrial thickness were important. The F1 score of the constructed AI model was good. In other words, the false-positive rate and FNR were balanced. Moreover, the same predictors were chosen using AI to guess which patients will be treated by surgery (hysterectomy), but the AI's accuracy for treatment was lower than the diagnostic model's accuracy.

Clinical implications

There are some important implications of our findings; the variability in the AUB and PMB profiles means that the evaluation of endometrial cancer is not always consistent, which is especially true among Black women and women from minority groups.[9] AI may be particularly helpful for these women. Age and BMI can be used as a self-monitoring system to analyze and understand one's risks, but women should still have the opportunity to approach an ultrasonographer. New social plans can be developed to provide at least 1 sonographer in areas where needed.

Input from the patient ensures patient engagement and motivation. This may help patients understand their risks and protective interventions, such as weight loss, diet, and bariatric surgery.[9] Digital tools and AI can help promote healthy interventions and protective measures by gamification. AI could help to overcome some barriers and inequalities for women's health, such as by helping to identify women at risk.

Strengths and limitations

One of the primary limitations of our study was the lack of external validation. In addition, it could be argued that a larger database could have been used. To analyze all potential predictors in univariate analyses, an event with a prevalence of 10% would require a sample size of at least 1200. Our study was based on data from 564 patients. The analysis was performed using AI rather than regression analysis. Multiple regression models and combinations of models were analyzed using AI. Our analysis provided several important results, including the F1 score and precision. These results are vital for understanding the effectiveness of our methods. Previous studies on diagnostic test performance were based on standard diagnostic tests using sensitivity, specificity, accuracy, and the ROC curve. The reported prevalence of endometrial cancer was 8% to 10%, which indicates an uneven class distribution. The F1 score is a great parameter for evaluation because it shows the trade-off between false negatives and false positives (Figure 2). Our study had several strengths, including the elimination of interobserver and intraobserver variabilities and the fact that all patients had a reference test regardless of the index test.

Research implications

The AI model still requires the measurement of endometrial thickness. Imaging is a very popular subject of AI, and in further studies, AI can be developed to measure the endometrial thickness. Such a development, integrated into our AI model, may be used in mobile or public AI centers to evaluate and predict the risk of endometrial cancer in rural areas, without the need for an expert sonographer.

Conclusion

The current methods of managing abnormal bleeding and EIN are not successful and have not decreased the mortality rate. However, our study found that AI can be used to identify women at risk of EIN and endometrial cancer. Moreover, the model is not contingent on menopausal status or symptoms. This may be an advantage over the traditional methodology because many women, particularly Black women and women from minority groups, could not recognize them. In addition, we have proposed to include patients to provide age and BMI, and further improvement of AI to measure endometrial thickness may help to improve the healthcare for women in rural or minority communities. ■

REFERENCES

1. Henley SJ, Miller JW, Dowling NF, Benard VB, Richardson LC. Uterine cancer incidence and mortality - United States, 1999-2016. MMWR Morb Mortal Wkly Rep 2018;67:1333–8.
2. Yamagami W, Nagase S, Takahashi F, et al. Clinical statistics of gynecologic cancers in Japan. J Gynecol Oncol 2017;28:e32.
3. Saeaib N, Sriplung H, Pichatechaiyoot A, Bilheem S. Trends in incidence of uterine cancer in Songkhla, Southern Thailand. J Gynecol Oncol 2019;30:e22.
4. Walther-Antonio MRS, Mariani A. Are early screening biomarkers for endometrial cancer needed to reduce health disparities? Am J Obstet Gynecol 2019;220:408–9.
5. Doll KM, Khor S, Odem-Davis K, et al. Role of bleeding recognition and evaluation in black-white disparities in endometrial cancer. Am J Obstet Gynecol 2018;219. 593.e1–14.
6. Tarney CM, Tian C, Wang G, et al. Impact of age at diagnosis on racial disparities in endometrial cancer patients. Gynecol Oncol 2018;149:12–21.
7. Nicula R, Costin N. Management of endometrial modifications in perimenopausal women. Clujul Med 2015;88:101–10.
8. Kaspers M, Llamocca E, Quick A, Dholakia J, Salani R, Felix AS. Black and Hispanic women are less likely than white women to receive guideline-concordant endometrial cancer treatment. Am J Obstet Gynecol 2020;223. 398.e1–18.
9. DeStephano CC, Bakkum-Gamez JN, Kaunitz AM, Ridgeway JL, Sherman ME. Intercepting endometrial cancer: opportunities to expand access using new technology. Cancer Prev Res (Phila) 2020;13:563–8.
10. Park YR, Lee SW, Kim Y, et al. Endometrial thickness cut-off value by transvaginal ultrasonography for screening of endometrial pathology in premenopausal and postmenopausal women. Obstet Gynecol Sci 2019;62:445–53.
11. Wang R, Pan W, Jin L, et al. Artificial intelligence in reproductive medicine. Reproduction 2019;158:R139–54.
12. Emin EI, Emin E, Papalois A, Willmott F, Clarke S, Sideris M. Artificial intelligence in obstetrics and gynaecology: is this the way forward? In Vivo 2019;33:1547–51.
13. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. Lancet 2020;395:1579–86.
14. Akazawa M, Hashimoto K. Artificial intelligence in gynecologic cancers: current status and future challenges - a systematic review. Artif Intell Med 2021;120:102164.
15. Pergialiotis V, Pouliakis A, Parthenis C, et al. The utility of artificial neural networks and classification and regression trees for the prediction of endometrial cancer in postmenopausal women. Public Health 2018;164:1–6.
16. Sobczuk K, Sobczuk A. New classification system of endometrial hyperplasia WHO 2014 and its clinical implications. Prz Menopauzalny 2017;16:107–11.
17. Giannella L, Mfuta K, Setti T, Cerami LB, Bergamini E, Boselli F. A risk-scoring model for the prediction of endometrial cancer among symptomatic postmenopausal women with endometrial thickness >4 mm. BioMed Res Int 2014;2014:130569.
18. Smith-Bindman R, Weiss E, Feldstein V. How thick is too thick? When endometrial thickness should prompt biopsy in postmenopausal women without vaginal bleeding. Ultrasound Obstet Gynecol 2004;24:558–65.
19. Louie M, Canavan TP, Mansuria S. Threshold for endometrial sampling among postmenopausal patients without vaginal bleeding. Int J Gynaecol Obstet 2016;132:314–7.
20. Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med 2003;348:1625–38.
21. Parslov M, Lidegaard O, Klintorp S, et al. Risk factors among young women with endometrial cancer: a Danish case-control study. Am J Obstet Gynecol 2000;182:23–9.