

RESEARCH

Open Access



Application of machine learning techniques in the diagnosis of endometriosis

Ningning Zhao^{1†}, Ting Hao^{1†}, Fengge Zhang^{1†}, Qin Ni¹, Dan Zhu¹, Yanan Wang¹, Yali Shi¹ and Xin Mi^{1*}

Abstract

Objective The aim of this study is to assess the use of machine learning methodologies in the diagnosis of endometriosis (EM).

Methods This study included a total of 106 patients with EM and 203 patients with non-EM conditions (like simple cysts and simple uterine fibroids), all admitted to the Shunyi Women's and Children's Hospital of Beijing Children's Hospital between January 2017 and September 2022. All participants were free of comorbidities and their diagnoses were confirmed via postoperative pathology. Comparative analysis was conducted between the EM and non-EM groups. Baseline data were assessed, including white blood cell count, neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio, lymphocyte-to-monocyte ratio, mean platelet volume, hemoglobin, carbohydrate antigen 125 (CA125), carbohydrate antigen 199, coagulation parameters, and other serologic indicators. An optimal predictive model was developed using an artificial intelligence algorithm to determine the presence of EM. The objective is to provide new insights for the clinical diagnosis and treatment of EM.

Results The random forest algorithm demonstrated superior performance when compared to decision trees, LogitBoost, artificial neural networks, naïve Bayes, support vector machines, and linear regression in machine learning methods. Combining CA125 with the NLR yielded a better prediction of EM than using CA125 alone when applying the random forest algorithm. The accuracy of predicting EM with CA125 combined with NLR was 78.16%, with a sensitivity of 86.21% and an area under the curve (AUC) of 0.85 ($P < 0.05$). In contrast, using CA125 alone resulted in an EM prediction accuracy of 75.8%, with a sensitivity of 79.3% and an AUC of 0.82 ($P < 0.05$).

Conclusion The diagnostic value of serum CA125 combined with the NLR for EM is higher than that of serum CA125 alone. This finding indicates that NLR could serve as a new supplementary biomarker along with serum CA125 in the diagnosis of EM.

Keywords Endometriosis, Random forests, Serum markers, Machine learning techniques

[†]Ningning Zhao, Ting Hao and Fengge Zhang contributed equally to this study.

*Correspondence:

Xin Mi

mixin1964@126.com

¹Department of Obstetrics and Gynecology, Shunyi Women's and Children's Hospital of Beijing Children's Hospital, No. 1 of Shun Hong Road, Shunyi District, Beijing 101300, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Endometriosis (EM) is a prevalent benign condition affecting the reproductive system in women of child-bearing age, with a prevalence rate of 5–10% [1]. It is characterized by the ectopic presence of endometrial tissue outside the uterine cavity, which undergoes cyclic changes in sync with the menstrual cycle. The etiology of EM is multifactorial, involving sex hormones, immune response, inflammation, and genetic predisposition, although its specific pathogenesis remains unclear. The dominant theory, Sampson's theory of retrograde menstruation, posits that endometrial cells reflux into the pelvic cavity, where they adhere, invade, and undergo vascularization to implant, grow, and develop. The "theory of endometrium in situ" highlights the characteristics role of the endometrial tissue in its ectopic location. Additional theories include coelomic metaplasia, vascular and lymphatic transfer, and stem cell theory. Recent research indicates a strong association between EM and genetic factors, epigenetics, neovascularization, neural neovascularization, epithelial-mesenchymal transition, progesterone resistance, abnormal apoptosis, and inflammation [2].

The ovary is a common site for EM, often affecting one side. As the disease progresses, ectopic endometrial cysts form on the ovary, leading to increased bleeding and intracavitary pressure, especially near the ovarian surface. The cyst wall is prone to repeated rupture, causing the contents of the cyst to irritate the pelvic cavity, resulting in significant adhesions. This can manifest clinically as lower abdominal pain, pain during sexual intercourse, infertility, and other symptoms. Early diagnosis and treatment are critical to enhance the quality of life and fertility outcomes for patients with EM. The gold standard for EM diagnosis is the histologic examination of specimens obtained via invasive procedures such as laparoscopy or transabdominal surgery. These methods, however, are associated with risks, invasiveness, and high costs. Consequently, there is an urgent need for a simple, non-invasive diagnostic test for EM. The objective of this study is to assess the diagnostic value of tumor markers, coagulation markers, and inflammatory factors in EM, with the goal of identifying a new non-invasive diagnostic method for predicting EM.

Machine learning, an artificial intelligence discipline emerged from the confluence of multiple fields, integrating principles from probability theory, statistics, and logic. Contemporary machine learning research has yielded advanced algorithmic tools such as Bayesian methods, logistic regression, and neural networks. These tools are selected based on their suitability for specific application scenarios. For instance, recurrent neural networks are particularly effective for processing text data with sequential and logical order characteristics, while

convolutional neural networks are used for image recognition tasks [3]. Also, regression and clustering algorithms are well suited for data fitting and classification problems. Therefore, various machine learning methods are used in the diagnosis and prediction of EM, yielding diverse results [4].

In a study focused on the preoperative diagnosis and prognosis prediction of ovarian cancer using serum markers, seven machine learning models were assessed: gradient boosting machine (GBM), support vector machine (SVM), random forest (RF), conditional random forest (CRF), naïve Bayes (NB), neural network, and elastic net (EN). When compared to the traditional multiple logistic regression analysis (MLRA) statistical model, which has a prediction accuracy of 86.7%, the area under the curve (AUC) of each machine learning model exceeded this benchmark. Notably, the prediction accuracies of GBM, RF, and CRF were 93.7%, 92.4%, and 93.7%, respectively. These results indicate that machine learning models offer significant value in predicting ovarian cancer-related diseases, outperforming traditional statistical methods in this context [5]. In scenarios involving big data and high complexity among variables, traditional statistical approaches demonstrate limitations in data processing capabilities. Recent biomarker studies have identified serum caspase 3, Annexin A2 (ANXA2), and Soluble Fas Ligand (sFasL) as significant predictors of endometriosis severity [6]. These findings highlight the potential of combining multiple biomarkers for improved diagnostic accuracy. Unlike classical ROC analysis, which evaluates the diagnostic performance of individual markers, the machine learning approach can analyze multiple markers simultaneously and capture complex interactions between them, leading to improved diagnostic accuracy.

Several studies have compared the predictive performance of machine learning models against traditional statistical models in cervical cancer prognosis, assisted reproduction outcomes, maternal postpartum hemorrhage risk, and gestational diabetes risk. These studies consistently reveal that machine learning models demonstrate superior accuracy and higher AUC values compared to their traditional statistical counterparts [7–10]. In the diagnosis of EM, serum markers offer notable advantages such as non-invasiveness, ease of collection, rapid results, and high sensitivity. While carbohydrate antigen 125 (CA125) and carbohydrate antigen 199 (CA199) are frequently used to assist in EM diagnosis, their limited specificity and sensitivity result in elevated levels primarily observed only in severe cases. Recent studies have explored the diagnostic use of various biological markers such as CA125 and Human Epididymis Protein 4 (HE4), in EM diagnosis, although with unsatisfactory results [11]. However, emerging evidence

indicates that as EM progresses, there are discernible changes in hematological markers such as leukocytes, lymphocytes, neutrophils, and neutrophil-to-lymphocyte ratio (NLR) levels [12, 13]. Hence, there is a critical need to identify biomarkers with heightened sensitivity and specificity for individuals with EM, using machine learning modeling methods [14, 15].

Data & methods

General data

The data used in the study were derived from participants who had been hospitalized and had undergone surgery at Shunyi Women's and Children's Hospital of Beijing Children's Hospital between January 2017 and September 2022. These participants had received pathological diagnoses of EM, uterine fibroids, or simple ovarian cysts. Information regarding the study was conveyed to the patients and their families, and informed consent forms were signed.

Exclusion criteria:

1. Patients with concomitant malignant tumors and other organic lesions.
2. A history of previous infection and diabetes.
3. Exhibiting poor adherence to treatment and testing protocols.
4. Pregnant or lactating women.
5. Patients with neurological disorders, who were unable to communicate normally.

Inclusion criteria:

1. Underwent laparoscopic or open surgery at Shunyi Women's and Children's Hospital and received pathological diagnoses consistent with the diagnostic criteria for EM, ovarian cysts, or uterine fibroids.
2. Underwent preoperative blood and serum CA125 and CA199 testing.
3. Had not undergone hormone therapy.

Methods

A total of 106 patients diagnosed with EM (EM group) and 203 patients diagnosed with non-EM conditions (such as uterine fibroids and simple ovarian cysts, control group) and were admitted to Shunyi Women's and Children's Hospital of Beijing Children's Hospital between January 2017 and September 2022, were enrolled. All enrolled patients were aged 18 to 45 years old, were free of comorbidities, and postoperative pathological examinations confirmed the presence of EM, uterine fibroids, or simple cysts. Comparative analyses between the two groups were conducted based on baseline data, including white blood cell count (WBC), hematocrit value,

NLR, platelet-to-lymphocyte ratio (PLR), lymphocyte-to-monocyte ratio (LMR), mean platelet volume (MPV), hemoglobin (Hb), CA125, CA199, and coagulation parameters (activated partial thromboplastin time (APTT), prothrombin time (PT), thrombin time (TT), and fibrinogen (Fib)). These parameters were further analyzed for their correlation with EM.

CA125 and CA199 levels were measured using the cobas® 8000 chemiluminescence instrument manufactured by Roche, Switzerland, along with its respective kit. WBC, neutrophils, lymphocytes, NLR, MPV, Hb, and Fib levels were determined using the CA700 automatic coagulation analyzer produced by Sysmex Corporation, Japan, along with its corresponding kit from Sysmex Corporation, Japan.

Machine learning

A range of machine learning models such as RF, SVM, NB, multiple linear regression, LogitBoost, decision trees, neural networks, and other relevant features, were used. The model demonstrating the highest accuracy was selected for optimal feature targeting and subsequent model development.

Algorithm principle

The RF algorithm was used to develop an auxiliary diagnostic model for EM, using a dataset categorized into EM and non-EM conditions (including cysts and fibroids). Missing data were imputed using the mice v3.14 package in R v4.1.0, using the RF interpolation method with 5 iterations. Subsequently, separate training and test sets were constructed for the two datasets, with 70% of the dataset allocated to the training set and 30% to the test set. The training set for EM and non-EM comprised 200 samples (67 in the EM group and 133 in the control group), while the test set included 87 samples (29 in the EM group and 58 in the control group). RF model training was conducted using the caret package in R version 4.1.3 with 500 trees. The training process involved repeated 10-fold cross-validation, and the optimal model was selected based on the highest accuracy, along with default parameters. Important markers, represented as feature vectors (features), were identified based on their high-ranked importance in contributing to the prediction accuracy of EM. This feature selection process was conducted using the varImp function of the caret package. Subsequently, the RF model was reconstructed for two-by-two combinations of CA125 and the selected markers.

Following the construction of the prediction model, it was initially applied to the test set, and the receiver operating characteristic (ROC) curve was generated to compute the AUC value. The optimal threshold point on the ROC curve was determined based on Youden's index. Youden's index is calculated as Sensitivity + Specificity - 1.

Table 1 Accuracy, specificity, and sensitivity of each model

Model	Accuracy(%)	sensitivity(%)	specificity(%)	AUC
LogitBoost	78.71	65.52	74.14	0.759
cart tree	79.66	75.86	82.76	0.793
RandomForest	83.4	89.66	67.24	0.851
knn	79.33	86.21	67.24	0.798

Samples with a predicted probability greater than or equal to the threshold were classified as EM, while samples with probabilities lower than the threshold were classified as non-EM. Subsequently, various performance metrics such as accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and kappa coefficient were computed on the test set to assess the performance of the model.

Results

As depicted in Table 1, RF achieved the highest accuracy of 83.4%, followed by Decision Tree with 79.66%, K-Nearest Neighbors with 79.33%, and LogitBoost with 78.71%. Therefore, the RF model was selected as the optimal model.

In using the RF model for feature selection (Fig. 1), CA125, CA199, APTT, Hb, and NLR were identified as the optimal indicators.

As depicted in Table 2, the RF model proved most effective in predicting EM. Among the parameters assessed, the combination of CA125 and CA199 predicted EM with an accuracy of 79.31%, sensitivity of 86.2%, specificity of 75.8%, and an AUC of 0.84. The combination of CA125 and Hb exhibited the highest sensitivity at 93.10%, accuracy of 74.1%, and specificity of 65.5%. Also, the combination of CA125 and NLR achieved a maximum AUC of 0.850, with a cutoff of 0.247, an accuracy of 78.1%, sensitivity of 86.2%, and specificity of 74.1%.

In addition, we found that CA125 combined with Hb predicted EM with a specificity of 65.5% and an AUC of 0.84. Additionally, CA125 combined with APTT predicted EM with an accuracy of 78.1%, sensitivity of 75.8%, specificity of 79.3%, and an AUC of 0.78. Currently, consensus is lacking on whether APTT and Hb can be combined with CA125 to predict EM diagnosis.

As depicted in Fig. 2, CA199, Hb, NLR, and APTT were combined with CA125 to predict the ROC curves of EM.

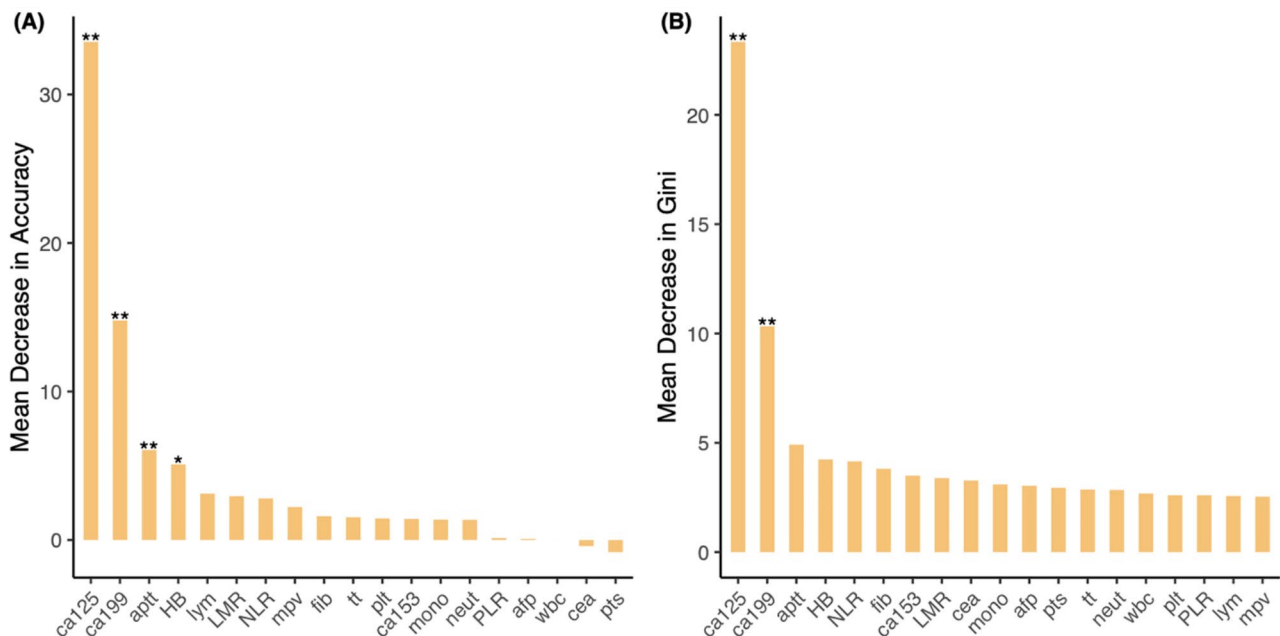


Fig. 1 Optimal feature metrics

Table 2 Predicting endometriosis using the RF algorithm for each feature data

Feature	Accuracy(%)	Sensitivity(%)	Specificity (%)	AUC	P	95%CI
ca125	75.80	79.30	74.10	0.822	0.041	0.655,0.844
ca125+ca199	79.31	86.21	75.86	0.841	0.006	0.6929,0.8725
ca125+APTT	78.16	75.86	79.31	0.789	0.0132	0.6802,0.8631
ca125+NLR	78.16	86.21	74.14	0.850	0.0132	0.6802,0.8631
ca125+HB	74.17	93.10	65.50	0.841	0.067	0.6425,0.8342

Note: area under the curve (AUC)

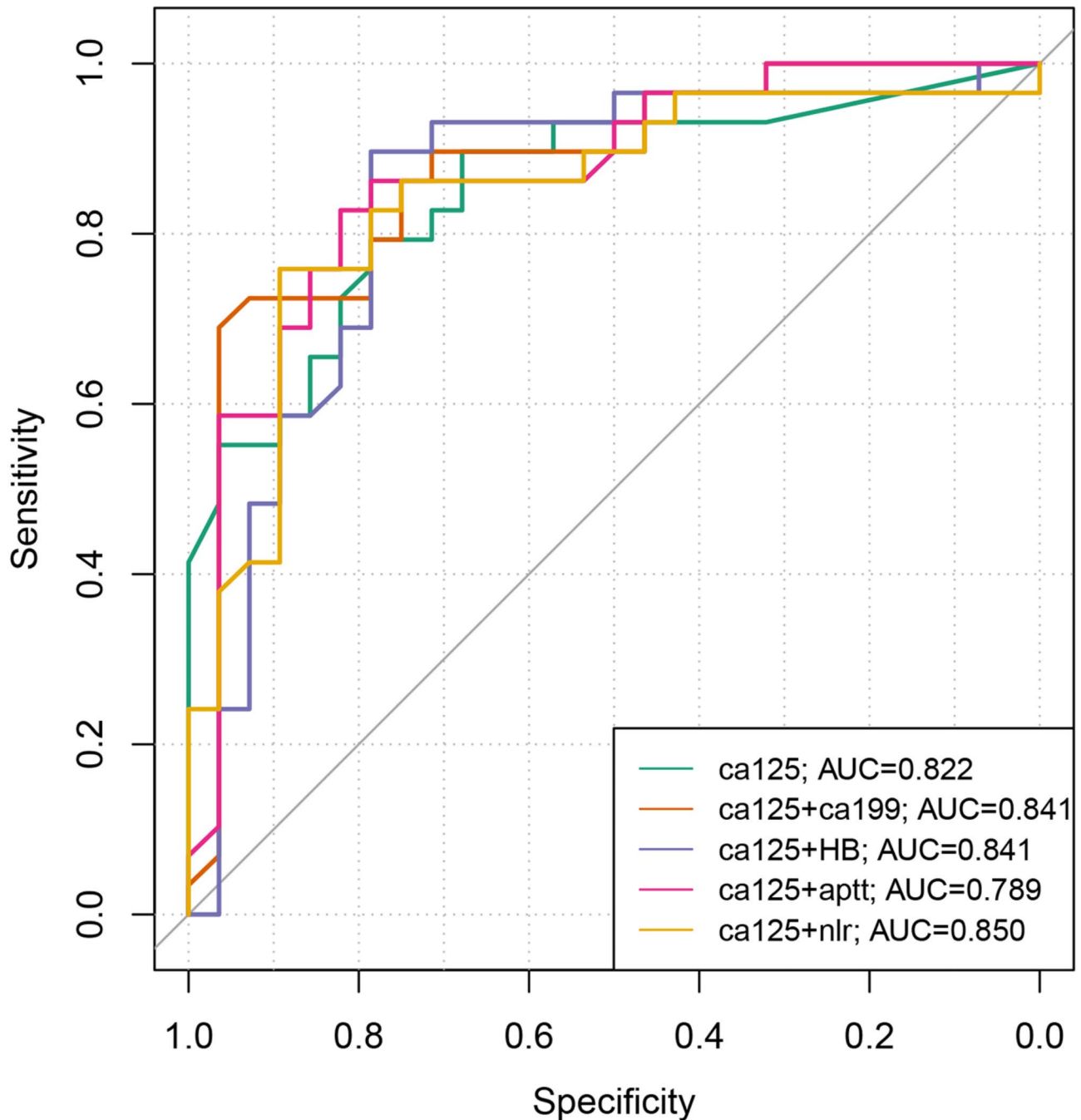


Fig. 2 ROC curve of CA125 combined with each feature

The AUC values indicate the effectiveness of each combination in predicting endometriosis. The combination of CA125 with NLR showed the highest AUC, indicating superior performance. Figure 3 displays the ROC curve for the combination of NLR and CA125 highlighting its effectiveness in predicting endometriosis. The midpoint of the curve at 0.247 indicates the threshold value that maximizes the sum of sensitivity and specificity, resulting in an optimal balance for diagnostic accuracy. The AUC

for this combination was 0.85, demonstrating a significant improvement over using CA125 alone.

Discussion

The cyclic rupture of ectopic cysts associated with EM, along with the subsequent repeated irritation caused by the cyst contents, induces a chronic inflammatory response within the body, thereby exacerbating the progression of EM [16, 17]. Consequently, it is crucial to

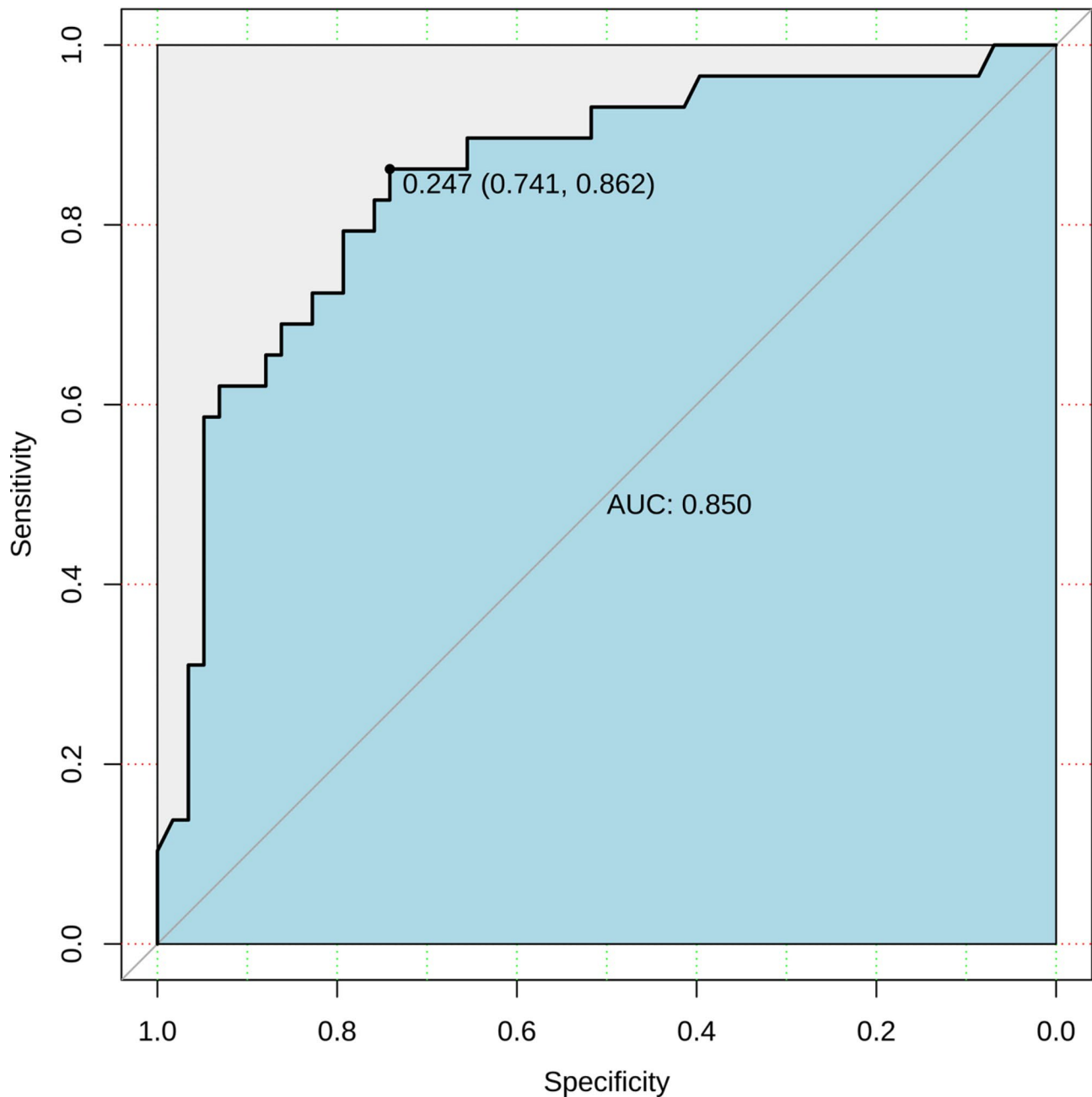


Fig. 3 ROC curve of CA125 combined with NLR

explore new methodologies for predicting the presence of EM and instituting timely treatment interventions to enhance the quality of life and reproductive health of individuals with EM.

Serum CA125, a tumor marker, is widely present in the epithelium and mesothelial cells of the mesonephric ducts. It exhibits a high concentration in various bodily fluids and is present in the cervical epithelium, endometrium, peritoneum, and fallopian tubes. Serving as a valuable indicator for the clinical diagnosis of numerous obstetrics and gynecology disorders, serum CA125

is highly expressed in both EM and adenomyosis. Serum CA199, a high molecular glycoprotein produced by tumor cells within the digestive tract, has been extensively investigated for its diagnostic and therapeutic significance in digestive tract pathologies. Over the past few years, some studies have discovered its expression in EM and mature cystic teratoma. However, the current understanding within the academic community regarding its relevance and diagnostic use for gynecological diseases remains relatively limited, lacking a unified and objective theoretical framework. Presently, CA125 and CA19-9 are

commonly used as adjuncts in the diagnostic assessment of EM [18]. However, due to the modest specificity and sensitivity, coupled with increased levels predominantly observed in severe cases, the NLR [19] emerges as a new inflammatory index capable of predicting and diagnosing immune system disorders. A recent study on tubo-ovarian abscess patients showed that higher NLR levels were associated with medical treatment failure, suggesting that NLR can be a useful marker in predicting treatment outcomes in gynecological conditions [20].

Routine inexpensive blood tests that are widely used can provide valuable insights into the inflammatory response of the body. Clinically, changes in individual blood cell counts can sensitively indicate inflammation, though their specificity is limited. Conversely, blood cell derivatives, involving at least two cell types, provide a more accurate reflection of the inflammatory state of the body. The NLR is significantly correlated with the levels of both cell types. Increased NLR levels can reflect the immune status and inflammatory response, indicating significant correlations with the diagnosis and prognosis of various immune disorders and malignant tumors, such as psoriasis, rectal cancer, and lung cancer based on numerous studies [21]. In this study, by comparing patients in the EM group with those in the control group, it was found that CA125 and NLR were significantly associated with the risk of developing ovarian EM. The combined prediction of CA125 and NLR was superior to the prediction using CA125 alone. It is important to note that CA125 can be elevated due to various causes such as pelvic inflammatory disease, menstruation, and other benign gynecological conditions [22]. This underscores the necessity of combining CA125 with other markers like NLR to improve diagnostic accuracy.

Our data demonstrated that the combination of NLR and CA125 was more sensitive (86.2%) than CA125 alone (79.3%) in diagnosing EM. Similarly, Tokmak et al. [23] reported that the combination of NLR and CA125 exhibited higher sensitivity (80%) and specificity (82%) compared to CA125 alone in diagnosing EM. Our findings indicated that the NLR-CA125 combination increased sensitivity with minimal change in specificity when compared to CA125 alone when differentiating EM from benign ovarian tumors or healthy controls. The ROC curves, sensitivity, and specificity of CA125 and NLR confirmed their use in diagnosing ovarian EM, with the AUC being 0.85. The combined assays significantly enhanced the detection rate of ovarian EM, achieving a sensitivity of 86.21%. Therefore, the combined detection of CA125 and NLR holds substantial value in diagnosing ovarian EM [16]. This study illustrated that NLR combined with CA125 has higher sensitivity than CA125 alone.

Inflammatory processes can initiate and promote coagulation, increasing the risk of bleeding, microvascular thrombosis, and organ dysfunction [24]. In the coagulation cascade reaction, activated platelets and tissue factor bind coagulation factors and thrombin to induce inflammation [25, 26]. Activated Fib also induces thrombin production, further activating chemokine production and macrophage adhesion [27]. It has been suggested that women with EM may exhibit a hypercoagulable and a hyperfibrinolysis state due to platelet aggregation at EM lesions [28, 29]. In patients with EM, platelet count and plasma Fib levels are elevated. Additionally, APTT and TT are decreased, while PT remains at normal levels. Paola et al. [30] demonstrated that APTT was reduced, while TT remained normal in patients with EM. In contrast, Guo et al. found that both TT and APTT were reduced, with PT maintaining normal levels [31]. Variations in coagulation parameters typically arise from using different reagents and manufacturers [28]. Consequently, further investigation into the coagulation function of patients with EM is necessary.

In our study, we demonstrated the significance of APTT as a complementary marker to CA125 in discriminating ovarian EM from non-EM cases. The combined diagnostic accuracy was 78.1%, with a sensitivity of 75.8% and a specificity of 79.3%. This combined accuracy was higher than that of CA125 alone in predicting EM, although the AUC for the combined markers was 0.789, which was lower than the AUC of 0.822 for CA125 alone.

Results by Moinia regarding Hb levels are consistent with other studies indicating that women with endometrial disease tend to have lower Hb concentrations [23]. Severe EM with low Hb levels may be linked to disruptions in erythrocyte regulation or iron metabolism. Parameters such as NLR, Hb levels, and neutrophil counts were effective diagnostic predictors of EM in the study conducted by Moinia [32].

These results indicate that EM is associated with an inflammatory process, implying that measuring these serum markers could provide a simple, non-invasive, inexpensive, and accessible diagnostic tool for clinicians.

In summary, results of the regression analysis conducted in this study confirmed that among several small molecule markers, CA125 and NLR were significantly associated with EM. ROC curve analysis revealed that NLR could complement CA125 in diagnosing EM. Combining these two markers enhanced the sensitivity of diagnosing ovarian EM compared to CA125 alone [33]. Despite this, serum CA125 remains a crucial marker in the diagnosis of EM. It has been suggested that platelets and PLR can predict EM when combined with OMA or uterine fibroids [16]. However, due to the single-center, retrospective design of this study, prospective,

multicenter clinical validation is required to confirm these findings before they can be applied in clinical practice.

As an auxiliary diagnostic tool, RF falls within the criteria of computer-assisted diagnosis and cannot entirely replace the judgment of clinicians. However, the diagnostic auxiliary model for EM established in this study, based on the Rf algorithm, can serve as a powerful tool for clinicians in diagnosing EM.

Conclusion

The diagnostic value of serum CA125 combined with NLR for EM is superior to that of serum CA125 alone. This indicates that NLR may serve as a new adjunctive biomarker for the diagnosis of EM. Women with ovarian EM exhibit hypercoagulability, likely due to the inflammation associated with EM. Consequently, APTT and Hb are also emerging as potential adjunctive biomarkers supporting CA125 in the diagnosis of EM. However, the findings in this area are not yet consistent.

Currently, the application of machine learning in EM remains in the preliminary exploration stage. While some progress has been made in constructing diagnostic models, their accuracy requires further validation. Translating theoretical research into practical clinical diagnostic tools continues to be challenging. Also, the research focus in constructing machine learning prediction models for EM has been relatively narrow. There are numerous prognostic outcomes that warrant further investigation, such as predicting infertility risk, recurrence risk after treatment, pregnancy prediction, and the malignancy rate in EM. At this stage, it is impractical to rely solely on machine learning models for the diagnosis of EM. However, using these models for patient self-testing and pre-screening triage is feasible and likely to become a focus of future research.

Abbreviations

WBC	leukocyte, white blood cell
NLR	neutrophils/lymphocytes
PLR	platelets/lymphocytes
LMR	lymphocytes/monocytes
MPV	mean platelet volume
Hb	Hemoglobin
CA125	Carbohydrate antigen 125
CA199	Carbohydrate antigen 199
APTT	Activated Partial Thromboplastin Time
PT	Prothrombin Time
TT	Thrombin Time
FIB	fibrinogen

Acknowledgements

We would like to acknowledge the hard and dedicated work of all the staff that implemented the intervention and evaluation components of the study.

Author contributions

Conception and design of the research: Xin Mi, Ningning Zhao Acquisition of data: Dan Zhu, Qin Ni, Yanan Wang, Yali Shi Analysis and interpretation of the data: Ting Hao, Fengge Zhang Statistical analysis: Ningning Zhao, Yali Shi Writing of the manuscript: Ningning Zhao Critical revision of the manuscript

for intellectual content: Xin Mi, Ting Hao, Fengge Zhang, Dan Zhu, Qin Ni, Yanan Wang All authors read and approved the final draft.

Funding

No external funding received to conduct this study.

Data availability

All data generated or analysed during this study are included in this article. Further enquiries can be directed to the corresponding author.

Declarations

Ethics approval and consent to participate

This study was conducted with approval from the Ethics Committee of Shunyi Women's and Children's Hospital of Beijing Children's Hospital. This study was conducted in accordance with the declaration of Helsinki. Written informed consent was obtained from all participants.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 17 June 2024 / Accepted: 26 August 2024

Published online: 05 September 2024

References

- Zullo F, Spagnolo E, Saccone G, Acunzo M, Xodo S, Ceccaroni M, Berghella V. Endometriosis and obstetrics complications: a systematic review and meta-analysis. *Fertil Steril*. 2017;108(4):667–72. e5.
- M Becker C, Bokor A, Heikinheimo O, Horne A, Jansen F, Kiesel L, King K, Kvaskoff M, Nap A, Petersen K, Saridogan E, Tomassetti C, Van Hanegem N, Vulliamoz N, Vermeulen N, Group E E G. ESHRE guideline: endometriosis. *Hum Reprod Open*. 2022;2022(2):hoac009.
- Lee KS, Ahn KH. Application of Artificial Intelligence in early diagnosis of spontaneous Preterm Labor and Birth. *Diagnostics (Basel)*, 2020, 10(9).
- Sivajohan B, Elgendi M, Menon C, Allaire C, Yong P, Bedaiwy MA. Clinical use of artificial intelligence in endometriosis: a scoping review. *NPJ Digit Med*. 2022;5(1):109.
- Kawakami E, Tabata J, Yanaihara N, Ishikawa T, Koseki K, Iida Y, Saito M, Komazaki H, Shapiro JS, Goto C, Akiyama Y, Saito R, Saito M, Takano H, Yamada K, Okamoto A. Application of Artificial Intelligence for Preoperative Diagnostic and Prognostic Prediction in Epithelial Ovarian Cancer based on blood biomarkers. *Clin Cancer Res*. 2019;25(10):3006–15.
- Kaya C, Alay I, Guraslan H, Gedikbasi A, Ekin M, Ertaş Kaya S, Oral E, Yasar L. The role of Serum Caspase 3 levels in prediction of endometriosis severity. *Gynecol Obstet Invest*. 2018;83(6):576–85.
- Matsuo K, Purushotham S, Jiang B, Mandelbaum RS, Takiuchi T, Liu Y, Roman LD. Survival outcome prediction in cervical cancer: Cox models vs deep-learning model. *Am J Obstet Gynecol*. 2019;220(4):381. e1–381 e14.
- Barnett-Itzhaki Z, Elbaz M, Buttermann R, Amar D, Amitay M, Racowsky C, Orvieto R, Hauser R, Baccarelli AA, Machtinger R. Machine learning vs. classic statistics for the prediction of IVF outcomes. *J Assist Reprod Genet*. 2020;37(10):2405–12.
- Venkatesh KK, Strauss RA, Grotegut CA, Heine RP, Chescheir NC, Stringer JSA, Stamilio DM, Menard KM, Jelovsek JE. Machine learning and statistical models to Predict Postpartum Hemorrhage. *Obstet Gynecol*. 2020;135(4):935–44.
- Ye Y, Xiong Y, Zhou Q, Wu J, Li X, Xiao X. Comparison of Machine Learning Methods and Conventional Logistic Regressions for Predicting Gestational Diabetes Using Routine Clinical Data: A Retrospective Cohort Study. *J Diabetes Res*, 2020, 2020: 4168340.
- Kundur M, Bhati P, Girish BK, Sheejamol VS, Nair IR, Pavithran K, Rajanbabu A. Endometriosis in clear cell and endometrioid carcinoma ovary: its impact on clinicopathological characteristics and survival outcomes. *Ecanermedicscience*. 2023;17:1591.
- Ozdemir A, Baran E, Kutu M, Celik S, Yilmaz M. Could systemic immune inflammation index be a new parameter for diagnosis and disease

- activity assessment in systemic lupus erythematosus? *Int Urol Nephrol*. 2023;55(1):211–6.
13. Soliman WM, Sherif NM, Ghanima IM, El-Badawy MA. Neutrophil to lymphocyte and platelet to lymphocyte ratios in systemic lupus erythematosus: relation with disease activity and lupus nephritis. *Reumatol Clin (Engl Ed)*. 2020;16(4):255–61.
 14. Blass I, Sahar T, Shraibman A, Ofer D, Rappoport N, Liniat M. Revisiting the risk factors for endometriosis: a Machine Learning Approach. *J Pers Med*, 2022, 12(7).
 15. Brosens I, Puttemans P, Benagiano G. Endometriosis: a life cycle approach? *Am J Obstet Gynecol*. 2013;209(4):307–16.
 16. Jing X, Li C, Sun J, Peng J, Dou Y, Xu X, Ma C, Dong Z, Liu Y, Zhang H, Shao Q, Zhang H, Wang L, Zhang Y, Qu X. Systemic inflammatory response markers Associated with Infertility and Endometrioma or Uterine Leiomyoma in Endometriosis. *Ther Clin Risk Manag*. 2020;16:403–12.
 17. Zhao K, Qu P. Noninvasive evaluation of ovarian endometriosis: a single-center experience. *Ann Palliat Med*. 2021;10(4):4728–35.
 18. Shen A, Xu S, Ma Y, Guo H, Li C, Yang C, Zou S. Diagnostic value of serum CA125, CA19-9 and CA15-3 in endometriosis: a meta-analysis. *J Int Med Res*. 2015;43(5):599–609.
 19. Turgut A, Hocaoglu M, Ozdamar O, Usta A, Gunay T, Akdeniz E. Could hematologic parameters be useful biomarkers for the diagnosis of endometriosis? *Bratisl Lek Listy*. 2019;120(12):912–8.
 20. Alay I, Kaya C, Karaca I, Eren E, Hosgoren M, Aslanova F, Cengiz H, Ekin M, YaSar L. The effectiveness of neutrophil to lymphocyte ratio in prediction of medical treatment failure for tubo-ovarian abscess. *J Obstet Gynaecol Res*. 2019;45(6):1183–9.
 21. Park W, Lopes G. Perspectives. Neutrophil-to-lymphocyte ratio as a potential biomarker in Immune checkpoint inhibitor for non-small-cell Lung Cancer. *Clin Lung Cancer*. 2019;20(3):143–7.
 22. Kaya C, Cengiz H, Ekin M, Yasar L. Transverse vaginal septum: a benign reason for elevated serum CA 19–9 and CA 125 levels. *Arch Gynecol Obstet*. 2012;286(3):821–3.
 23. Tokmak A, Yildirim G, Oztas E, Akar S, Erkenekli K, Gulsen P, Yilmaz N, Ugur M. Use of Neutrophil-to-lymphocyte ratio combined with CA-125 to Distinguish Endometriomas from other Benign Ovarian cysts. *Reprod Sci*. 2016;23(6):795–802.
 24. Levi M, Van Der Poll T. Coagulation and sepsis. *Thromb Res*. 2017;149:38–44.
 25. Hirota Y, Osuga Y, Hirata T, Harada M, Morimoto C, Yoshino O, Koga K, Yano T, Tsutsumi O, Taketani Y. Activation of protease-activated receptor 2 stimulates proliferation and interleukin (IL)-6 and IL-8 secretion of endometriotic stromal cells. *Hum Reprod*. 2005;20(12):3547–53.
 26. Hirota Y, Osuga Y, Hirata T, Yoshino O, Koga K, Harada M, Morimoto C, Nose E, Yano T, Tsutsumi O, Taketani Y. Possible involvement of thrombin/protease-activated receptor 1 system in the pathogenesis of endometriosis. *J Clin Endocrinol Metab*. 2005;90(6):3673–9.
 27. Szaba FM, Smiley ST. Roles for thrombin and fibrin(ogen) in cytokine/chemokine production and macrophage adhesion in vivo. *Blood*. 2002;99(3):1053–9.
 28. Ding S, Lin Q, Zhu T, Li T, Zhu L, Wang J, Zhang X. Is there a correlation between inflammatory markers and coagulation parameters in women with advanced ovarian endometriosis? *BMC Womens Health*. 2019;19(1):169.
 29. Ding D, Liu X, Duan J, Guo SW. Platelets are an unindicted culprit in the development of endometriosis: clinical and experimental evidence. *Hum Reprod*. 2015;30(4):812–32.
 30. Vigano P, Ottolina J, Sarais V, Rebonato G, Somigliana E, Candiani M. Coagulation Status in Women with endometriosis. *Reprod Sci*. 2018;25(4):559–65.
 31. Ng VL. Prothrombin time and partial thromboplastin time assay considerations. *Clin Lab Med*. 2009;29(2):253–63.
 32. Moini A, Ghanaat M, Hosseini R, Rastad H, Hosseini L. Evaluating hematological parameters in women with endometriosis. *J Obstet Gynaecol*. 2021;41(7):1151–6.
 33. Tang T, Lai H, Huang X, Gu L, Shi H. Application of serum markers in diagnosis and staging of ovarian endometriosis. *J Obstet Gynaecol Res*. 2021;47(4):1441–50.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.