International Journal of Women's Health

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Extracellular Trap-Related Genes as Potential Diagnostic Biomarkers for Endometriosis

Ya-Xing Fang[1–3,*], Yu-Feng He[1–3,*], Bing-Bing Wang[1–3,*], Juan He[1,2], Yang Dong[1,2], Guo Chen[1–3], Shu-Guang Zhou[3–6]

[1]Department of Obstetrics and Gynecology, Maternity and Child Healthcare Hospital Affiliated to Anhui Medical University, Hefei, Anhui, 230001, People's Republic of China; [2]Department of Gynecology, Anhui Women and Children's Medical Center, Hefei, Anhui, 230001, People's Republic of China; [3]Department of Obstetrics and Gynecology, Anhui Medical University, Hefei, Anhui, 230032, People's Republic of China; [4]Department of Obstetrics and Gynecology, Anhui Maternal and Child Health Hospital, Hefei, Anhui, 230051, People's Republic of China; [5]Department of Obstetrics and Gynecology, Anhui Provincial Children's Hospital, Hefei, Anhui, 230051, People's Republic of China; [6]School of Public Health, Harbin Medical University, Harbin, Heilongjiang, 150081, People's Republic of China

*These authors contributed equally to this work

Correspondence: Shu-Guang Zhou; Guo Chen, Email zhoushuguang@ahmu.edu.cn; chenguo@ahmu.edu.cn

Background: Endometriosis (EM), a prevalent gynecological disorder in reproductive-age women, lacks reliable noninvasive diagnostic tools. EM may be detected by neutrophil extracellular traps (NETs), which are essential to inflammation and immunological regulation. This research utilized 13 machine learning algorithms to improve the predictive precision of the diagnostic model and pinpointed four potential biomarkers that could aid in the diagnosis of EM.

Methods: Using the GSE141549 dataset, we combined differentially expressed genes with NETs-related markers to identify key genes linked to EM. We performed functional analysis to understand their biological roles. Through 13 machine learning methods, we built 107 different models and selected four central genes: CEACAM1, FOS, PLA2G2A, and THBS1. We developed a diagnostic model, evaluating its performance with ROC curves, calibration plots, and decision curve analysis; its robustness was rigorously assessed through 10-fold cross-validation.

Results: The four-gene model demonstrated superior performance, with high accuracy in the training set (AUC: 0.962), robust generalizability in cross-validation (mean AUC: 0.975) and external cohorts, confirmed clinical utility, and consistent gene expression across multiple datasets.

Conclusion: This study identifies CEACAM1, FOS, PLA2G2A, and THBS1 as promising biomarkers for endometriosis. Their association with immune infiltration may improve early detection strategies, highlighting the potential for developing non-invasive diagnostic tools for EM in the future.

Keywords: endometriosis, neutrophil extracellular traps, machine learning, prediction, diagnosis

Introduction

Endometriosis (EM) is a prevalent gynecological disorder characterized by the atypical growth of endometrial-like tissue outside of the uterine cavity. This condition occurs when tissue similar to the lining of the uterus, known as the endometrium, develops in areas beyond the uterus, which can lead to a variety of symptoms and complications for those affected. Endometriosis is recognized as a significant health concern among individuals with uterine anatomy, often impacting their quality of life and reproductive health.[1,2] Often, EM is linked with infertility, chronic pelvic pain, and various other symptoms.[3,4] Its definite diagnosis is made through invasive laparoscopic surgery. The peak occurrence happens in individuals aged 25–45 years, with a postponement of eight to ten years in obtaining a definitive diagnosis.[5,6] Driven by progress in molecular biology and bioinformatics methods, the amalgamation of multi-omics technologies has facilitated an in-depth analysis of genome, metabolome and transcriptome big data associated with endometriosis (EM),

Received: 27 July 2025
Accepted: 5 November 2025
Published: 19 November 2025

consequently opening up a new channel for EM screening and early diagnostic marker discovery within the context of precision medicine.[7,8]

There is growing evidence indicating that neutrophil extracellular traps (NETs) are significantly involved in the pathogenesis of EM.[9,10] NETs are fibrous structures expelled by neutrophils, composed mainly of DNA and various bactericidal proteins such as histones, granule proteases (NE, MPO), and lactoferrin.[11] In the immune microenvironment of EM, NETs are postulated to promote disease progression by altering the local immune landscape.[12] However, the specific gene expression patterns driven by NETs in EM remain poorly characterized, posing a challenge for diagnosis. To systematically address this, we applied machine learning (ML) to identify NET-related diagnostic biomarkers. ML technologies have shown superior performance in biomedical pattern recognition and data mining,[13,14] particularly when multiple algorithms are integrated to enhance predictive accuracy.[15]

In this study, we employed a comprehensive computational strategy by constructing 107 distinct models from combinations of 13 classical machine learning algorithms. This multi-model approach allowed us to rigorously screen for genes associated with NET-related features in EM. Our analysis identified four genes—CEACAM1, FOS, PLA2G2A, and THBS1—as core diagnostic biomarkers closely linked to NETs. The optimal diagnostic model, which integrated the Stepglm [backward] and Random Forest algorithms, achieved an exceptional AUC of 0.976, demonstrating high discriminative power. Unlike conventional single-model analyses, our multi-algorithm ensemble method mitigates model-specific bias and captures complex biological signals more reliably.[16] Additionally, the study demonstrated the connections between immune cells and CEACAM1, FOS, PLA2G2A, and THBS1, providing new information on the immunological processes behind EM.

Materials and Methods

Data Source

The current study utilized gene expression datasets sourced from the GSE141549, GSE7305, GSE23339, and GSE25628 series, which were retrieved from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). Specifically, the GSE141549 dataset, based on the GPL13376 platform, included 179 endometriosis (EM) cases and 43 controls; GSE7305, based on the GPL570 platform, included 10 EM cases and 10 controls; GSE23339, based on the GPL6102 platform, included 10 EM cases and 9 controls; and GSE25628, based on the GPL571 platform, included 16 EM cases and 6 control samples. All data were derived from Homo sapiens. In this study, GSE141549 was used as the training set, while GSE7305, GSE23339, and GSE25628 served as the validation sets. Furthermore, a review of recent literature led to the selection of 271 genes associated with neutrophil extracellular traps (NETs) for further analysis. The flowchart depicting the study design is shown in Figure 1.

Screening DE-NETRGs

A comprehensive differential analysis employing a Bayesian model was performed using the limma package (version 3.50.0) in R to examine transcriptomic differences between individuals with endometriosis (EM) and healthy control participants. Differentially expressed genes (DEGs) were identified based on two strict criteria: an absolute log-fold change log2(fc) greater than 1.5 in order to generate more differential genes for subsequent analysis, indicating biologically relevant gene expression changes, and an FDR-adjusted P-value <0.05, with correction applied via the Benjamin-Hochberg method to ensure statistical validity. In order to delve deeper into the functional significance of these DEGs, a co-occurrence analysis of gene sets was conducted utilizing the Venn Diagram package (version 1.11), identifying common gene groups between the DEGs and a predefined gene set related to neutrophil extracellular traps (NETs) (n = 271). This approach allowed for the identification of DE-NETRGs that are potentially implicated in the pathological mechanisms of endometriosis.

Functional Enrichment Analysis

The present study employed functional enrichment analysis within the Gene Ontology (GO) framework and conducted pathway analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG) to explore the biological functions and

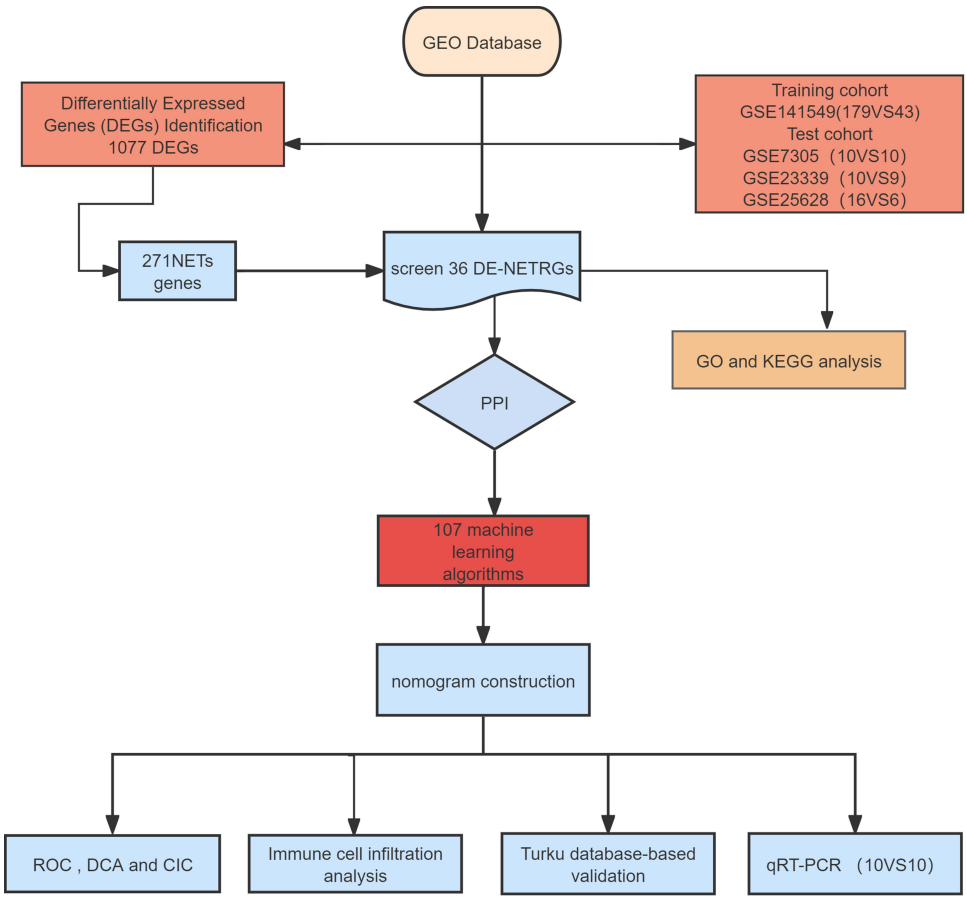

Figure 1 A flowchart consistent with our research findings.

pathways linked to DE-NETRGs. The analysis of gene ontology (GO) was divided into three categories: biological processes (BP), molecular functions (MF), and cellular components (CC). In contrast, the KEGG analysis sought to illuminate the biological significance of genes on both a molecular scale and a larger organizational framework. A hypergeometric test was employed for enrichment analysis, establishing a significance threshold of a P-value under 0.05. The "cluster Profiler" package in R was utilized for annotation and visualization, while the "GO plot" tool was used to present the GO enrichment findings.

Strategy for Optimizing PPI Network Topology Parameters for EM-Related Hub Gene Screening

The STRING database (https://string-db.org) was used as a resource for constructing a network of protein-protein interactions (PPI), illustrating the links among DE-NETRGs. An interaction confidence criterion of 0.400 was established for the analysis. Subsequently, the PPI network was constructed and visualized with Cytoscape software. To identify key genes within the network, the top 10 central genes were selected using the CytoHubba plugin in conjunction with eight distinct topological methods (e.g., degree, eccentricity, closeness), which clarified their essential roles in the network.

Utilization of Machine Learning Techniques for Biomarker Screening

The research employed 13 traditional machine learning algorithms to create an exceptionally effective diagnostic model for endometriosis (EM). These algorithms included least absolute shrinkage and selection operator (Lasso), stepwise generalized linear model (Stepglm), generalized linear model boost (glmBoost), support vector machine (SVM), Ridge regression, elastic net regression (Enet), partial least squares regression (plsRglm), random forest (RF), linear

discriminant analysis (LDA), gradient boosting machine (GBM), Eval algorithm, extreme gradient boosting (XGBoost), and Naive Bayes. This study constructed 107 machine learning feature combination models by integrating the multivariate algorithm strategy system. Based on the area under the curve (AUC) evaluation framework of the receiver operating characteristic feature curve (ROC), the model architecture with the optimal discriminative efficiency was screened to minimize the prediction bias. Finally, by putting together the outputs from 107 combination models, DE-NETRGs, which are differentially expressed neutrophil extracellular trap-related genes, were found to be potentially useful diagnostic biomarkers for EM.

Construction and Validation of the Endometriosis Diagnostic Nomogram

Nomogram plots were created using multivariate logistic regression models, with regression coefficients reflecting the extent to which each predictor influenced the dependent variable. This method creates a line segment plot with a specific scale by combining many influencing factors and assigning a score to each one. The nomogram effectively converts complex multivariate logistic regression models into an understandable graphical representation, increasing the readability of prediction findings and simplifying clinical evaluation. This study used multivariate logistic regression to create a diagnosis model for endometriosis. The model's discriminatory ability was assessed using ROC curve analysis, and the calibration curve was used to establish the compatibility of projected outcomes with actual observations. Furthermore, the model's net benefit advantage within the clinical intervention threshold range was measured using a two-dimensional decision and clinical effect curve.

Assessment of Model Generalization Performance Using K-Fold Cross-Validation

To rigorously evaluate the generalization performance and stability of the diagnostic model, we employed 10-fold cross-validation. The dataset was randomly partitioned into 10 equally sized subsets. The model was trained and validated 10 times, each time using a different subset as the validation set and the remaining 9 subsets as the training set. This process yielded 10 different performance estimates (e.g., 10 AUC values). To determine whether the model's performance was significantly better than a random classifier, a one-sample *t*-test was conducted on these 10 AUC values against the null hypothesis that the mean AUC equals 0.5. The results of the *t*-test provide statistical evidence for the model's robust discriminatory power.

Evaluation of Invasive Immune Cells and Correlation Analysis

The study measured the existence of 22 immune cell types using a complicated algorithm and examined the variability of the immune microenvironment using 1000 Monte Carlo simulations in conjunction with the expression profile of certain genes (Leukocyte signature matrix 22). The LM 22 matrix offers information on gene expression signatures for 22 distinct immune cells. These data have been widely confirmed and used in the analysis of immune cell infiltration, allowing for reliable estimation of the immune cell proportion, simplification of the analysis flow, and improved analysis accuracy. A permutation test validated the biological basis for the cell proportion distribution ($P < 0.05$). A Spearman correlation matrix, which represents the amount of immune infiltration and marker expression ($|r| > 0.3$), was also constructed using R's Corr plot package. The ggplot2 software was used to depict the linear regression trend between important biomarker expression levels (e.g., CEACAM1) and dendritic cell infiltration ($R^2 = 0.42$).

Collection of Human Tissues

The objective of this research is to elucidate the expression patterns of specific DE-NETRGs in endometriosis. This study utilized ectopic endometrial tissues from ten patients with ovarian chocolate cysts (experimental group, n=10) and normal endometrial tissues from ten patients undergoing surgery for uterine fibroids (control group, n=10). Patients with severe systemic diseases or recent hormone therapy were excluded. The study protocol for human tissue collection was approved by the Ethics Committee of the Lin Quan Branch of Anhui Provincial Women's and Children's Medical Center (Ethics No: PJ-KY20240923-01). Informed consent was obtained from all participants prior to tissue collection.

Table 1 Primer Information

Primer Name	Primer Sequences (5′→3′)	Amplification Product (bp)
H-DSG-2-F	TCTTCTAGGCAGGCGCAAAA	112
H-DSG-2-R	TGGTTGGTGGCATAGTGGAC	
h-FOS-F	GGGGCAAGGTGGAACAGTTAT	126
h-FOS-R	CCGCTTGGAGTGTATCAGTCA	
H-THBS1-F	TGCTATCACAACGGAGTTCAGT	108
H-THBS1-R	GCAGGACACCTTTTTGCAGATG	
H-PLA2G2A-F	GAAAGGAAGCCGCACTCAGTT	122
H-PLA2G2A-R	CAGACGTTTGTAGCAACAGTCA	
H-CEACAM1-F	AACCAAAGCGACCCCATCAT	195
H-CEACAM1-R	GTGCTCTGTGAGATCACGCT	
H-GAPDH-F	GGAGCGAGATCCCTCCAAAAT	197
H-GAPDH-R	GGCTGTTGTCATACTTCTCATGG	

Quantitative Real-Time Polymerase Chain Reaction

We collected tissue blocks (~23 mm^3) from 10 endometriosis and 10 control samples. Tissues were homogenized, and total RNA was extracted using 1 mL of Trizol reagent, followed by chloroform phase separation and isopropanol precipitation. The RNA pellet was washed, dried, and dissolved in DEPC-water. cDNA was synthesized from 1 μg of total RNA using the HiScript II Q RT SuperMix. For qPCR, the reaction mixture was prepared with SYBR Master Mix, primers, and cDNA template. The thermal cycling protocol consisted of an initial denaturation at 95°C for 5 min, followed by 40 cycles of 95°C for 15s, 58°C for 15s, and 72°C for 20s, with a final melting curve analysis. Primer sequences are listed in Table 1. Detailed protocols are available in Supplementary File S1.

Validation Using the Turku Database

This research employed the Turku database (https://endometdb.utu.fi), which amalgamates clinical data and gene expression patterns from 115 endometriosis patients and 53 normal control samples. This database comprises gene expression patterns for the gathered samples, facilitating additional evaluation of the endometriosis diagnostic model's reliability and resilience.

Statistical Analysis

The data analysis for this study was conducted using the R program (v4.4.2) platform, utilizing the integrated statistics tools for data processing. Differences among independent samples were assessed through the Wilcoxon rank-sum test tailored for data that do not follow a normal distribution. All statistical inferences employed two-tailed tests, and a significance threshold was set at 0.05.

Results

Screening of DE-NETRGs

This study retrospectively analyzed gene expression data from 179 endometriosis patients and 43 control samples in the GSE 141549 dataset. Firstly, examination of the distribution attributes of endometriosis patients and healthy controls with principal component analysis (PCA) (Figure 2A). The GSE 141549 dataset comprises 179 samples of endometriosis and 43 samples from normal controls, where the control group is represented in red and the endometriosis group is shown in

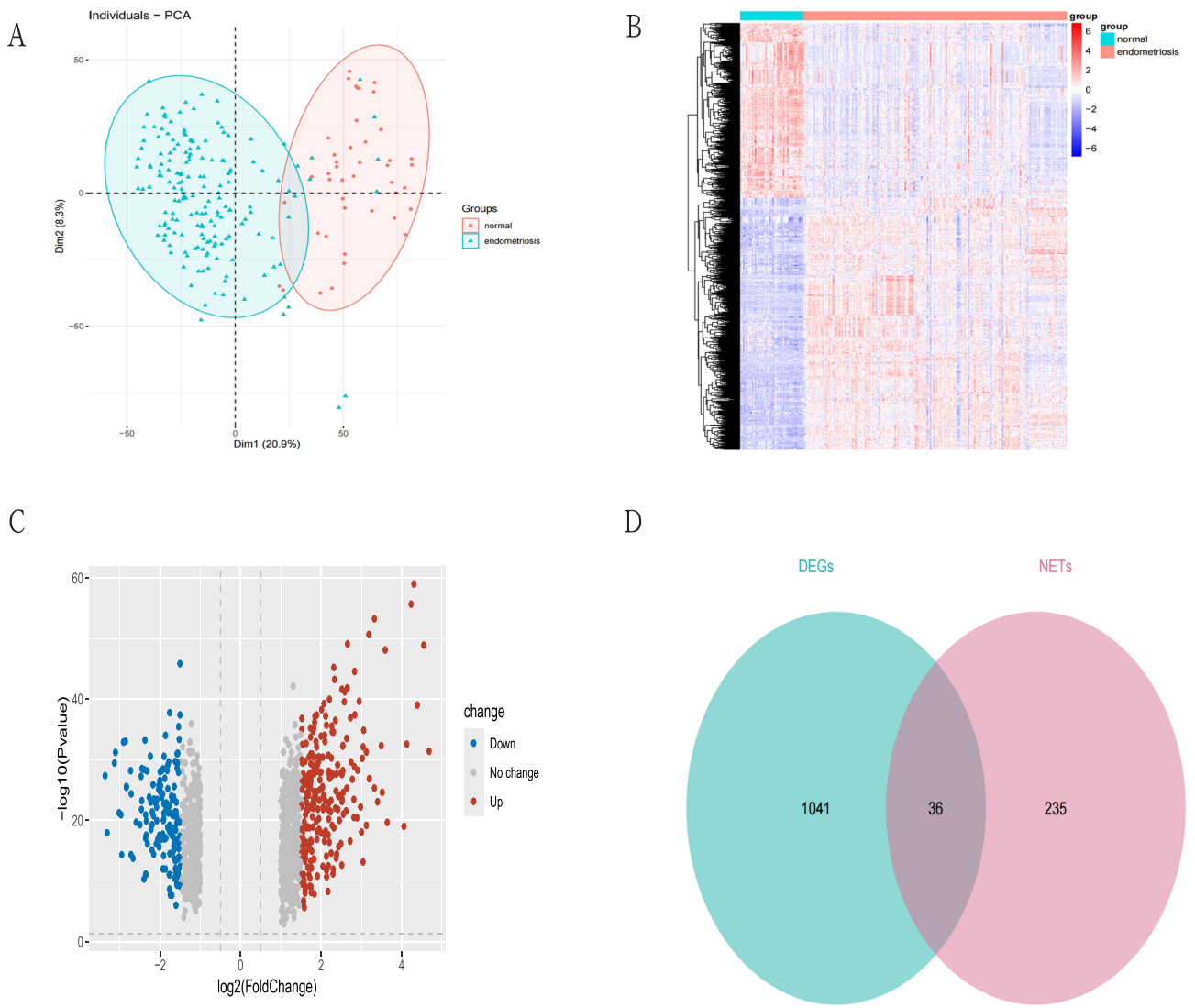

Figure 2 Screening and visualization of the DE-NETRGs. (**A**) Principal component analysis: red represents the normal group, and blue represents the endometriosis group. (**B**) Heatmap of DEG: red indicates high expression and blue indicates low expression. (**C**) Volcano plot: Gray indicates genes not differentially expressed, and red and blue indicate up- and down-regulated DEGs, respectively. (**D**) Venn diagram: 36 DE-NETRGs are shown.

blue. After removing the batch effect, the sample distribution map demonstrated clear separation between the two datasets, thereby enhancing the reliability of the analysis. Subsequently, 1077 DEGs were identified from the training set (GSE 141549) (Figure 2B and C). To examine the interaction between these DEGs and 271 neutrophil extracellular traps (NETs), gene set intersection analysis using a Venn diagram identified 36 common DE-NETRGs (Figure 2D).

Functional Analysis of the DE-NETRGs

Through the GO/KEGG combined enrichment analysis system, the functional characteristics of DE-NETRGs were analyzed and clarified, and the potential biological regulatory network of them was elucidated. The Gene Ontology functional enrichment study indicated that, within the biological process (BP) category, DE-NETRGs exhibited high enrichment in many biological processes, such as wound healing, fluid homeostasis, leukocyte migration, coagulation and hemostasis, and leukocyte-cell adhesion. Cellular component (CC) analysis showed significant enrichment of DE-NETRGs in collagen-rich extracellular matrices, secretory granules, and cytoplasmic vesicles. Regarding molecular function (MF), notable enrichments were observed in binding to glycosaminoglycans, structural elements of the

extracellular matrix, and activities related to receptor-ligand interactions (Figure 3A). Circle plots for enrichment analysis illustrate the top 10 GO terms that are significantly enriched in DE-NETRGs. The concepts encompass the control of bodily fluid concentrations, movement of leukocytes, blood clotting, hemostatic processes, coagulation, mechanisms of defense against bacterial infections, healing of wounds, reaction to lipopolysaccharides, regulation of vascular development, and response to molecules originating from bacteria (Figure 3B).

Furthermore, KEGG pathway analysis also shows that DE-NETRGs are closely linked to a number of immune-related signaling pathways. This includes subjects pertaining to malaria, lipid metabolism and atherosclerosis, Chagas disease, rheumatoid arthritis, the signaling pathway that involves IL-17, the AGE-RAGE pathway linked to complications from diabetes, the PI3K-Akt signaling pathway, the TNF signaling pathway, the formation of neutrophil extracellular traps, along with the interactions between cytokines and their receptors (refer to Figure 3C).

Protein–Protein Interaction (PPI) Network

Utilizing the STRING 11.5 platform, a network of protein-protein interactions was established for the 36 DE-NETRGs. The visualized topological structure was generated using Cytoscape (version 3.10.1), with the results presented in Figure 4A. To further assess the interactions between these 36 genes, we employed the Cytoscape plugin CytoHubba

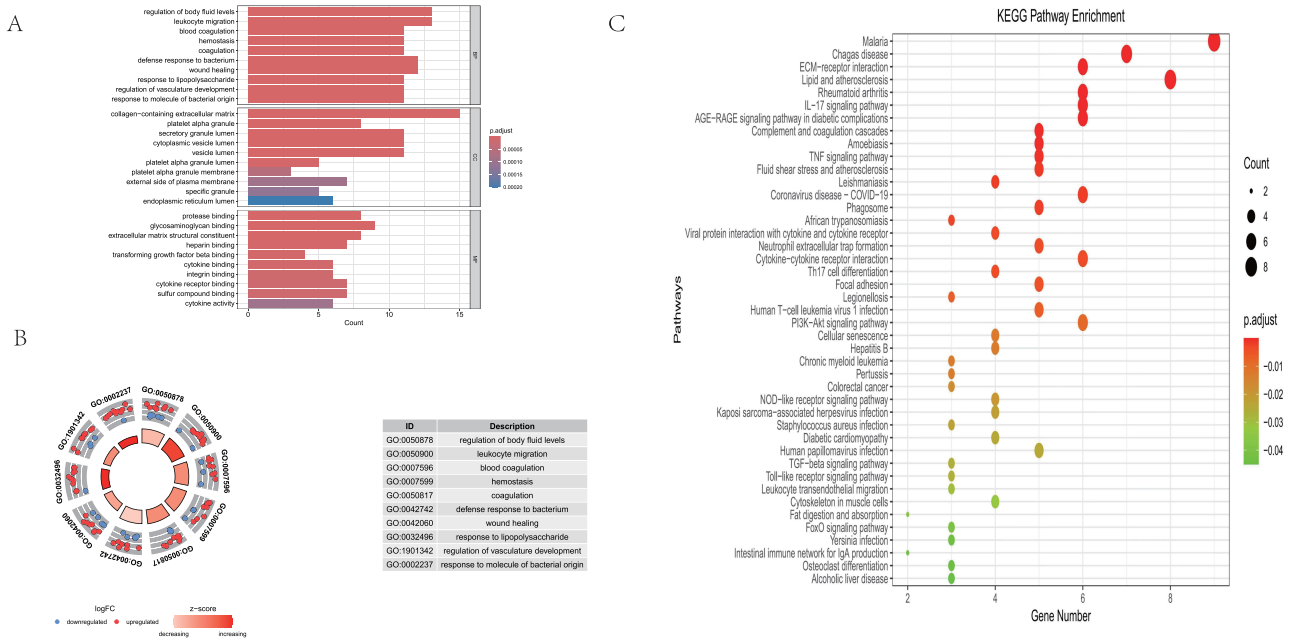

Figure 3 GO& KEGG enrichment analysis of DE-NETRGs. (**A**) GO results of the analysis of DE-NETRGs to show the enrichment results of BP, CC, and MF. (**B**) GO enrichment analysis circle diagram and the top 10 GO enrichment analyses. (**C**) Results of the KEGG enrichment analysis of the DE-NETRGs.

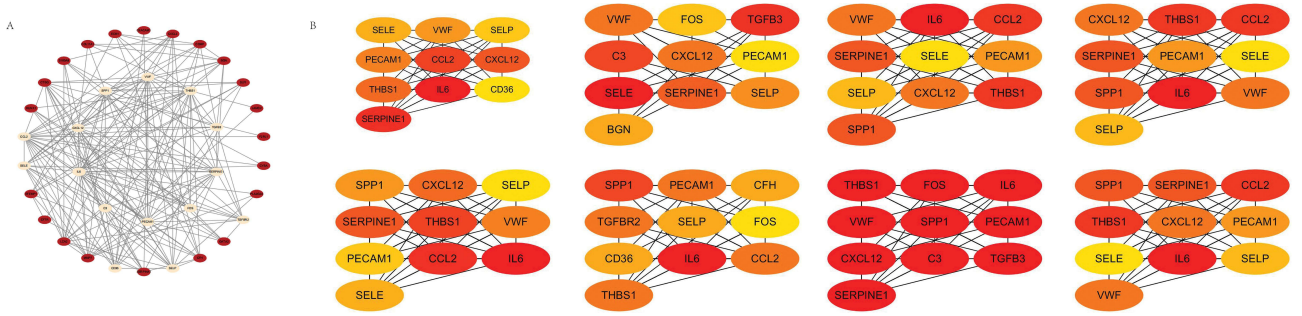

Figure 4 PPI network for DE-NETRGs. (**A**) Cytoscape was used to visualize the PPI network of 36 DE-NETRGs. (**B**) Top ten genes screened using 8 algorithms by Cytohubba plug-in (MCC, DMNC, MNC, Degree, EPC, Bottleneck, Eccentricity, Closeness).

https://doi.org/10.2147/IJWH.S556544

(version 0.1) and applied 8 different algorithms to identify key hub genes. Figure 4B displays the findings of the CytoHubba analysis.

Utilization of 107 Machine Learning Models to Determine Principal Diagnostic Biomarkers

This work applied 13 distinct machine learning algorithms, which were randomly mixed to create 107 machine learning models, utilized to discover significant signature genes from 36 DE-NETRGs linked to endometriosis (EM). The predictive efficacy of each algorithm combination was evaluated using data from the test set and three validation sets, employing the area under the curve (AUC) as the primary metric to quantify the performance of different model combinations. In the screening procedure, the optimal algorithm combination, Stepglm [backward] and RF, produced an AUC score of 0.962. Four potentially important genes—CEACAM1, FOS, PLA2G2A, and THBS1—were identified as putative diagnostic biomarkers for endometriosis based on the results of this combination (Figure 5).

Validation of the Differential Expression Characteristics and Diagnostic Efficacy of Candidate Biomarkers for Endometriosis

The GSE 141549 dataset was employed in this study as the training set to examine the expression variations of key genes between the control group and the endometriosis (EM) group. According to the analysis results (Figure 6A–D), the expression levels of CEACAM1, FOS, PLA2G2A, and THBS1 ($P < 2.22e-16$) were significantly higher in the EM group compared to the control group. Also, the expression patterns of these hub genes were the same in three separate validation datasets (GSE 7305, GSE 23339, and GSE 25628) as they were in the training set (GSE 141549). This shows that gene expression stays the same across datasets.

In order to assess the diagnostic potential of these genes related to endometriosis, we created ROC curves and examined their effectiveness in the designated test datasets (see Figure 7A–D). The ROC analysis results revealed the following: In the GSE 7305 dataset, the AUC values were 1.000 for CEACAM1, 0.800 for FOS, 0.990 for PLA2G2A, and 0.940 for THBS1; in the GSE 23339 dataset, the AUC values were 1.000 for CEACAM1, 0.756 for FOS, 0.900 for PLA2G2A, and 0.889 for THBS1; and in the GSE 25628 dataset, the AUC values were 0.833 for CEACAM1, 0.802 for FOS, 0.771 for PLA2G2A, and 0.885 for THBS1. These AUC values from all validation datasets indicate that these genes exhibit strong predictive power for the diagnosis of endometriosis.

Construction and Validation of the Diagnostic Nomogram

To enhance the clinical applicability of the four DE-NETRG biomarkers, we constructed a nomogram based on the expression data of these biomarkers (Figure 8A). This nomogram is designed to predict the probability of endometriosis (EM) by calculating patient scores. The calibration curve (Figure 8B) was employed to evaluate the model's predictive accuracy, while the ROC curve further corroborated its diagnostic efficacy. The results showed that the AUC values for all model genes surpassed 0.7 in both the training and validation datasets, with the highest AUC reaching 0.976 (Figure 8C). This indicates that the model offers higher accuracy and specificity compared to single-gene models in predicting EM.

Furthermore, we performed decision curve analysis (DCA) and clinical impact curve (CIC) analysis (Figures 9A and B), both of which demonstrated that the diagnostic prediction for EM patients provided greater net benefit and diagnostic accuracy. The nomogram introduced in this research demonstrates remarkable predictive capability and possesses significant clinical value, serving as a useful instrument for the early detection and individualized therapy of endometriosis.

Robustness Assessment and Model Comparison via K-Fold Cross-Validation

To systematically evaluate the robustness and generalizability of the predictive models, we employed 10-fold cross-validation combined with paired t-tests to compare the integrated multi-gene model (Genes Combined) against individual gene models. The results (see in Table 2) demonstrated that the integrated model exhibited superior discriminatory performance and stability, achieving a mean AUC of 0.975 (±0.033), which was significantly higher than all single-gene

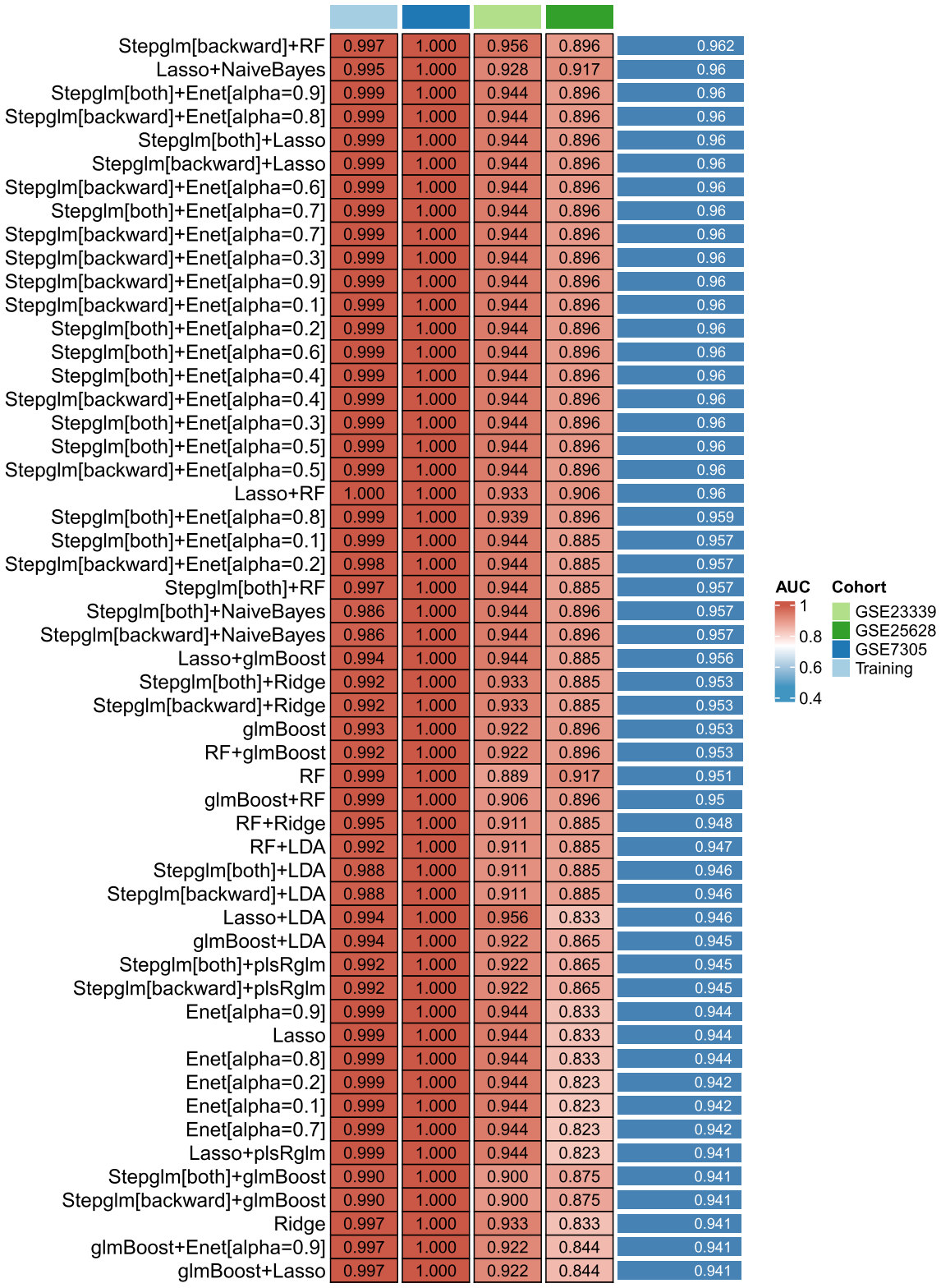

Figure 5 Continued.

https://doi.org/10.2147/IJWH.S556544

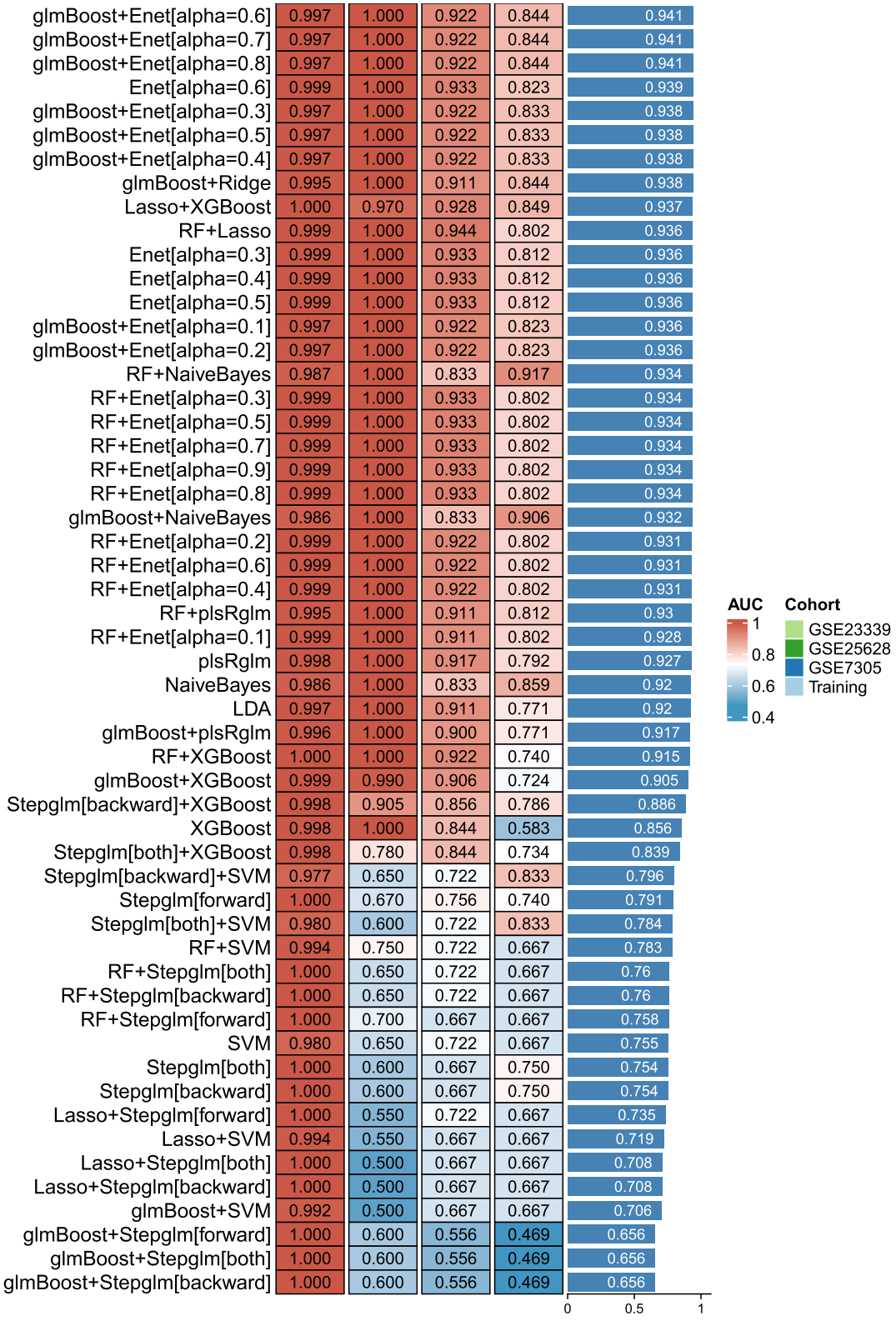

Figure 5 107 models were constructed for 13 machine learning algorithms. Evaluation of 107 models combining 13 algorithms with feature selection methods, assessed by 5-fold cross-validation on training cohort GSE7305. Points indicate mean AUC values for each model across three cohorts (GSE23339, GSE25628, GSE7305). Models are ranked left-to-right by training set AUC. The model Stepglm[backward]+RF demonstrated superior and stable performance (AUC = 0.962) in both training and validation, identifying it as the optimal model for further analysis.

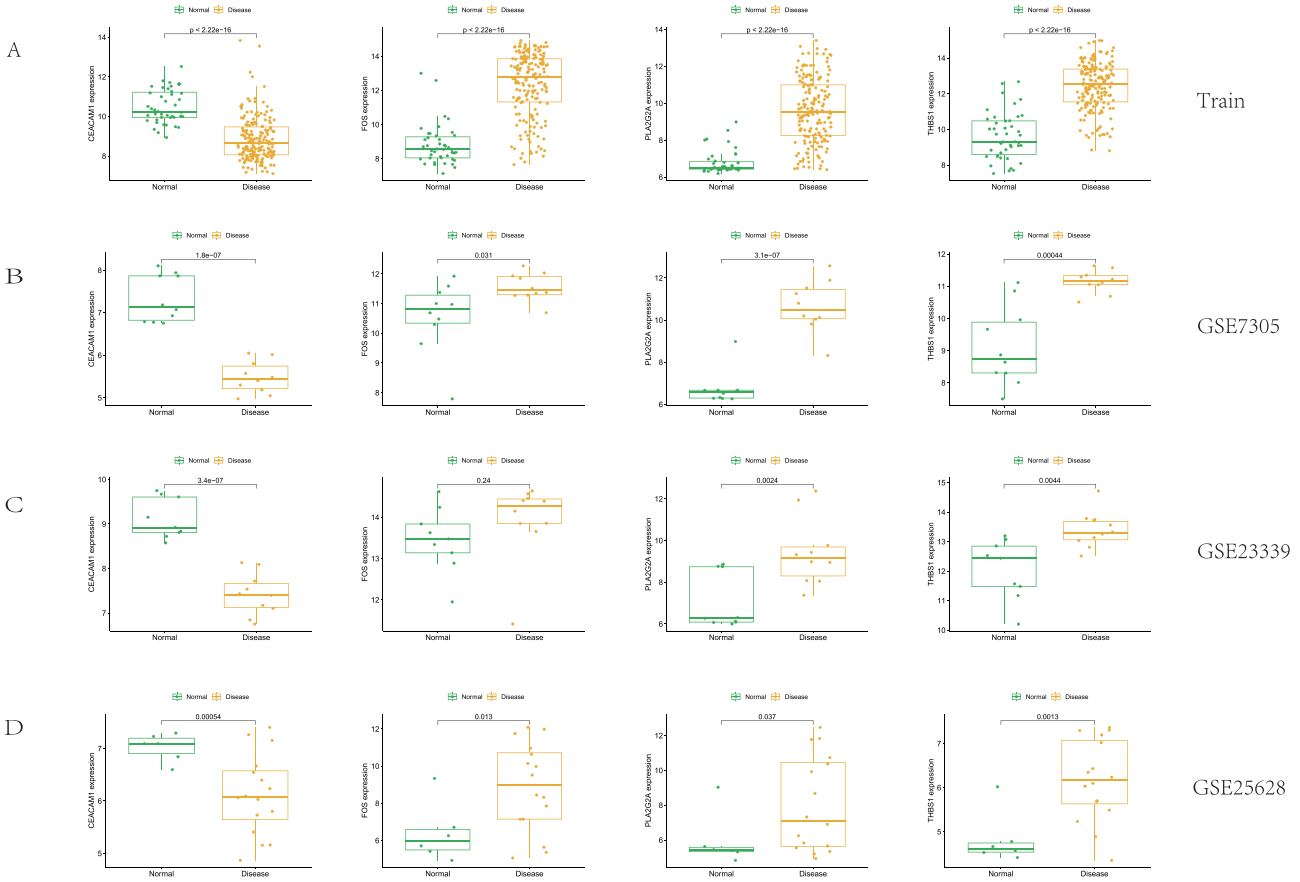

Figure 6 Validation of key biomarker expression across cohorts. (**A–D**) Expression levels of four candidate biomarkers (CEACAM1, FOS, PLA2G2A, THBS1) identified by the optimal model are shown between normal and disease groups in the training cohort (GSE7305) and two independent validation cohorts (GSE23339, GSE25628). Box plots depict the distribution of expression values, and statistical significance was determined by the Wilcoxon test (p-values indicated). The consistent and significant dysregulation of these genes across all cohorts reinforces their biological relevance and supports the robustness of the feature selection process.

models (all $p < 0.05$). The paired *t*-test further confirmed that the integrated model showed statistically significant improvements compared to FOS ($p = 0.015$), THBS1 ($p = 0.049$), and CEACAM1 ($p = 0.001$). Although the difference compared to PLA2G2A ($p = 0.204$) did not reach statistical significance, the observed increase in AUC coupled with greater stability (lower SD) still supports the overall advantage of the integrated model (Figure 9C). These findings indicate that the multi-gene combination strategy significantly enhances the discriminative power and generalizability of the diagnostic model for endometriosis, highlighting its strong potential for clinical translation.

The Relationship Between Diagnostic Biomarkers of DE-NETRGs and the Infiltration of Immune Cells in EM

Immune dysfunction is a key contributor to endometriosis (EM). Using the CIBERSORT algorithm, we characterized the immune cell landscape in EM tissues, which revealed a distinct infiltration pattern compared to controls. Specifically, EM tissues exhibited increased proportions of T follicular helper cells, eosinophils, M1 macrophages, naive B cells, and activated mast cells, but decreased levels of naive CD4 T cells, resting memory CD4 T cells, NK cells (both activated and resting), and monocytes (Figure 10A–D). Furthermore, correlation analysis identified significant associations between the four key hub genes (CEACAM1, FOS, PLA2G2A, THBS1) and specific immune cell subsets (Figure 10E). Notably, CEACAM1 and FOS/THBS1 showed opposing correlation trends with several immune cells (e.g., resting NK cells, activated mast cells), suggesting their potential roles in regulating the immune microenvironment. PLA2G2A was

https://doi.org/10.2147/IJWH.S556544

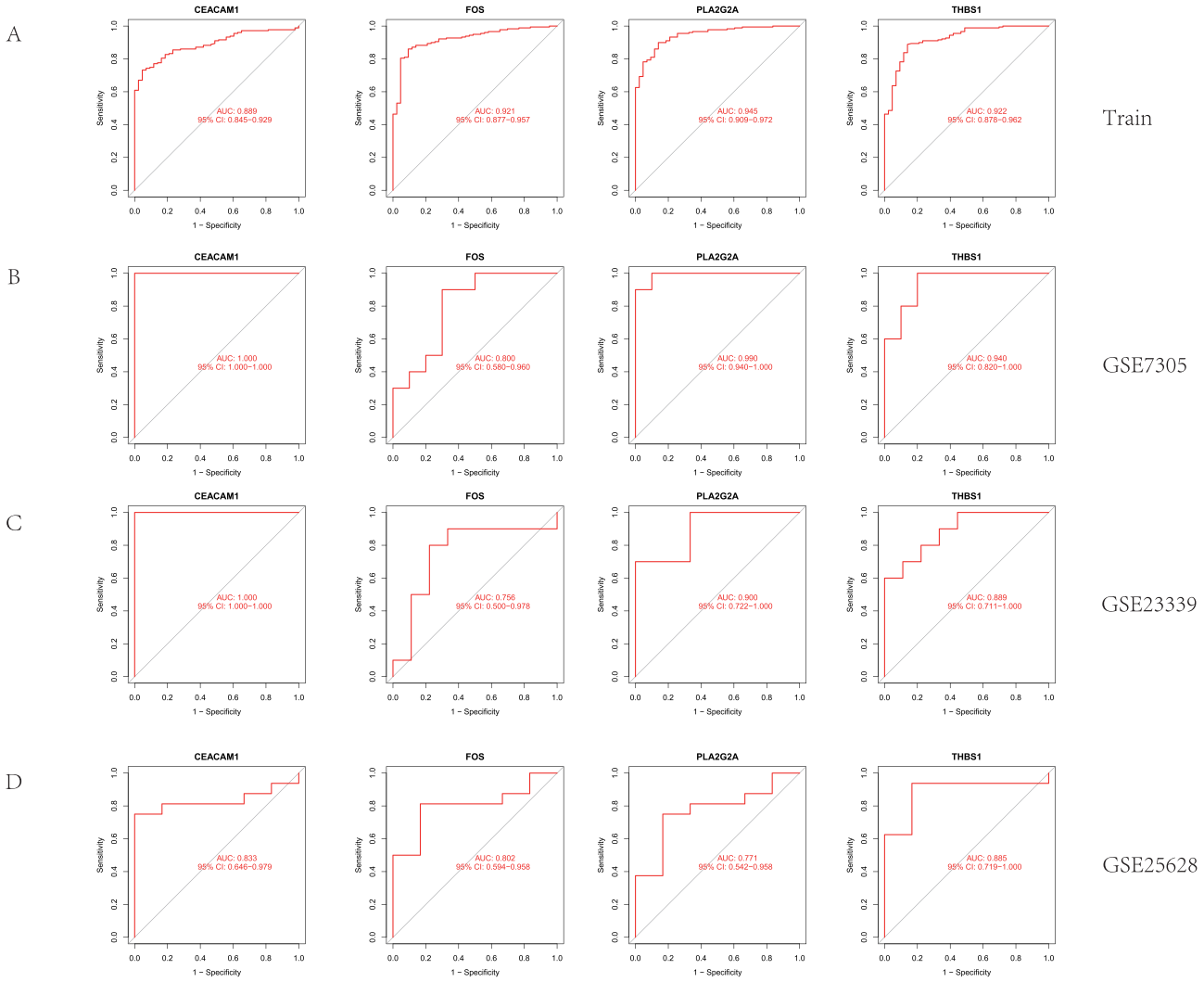

Figure 7 Diagnostic performance of the four candidate biomarkers. (**A**–**D**) Receiver Operating Characteristic (ROC) curves illustrate the ability of CEACAM1, FOS, PLA2G2A, and THBS1 to distinguish between disease and normal samples within the training (GSE7305) and two independent validation cohorts (GSE23339, GSE25628). The Area Under the Curve (AUC) values and their 95% confidence intervals (95% CI) are indicated for each gene in each cohort. All four biomarkers demonstrate robust discriminatory power (AUC > 0.77 across all tests).

positively correlated with naive B cells and activated mast cells. These findings highlight the disrupted immune landscape in EM and point to these hub genes as potential immunomodulatory targets.

Examination of the PCR Verification Outcomes and the Validation of the Turku Online Database

The validation process of the clinical sample-based qPCR method confirmed that the model's core gene expression signature could be used again and again. The expression patterns of CEACAM1 ($p < 0.05$), FOS ($p < 0.01$), PLA2G2A ($p < 0.05$), and THBS 1 ($p < 0.001$) were consistent with the expression patterns in the training dataset (Figure 11A–D). Furthermore, we used the Turku online endometriosis database to assess the confidence of our findings. The expression trajectories of all four genes coincided with the ones seen in the training cohort (Figure 12A–D). Again, CEACAM1 was verified as a protective gene for EM, and the remaining three genes were potential risk biomarkers for EM.

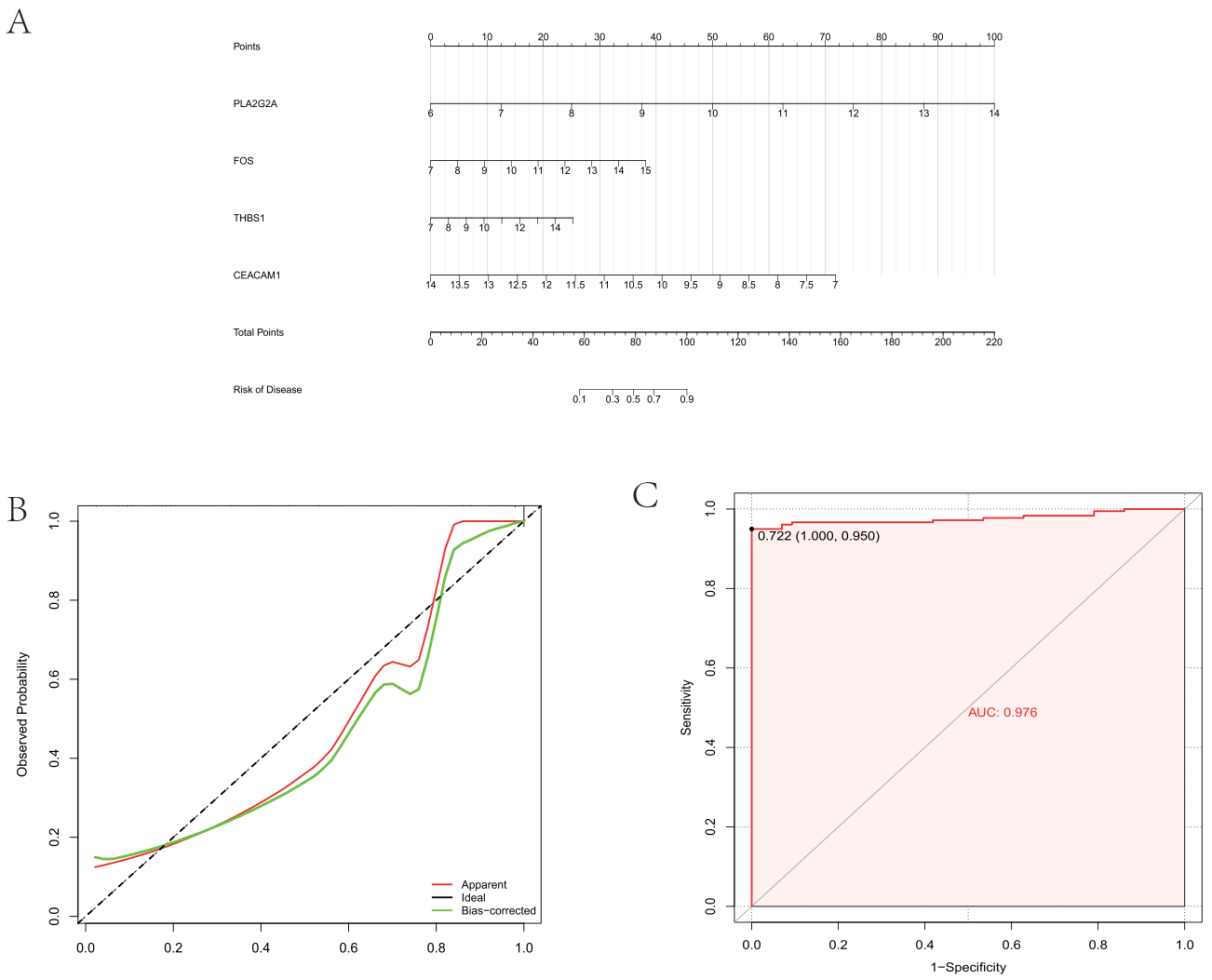

Figure 8 Development and validation of a clinical nomogram based on the four-gene signature. (**A**) The Nomogram for predicting individual disease risk using expression levels of PLA2G2A, FOS, THBS1, and CEACAM1. Points from each gene are summed and projected to estimate disease probability. (**B**) Calibration curve assessing prediction accuracy (n=222). The bias-corrected line (bootstrap repetition B=1000) closely follows the ideal line, indicating excellent agreement between predicted and observed outcomes (mean absolute error = 0.025). (**C**) ROC of this subject was 0.976. The coordinate point (1.000, 0.950) on the curve suggests the existence of an optimal cut-off value, which can achieve a sensitivity of 0.722 when the specificity is 1.000.

Discussion

Endometriosis, as a typical gynecological disorder in reproductive-aged women, has not yet fully clarified its multifactorial pathogenic characteristics and pathological mechanisms. Current clinical diagnostic protocols are varied.[17] Pelvic ultrasound can identify endometriomas,[18] but its diagnostic accuracy is highly operator-dependent. Magnetic resonance imaging (MRI) provides enhanced sensitivity in identifying deep infiltrated endometriosis and rectosigmoid lesions.[19] Although computed tomography (CT) is useful for assessing pleural and ureteral involvement, it is less effective than MRI in visualizing pelvic lesions.[20] While these imaging methods can be helpful, surgery is still the best way to diagnose a problem, with histopathological confirmation of lesions being the gold standard. However, surgery is invasive and may lead to a reduction in ovarian reserve function.[21] Endometriosis has profound implications for fertility, with research indicating that 30% to 40% of patients also experience infertility.[22] Suboptimal medical management may exacerbate fertility challenges and contribute to the development of chronic pelvic pain.[23] The initial phase of endometriosis is strongly linked to an unfavorable reproductive outlook.[24] Recent studies indicate that NETs have a

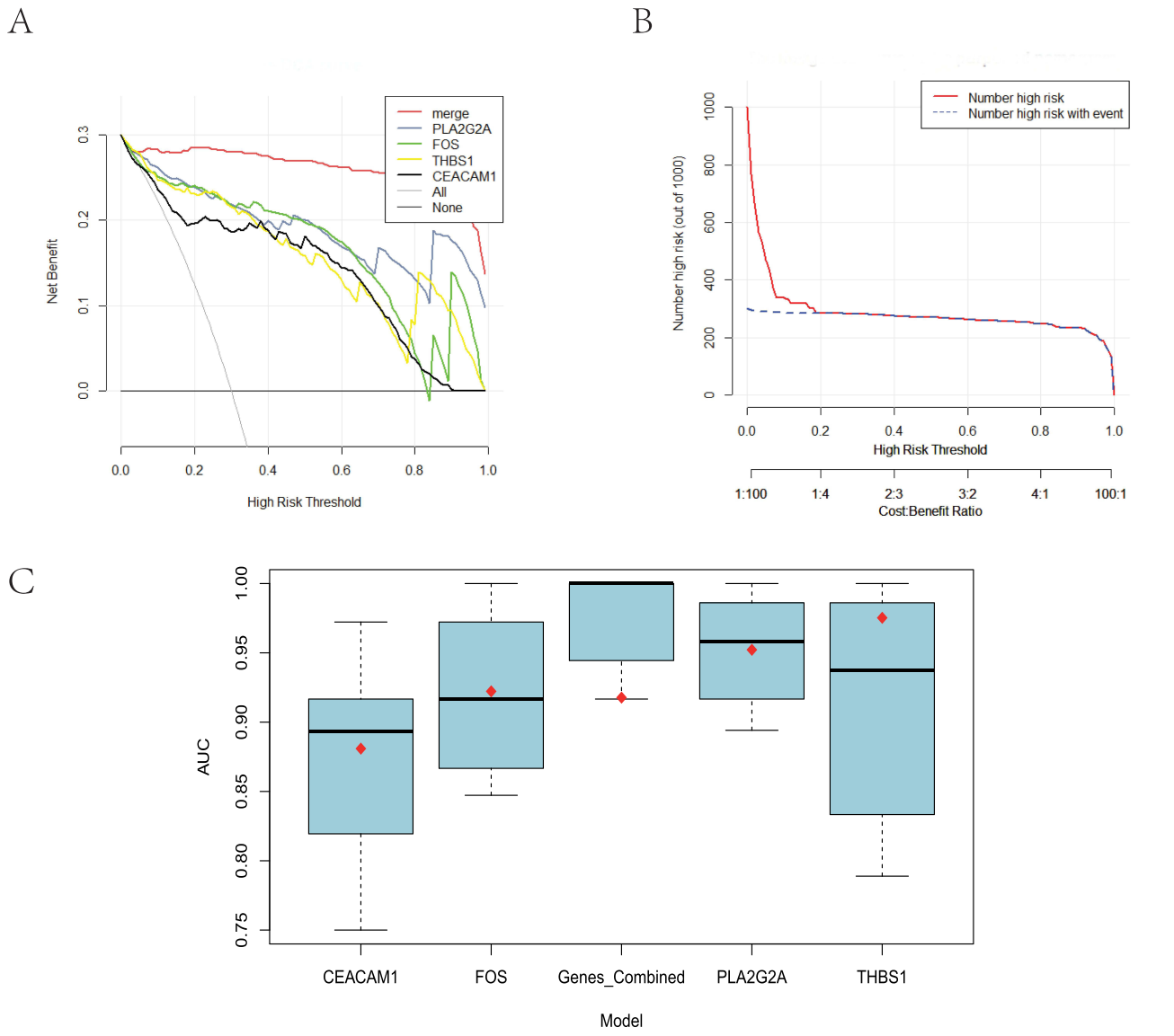

Figure 9 DCA, CIC and Cross-Validation AUC Distribution. (**A**) DCA validation indicated that the composite model of four genes was markedly superior to that of a single gene. (**B**) The CIC curve indicated that the diagnosis based only on the four genes exhibited greater accuracy within the 0.4–0.6 range on the X-axis, whereas the model collaboratively developed from the four genes had enhanced accuracy in the 0.2–0.4 region, thereby greatly augmenting model performance. (**C**) This plot illustrates the distribution of AUC values derived from cross-validation, ranging from 0.75 to 0.95. Evaluated markers and models include: CEACAM1, FOS, PLA2G2A, THBS1, a combined gene set (Genes Combined), and the final model.

significant regulatory function in the pathophysiological processes associated with endometriosis. Since NETs are important parts of innate immunity, when they are not working properly, they can make diseases worse by changing the pathways for inflammation, blood vessel growth, and tissue fibrosis.[25,26]

Table 2 Performance Comparison of the Integrated Gene Model (Genes Combined) and Individual Gene Markers Based on K-Fold Cross-Validation

Model	Mean AUC (±SD)	Mean Accuracy	Mean Sensitivity	Mean Specificity	P-value (vs Genes Combined)
Genes Combined	0.975 (±0.033)	0.937	0.956	0.86	–
PLA2G2A	0.952 (±0.038)	0.896	0.927	0.76	0.204
FOS	0.922 (±0.058)	0.878	0.928	0.68	0.015*
THBS1	0.918 (±0.076)	0.866	0.927	0.605	0.049*
CEACAM1	0.881 (±0.070)	0.847	0.955	0.4	0.001**

Notes: *P < 0.05, **P < 0.01.
Abbreviations: AUC, Area Under the Curve; SD, Standard Deviation.

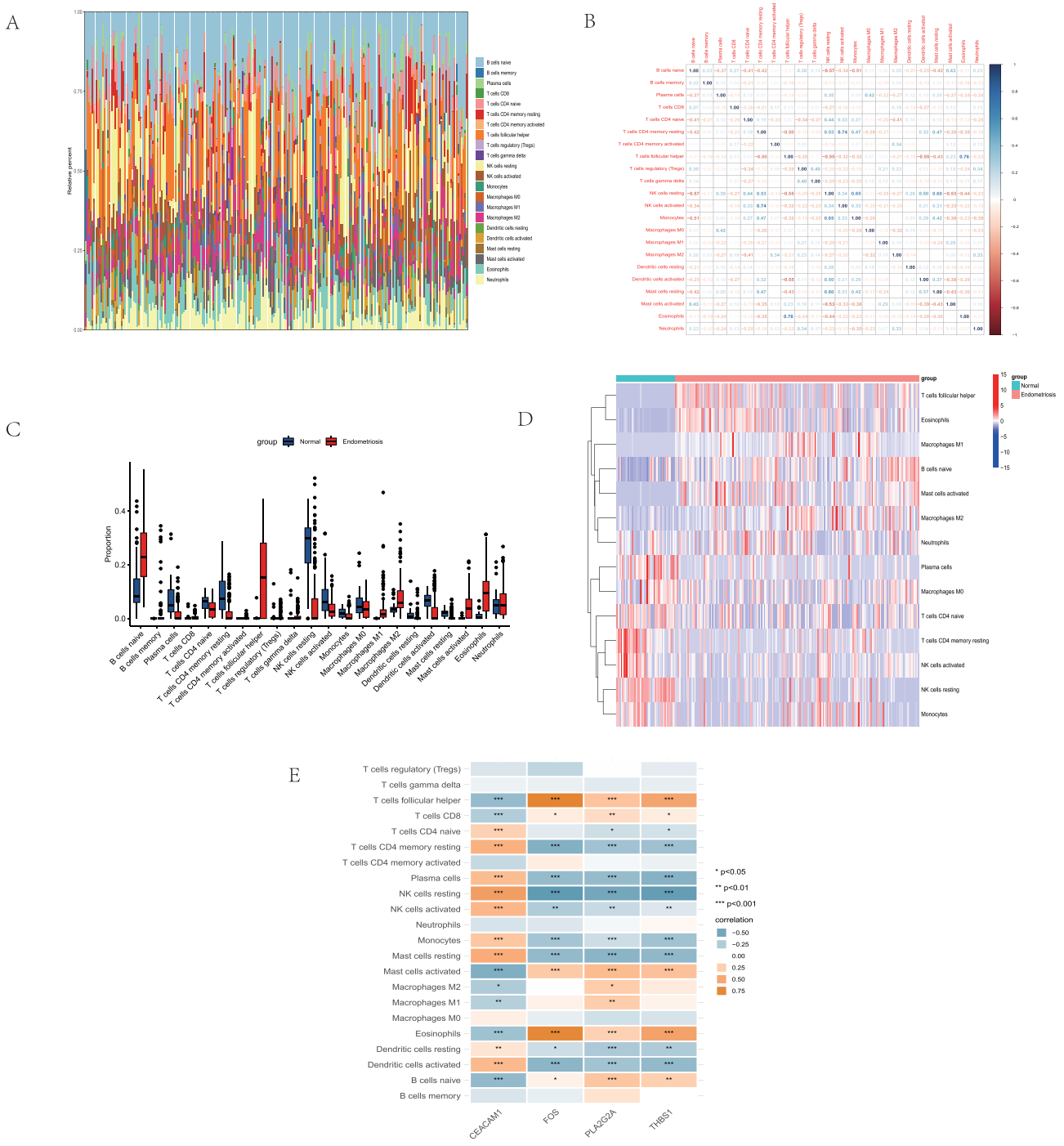

Figure 10 Analysis of the immune cell infiltration. (**A**) Relative proportion of 22 immune cell subtypes were deconvoluted by CIBERSORT from the gene expression signature of 1077 endometriosis-related differentially expressed genes. The microenvironment is characterized by elevated immunosuppressive populations, such as M2 macrophages and regulatory T cells (Tregs). (**B**) Pairwise Pearson correlation coefficients among 22 immune cell subtypes are displayed. The analysis, based on 1077 endometriosis-related differentially expressed genes, uses a color scale (red: positive; blue: negative) to represent correlation strength. (**C** and **D**) Select immune cell subsets (e.g., Tfh cells, M1/M2 macrophages, eosinophils) show differential infiltration levels between conditions, as identified by deconvolution analysis. (**E**) Correlation matrix illustrate the relationships between the expression levels of four candidate genes (CEACAM1, FOS, PLA2G2A, THBS1) and the relative abundance of 22 immune cell subtypes. Pearson correlation coefficients and statistical significance (*p < 0.05, **p < 0.01, ***p < 0.001) are shown. Notably, FOS, PLA2G2A and THBS1 (pro-inflammatory alarmins) show strong positive correlations with T cells follicular helper and Eosinophils, while CEACAM1 is associated with immunosuppressive cells like NK cells resting, suggesting distinct roles in shaping the endometriotic immune microenvironment.

https://doi.org/10.2147/IJWH.S556544

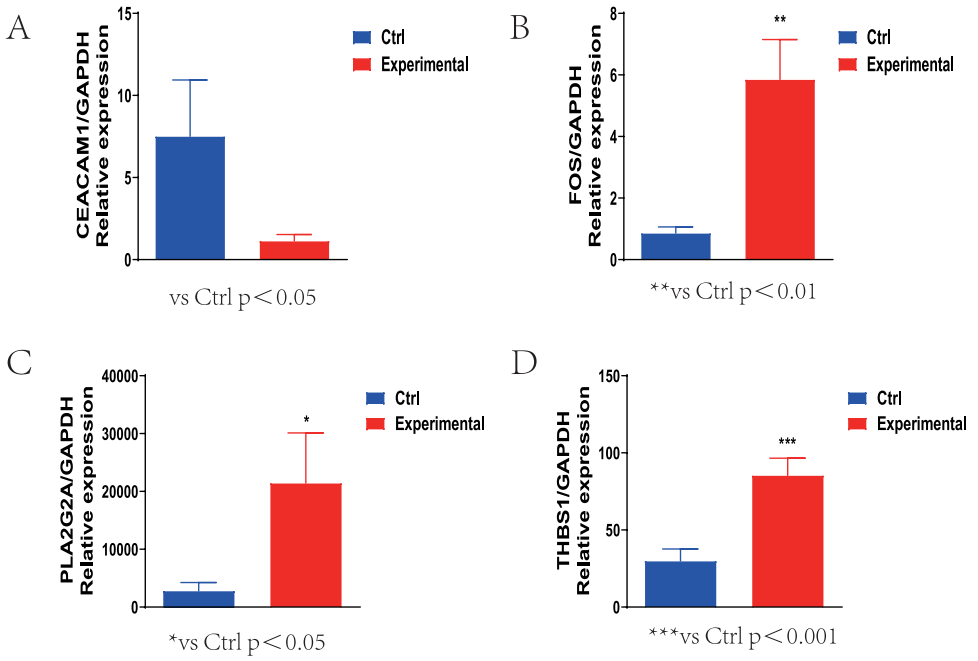

Figure 11 PCR experiments verified the expression levels of the four hub genes. (**A–D**) Distribution of the four hub genes in the control and experimental groups. (*p<0.05: Notable, **p<0.01: Exceptionally notable, ***p<0.001: Unusually prominent).

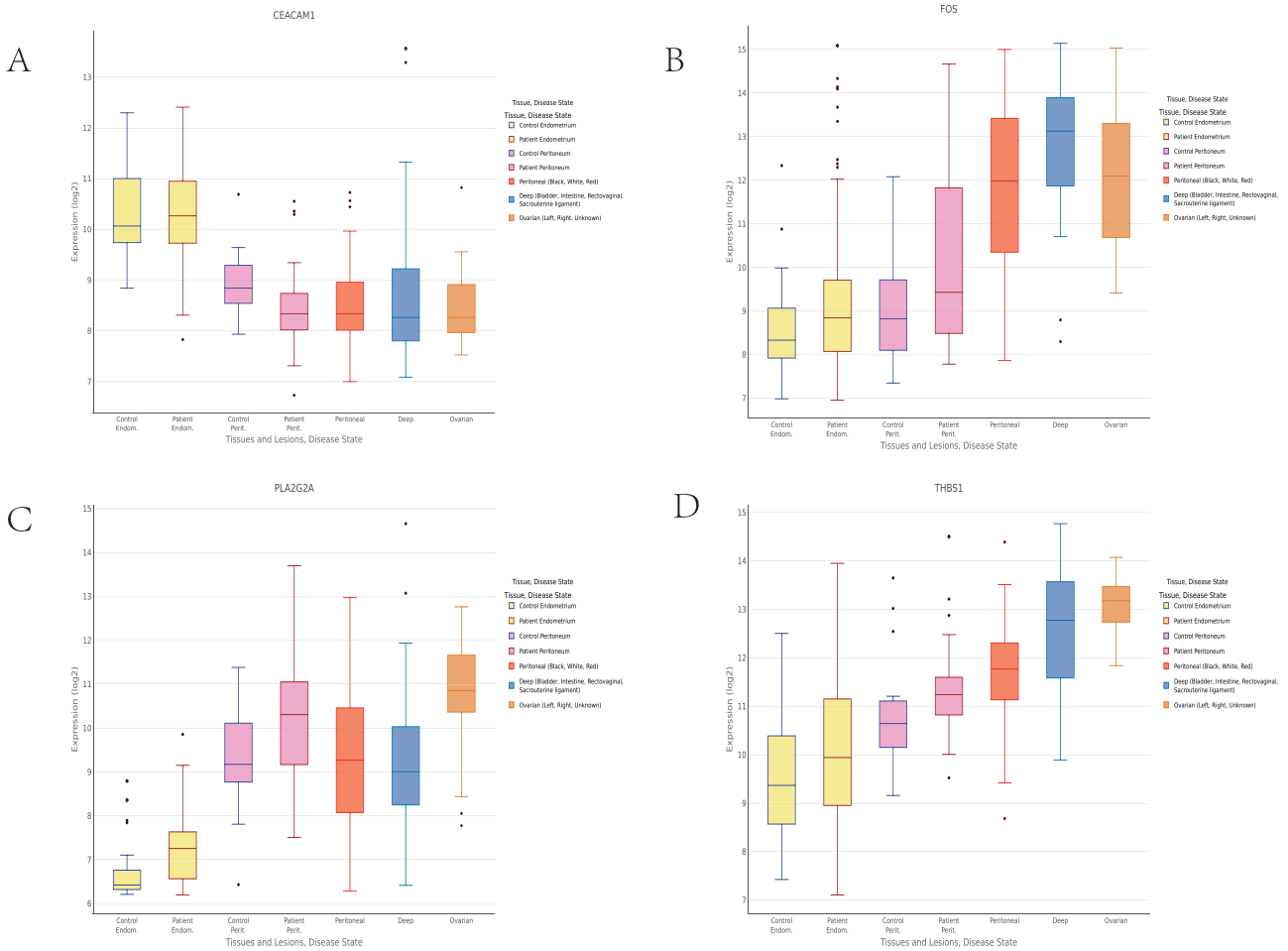

Figure 12 Online endometriosis database Turku was used to evaluate the reliability of our results. (**A–D**) The expression level of four diagnostic biomarkers respectively.

Neutrophil extracellular trap (NET) formation, a process known as NETosis, is a regulated cell death mechanism by which neutrophils release decondensed chromatin and antimicrobial proteins to immobilize and eliminate pathogens.[27,28] While this response is crucial for host defense, NET components can also function as damage-associated molecular patterns (DAMPs). Consequently, excessive or dysregulated NETosis can exacerbate tissue damage by promoting reactive oxygen species (ROS) generation and pro-inflammatory cytokine production via pathways such as TLR2/4/9 and inflammasome activation.[29–31] This dual role links NETs to the pathogenesis of various inflammatory and autoimmune diseases.[32,33] Specifically in endometriosis, recent evidence shows elevated NET levels in both human patients and mouse models.[10] NETs are implicated in promoting angiogenesis, fibrosis, and the survival of ectopic endometrial lesions through mechanisms involving VEGF, TGF-β, and HMGB1-mediated immune cell recruitment.[34,35] Therefore, to elucidate novel pathophysiological insights and therapeutic targets, we conducted a preliminary investigation into NET-related genes in endometriosis.

We utilized datasets GSE141549, GSE7305, GSE23339, and GSE25628 obtained from the GEO database and gathered genes associated with neutrophil extracellular traps (NETs) from recent scientific publications. Functional enrichment analysis indicated that the NETs-related DEGs were primarily enriched in biological processes and pathways; this encompasses the processes of healing wounds, maintaining body fluid balance, the extracellular matrix rich in collagen, structural elements of the extracellular matrix, the functioning of receptor-ligand interactions, the relationship between ECM and receptors, rheumatoid arthritis, as well as the signaling pathway involving IL-17. Neutrophil extracellular traps (NETs) play a vital role in both healing wounds and responding to microbial threats. NETs are crucial for wound healing and antimicrobial responses; however, their excessive formation or inadequate clearance can lead to pathological conditions.[36] Research indicates that NETs (neutrophil extracellular traps) can facilitate gallstone development and the onset of acute pancreatitis, as well as influence the healing process of both early and delayed postoperative wounds. Furthermore, NETs may intensify intravascular coagulation, as well as promote cancer proliferation and metastasis.[37] These processes encompass cellular migration and invasion, significantly contributing to the pathophysiology of EM.[38,39] These findings indicate that NETs may serve as novel diagnostic and therapeutic targets for endometriosis.

Recent advancements in machine learning have led to significant progress in various fields. In this study, we integrated 13 machine learning algorithms to generate 107 distinct models for screening biomarkers associated with endometriosis (EM) and neutrophil extracellular traps (NETs). The evaluation of these algorithms was conducted by determining the area beneath the curve (AUC). The combination yielding the highest AUC value—Stepglm [backward] and RF—was chosen, resulting in the identification of four pivotal genes (CEACAM1, FOS, PLA2G2A, THBS1) as potential candidate biomarkers for constructing a diagnostic model for EM. The model was represented through a nomogram, a graphical instrument that assesses the likelihood of clinical occurrences in an individual based on biological or clinical information. Each variable is allocated a distinct weight, and the final outcome is derived by computing the aggregate score. The model's strong predictive abilities were illustrated by the receiver operating characteristic (ROC) curve, while the calibration curve confirmed the correspondence between the actual data and the predictions made. Furthermore, decision curve analysis (DCA), as introduced by Vickers and Elkin,[40] was utilized to evaluate the clinical value of the prediction model based on threshold probability. The threshold possibility denotes the least likelihood needed to initiate medical intervention, while total benefit represents the proportion of true positives compared to false positives. The clinical applicability of the diagnostic model was further validated by graphically analyzing the net benefit across different threshold values. Moreover, our cross-validation analyses provide robust evidence for the enhanced generalizability and discriminatory power of the multi-gene model, underscoring its potential for more accurate diagnosis of endometriosis.

This study provides pioneering insights into the immune microenvironment of endometriosis by focusing on a novel set of biomarkers—CEACAM1, FOS, PLA2G2A, and THBS1. In contrast to the extensively studied but limited conventional markers like CA-125, miRNAs, HMGB1, NGF, and others,[41] our work shifts the paradigm by delineating how these under-investigated genes interact with specific immune cells, offering a fresh perspective on the disease's immunological mechanisms.

Specifically, we found CEACAM1 to be downregulated in endometriotic tissues, correlating positively with NK cell activity. This supports the notion that impaired immune surveillance, potentially due to low CEACAM1 expression, facilitates the survival of ectopic lesions.[42] Conversely, FOS, PLA2G2A, and THBS1 were upregulated and associated with the activation of T follicular helper cells, mast cells, and eosinophils. This pattern suggests a pro-inflammatory shift. Notably, the connection between FOS and neutrophil-related inflammation,[43] along with PLA2G2A's known role in enhancing NET formation,[44] provides a potential link to NETosis-driven pathology in endometriosis. Furthermore, THBS1's function in angiogenesis and neutrophil aggregation may contribute to the vascularization and chronic inflammation of lesions.[45]

The innovative multi-gene model proposed here, centered on CEACAM1, FOS, PLA2G2A, and THBS1, moves beyond single-marker approaches. By framing these genes as key regulators of a dysregulated immune network, our work not only offers promising biomarkers for early detection but also opens new avenues for developing targeted immunomodulatory therapies for endometriosis.

This research employed a combination of 107 machine learning models, which is an unusual approach in the current literature. Several studies[46,47] typically concentrate on a single machine learning technique, such as support vector machines or random forests, for developing diagnostic models. In this study, the combination of Stepglm [backward] and the RF algorithm was utilized, achieving effective biomarker discovery through the collaborative optimization of statistical methods and machine learning. Initially, Stepglm eliminated unnecessary genes by applying the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC),[48] which contributed to reducing data dimensionality and alleviating multicollinearity issues. Subsequently, the RF algorithm employed decision trees to evaluate gene importance and identify nonlinear interaction relationships. The dual mechanism of out-of-bag error verification in Stepglm's statistical screening and RF significantly enhanced the model's robustness. By applying these methods, a preliminary investigation into the relationship between NETs and EM offers a novel perspective on the pathophysiology of the disease and may facilitate the development of more accurate multi-association early diagnosis techniques.

This study still has certain limitations. Although the model demonstrated good diagnostic efficacy in both the training set and the external validation set, the limited sample size may lead to an overestimation of the model's generalization ability and potential overfitting risk. Moreover, the accessibility of public databases, the small scale of clinical validation cohorts, data heterogeneity (such as differences in the anatomical location of ectopic lesions), and the biological misrepresentations limitations of endometrial tissue validation (lack of multi-dimensional validation such as circulating blood samples, etc.) may all affect the generalization of the results. It is worth noting that the molecular mechanisms of key genes (CEACAM1, FOS, PLA2G2A, THBS1) in EM have not been verified through in vitro and in vivo functional experiments (such as gene knockout/expression models) in this study. Future studies should prioritize three key directions to advance this work. Firstly, large-scale, multicenter collaborations are essential to validate the model's robustness across diverse populations and incorporate a broader range of biological specimens, such as peripheral blood, peritoneal fluid, urine, saliva, and cervical swabs. Secondly, functional studies are needed to decipher the pathological roles of the identified genes. Finally, and most critically for clinical translation, efforts should focus on developing a non-invasive diagnostic strategy based on peripheral blood biomarkers, paving the way for a multimodal tool that integrates NETs-related genes with conventional markers.

Conclusion

We created 107 machine learning models using a mix of 13 machine learning methods. Furthermore, we used three separate datasets from GEO to assess the best models. When used as a diagnostic tool for endometriosis, the biomarker combination of CEACAM1, FOS, PLA2G2A, and THBS1 can increase the diagnostic efficiency. Compared to well-established but limited biomarkers such as CA-125, the association of our panel with immune cell infiltration provides a more disease-specific pathophysiological rationale, which may translate into superior diagnostic specificity—particularly in distinguishing endometriosis from other pelvic pathologies. We propose that the principal clinical value of these biomarkers lies in their integration with established markers like CA-125. The development of a combinatorial diagnostic model that leverages the strengths of both conventional and novel biomarkers represents a critical next step. Such a tool

is anticipated to be more accurate and non-invasive, ultimately enabling earlier intervention and improved patient outcomes.

Abbreviations

EM, Endometriosis; NETs, Neutrophil extracellular traps; GEO, Gene Expression Omnibus; DEGs, differentially expressed genes; DE-NETRGs, differentially expressed neutrophil extracellular trap-related genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, Protein-Protein Interaction Networks; Lasso, Least Absolute Shrinkage and Selection Operator; Stepglm, Stepwise generalized linear model; glmboost, gradient boosting with component-wise linear models; SVM, Support Vector Machine; Enet, Efficient Neural Network; plsRglm, Partial Least Squares Regression; RF, random forest; LDA, Linear Discriminant Analysis; GBM, Gradient Boosting Machine; XGBoost, Extreme Gradient Boosting; ROC, Receiver Operating Characteristic Curve; CIC, clinical impact curve; DCA, decision curve analysis; ML, Machine Learning; BP, biological process; MF, molecular function; CC, cellular component; AUC, Area Under the Curve; SD, Standard Deviation; qPCR, Quantitative Real-Time Polymerase Chain Reaction; PCA, Principal Component Analysis; MRI, Magnetic Resonance Imaging; CT, Computed Tomography; DAMPs, Damage-Associated Molecular Patterns; ROS, Reactive Oxygen Species; AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; Ca-125, Cancer Antigen 125; NK, Natural Killer.

Data Sharing Statement

We conducted this study using public databases, and all of the data from the above studies is available in two databases: GEO (http://www.ncbi.nlm.nih.gov/geo/) and Gene Cards (https://www.genecards.org/).

Ethical Certification

This study adhered to the Helsinki Declaration. The Ethics Committee of the Lin Quan Branch of Anhui Women and Children's Medical Center sanctioned the study protocol (ethics number: PJ-KY20240923-01). Research compliance and academic authority were established via the Chinese Clinical Registration Center (http://www.chictr.org.cn): ChiCTR2500096203, and the National Medical Research Registration and Filing Information System (https://www.medicalresearch.org.cn/login): MR-34-25-000220. All participants were appraised of the study methods and objectives, and the research was done subsequent to their signing of the informed consent form.

Informed Consent

Patients and their families could be actively involved in the trial. With a complete understanding of the treatment, all patients will sign an "informed consent form" and provide informed consent to the trial process. Maximal protection of participants' rights, interests, safety, and health is the basis of a clinical trial protocol.

Acknowledgment

We express our gratitude to the authors for contributing their platforms and invaluable datasets to the GEO and Gene Cards library. Gratitude is extended to all family and patients who contributed to the research.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising, or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by the Clinical Research Project of the Medical and Health Technology Development Research Centre of the National Health Commission (WKZX2024DN0144, WKZX2024GQ0315), the Hefei Municipal Health and

Health Applied Medical Scientific Research Project (Hwk2023zd001), and the Anhui Province Health Research Project Qi Lu Tumor Special Key Project (AHWJ2023BAa10009).

Disclosure

The authors confirm that there is no financial or commercial connection that may be seen as a potential conflict of interest in this work.

References

1. Lamceva J, Uljanovs R, Strumfa I. The main theories on the pathogenesis of endometriosis. *Int J Mol Sci*. 2023;24(5):4254. doi:10.3390/ijms24054254
2. Horne AW, Missmer SA. Pathophysiology, diagnosis, and management of endometriosis. *BMJ*. 2022;379:e070750. doi:10.1136/bmj-2022-070750
3. Küpker W, Aizpurua J, Felberbaum RE, Diedrich K. Endometriose und Infertilität. *Die Gynäkologie*. 2024;57(3):131–137. doi:10.1007/s00129-024-05203-3
4. Helbert S, Morisse M. [Endometriosis and pain]. *Rev Infirm*. 2024;73(299):24–28. doi:10.1016/j.revinf.2024.01.010
5. Smolarz B, Szyłło K, Romanowicz H. Endometriosis: epidemiology, classification, pathogenesis, treatment and genetics (Review of Literature). *Int J Mol Sci*. 2021;22(19):10554. doi:10.3390/ijms221910554
6. Ahn SH, Singh V, Tayade C. Biomarkers in endometriosis: challenges and opportunities. *Fertil Steril*. 2017;107(3):523–532. doi:10.1016/j.fertnstert.2017.01.009
7. Giudice LC. Clinical practice. Endometriosis. *N Engl J Med*. 2010;362(25):2389–2398. doi:10.1056/NEJMcp1000274
8. Pandey S. Metabolomics for the identification of biomarkers in endometriosis. *Arch Gynecol Obstetrics*. 2024;310(6):2823–2827. doi:10.1007/s00404-024-07796-5
9. Munrós J, Tàssies D, Reverter JC, et al. Circulating neutrophil extracellular traps are elevated in patients with deep infiltrating endometriosis. *Reprod Sci*. 2019;26(1):70–76. doi:10.1177/1933719118757682
10. Sun Y, Cai J, Zhang Y, Bao S. A high concentration of neutrophil extracellular traps is observed in humans and mice suffering from endometriosis. *J Reprod Immunol*. 2025;167:104414. doi:10.1016/j.jri.2024.104414
11. Wang Y, Du C, Zhang Y, Zhu L. Composition and function of neutrophil extracellular traps. *Biomolecules*. 2024;14(4):416. doi:10.3390/biom14040416
12. Wilson TR, Peterson KR, Morris SA, Kuhnell D, Kasper S, Burns KA. Neutrophils initiate pro-inflammatory immune responses in early endometriosis lesion development. *JCI Insight*. 2025;10. doi:10.1172/jci.insight.186133
13. Ahsan MM, Luna SA, Siddique Z. Machine-learning-based disease diagnosis: a comprehensive review. *Healthcare*. 2022;10(3):541. doi:10.3390/healthcare10030541
14. Sahriar S, Akther S, Mauya J, et al. Unlocking stroke prediction: harnessing projection-based statistical feature extraction with ML algorithms. *Heliyon*. 2024;10(5):e27411. doi:10.1016/j.heliyon.2024.e27411
15. Mavrogiorgou A, Kiourtis A, Kleftakis S, Mavrogiorgos K, Zafeiropoulos N, Kyriazis D. A catalogue of machine learning algorithms for healthcare risk predictions. *Sensors*. 2022;22(22):8615. doi:10.3390/s22228615
16. Ng S, Masarone S, Watson D, Barnes MR. The benefits and pitfalls of machine learning for biomarker discovery. *Cell Tissue Res*. 2023;394(1):17–31. doi:10.1007/s00441-023-03816-z
17. Allaire C, Bedaiwy MA, Yong PJ. Diagnosis and management of endometriosis. *Cmaj*. 2023;195(10):E363–e371. doi:10.1503/cmaj.220637
18. Matsuda S, Ichikawa M, Kaseki H, et al. Accuracy of transvaginal ultrasonographic diagnosis of retroflexed uterus in endometriosis, with magnetic resonance imaging as reference. *J Nippon Med Sch*. 2023;90(1):26–32. doi:10.1272/jnms.JNMS.2023_90-106
19. Zhang X, Li M, Tang Z, Li X, Song T. Differentiation between endometriosis-associated ovarian cancers and non- endometriosis-associated ovarian cancers based on magnetic resonance imaging. *Br J Radiol*. 2021;94(1125):20201441. doi:10.1259/bjr.20201441
20. Rousset P, Rousset-Jablonski C, Alifano M, Mansuet-Lupo A, Buy JN, Revel MP. Thoracic endometriosis syndrome: CT and MRI features. *Clin Radiol*. 2014;69(3):323–330. doi:10.1016/j.crad.2013.10.014
21. Luna Russo MA, Chalif JN, Falcone T. Clinical management of endometriosis. *Minerva Ginecol*. 2020;72(2):106–118. doi:10.23736/S0026-4784.20.04544-X
22. Sokteang S, Tran C, Ou P, Ouk C, Pirtea P, de Ziegler D. Clinical management of infertility associated with endometriosis. *J Obstet Gynaecol Can*. 2024;46(6):102409. doi:10.1016/j.jogc.2024.102409
23. Pereira A, Herrero-Trujillano M, Vaquero G, et al. Clinical management of chronic pelvic pain in endometriosis unresponsive to conventional therapy. *J Pers Med*. 2022;12(1):101. doi:10.3390/jpm12010101
24. Leone Roberti Maggiore U, Chiappa V, Ceccaroni M, et al. Epidemiology of infertility in women with endometriosis. *Best Pract Res Clin Obstet Gynaecol*. 2024;92:102454. doi:10.1016/j.bpobgyn.2023.102454
25. Wang D, Zhao Y, Zhou Y, Yang S, Xiao X, Feng L. Angiogenesis-an emerging role in organ fibrosis. *Int J Mol Sci*. 2023;24(18):14123. doi:10.3390/ijms241814123
26. Zhang A, Zou X, Yang S, Yang H, Ma Z, Li J. Effect of NETs/COX-2 pathway on immune microenvironment and metastasis in gastric cancer. *Front Immunol*. 2023;14:1177604. doi:10.3389/fimmu.2023.1177604
27. Brinkmann V, Reichard U, Goosmann C, et al. Neutrophil extracellular traps kill bacteria. *Science*. 2004;303(5663):1532–1535. doi:10.1126/science.1092385
28. Ku TH, Ram-Mohan N, Zudock EJ, Abe R, Yang S. Neutrophil extracellular traps have active DNAzymes that promote bactericidal activity. *Nucleic Acids Res*. 2025;53(3). doi:10.1093/nar/gkae1262
29. Block H, Rossaint J, Zarbock A. The fatal circle of NETs and NET-associated DAMPs contributing to organ dysfunction. *Cells*. 2022;11(12):1919. doi:10.3390/cells11121919

30. Raghavan P, Perez CA, Sorrentino TA, Kading JC, Finbloom JA, Desai TA. Physicochemical design of nanoparticles to interface with and degrade neutrophil extracellular traps. *ACS Appl Mater Interfaces*. 2025;17(6):8862–8874. doi:10.1021/acsami.4c17324
31. Morales-Primo AU, Becker I, Zamora-Chimal J. Neutrophil extracellular trap-associated molecules: a review on their immunophysiological and inflammatory roles. *Int Rev Immunol*. 2022;41(2):253–274. doi:10.1080/08830185.2021.1921174
32. Berthelot JM, Le Goff B, Neel A, Maugars Y, Hamidou M. NETosis: at the crossroads of rheumatoid arthritis, lupus, and vasculitis. *Joint Bone Spine*. 2017;84(3):255–262. doi:10.1016/j.jbspin.2016.05.013
33. Petretto A, Bruschi M, Pratesi F, et al. Neutrophil extracellular traps (NET) induced by different stimuli: a comparative proteomic analysis. *PLoS One*. 2019;14(7):e0218946. doi:10.1371/journal.pone.0218946
34. Griffin N, Rowe CW, Gao F, et al. Clinicopathological significance of nerves in esophageal cancer. *Am J Pathol*. 2020;190(9):1921–1930. doi:10.1016/j.ajpath.2020.05.012
35. Giacomini E, Scotti GM, Vanni VS, et al. Global transcriptomic changes occur in uterine fluid-derived extracellular vesicles during the endometrial window for embryo implantation. *Hum Reprod*. 2021;36(8):2249–2274. doi:10.1093/humrep/deab123
36. Hidalgo A, Libby P, Soehnlein O, Aramburu IV, Papayannopoulos V, Silvestre-Roig C. Neutrophil extracellular traps: from physiology to pathology. *Cardiovasc Res*. 2022;118(13):2737–2753. doi:10.1093/cvr/cvab329
37. James P, Kaushal D, Beaumont Wilson R. NETosis in surgery: pathophysiology, prevention, and treatment. *Ann Surg*. 2024;279(5):765–780. doi:10.1097/SLA.0000000000006196
38. Szymański M, Bonowicz K, Antosik P, et al. Role of cyclins and cytoskeletal proteins in endometriosis: insights into pathophysiology. *Cancers*. 2024;16(4):836. doi:10.3390/cancers16040836
39. Kapoor R, Stratopoulou CA, Dolmans MM. Pathogenesis of endometriosis: new insights into prospective therapies. *Int J Mol Sci*. 2021;22 (21):11700. doi:10.3390/ijms222111700
40. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26(6):565–574. doi:10.1177/0272989X06295361
41. Tian Z, Chang XH, Zhao Y, Zhu HL. Current biomarkers for the detection of endometriosis. *Chin Med J*. 2020;133(19):2346–2352. doi:10.1097/CM9.0000000000001063
42. Ahn SH, Khalaj K, Young SL, Lessey BA, Koti M, Tayade C. Immune-inflammation gene signatures in endometriosis patients. *Fertil Steril*. 2016;106(6):1420–1431.e1427. doi:10.1016/j.fertnstert.2016.07.005
43. Hackert NS, Radtke FA, Exner T, et al. Human and mouse neutrophils share core transcriptional programs in both homeostatic and inflamed contexts. *Nat Commun*. 2023;14(1):8133. doi:10.1038/s41467-023-43573-9
44. Iyer A, Lim J, Poudyal H, et al. An inhibitor of phospholipase A2 group IIA modulates adipocyte signaling and protects against diet-induced metabolic syndrome in rats. *Diabetes*. 2012;61(9):2320–2329. doi:10.2337/db11-1179
45. Seif K, Alidzanovic L, Tischler B, et al. Neutrophil-mediated proteolysis of thrombospondin-1 promotes platelet adhesion and string formation. *Thromb Haemost*. 2018;118(12):2074–2085. doi:10.1055/s-0038-1675229
46. Yang B, Zhong J, Yang Y, Xu J, Liu H, Liu J. Machine learning constructs a diagnostic prediction model for calculous pyonephrosis. *Urolithiasis*. 2024;52(1):96. doi:10.1007/s00240-024-01587-y
47. Wu JY, Lin Y, Lin K, Hu YH, Kong GL. [Predicting prolonged length of intensive care unit stay via machine learning]. *Beijing Da Xue Xue Bao Yi Xue Ban*. 2021;53(6):1163–1170. Dutch. doi:10.19723/j.issn.1671-167X.2021.06.026
48. Agegnehu CD, Merid MW, Yenit MK. Incidence and predictors of virological failure among adult HIV patients on first-line antiretroviral therapy in Amhara regional referral hospitals; Ethiopia: a retrospective follow-up study. *BMC Infect Dis*. 2020;20(1):460. doi:10.1186/s12879-020-05177-2