

Metabolomics Analysis-Based Machine Learning for Endometrial Cancer Diagnosis: Integration of Biomarker Discovery and Explainable Artificial Intelligence

✉ Fatma Hilal Yagin,^{1,2} ✉ Abdolvahap Pinar³

¹Department of Biostatistics, Faculty of Medicine, Malatya Turgut Özal University, Malatya, Türkiye

²Department of Computer Science, Lakehead University, Thunder Bay, Canada

³Rectorate Unit, Adıyaman University, Adıyaman, Türkiye

ABSTRACT

Objective: Endometrial cancer (EC) is the most frequent gynecological malignancy in women worldwide. This study aims to develop a predictive model integrating machine learning (ML) approaches with explainable artificial intelligence (XAI) using metabolomics panel data for significant biomarker discovery in EC.

Materials and Methods: This study applied metabolomics and XAI to uncover diagnostic biomarkers for EC, the most common gynecologic malignancy. A total of 191 EC cases and 204 controls were analyzed using mass spectrometry. ML and XAI techniques were incorporated, including SHapley Additive exPlanation, Random Forest, BaggedCART, LightGBM, Adaptive Boosting, and Extreme Gradient Boosting.

Results: Statistically significant differences (adjusted $p < 0.05$) were found in 25 metabolites. Effect sizes (ES) of $m/z=219.125$ (ES=1.516), $m/z=672.6961$ (ES=0.913), and $m/z=203.1564$ (ES=0.839) were notably large, suggesting strong discriminatory ability. These metabolites are involved in lipid dysregulation, steroid hormone pathways, and oxidative stress, reflecting cancer-specific metabolic reprogramming. The ML models, particularly LightGBM, demonstrated high accuracy and good calibration. After training with the final feature dataset, SHapley Additive exPlanations (SHAP) analysis identified $m/z=219.125$, $m/z=672.6961$, and $m/z=127.0769$ as the top contributing features, aligning with their biological impact on EC pathogenesis.

Conclusion: This study suggests non-invasive biomarkers for early detection of EC screening, highlighting the heterogeneity of metabolic adaptation in EC and the need for multi-omics approaches to understand disease mechanisms. Limitations include diverse cohorts and reliance on tandem mass spectrometry. Nonetheless, these findings represent a step forward in precision oncology.

Keywords: Biomarker discovery, endometrial cancer, explainable artificial intelligence (AI), machine learning, metabolomics.



Cite this article as:

Yagin FH, Pinar A. Metabolomics Analysis-Based Machine Learning for Endometrial Cancer Diagnosis: Integration of Biomarker Discovery and Explainable Artificial Intelligence.

JClinPractRes2025;47(5):503–511.

Address for correspondence:

Fatma Hilal Yagin.
Department of Biostatistics,
Malatya Turgut Özal University
Faculty of Medicine, Malatya,
Türkiye

Phone: +90 555 754 53 34

E-mail: hilal.yagin@gmail.com

Submitted: 03.05.2025

Revised: 13.08.2025

Accepted: 26.09.2025

Available Online: 23.10.2025

Erciyes University Faculty of
Medicine Publications -
Available online at www.jcpres.com

INTRODUCTION

Endometrial cancer (EC) is the most common gynecologic malignancy in women worldwide, with an estimated 417,000 new cases and 97,000 deaths per year.¹ The incidence of the disease is



Copyright © Author(s)
This work is licensed under
a Creative Commons
Attribution-NonCommercial
4.0 International License.

increasing, particularly in developed countries, due to rising rates of obesity, physical inactivity, and aging populations. Approximately 40–50% of EC cases result from obesity, through its association with unopposed estrogen levels and potentially insulin resistance.² This trend is concerning, as younger patients often present with aggressive EC subtypes and poor prognosis. Ultrasound-based distinction is challenging; therefore, early and precise diagnosis is critical. The CA-125 biomarker has limited diagnostic utility, with an area under the curve (AUC) of 0.610–0.684.³ This limitation underscores the urgent need for novel diagnostic tools to improve sensitivity and specificity.

Metabolomics investigates tumor metabolic patterns in EC using tools such as the National Metabolomics Data Repository (NMDR) 2024, revealing increased glucose utilization and dysregulation of amino acids and lipids. Lipidomics has identified early-warning metabolites, such as phosphatidylcholines and sphingomyelins, which may indicate endothelial failure before symptoms appear. Studies show that EC cancers consume more branched-chain amino acids and generate excess glycolytic byproducts, thereby altering pathways and enhancing resistance to apoptosis.^{4,5}

Recent investigations have confirmed the promise of metabolomics in EC diagnosis. PELDI-MS (ProteinChip®-based Surface-Enhanced Laser Desorption/Ionization Mass Spectrometry) has identified serum metabolic fingerprints with an AUC of 0.87–0.93, while other studies have reported alterations in glutamine and amino acid metabolism associated with prognosis and treatment outcomes. Machine learning, namely AutoML-XAI, has expanded these discoveries by emphasizing transparent and effective computational models.^{3,6,7} Our study builds directly upon these advancements by applying a suite of robust ML models and XAI to an independent cohort, aiming to both validate and refine the search for the most impactful metabolic biomarkers for EC. This biological plausibility is grounded in the well-characterized metabolic reprogramming that underpins EC pathogenesis. Metabolic disturbances in EC are driven by obesity and unopposed estrogen. Estrogen enhances lipid synthesis, while insulin resistance associated with obesity alters glucose and fatty acid metabolism. Metabolomics may reveal dysfunctions in pathways and biomarkers unique to diseases by detecting lipogenesis, hormone imbalance, and oxidative stress.^{8,9}

Research demonstrated the ability of serum metabolic fingerprints (SMFs), tested with PELDI-MS mass spectrometry, to distinguish EC conditions from non-EC conditions.³ More recent studies show that understanding how EC patients alter sugar and fat metabolism helps doctors develop better personalized diagnoses.¹⁰ The metabolic pathways of

KEY MESSAGES

- Robust metabolic biomarkers for endometrial cancer diagnosis were identified using a machine learning-integrated explainable artificial intelligence model.
- The LightGBM machine learning model offered reliable diagnostic potential.
- SHAP (SHapley Additive exPlanations) analysis supported the diagnostic value of the biomarkers.

diseases have been extensively studied through both liquid chromatography-mass spectrometry (LC-MS) and nuclear magnetic resonance (NMR) instruments, supported by automated data analysis that helps reveal disease markers. By combining powerful testing systems and information tools, researchers can better understand endometrial cancer metabolism and develop better healthcare solutions.¹¹ The integration of powerful analytic tools with XAI methods helps scientists interpret complex data findings and better understand their biomarker results. XAI methodologies, such as SHAP (Shapley Additive Explanations), quantify the contribution of several metabolites to diagnostic predictions. This facilitates the connection between data-driven insights and clinical applicability. XAI methods boost how well we predict and understand information in SMFs to create more accurate metabolic biomarker panels. Studies show that using XAI alongside metabolomics helps doctors understand disease processes and make better treatment choices.^{7,12}

Although previous research provides a foundation, many gaps remain. Many metabolomic models function as “black boxes,” restricting biological understanding, while approaches such as LightGBM and XGBoost are underexplored with EC data. The transition from *m/z* patterns to biological significance is often neglected. To overcome this, we incorporate XAI from the start, using SHAP to quantify metabolite contributions and ensure interpretability. We compare Random Forest, BaggedCART, AdaBoost, XGBoost, and LightGBM to determine the most robust model, with the goal of developing not only a dependable diagnostic tool but also an interpretable framework that emphasizes critical metabolic indicators and promotes clinical validation.

MATERIALS AND METHODS

Dataset, Related Factors, and Ethics Approval

This study used open-access metabolomics data³ from a previous case-control study, with 191 EC participants and 204 non-EC participants. A total of 272 metabolites were included in the analyses. Serum metabolic fingerprints were examined using particle-enhanced laser desorption/ionization mass

spectrometry. Advanced techniques, including liquid chromatography-mass spectrometry and nuclear magnetic resonance, were evaluated to ensure thorough metabolic profiling, yielding reliable and reproducible results. Data were obtained via the National Institutes of Health (NIH) Common Fund's Metabolomics Workbench, Project ID PR001898.³ This study was approved by the İnönü University Health Sciences Non-Interventional Clinical Research Ethics Committee on 03-12-2024 (approval number: 2024/6840). All procedures were conducted in accordance with the ethical standards outlined in the Declaration of Helsinki.

Sample Collection and Preparation

Venipuncture was used to obtain serum samples, which were centrifuged at 2,000 g for 10 minutes, stored at -80°C, and exposed to one freeze-thaw cycle prior to SMF analysis to ensure metabolite stability. The decision was made to conceal any information on the SMF research from pathologists to reduce the likelihood of bias. To carry out the procedure, 1.5 microliters of serum diluted tenfold was placed onto a 384 polished steel plate. After the serum dried, it was prepared for PELDI-MS analysis in accordance with established methods for mass spectrometry-based metabolomics.^{3,5}

Mass Spectrometry Analysis

The mass spectrometry analysis was conducted in positive ion mode using a Bruker Autoflex Speed TOF/TOF (time-of-flight/time-of-flight). To achieve accurate detection, reference metabolites such as alanine, proline, glucose, lactate, and citrate were used for *m/z* calibration. The matrix solutions were prepared by dissolving 10 mg of organic matrices in 1 mL of TA30 (acetonitrile: 0.1% trifluoroacetic acid [TFA], 3:7). Each sample underwent quality control and was examined at 1000 Hz with 2000 laser shots per run to increase consistency. Raw spectra were cleaned of noise and peak deviations before normalization. To maintain consistency, samples were analyzed in technical replicates with signal intensities calibrated against internal standards.^{3,13}

Statistical Analysis

Quantitative variables were summarized using mean±standard deviation (SD) and 95% confidence intervals. The Shapiro-Wilk test was applied to assess normality. For data following a regular distribution, differences between two means were tested for significance (independent samples test). Three basic parameters are proposed for effect sizes: low ($d < 0.5$), medium ($0.5 \leq d < 0.8$), and high ($d \geq 0.8$).¹⁴ False Discovery Rate-adjusted (FDR-adjusted) *p*-values are reported. All hypothesis tests yielded statistically significant results at $p < 0.05$. Statistical analyses were conducted using SPSS 28.0 (IBM Corp., Armonk, NY, United States).

Machine Learning and XAI Methods

The combination of ML and XAI approaches is transforming diagnostic and therapeutic procedures in healthcare. Applications of ML algorithms are numerous and include genetic data for personalized treatment planning, medical imaging data for early disease detection, and the identification of risk factors in electronic health records. However, because these models are so complex, understanding how they make decisions is a challenging task. XAI approaches provide reliable explanations to physicians and patients by clarifying the decision-making processes of ML models. They help clinicians make decisions by indicating which factors indicate the presence of disease or determine treatment response in the analysis of medical data.¹⁵

The goal of this work was to create an interpretable EC prediction model by combining SHAP and machine learning (ML) methods. Because metabolomic datasets are high-dimensional and prone to overfitting, the least absolute shrinkage and selection operator (LASSO) with L1 regularization was employed for variable selection, dimensionality reduction, and biomarker identification. We evaluated ensemble algorithms (AdaBoost, BaggedCART, LightGBM, Random Forest, and XGBoost) using accuracy, F1 score, receiver operating characteristic (ROC) AUC, sensitivity, specificity, and Brier score. A holdout validation using 100 iterations (4:1 train-test split) yielded reliable performance estimates. SHAP analysis evaluated each metabolite's contribution, ensuring that the models were both predictive and interpretable.

Adaptive Boosting (AdaBoost)

The goal of AdaBoost, a classic ensemble learning method, is to improve prediction accuracy through the use of adaptive model combinations. An essential premise of AdaBoost is the repeated training of weak learners, with each assigned a weight according to its performance. AdaBoost increases the weights of previously misclassified instances in each iteration, pushing the algorithm to focus on the most challenging data points. The final prediction is a weighted average of the individual weak learners' outputs, with higher-performing weak learners exerting greater influence.^{15,16}

BaggedCART

Bagging is an ensemble approach that enhances the stability and accuracy of machine learning systems. In this work, the bagged CART method involved building several CART models from distinct subsets of the training data. These subsets were generated by randomly sampling with replacement from the original dataset, a procedure known as bootstrapping. Each

CART model was trained on a bootstrap sample before being integrated into an ensemble model. The final prediction was obtained by averaging the predictions of different CART models, thereby lowering variance and increasing the model's resilience.¹⁷

Light Gradient Boosting Machine (LightGBM)

LightGBM employs a histogram-based technique, in which data is separated into histograms during tree construction to enable more efficient computation of gradient and Hessian values. This method uses substantially less memory and speeds up the training process. LightGBM includes special features such as "Gradient-based One-Side Sampling" and "Exclusive Feature Bundling" to improve training efficiency and effectively handle high-dimensional data. LightGBM is a popular choice for real-time and large-scale machine learning applications because of its ability to handle large datasets efficiently while minimizing memory consumption.¹⁸

Random Forest

Random Forest is a set of decision trees created using a bootstrapped version of the training dataset. Each decision tree in this approach is built via recursive partitioning, which involves continuously applying the same node-splitting procedure, beginning with the root node, until the given stopping criteria are fulfilled. The predictive power of RF derives from the aggregation of many weak learners, and performance improves dramatically when the correlation between trees in the forest is low. This method seeks to create a more general and resilient model by reducing the problem of excessive variation in individual decision trees.^{19,20}

Extreme Gradient Boosting (XGBoost)

Chen and Guestrin created the XGBoost algorithm, an improved gradient boosting approach similar to Gradient Boosting (GB) decision trees and machines. XGBoost generates parallel trees rapidly and reliably, utilizing a distinct regularized boosting technique that distinguishes it from traditional Gradient Boosting. It effectively combines weak learners to improve group performance by integrating Gradient Boosting with advanced techniques. Its adaptability and scalability make it a favored method for enhancing model precision compared to conventional methods.²¹

SHapley Additive exPlanation (SHAP)

SHapley Additive exPlanation developed by Lundberg and Lee, has become a prominent technique for elucidating ML model predictions by employing Game Theory methodologies.²² SHAP explains ML predictions by assigning a SHAP value to each feature, indicating how it contributes to the model's output. It quantifies the effect of individual features using SHAP

values, making it especially helpful for sophisticated black-box models. This makes SHAP an important tool for evaluating model behavior and supporting reliable predictions.²³

RESULTS

This work emphasizes the uniqueness of discovering and contextualizing a metabolite panel with high diagnostic potential for EC. Aside from demonstrating metabolic reprogramming, it enhances previous research by identifying critical diagnostic metabolites within an ML framework and suggesting their functional roles in EC pathophysiology. The alignment of large effect sizes, high SHAP values, and physiologically plausible identities for top metabolites such as $m/z=219.125$, $m/z=672.6961$, and $m/z=127.0769$ highlights their particular relevance. Comparative statistical findings between EC and Non-EC groups are shown in Table 1. For each m/z metabolic value examined, the mean and SD, along with the 95% confidence interval (CI), effect size (ES), and adjusted p-value were also calculated. For most metabolite variables (m/z) in Table 1, statistically significant differences exist between EC and Non-EC groups (adjusted $p<0.05$). In particular, $m/z=109.195$, $m/z=122.0264$, $m/z=127.0769$, $m/z=154.1687$, $m/z=156.1351$, $m/z=169.0587$, $m/z=192.1574$, $m/z=203.1564$, $m/z=219.125$, $m/z=232.941$, $m/z=242.1291$, $m/z=272.9065$, $m/z=285.9427$, $m/z=304.9443$, $m/z=370.9469$, $m/z=467.3484$, $m/z=672.6961$, $m/z=673.7107$, $m/z=854.0047$, $m/z=879.0429$, and $m/z=908.2128$ all showed adjusted $p<0.05$. These results indicate that metabolites differ significantly between EC and Non-EC groups. When the effect sizes related to metabolites are examined in the table, the values obtained (ES=1.516 for $m/z=219.125$, 0.913 for $m/z=672.6961$, 0.839 for $m/z=203.1564$, and 0.808 for $m/z=192.1574$) demonstrate that these metabolites strong explanatory power for differences between the groups. Medium or small ES values are observed for other metabolite variables (Table 1).

Table 2 shows that there are significant differences in prediction performance among the algorithms studied. Overall, LightGBM was the best model, with the highest accuracy (0.888), F1 score (0.888), ROC AUC (0.955), and Brier score (0.0881). This demonstrates its strength in distinguishing between different types of data and calibrating probabilities. XGBoost was not far behind, with 0.886 accuracy, 0.886 F1 score, 0.951 ROC AUC, and a low Brier score of 0.090. It was particularly strong in terms of sensitivity (0.892), achieving the highest value among all models. Random Forest also achieved competitive results. While it showed balanced yet slightly less consistent performance compared to gradient boosting techniques, it still reached accuracy (0.872) and ROC AUC (0.950) values, along with a Brier score of 0.098. With a Brier score of 0.214, AdaBoost was clearly less effective in probability calibration than the other two models (0.885

Table 1. Quantitative metabolite analysis in endometrial cancer (EC) vs. non-EC groups							
Variables	Factor				ES	Status	adj. p*
	EC		Non-EC				
	Mean±SD	95.0% CI	Mean±SD	95.0% CI			
m/z=106.0371	3203.7±1172.5	(5400.9, 4723)	2982.1±1159.2	(4919, 4449.4)	–	–	0.080
m/z=109.195	747.7±415.2	(1685.2, 1235.8)	1090.5±555.5	(2367.3, 1803.8)	0.776	Middle	<0.001
m/z=122.0264	4242.5±2466.4	(10213, 8165.7)	3167±1816.2	(6803, 5692.4)	0.423	Low	0.012
m/z=127.0769	492.9±787.5	(1973.2, 1017.8)	243.2±133.2	(560.5, 403.8)	0.513	Middle	<0.001
m/z=134.0109	1118.9±592.8	(2597.2, 1959.3)	1319.9±740.8	(3494.7, 2140.1)	0.272	Low	0.011
m/z=135.0956	15467.1±8206.6	(30402.8, 26141.7)	17639.5±16856.3	(45244.1, 30609.9)	–	–	0.975
m/z=154.1687	590.1±542.3	(1966.1, 1168.6)	349.4±222.6	(880.7, 560.9)	0.73	Middle	<0.001
m/z=156.1351	324.4±250.8	(981.8, 619.8)	213.9±125.5	(518.2, 376)	0.534	Middle	<0.001
m/z=169.0587	1525.5±1008.3	(3730.2, 2676.2)	1239±918.2	(3559.4, 1981.2)	0.388	Low	<0.001
m/z=184.1607	139.8±88.1	(374, 235.3)	145.9±97	(377.7, 266.5)	–	–	0.940
m/z=186.0822	253.3±142.2	(654.8, 429.6)	258.6±139.5	(556.9, 466.6)	–	–	0.775
m/z=192.1574	869.8±810.4	(2477.1, 1920.5)	476.6±401.4	(1236.4, 810)	0.808	High	<0.001
m/z=203.1564	29840.7±28991.6	(123856.1, 59329.1)	12126.7±11039.4	(39883.6, 25394.3)	0.839	High	<0.001
m/z=219.125	3229.4±2384.8	(9024.7, 6598.3)	1086.5±840.7	(2874.5, 1833.7)	1.516	High	<0.001
m/z=232.941	279.9±209.8	(895, 652.6)	158.2±84.8	(349.1, 275.3)	0.68	Middle	<0.001
m/z=236.1871	81.3±45.5	(168.3, 136.4)	71.7±32	(141.9, 116.2)	–	–	0.055
m/z=242.1291	144.8±100.7	(375.6, 259.6)	101.1±55.9	(234.1, 179.4)	0.56	Middle	<0.001
m/z=261.942	77.9±34.4	(160.1, 126.9)	74.5±33.8	(141.4, 108)	–	–	0.372
m/z=272.9065	249.8±126.7	(518.7, 435.1)	203.1±103.4	(417.8, 340.6)	0.395	Low	<0.001
m/z=285.9427	41.4±17.1	(82.2, 63.9)	32.3±12.2	(55.6, 49.9)	0.584	Middle	<0.001
m/z=304.9443	113±68.9	(302, 210.3)	85.8±43.7	(184.6, 129.6)	0.415	Low	<0.001
m/z=317.2478	41±16.1	(78.3, 60.5)	45.7±24.3	(89, 72.6)	–	–	0.108
m/z=319.2424	38.8±15.9	(76.4, 61)	39.7±13.9	(68.6, 60)	–	–	0.456
m/z=325.3013	80.4±50.8	(198.1, 164.9)	89.5±51.8	(210.5, 160.5)	0.245	Low	0.023
m/z=327.3211	104.8±84.8	(351.8, 214.7)	111.4±67.4	(251.3, 205.1)	0.232	Low	0.031
m/z=329.3345	41.3±22.4	(104.3, 76.3)	47.7±24.1	(114.3, 86.2)	0.303	Low	<0.001
m/z=333.2464	35.5±14.7	(60.9, 49.9)	35.7±12.8	(69.1, 53.5)	–	–	0.940
m/z=370.9469	58.9±61	(269.6, 122.9)	42.1±80.9	(125.1, 68.6)	0.549	Middle	<0.001
m/z=450.4462	3359.2±2313.3	(8193, 7015.8)	3320.2±2230.2	(9207.1, 6720.6)	–	–	0.975
m/z=467.3484	87.4±53.8	(245.8, 169.8)	65±36.5	(150.3, 125.2)	0.429	Middle	<0.001
m/z=672.6961	142±94.1	(367.8, 278.1)	258.7±177.5	(768.7, 505.2)	0.913	High	<0.001
m/z=673.7107	64.2±36.1	(162.6, 122.2)	97±57.3	(236, 170.8)	0.732	Middle	<0.001
m/z=854.0047	81.7±114.7	(554.8, 191.7)	62.2±138.4	(285.7, 131.8)	0.441	Low	<0.001
m/z=879.0429	938.1±1384.8	(4887.7, 2093.4)	547.1±952.4	(2890.8, 1181.2)	0.475	Low	<0.001
m/z=908.2128	67.8±87.1	(342.5, 168.7)	36.1±64.3	(100.4, 68.2)	0.646	Middle	<0.001
*: Significance test of the difference between two means (independent samples test); EC: Endometrial cancer; SD: Standard deviation; CI: Confidence interval; ES: Effect size.							

Table 2. Machine learning model evaluation for gynecologic cancer diagnosis

Model	Accuracy	F1 score	ROC AUC	Sensitivity	Specificity	Brier score
AdaBoost	0.885	0.885	0.943	0.885	0.887	0.214
BaggedCART	0.860	0.861	0.931	0.870	0.850	0.104
LightGBM	0.888	0.888	0.955	0.890	0.886	0.0881
Random Forest	0.872	0.873	0.950	0.881	0.863	0.098
XGBoost	0.886	0.886	0.951	0.892	0.879	0.090

ROC: Receiver operating characteristic; AUC: Area under the curve.

and 0.943). BaggedCART received the lowest scores across all parameters, including accuracy (0.860), F1 score (0.861), and ROC AUC (0.931), showing that it was less predictive than the other ensemble models. In general, XGBoost and Random Forest were good alternatives, while AdaBoost had average results. However, LightGBM always outperformed all other models on important evaluation criteria, making it the best method for using metabolomic biomarkers to diagnose gynecologic cancer (Table 2).

The visualization of the average SHAP values is presented in Figure 1. When the figure is examined, the biomarker $m/z=219.125$ is shown to have the greatest effect on the model output. This indicates that the relevant biomarker has the strongest influence on the model's predictions. The biomarkers $m/z=672.6961$ and $m/z=127.0769$ were also identified as factors with significant effects on the model output. Other biomarkers had comparatively smaller effects (Fig. 1).

SHAP analysis of the XAI model utilizing Metabolomics Data Repository data is shown in Figure 2. The biomarker $m/z=219.125$ had the greatest influence on predictions, suggesting a diagnostic role, followed by $m/z=672.6961$ and $m/z=127.0769$. In contrast, $m/z=106.0371$ had little impact. These findings imply that XAI and metabolomics may help identify the most important biomarkers for gynecologic oncology diagnosis (Fig. 2).

DISCUSSION

This study demonstrates substantial differences between EC and Non-EC metabolic profiles and shows that ML has the capability to utilize these biomarkers as diagnostic indicators. The integration of metabolomics data with XAI not only defines the biological and clinical relevance of important metabolites but also creates an effective diagnostic system for gynecologic oncology. This study identifies significant metabolic breakdowns between EC and Non-EC groups, making metabolomics a promising diagnostic evaluation method in gynecologic oncology. The discriminatory power of these metabolites was particularly strong, as they exhibited large ESs with $m/z=219.125$ (ES=1.516), $m/z=672.6961$ (ES=0.913), and $m/z=203.1564$

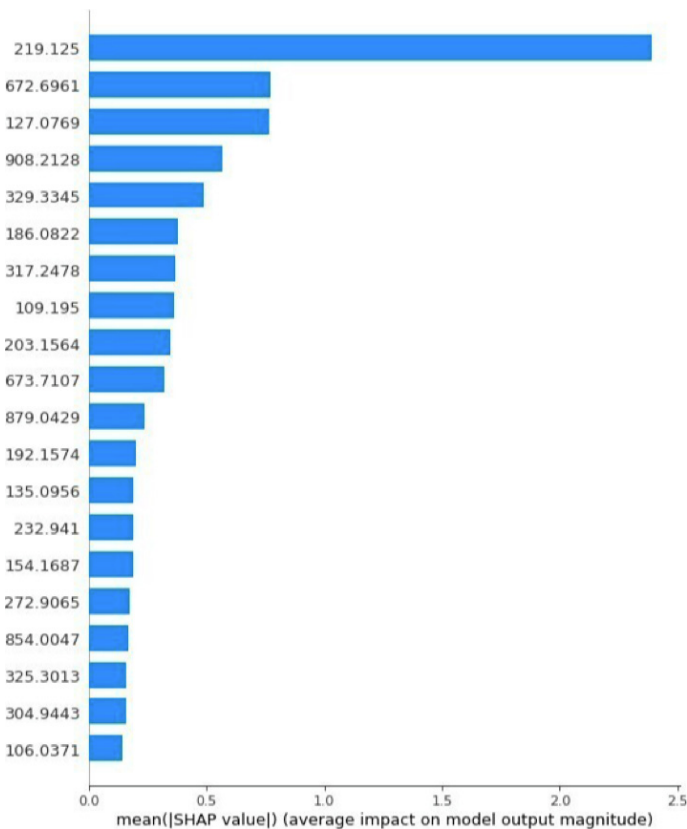


Figure 1. Effects of biomarkers on model output with mean SHAP (SHapley Additive exPlanations) values.

(ES=0.839). The data support existing evidence showing that cancer is characterized by metabolic reprogramming, involving lipids, amino acids, and energy production pathways.²⁴ Elevated $m/z=219.125$ in EC patients indicates dysregulated lipid synthesis that supports membrane biogenesis in rapidly proliferating cells. Disease progression is further associated with $m/z=672.6961$ and $m/z=203.1564$, which are linked to altered steroid hormone pathways and oxidative stress. While high-ES metabolites characterize essential EC processes, signals with moderate or low ES, such as $m/z=122.0264$ (ES=0.423), are likely to reflect patient-specific or secondary metabolic alterations.

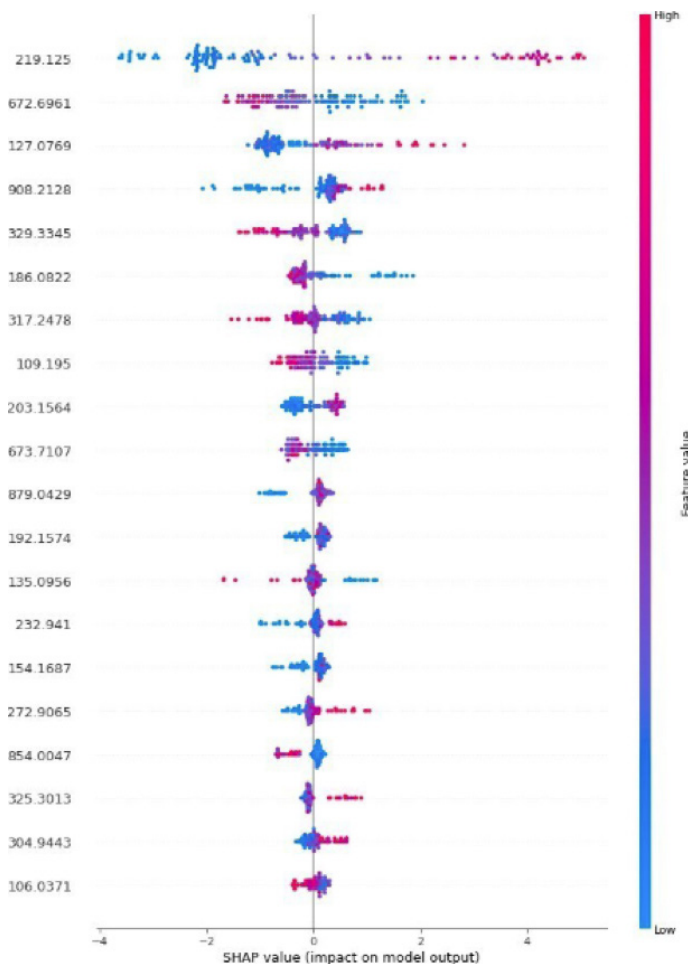


Figure 2. Biomarker importance analysis with SHAP (SHapley Additive exPlanations) values.

These results highlight the complexities of metabolism and the need to use several analytical methods to understand its role in EC pathophysiology.²⁵

The analysis demonstrated extensive variation between EC and Non-EC cohorts in terms of metabolite quantities, especially for metabolites such as m/z=219.125, m/z=672.6961, and m/z=127.0769 (Table 1). The strong discriminative power of metabolites was confirmed by the findings, as they displayed $p < 0.001$ significance with effect sizes reaching 1.516 (e.g., m/z=219.125) and higher. Metabolites with large effect sizes appear to influence EC pathophysiology. The strongest impact was seen at m/z=219.125, most likely related to lipids or lipid-derived compounds involved in membrane formation and signaling. m/z=672.6961 may correspond to nucleotide derivatives or glycosylated chemicals associated with tumor development and metastasis. The variability of metabolomic profiles shows that even minor indicators may improve diagnostic accuracy.^{26–28}

The evaluation of ML models showed that LightGBM outperformed other programs with an accuracy rate of 0.888 and a ROC AUC score of 0.955 (Table 2). Its combination of excellent performance and a minimal Brier score (0.0881) confirms the accuracy of LightGBM for using metabolomics data to predict EC occurrence. LightGBM demonstrates clinical suitability for oncology diagnostics thanks to its high sensitivity (0.890) and high specificity (0.886). Figures 1 and 2 from the SHAP analysis show that m/z=219.125 and m/z=672.6961, together with m/z=127.0769, function as the top-contributing factors for model prediction results. The findings from Table 1 support the biological significance of these metabolites, as they demonstrated statistical significance in the results. The metabolite m/z=127.0769 represents amino acid derivatives associated with tryptophan and kynurenine, which play essential roles in cancer immune evasion processes. The alignment between statistical results and features identified by ML models supports the theory that these metabolites play a critical role in EC pathogenesis.^{29,30} The exceptional results achieved by LightGBM stem from its rapid processing of high-dimensional sparse datasets, in addition to its overfitting prevention through leaf-wise tree building and histogram splitting. Model calibration excellence is indicated by the low Brier score of 0.0881, which makes the model suitable for clinical decision-making applications.³¹

The SHAP analysis delivered essential information about which metabolic factors influence model predictions. The metabolite m/z=219.125 proved to be the most crucial biomarker, with m/z=672.6961 and m/z=127.0769 ranking closely after it. This prioritization aligns with each biomarker's strong effect significance and statistical measurements, which strengthens their prospect of becoming diagnostic markers. The SHAP results also showed that metabolite m/z=106.0371 registered minimal changes during the prediction process, pointing to its insignificant diagnostic value for EC. Explaining model predictions is critical for clinical application, as it converts ML outputs into actionable information. The consistent finding of m/z=219.125 highlights its potential as a significant EC screening and treatment biomarker, with previous research associating values above this range with phospholipids and prostaglandins involved in inflammation and vascular development. Similarly, m/z=672.6961 corresponds to glycosphingolipids that are overexpressed in cancer and may serve as therapeutic targets or monitoring indicators. Integrating XAI with metabolomics improves interpretability, guiding doctors toward the most important metabolites for future investigation. SHAP also revealed that m/z=106.0371 had minimal influence, despite a p-value of 0.06, emphasizing the need to combine statistical and ML approaches for biomarker selection.^{32,33} The detected metabolites suggest the possibility of developing non-invasive

screening techniques for EC, minimizing dependence on biopsies and imaging. In particular, $m/z=219.125$ has the potential to be used as a blood marker to monitor therapy response and reoccurrence. The involvement of lipid and steroid pathways also indicates treatment options, which are validated by preclinical models targeting fatty acid synthase (FASN). Future research should confirm these metabolites in larger cohorts and define their biological significance using platforms such as MetaboAnalyst.^{34,35}

Our findings demonstrated varying alignment between univariate significance (Table 1) and SHAP-based feature importance. This recurrent mismatch in ML biomarker investigations highlights the importance of methodological complementarity: univariate tests detect mean differences, while ML models account for correlations and non-linear effects. For example, $m/z=203.1564$ ($ES=0.839$) may appear less relevant in SHAP if paired with other features, whereas $m/z=127.0769$ scores highly in SHAP, indicating a key role when combined with other metabolites. As a result, ML/XAI goes beyond basic differences to identify feature combinations with the greatest predictive power, providing a more clinically relevant biomarker panel than univariate analysis alone.

Several limitations should be noted. First, the sample size is sufficient for initial analysis but inadequate for broad generalization. Second, the lack of major confounders, including Body Mass Index (BMI), age, menopausal status, and comorbidities, restricts interpretation. Because obesity has a significant impact on the metabolome, the model may capture obesity-related signals rather than cancer-specific ones. Future research should recruit cohorts matched for these variables or use statistical approaches to isolate EC's independent influence. Third, the cross-sectional design makes causal inference impossible; therefore, longitudinal investigations are required to determine whether metabolic alterations occur before or after EC development. The model interpretability enhancement from SHAP analysis also falls short of resolving data biases and confounding variables, including age, BMI, and comorbidities, which should be incorporated into future models.³⁶

CONCLUSION

This research demonstrates how metabolomics can be combined with ML techniques to enhance detection strategies for EC. The identification of essential biomarkers using XAI validation now enables precision oncology to develop accurate tools that are also interpretable in practical use. Extending research into biological functions of the newly discovered metabolites will strengthen the connection between algorithmic detection and medical implementation.

Ethics Committee Approval: The The Inönü University Health Sciences Non-Interventional Clinical Research Ethics Committee granted approval for this study (date:03.12.2024,number: 2024/6840).

Informed Consent: This study used an open-access dataset. As no individual patient data were directly collected, informed consent was not required.

Conflict of Interest: The authors have no conflict of interest to declare.

Financial Disclosure: The authors declared that this study has received no financial support.

Use of AI for Writing Assistance: Not declared.

Author Contributions: Concept – FHY, AP; Design – FHY; Supervision – FHY; Analysis and/or Interpretation – FHY, AP; Literature Search – FHY; Writing – FHY; Critical Reviews – FHY, AP.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Sung H, Ferlay J, Siegel R, Laversanne M, Soerjomataram I, Jemal A. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin.* 2021;71(3):209-49. [\[CrossRef\]](#)
2. Calle E, Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. *Nat Rev Cancer.* 2004;4(8):579-91. [\[CrossRef\]](#)
3. Liu W, Ma J, Zhang J, Cao J, Hu X, Huang Y. Identification and validation of serum metabolite biomarkers for endometrial cancer diagnosis. *EMBO Mol Med.* 2024;16(4):988-1003. [\[CrossRef\]](#)
4. Zhang X, Lu J, Abudukeyoumu A, Hou D, Dong J, Wu J. Glucose transporters: Important regulators of endometrial cancer therapy sensitivity. *Front Oncol.* 2022;12. [\[CrossRef\]](#)
5. Zhai L, Yang X, Cheng Y, Wang J. Glutamine and amino acid metabolism as a prognostic signature and therapeutic target in endometrial cancer. *Cancer Med.* 2023;12:16337-58. [\[CrossRef\]](#)
6. Lou Y, Jiang F, Guan J. The effect of lipidomes on the risk of endometrioid endometrial cancer: a Mendelian randomization study. *Front Oncol.* 2024. [\[CrossRef\]](#)
7. Bifarin O, Fernández F. Automated machine learning and explainable AI (AutoML-XAI) for metabolomics: improving cancer diagnostics. *J Am Soc Mass Spectrom* 2024;35(6):1089-100. [\[CrossRef\]](#)
8. Cazzaniga M, Bonanni B. Relationship between metabolic reprogramming and mitochondrial activity in cancer cells. Understanding the anticancer effect of metformin and its clinical implications. *Anticancer Res.* 2015;35(11):5789-96.

9. Londero A, Parisi N, Tassi A, Bertozzi S, Cagnacci A. Hormone replacement therapy in endometrial cancer survivors: a meta-analysis. *J Clin Med*. 2021;10(14):3165. [\[CrossRef\]](#)
10. Marin A, Vladareanu R, Petca A, Filipescu A. The Importance of Metabolic Factors in Endometrial Cancer: Evaluating the Utility of the Triglyceride-to-Glycemia Index and Triglyceride-to-High-Density Lipoprotein Ratio As Biomarkers. *Cureus*. 2024;16(6). [\[CrossRef\]](#)
11. Ashaq I, Sheikh A, Beri N, Mittal S, Naim M, Alam O. Biomarkers in disease diagnosis and the role of LC-MS and NMR: A Review. *Curr Pharm Des*. 2022.
12. Ghasemi A, Hashtarkhani S, Schwartz D, Shaban-Nejad A. Explainable artificial intelligence in breast cancer detection and risk prediction: A systematic scoping review. *Cancer Innov*. 2024;3. [\[CrossRef\]](#)
13. Zhang N, Chen Q, Zhang P, Zhou K, Liu Y, Wang H. Quartet metabolite reference materials for inter-laboratory proficiency test and data integration of metabolomics profiling. *Genome Biol*. 2024;25(1):34. [\[CrossRef\]](#)
14. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*: Routledge; 2013. [\[CrossRef\]](#)
15. Guldogan E, Yagin F, Pinar A, Colak C, Kadry S, Kim J. A proposed tree-based explainable artificial intelligence approach for the prediction of angina pectoris. *Sci Rep*. 2023;13(1):22189. [\[CrossRef\]](#)
16. Colakovic I, Karakatić S. Adaptive Boosting Method for Mitigating Ethnicity and Age Group Unfairness. *SN Comput Sci*. 2023;5(1):10. [\[CrossRef\]](#)
17. Zhang T, Fu Q, Wang H, Liu F, Wang H, Han L. Bagging-based machine learning algorithms for landslide susceptibility modeling. *Nat Hazard*. 2022;110(2):823–46. [\[CrossRef\]](#)
18. Rufo D, Debelee T, Ibenthal A, Negera W. Diagnosis of diabetes mellitus using gradient boosting machine (LightGBM). *Diagnostics*. 2021;11(9):1714. [\[CrossRef\]](#)
19. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5–32. [\[CrossRef\]](#)
20. Zhang H, Singer B. *Recursive Partitioning and Applications*: Springer Science & Business Media; Berlin; 2010. [\[CrossRef\]](#)
21. Zhou J, Qiu Y, Khandelwal M, Zhu S, Zhang X. Developing a hybrid model of Jaya algorithm-based extreme gradient boosting machine to estimate blast-induced ground vibrations. *Int J Rock Mech Mining Sci*. 2021;145:104856. [\[CrossRef\]](#)
22. Shapley L. A value for n-person games. *Contributions to the Theory of Games*. Princeton: Princeton University Press; 1953. [\[CrossRef\]](#)
23. Van den Broeck G, Lykov A, Schleich M, Suci D. On the tractability of SHAP explanations. *J Artif Intel Res*. 2022;74:851–86. [\[CrossRef\]](#)
24. Pavlova N, Thompson C. The emerging hallmarks of cancer metabolism. *Cell Metab*. 2016;23(1):27–47. [\[CrossRef\]](#)
25. Kvaskoff M, Mahamat-Saleh Y, Farland L, Shigesu N, Terry K, Harris H. Endometriosis and cancer: a systematic review and meta-analysis. *Hum Reprod Update*. 2021;27(2):393–420. [\[CrossRef\]](#)
26. Cheng S, Shah S, Corwin E, Fiehn O, Fitzgerald R, Gerszten R. Potential impact and study considerations of metabolomics in cardiovascular health and disease: a scientific statement from the American Heart Association. *Circulation: Cardiovasc Gen*. 2017;10(2):e000032. [\[CrossRef\]](#)
27. Zhang A, Sun H, Yan G, Wang P, Wang X. Metabolomics for biomarker discovery: moving to the clinic. *BioMed Res Int*. 2015;2015(1):354671. [\[CrossRef\]](#)
28. Fiehn O. Metabolomics by gas chromatography-mass spectrometry: Combined targeted and untargeted profiling. *Curr Protoc Mol Biol*. 2016;114(1):30.4.1–4.2. [\[CrossRef\]](#)
29. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K. A guide to deep learning in healthcare. *Nat Med*. 2019;25(1):24–39. [\[CrossRef\]](#)
30. Platten M, Nollen E, Röhrig U, Fallarino F, Opitz C. Tryptophan metabolism as a common therapeutic target in cancer, neurodegeneration and beyond. *Nat Rev Drug Discov*. 2019;18(5):379–401. [\[CrossRef\]](#)
31. Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Front Neurorobot*. 2013;7:21. [\[CrossRef\]](#)
32. Hilvo M, Denkert C, Lehtinen L, Müller B, Brockmöller S, Seppänen-Laakso T. Novel theranostic opportunities offered by characterization of altered membrane lipid metabolism in breast cancer progression. *Cancer Res*. 2011;71(9):3236–45. [\[CrossRef\]](#)
33. Carubia J, Robert K, Macala L, Kirkwood J, Varga J. Gangliosides of normal and neoplastic human melanocytes. *Biochem Biophys Res Commun*. 1984;120(2):500–4. [\[CrossRef\]](#)
34. Flavin R, Peluso S, Nguyen P, Loda M. Fatty acid synthase as a potential therapeutic target in cancer. *Future Oncol*. 2010;6(4):551–62. [\[CrossRef\]](#)
35. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res*. 2018;46(W1):W486–W94. [\[CrossRef\]](#)
36. Wishart D. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat Rev Drug Disc*. 2016;15(7):473–84. [\[CrossRef\]](#)