Original Article

ANATOLIAN CURRENT MEDICAL

DOI: 10.38053/acmj.1751000

Anatolian Curr Med J. 2025;7(5):687-694

Effectiveness of artificial intelligence algorithms in predicting progression-free survival in epithelial ovarian cancer patients

Aysun Alcı[1], Necim Yalçın[2], Mustafa Gökkaya[3], Gülsüm Ekin Sarı[4], Fatih İkiz[5], Işın Üreyen[4], Tayfun Toptaş[4]

[1]*Department of Gynecologic Oncology, Kahramanmaraş Necip Fazıl City Hospital, Kahramanmaraş, Turkiye*

[2]*Department of Gynecologic Oncology, Batman Training and Research Hospital, Batman, Turkiye*

[3]*Department of Gynecologic Oncology, Gaziantep City Hospital, Gaziantep, Turkiye*

[4]*Department of Gynecologic Oncology, Antalya Training and Research Hospital, University of Health Sciences, Antalya, Turkiye*

[5]*Department of Emergency Medicine, Beyhekim Training and Research Hospital, University of Health Sciences, Konya, Turkiye*

Cite this article as: Alcı A, Yalçın N Gökkaya M, et al. Effectiveness of artificial intelligence algorithms in predicting progression-free survival in epithelial ovarian cancer patients. *Anatolian Curr Med J.* 2025;7(5):687-694.

Received: 25.07.2025 • **Accepted**: 08.09.2025 • **Published**: 15.09.2025

ABSTRACT

Aims: This study aimed to assess the predictive performance of artificial intelligence–based models in estimating progression-free survival (PFS) in patients with epithelial ovarian cancer and to compare various interpretable machine learning approaches.

Methods: Between January 2015 and December 2020, a total of 167 patients who underwent surgical intervention at the Gynaecological Oncology Department of Antalya Training and Research Hospital were retrospectively included in the study if their data were complete. Clinical data were analysed, and the dataset was randomly divided into a training group (n=117; 75%) and a validation group (n=42; 25%). A machine learning (ML) analysis was conducted using the eight most relevant and widely applied algorithmic models for this study design. Model development time, mean absolute error (MAE), root mean square error (RMSE), and correlation coefficient (CC) were evaluated.

Results: Random Forest demonstrated the highest accuracy (MAE=16.45, CC=0.571, RMSE=20.98, time=0.03) and thus became the focus of subsequent analyses. Other algorithms included Linear Regression, Bootstrap Aggregating, Additive Regression, Random Committee, and Regression by Discretization (CC=0.533, 0.492, 0.449, 0.408, and 0.382, respectively). For Random Forest, a moderate correlation was observed between actual and predicted PFS values (CC=0.4–0.6), indicating moderate predictive performance.

Conclusion: The findings of this study demonstrate that machine learning models, particularly Random Forest, can achieve moderate yet clinically relevant prognostic performance based on routinely collected clinical data. In particular, Random Forest demonstrates potential clinical value in guiding patient follow-up strategies and supporting individualized management in ovarian cancer, although further research is required to enhance its clinical validity and applicability.

Keywords: Artificial intelligence, deep learning in gynecologic oncology, epithelial ovarian neoplasms

INTRODUCTION

Ovarian cancer is the eighth most common cancer among women worldwide, accounting for approximately 3.7% of diagnoses and 4.7% of cancer-related deaths, although its incidence varies significantly across regions.[1] The current standard treatment for epithelial ovarian cancer consists of maximal cytoreductive surgery followed by platinum-based chemotherapy, with the possible addition of maintenance therapies such as bevacizumab and/or Poly (ADP-ribose) polymerase (PARP-1) inhibitors.[2,3] Conventional prognostic tools rely on parameters such as tumor stage, histology, patient age, comorbidities, and the extent of cytoreduction; however, these factors often fail to capture the complexity and heterogeneity of the disease.[4] Recent advances in machine learning (ML) have introduced novel approaches for improving prognostic accuracy in oncology. By applying sophisticated analytical methods to large and complex datasets, ML can identify patterns that remain undetected by traditional statistical techniques. The effectiveness of methods such as supervised learning, neural networks, and ensemble approaches has already been demonstrated in various cancer types, highlighting their potential to outperform conventional models in outcome prediction. Consequently, the integration of artificial intelligence (AI) and ML into oncology holds promise for enhancing diagnostic and prognostic accuracy and for enabling more personalized and effective treatment strategies.[5-7]

Corresponding Author: Aysun Alcı, aysun_alci@hotmail.com

With the increasing emphasis on personalized medicine, the need for reliable predictive tools tailored to individual patient characteristics has become more pressing. This is particularly important for heterogeneous diseases such as epithelial ovarian cancer, where traditional statistical models are still difficult to apply in practice. In this context, the aim of our study was to assess the potential of AI based models to provide patient-specific prognostic information, which is critical for the clinical management of these tumors.[8,9] Specifically, we compared the performance of ML methods in predicting progression-free survival (PFS) in patients with epithelial ovarian cancer using preoperative, intraoperative, and postoperative clinical variables. This evaluation seeks to provide a more comprehensive understanding of both the advantages and limitations of AI driven approaches in clinical oncology, and to offer insights that may inform future strategies for clinical practice.

METHODS

Ethics

This retrospective analysis was approved by the Scientific Ethics Committee for Medical Researches at Antalya Training and Research Hospital in Turkiye (Date: 07.11.2024, Decision No: 17/2), was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki.

Study Design and Patients

The data utilized in this study were fully anonymized prior to analysis and contained no personally identifiable information. We reviewed the records of consecutive adult patients (aged 18 years and above) who underwent surgery for epithelial ovarian cancer at our institution between January 2015 and December 2020. Patients with other gynaecological malignancies or with benign histopathological findings were excluded from the analysis. Individuals diagnosed with nonovarian gynaecological malignancies or benign pathological entities were not included in the final analysis.

Data Collection

Demographic and clinical characteristics were recorded for all eligible patients, including age, diabetes mellitus status, Eastern Cooperative Oncology Group (ECOG) performance score, and the presence of significant cardiovascular or pulmonary comorbidities. The extent of intra-abdominal disease was documented in accordance with standard clinical guidelines, and any evidence of extra-abdominal spread on preoperative imaging was noted. In line with evidence-based recommendations from randomized controlled trials,[10,11] patients considered suitable for interval debulking surgery initially received neoadjuvant chemotherapy (NACT). Surgical cytoreduction was performed approximately 21 days after the final NACT cycle. The total number of NACT cycles, typically three, four, or six, was carefully recorded for each patient.

Detailed intraoperative findings were also documented, including the type and extent of surgical procedures performed (e.g., multiorgan resections when applicable), the need for intraoperative blood transfusion, and the degree of cytoreduction achieved, categorized according to residual tumor size after debulking. Postoperative outcomes were assessed by recording intensive care unit (ICU) admissions and length of hospital stay.

Preoperative serum CA-125 levels were obtained for all patients. Pathological and adjuvant treatment data were likewise collected, including International Federation of Gynecology and Obstetrics (FIGO) stage, histological subtype, and the number of lymph nodes resected. Receipt of adjuvant chemotherapy and the number of cycles administered were also recorded. Disease status at the first post-treatment evaluation (e.g., no evidence of disease vs. residual disease) was assessed. Postoperative complications were graded using the Clavien–Dindo classification system.[12] Finally, the interval between debulking surgery and initiation of adjuvant therapy was documented for each patient.

Machine Learning Model Development

The dataset was randomly divided into two subsets: approximately 75% of the patient records were allocated for model training, and the remaining 25% were reserved for testing. To ensure an optimal partitioning strategy, multiple train-test ratios were evaluated (10%, 20%, 25%, 40%, and 50%). Among these, the 25% test set provided the best balance between model development and evaluation, resulting in 117 patients in the training set and 42 in the test set. Model construction including feature selection and algorithm training was performed exclusively on the training dataset, while the test set was retained for independent validation. The distribution of outcome classes (group 1 and group 2) was assessed using z-tests, which confirmed no statistically significant imbalance between subsets. To ensure the stability and generalizability of our models, we evaluated them using a repeated random sub-sampling validation strategy. We performed 100 iterations of partitioning the dataset. In each iteration, the data was randomly split into a training set (75% of patients, n=117) and a test set (25% of patients, n=42). To prevent distributional bias, the splits were stratified to maintain the same proportion of outcome classes (group 1 and group 2) in both the training and test sets as in the original cohort. Missing values were handled internally by the classifier's default method, which distributes instances with unknown values fractionally across the branches of the decision trees based on the observed training data distribution. The feature of importance have defined with Shapley Additive Explanations (SHAP) values, using Python version 3.14 (**Figure 1, 2**).

Eight ML algorithms were applied, selected based on their prevalence in the literature and relevance to the classification task, and implemented using the Waikato Environment for Knowledge Analysis (WEKA), version 3.8.6. Following training, predictive performance was evaluated on the test set using classification accuracy along with additional indicators of predictive strength. To identify the most effective model, performance metrics including mean absolute error (MAE), root mean squared error (RMSE), and Pearson's correlation coefficient (CC) were calculated, while model calibration and fit were examined through statistical comparisons between

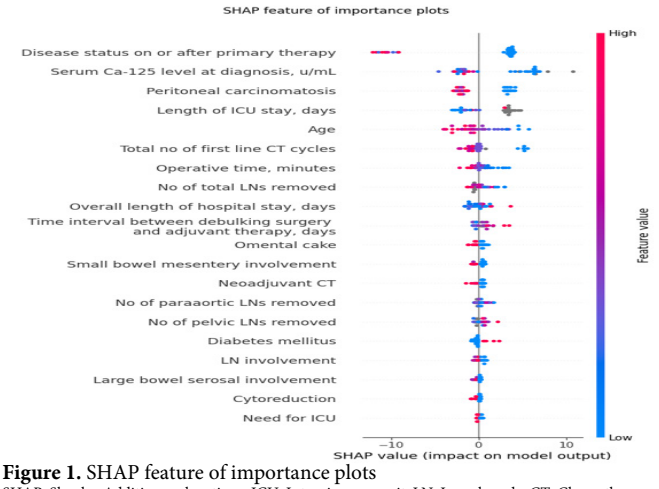

Figure 1. SHAP feature of importance plots
SHAP: Shapley Additive explanations, ICU: Intensive care unit, LN: Lymph node, CT: Chemotherapy

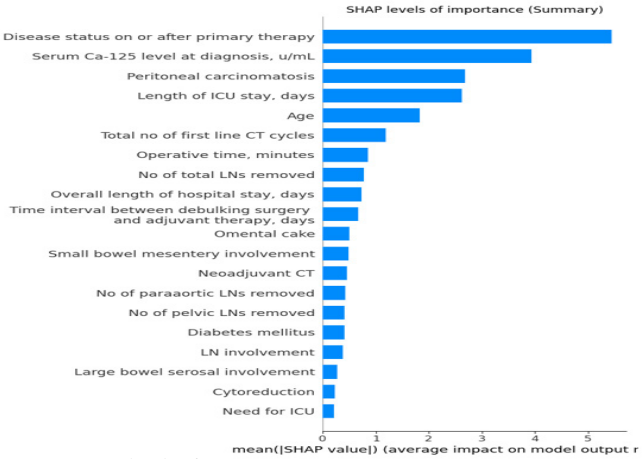

Figure 2. SHAP levels of importance
SHAP: Shapley Additive explanations, ICU: Intensive care unit, LN: Lymph node, CT: Chemotherapy

predicted and observed results in the test cohort. Random Forest, the best-performing algorithm, was carried out with 100 trees and a tree depth value of 10.

Statistical Analysis

All supplementary statistical analyses were performed using IBM SPSS Statistics (version 27.0; IBM Corp., Chicago, IL, USA), underscoring the use of this software for rigorous data analysis. The normality of continuous data distributions was evaluated with the Kolmogorov–Smirnov test, supplemented by visual inspection methods such as histograms, Q–Q plots, and examination of skewness and kurtosis. Categorical variables were presented as frequencies with corresponding percentages for clarity. Pearson's CC was employed to examine associations between continuous variables. To evaluate the concordance between predicted and observed values of the primary outcome, a paired-samples t-test was performed. All ML computations were executed on a Windows 11 system equipped with an Intel Core i7 CPU, 16 GB RAM, and an NVIDIA GeForce GTX 1660 Ti graphics card (8 GB of memory). A two-tailed significance level of 5% (α=0.05) was applied to all statistical tests, and p-values below this threshold were considered statistically significant.

RESULTS

Figure 3 shows the inclusion of 167 patients who underwent surgery for epithelial ovarian carcinoma during the study period. As presented in Table 1, 2, the mean patient age was 58±11 years. Notably, 36.5% of the cohort (n=61) had stage III disease, while 74.9% (n=125) exhibited high-grade tumor histology. A total of eight widely used ML algorithms were tested, and their predictive performance was evaluated. The algorithms utilised in this study were Random Forest, Multilayer Perceptron, Linear Regression, Support Vector Regression, Additive Regression, Bootstrap Aggregating (bagging), Random Committee and Regression by Discretisation. The Multilayer Perceptron and Support Vector Regression algorithms exhibited the lowest performance in terms of CC (0.1543, 0.1997). Following a rigorous evaluation process, the Random Forest algorithm was identified as the most effective algorithm and thus became the focus of subsequent research. The CC of the Random Forest algorithm was 0.5731, with a MAE of 16.45 and a RMSE of 20.98. The time required to create the model was 0.03 seconds. The remaining algorithms were Linear Regression, Bootstrap Aggregating (bagging), Additive Regression, Random Committee, and Regression by Discretization (CCs: 0.5326, 0.4915, 0.4491, 0.4077, 0.3817) (Table 3). A statistical analysis of actual and predicted PFS was performed to determine the success rate of the best performing Random Forest algorithm. A moderately significant correlation was found between actual and predicted PFS ($p<0.001$ and CC=0.573). In addition, an analysis of the difference between the actual and predicted PFS values was performed and no statistically significant difference was found (the difference between the actual and predicted values was very small and the p-value was greater than 0.946) (Table 4, Figure 3). This study highlights that the Random Forest algorithm provides the highest prediction accuracy compared to the other models tested. The moderate correlation observed between actual and PFS values (CC=0.573, $p<0.001$), despite the model demonstrating meaningful predictive capacity, indicates that its performance is not yet optimal for clinical application. Importantly, the absence of statistically significant differences between actual and predicted values (p=0.946) further supports the model's validity. These findings suggest that Random Forest could serve as a promising foundation for clinical prognostic tools in epithelial ovarian cancer. However, future studies with larger cohorts, more diverse datasets, and refined parameter optimisation are necessary to improve prediction accuracy and clinical applicability.

The Bland–Altman analysis demonstrated an overall acceptable agreement between the actual and predicted PFS values. The mean difference was close to zero, indicating the absence of a systematic bias in the predictions. Most of the data points were within the 95% limits of agreement, reflecting a generally reliable concordance between the model outputs and observed outcomes. However, a wider spread of differences was noted at higher mean PFS values, suggesting reduced reliability of the model in patients with longer

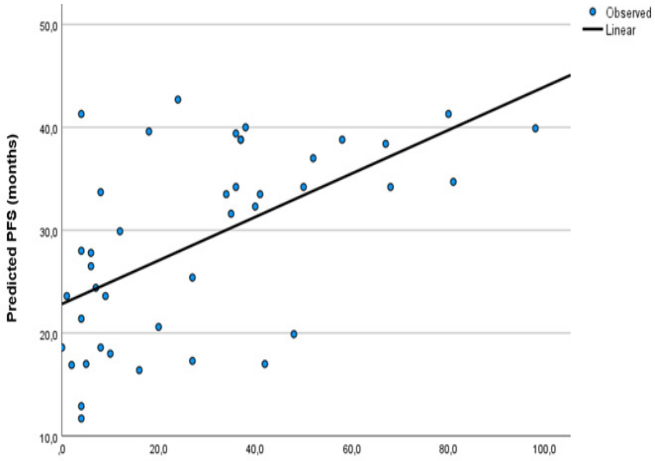

Figure 3. Correlation relationship between the actual and predicted PFS values
PFS: Progression-free survival

survival times. These findings imply that while the Random Forest algorithm provides statistically consistent predictions, its clinical applicability may be limited, particularly for cases with extended PFS (**Figure 4**).

DISCUSSION

To the best of our knowledge, this study is among the few that comprehensively incorporate demographic data, intraoperative and perioperative findings, and adjuvant treatment responses to evaluate the predictive power of ML models for PFS in patients undergoing surgery for epithelial ovarian cancer. Among the algorithms tested, Random Forest demonstrated the highest performance, while Multilayer Perceptron and Support Vector Regression showed the lowest performance, with CCs of 0.1543 and 0.1997, respectively. Random Forest achieved the best results, with a CC of 0.5731,

Table 1. Summary of the distribution of qualitative parameters

Parameter		n (%)	Parameter		n (%)
ECOG-PS	0-1	129 (77.25%)	Large bowel serosal invasion	No	97 (58.08%)
	≥2	38 (22.75%)		Localized foci	40 (23.95%)
Major cardiac comorbidities	No	128 (76.65%)		Diffuse. military	30 (17.96%)
	Yes	39 (23.35%)	Large bowel mesentery invasion	No	75 (44.91%)
Major pulmonary comorbidities	No	153 (91.62%)		Localized foci	54 (32.34%)
	Yes	14 (8.38%)		Diffuse. military	38 (22.75%)
Diabetes mellitus	No	138 (82.63%)	Spleen metastasis	No	149 (89.22%)
	Yes	29 (17.37%)		Yes	18 (10.78%)
Neoadjuvant KT	No	109 (65.27%)	Liver metastasis	No	150 (89.82%)
	3 cycles	22 (13.17%)		Any surface lesion	11 (6.59%)
	4 cycles	25 (14.97%)		Parencyhmal	6 (3.59%)
	≥6 cycles	11 (6.59%)	Pleural effusion	No	133 (79.64%)
Ascite	No	108 (64.67%)		Yes	34 (20.36%)
	Small volume	27 (16.17%)	Pleural or pulmonary nodule	No	157 (94.01%)
	Large volume	32 (19.16%)		Yes	10 (5.99%)
Omental cake	No	106 (63.47%)	Mediastinal and or paracardiac LN	No	145 (86.83%)
	Yes	61 (36.53%)		Yes	22 (13.17%)
Peritoneal carcinomatosis	No	67 (40.12%)	Inguinal LN	No	164 (98.2%)
	Localized foci	34 (20.36%)		Yes	3 (1.8%)
	Diffuse, military	66 (39.52%)	Supraklavikular LN	No	164 (98.2%)
Diyaphragmatic disease	No	120 (71.86%)		Yes	3 (1.8%)
	Localized foci	11 (6.59%)	Cytoreduction	Maximal (no visible%)	119 (71.26%)
	Diffuse, military	36 (21.56%)		Optimal (<1 cm%)	36 (21.56%)
Small bowel serosal invasion	No	127 (76.05%)		Suboptimal (≥1 cm%)	12 (7.19%)
	Localized foci	18 (10.78%)	Intestinal resection	No	128 (76.65%)
	Diffuse, military	22 (13.17%)		Yes	39 (23.35%)
Small bowel mesentery invasion	No	110 (65.87%)	Small bowel resection	No	158 (94.61%)
	Localized foci	17 (10.18%)		Yes	9 (5.39%)
	Diffuse, military	40 (23.95%)	Colorectal anastomozis	No	137 (82.04%)
				Yes	30 (17.96%)

The table continues

Table 1. Summary of the distribution of qualitative parameters (The table continues)

Pelvic peritonectomy	No	82 (49.1%)	FIGO stage	I	33 (19.76%)
	Yes	85 (50.9%)		II	16 (9.58%)
Paracolic peritonectomy	No	117 (70.06%)		III	61 (36.53%)
	Yes	50 (29.94%)		IV	57 (34.13%)
Diafragm peritonectomy	No	139 (83.23%)	LN involvement	0	70 (41.92%)
	Yes	28 (16.77%)		1	40 (23.95%)
Splenectomy and or distal pancreatectomy	No	147 (88.02%)		2	57 (34.13%)
	Yes	20 (11.98%)	Adjuvant therapy	No	18 (10.78%)
Lymphadenectomy	No	55 (32.93%)		Yes	142 (85.03%)
	Selective LN debulking	5 (2.99%)		Lost followup	0 (0%)
	Systemic pelvic-paraaortic	107 (64.07%)		Death before adjuvant therapy	7 (4.19%)
Intraop. complication	No	150 (89.82%)	Total no of first line chemotherapy cycles ¥	0	12 (7.41%)
	Yes	17 (10.18%)		1-6 cycles of chemotherapy	93 (57.41%)
Intraop need for bloood transfusion	No	79 (47.31%)		7-8 cycles of chemotherapy	49 (30.25%)
	Yes	88 (52.69%)		Death before adjuvant therapy.lost to follow up	8 (4.94%)
Needfor ICU	No	68 (40.72%)	Disease status on or after primary therapy	Complete response	129 (77.25%)
	Yes	99 (59.28%)		Partial response	22 (13.17%)
Postoperative any adverse event including deaths	No	86 (51.5%)		Stable disease	5 (2.99%)
	Yes	81 (48.5%)		Progression	4 (2.4%)
Clavien Dindo classification of surgical advers events	No	84 (50.3%)		Death before completion of primary therapy	7 (4.19%)
	Grade 1	19 (11.38%)		Lost to follow up	0 (0%)
	Grade 2	31 (18.56%)			
	Grade 3	17 (10.18%)			
	Grade 4	10 (5.99%)			
	Grade 5	6 (3.59%)			
Tumor histotype	High grade	125 (74.9%)			
	Others	42 (25.1%)			

¥ Data is missing for five patients. The distribution percentages were calculated based on a sample of 162 patients. ECOG PS: Eastern Cooperative Oncology Group Performance Status, ICU: Intensive care unit, NACT: Neoadjuvant chemotherapy

Tablo 2. General distribution patterns of the quantitative attributes used in the ML models

Parameter	Minimum	Maximum	Median	Mean	SD
Age	22	82	57	58	11
Serum CA-125	3.4	25801.0	614.0	1548.4	3436.0
LOSH-ICU	0	38	1	2	4
LOSH (overall)	1	77	11	14	9
No. of pelvic LNs removed	12	69	29	31	12
No. of paraaortic LNs removed	8	73	26	27	13
No. of total LNs removed	15	129	56	58	19
Time interval between debulking surgery and adjuvant therapy, days	14	99	36	38	14
Recurrence time, months	22	82	57	58	11

LOSH-ICU: Length Of stay hospital ıntensive care unit, LN: Lymph node

a MAE of 16.45, a RMSE of 20.98, and a model-building time of only 0.03 seconds, thereby emerging as the most effective algorithm and warranting further investigation. In comparison, Linear Regression, Bootstrap Aggregating Additive Regression, Random Committee, and Regression by Discretization yielded CC of 0.5326, 0.4915, 0.4491, 0.4077, and 0.3817, respectively. Correlation analysis between actual and predicted PFS values for Random Forest indicated a moderate correlation (r=0.4–0.6), suggesting moderate predictive accuracy. The ability of Random Forest to reduce overfitting by combining multiple models and capture complex, non-linear relationships between features is the reason for its strong performance.[13,14] Recent studies have shown that using radiomic and multi-omic data alongside clinical information improves the accuracy of ovarian cancer predictions. For example, Jian et al.[15] developed a Random Forest model that combined imaging data with clinical information and achieved a 77.2% AUC, outperforming models based solely on clinical

Table 3. Evaluation of ML model performance and prediction capabilities for algorithms used

		Evaluation of model performance		
ML algorithms	**Time taken to build model**	**Mean absolute error**	**Root mean squared error**	**Correlation coefficient†**
Random Forest	0.03 sec	16.45	20.98	0.5731
Multilayer perceptron	0.02 sec	27.30	33.70	0.1543
Linear regression	0.02 sec	18.16	21.47	0.5326
Support vector regression	0.03 sec	25.25	29.62	0.1997
Additive regression	0.02 sec	21.47	25.83	0.4491
Bootstrap aggregating (Bagging)	0.03 sec	17.46	21.67	0.4915
Random Committee	0.03 sec	18.5026	24.0229	0.4077
Regression by discretization	0.08 sec	18.5396	26.6757	0.3817

*Training and validation split: %75 and %25, †Correlation between the actual and the predicted PFS data, ML: Machine learning, PFS: Progression-free survival

Table 4. Evaluating the correlation between the actual and predicted progression-free survival (PFS) values, as well as exploring differences within the PFS values

	Correlation*		Difference within PFS values**	
	Coefficient (r)	**p value**	**t**	**p value**
Actual PFS Predicted PFS	0.573	<0.001	-0.068	0.946

*Pearson's correlation analysis **Paired sample t-test. PFS: Progression-free survival (months)

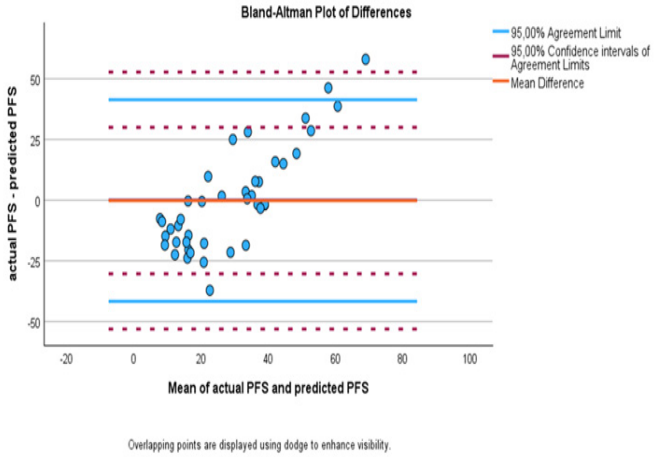

Figure 4. Bland–Altman plot of differences between actual and predicted PFS values

data. Similarly, Laios et al.[16] used ML in advanced high grade serous ovarian cancer, emphasised the importance of feature selection, and predicted 2 year survival with approximately 73% accuracy using support vector machines and ensemble models.These results demonstrate that radiomic features and accurate feature selection provide added value beyond traditional clinical methods. Furthermore, systematic reviews highlight the importance of combining different types of data. For example, Piedimonte et al.[17] reported that radiomics based ML models achieved AUC values ranging from 0.77 to 0.93 in various studies. Radiogenomic models, which combine imaging data with molecular profiles, have yielded even more promising results. A meta-analysis by Maiorano et al.[18] revealed that the AUC value of these models can reach up to 0.975, while Zeng et al.[19] deep learning-based model, which combines multi-centre imaging and genetic data, achieved the highest reported prediction accuracy (AUC=0.975).

Beyond imaging methods, multi-omic-based models have also made significant contributions. Wu et al.[20] developed the AI assisted prognostic index AIDPI by integrating transcriptomic and clinical data, demonstrating that this index improves patient risk classification. Similarly, Chen et al.[21] introduced the CSOARG model based on eight gene expression signatures, which achieved an AUC value of 0.68 in five-year survival prediction. These gene-based signatures provide valuable insights into biological risk profiles, treatment responses, and the tumours immune microenvironment. Jiang et al.[22] contributed to the field with AUTOSurv, an interpretable deep learning based platform that combines clinical, gene expression, and miRNA data, reporting that this platform outperforms traditional ML methods. Overall, these studies demonstrate that analysing multidimensional data with deep neural networks can provide accuracy beyond classical methods.

In this context, our study highlights the limitations of models based solely on clinical parameters. Using the Random Forest algorithm, a statistically significant but moderate correlation (Pearson r=0.573, p<0.001) was obtained between actual and predicted PFS. Furthermore, the absence of a significant difference between predicted and observed PFS values (p=0.946) supports the model's consistency. Although the findings demonstrate the predictive power of our ML model based on clinical data, they also indicate that its performance is limited compared to radiomic and multi-omic approaches that better reflect tumour heterogeneity and treatment response.

Nevertheless, the main strength of our study lies in its practicality and accessibility. Since the model was developed solely on the basis of routinely collected clinical data, it does not rely on advanced radiomic analyses or expensive genomic technologies. This makes it especially valuable for resource-constrained settings, where the implementation of high-performance radiomic or multi-omic models may not be feasible. Thus, although integrative models can achieve higher accuracy, a clinically based model with acceptable predictive performance can still function as a rapid, cost-effective, and complementary decision-support tool in the management of epithelial ovarian cancer. In epithelial ovarian cancer, achieving R0 resection is recognized as a critical determinant of patient survival outcomes. A comprehensive study of 571

patients sought to develop an AI-based prediction model focused on estimating the probability of R0 resection. Using the eXtreme Gradient Boosting (XGBoost) algorithm, the model incorporated multiple variables primarily related to patient characteristics and surgical features. To enhance interpretability, SHAP were applied, enabling both global and local explanations of model predictions. The XGBoost framework demonstrated strong predictive accuracy, with an AUC of 0.866 (95% confidence interval [CI]: 0.80–0.93).[23]

Cox proportional hazards regression, a commonly used method in survival analysis, is a powerful tool for evaluating the effects of covariates on the hazard function. In clinical practice, tumour characteristics, patient demographics, and treatment responses often exhibit complex, non-linear relationships with survival outcomes. Researchers, aware of these limitations, are increasingly turning to alternative approaches that can capture such dynamics. ML-based methods offer the opportunity to model these non-linear interactions more effectively, revealing patterns and relationships that traditional techniques may overlook. The application of these methods holds promise for improving prognosis accuracy, enhancing patient outcomes, and supporting more informed clinical decision-making.[24,25]

Limitations

This study has some limitations. It was conducted in a single centre with a relatively small patient group, so larger multicentre studies are needed to confirm our results. Many advanced models in the literature show high performance, but these are often based on retrospective data and internal validation.[26] For safe use in clinical practice, prospective and external validation, as well as studies on their impact on clinical decision-making, are required. In the future, adding radiomic and molecular biomarkers to clinical models and applying better feature selection methods may improve prognostic accuracy. Our findings highlight the gap between simple clinical models and advanced approaches, and suggest that hybrid models combining both may provide a good balance between accuracy and practicality in ovarian cancer prognosis. Despite the creation of a forward-looking dataset, a limitation of our retrospective design is the relatively modest sample size, which may limit the generalisability of the findings. Despite the limitations mentioned above, the value of the study lies in the careful examination of the parameters.

CONCLUSION

As a result, this study demonstrated that the Random Forest algorithm yielded better prediction results than the other methods tested and was effective in processing complex clinical data. The CC (r=0.57) indicates moderate accuracy, but this still has clinical value. In oncology, even models with moderate accuracy can assist by classifying patients according to their risk of recurrence, guiding follow up programmes, and identifying patients who may benefit from early treatment or clinical trials. The model does not replace clinical judgement but can support decision-making in multidisciplinary care. Future studies should test larger patient cohorts, refine parameters, improve data processing, and combine clinical data with radiomic or genomic information to enhance accuracy and generalisability across different healthcare settings.

ETHICAL DECLARATIONS

Ethics Committee Approval

This retrospective analysis was approved by the Scientific Ethics Committee for Medical Researches at Antalya Training and Research Hospital (Date: 07.11.2024, Decision No: 17/2).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209-249. doi:10.3322/caac.21660
2. Shachar E, Raz Y, Rotkop G, et al. Cytoreductive surgery in advanced epithelial ovarian cancer: a real-world analysis guided by clinical variables, homologous recombination, and BRCA status. *Int J Gynecol Cancer.* 2025;35(6):101809. doi:10.1016/j.ijgc.2025.101809
3. Caruso G, Tomao F, Parma G, et al. Poly (ADP-ribose) polymerase inhibitors (PARPi) in ovarian cancer: lessons learned and future directions. *Int J Gynecol Cancer.* 2023;33(4):431-443. doi:10.1136/ijgc-2022-004149
4. Atallah GA, Kampan NC, Chew KT, et al. Predicting prognosis and platinum resistance in ovarian cancer: role of immunohistochemistry biomarkers. *Int J Mol Sci.* 2023;24(3):1973. doi:10.3390/ijms24031973
5. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. *Cancer Inform.* 2006;2:117693510600200030. doi:10.1177/117693510600200030
6. Fountzilas E, Pearce T, Baysal MA, Chakraborty A, Tsimberidou AM. Convergence of evolving artificial intelligence and machine learning techniques in precision oncology. *NPJ Digit Med.* 2025;8(1):75. doi:10.1038/s41746-025-01471-y
7. Abbas S, Asif M, Rehman A, Alharbi M, Khan MA, Elmitwally N. Emerging research trends in artificial intelligence for cancer diagnostic systems: a comprehensive review. *Heliyon.* 2024;10(17):e36743. doi:10.1016/j.heliyon.2024.e36743
8. Enshaei A, Robson C, Edmondson R. Artificial intelligence systems as prognostic and predictive tools in ovarian cancer. *Ann Surg Oncol.* 2015;22:3970-3975. doi:10.1245/s10434-015-4475-6
9. Krzyszczyk P, Acevedo A, Davidoff EJ, et al. The growing role of precision and personalized medicine for cancer treatment. *Technol Cancer Res Treat.* 2018;6(03n04):79-100. doi:10.1142/S2339547818300020
10. Fagotti A, Ferrandina MG, Vizzielli G, et al. Randomized trial of primary debulking surgery versus neoadjuvant chemotherapy for advanced epithelial ovarian cancer (SCORPION-NCT01461850). *Int J Gynecol Cancer.* 2020;30(11):1657-1664. doi:10.1136/ijgc-2020-001640

11. Van Meurs HS, Tajik P, Hof MH, et al. Which patients benefit most from primary surgery or neoadjuvant chemotherapy in stage IIIC or IV ovarian cancer? An exploratory analysis of the European Organisation for Research and Treatment of Cancer 55971 randomised trial. *Eur J Cancer*. 2013;49(15):3191-3201. doi:10.1016/j.ejca.2013.06.013
12. Clavien PA, Barkun J, De Oliveira ML, et al. The Clavien-Dindo classification of surgical complications: five-year experience. *Ann Surg*. 2009;250(2):187-196. doi:10.1097/SLA.0b013e3181b13ca2
13. Biau G, Scornet E. A random forest guided tour. *Test*. 2016;25(2):197-227.
14. Fernández-Delgado M, Cernadas E, Barro S, Amorim D. Do we need hundreds of classifiers to solve real world classification problems? *J Machine Learning Res*. 2014;15(1):3133-3181.
15. Jian L, Chen X, Hu P, et al. Predicting progression-free survival in patients with epithelial ovarian cancer using an interpretable random forest model. *Heliyon*. 2024;10(15):e35344. doi:10.1016/j.heliyon.2024.e35344
16. Laios A, Katsenou A, Tan YS, et al. Feature selection is critical for 2-year prognosis in advanced stage high grade serous ovarian cancer by using machine learning. *Cancer Control*. 2021;28:10732748211044678. doi:10.1177/10732748211044678
17. Piedimonte G, Fusco R, Rizzo A, et al. Machine learning models based on radiomics for the prediction of ovarian cancer outcomes: a systematic review. *Cancers (Basel)*. 2025;17(3):336. doi:10.3390/cancers17030336
18. Maiorano BA, Rabaiotti E. Prognostic and predictive potential of radiogenomics in epithelial ovarian cancer: a systematic review and meta-analysis. *Onco*. 2025;6(4):84. doi:10.3390/onco6040084
19. Zeng S, Wang XL, Yang H. Radiomics and radiogenomics: extracting more information from medical images for the diagnosis and prognostic prediction of ovarian cancer. *Mil Med Res*. 2024;11(1):77. doi:10.1186/s40779-024-00580-1
20. Wu J, Zhao L, Li Y, et al. AIDPI: a transcriptomic and multi-omics prognostic model for ovarian cancer. *Front Genet*. 2025;16:12268085.
21. Chen X, Liu Y, Zhang H, et al. Construction and validation of a CSOARG gene signature for prognosis in ovarian cancer. *Front Oncol*. 2025;15:1592426. doi:10.3389/fonc.2025.1592426
22. Jiang Y, Yang C, Xu W, et al. AUTOSurv: deep learning framework integrating multi-omics and clinical data for ovarian cancer survival prediction. *NPJ Precis Oncol*. 2024;8:37. doi:10.1038/s41698-023-00494-6
23. Laios A, Kalampokis E, Johnson R, et al. Explainable artificial intelligence for prediction of complete surgical cytoreduction in advanced-stage epithelial ovarian cancer. *J Pers Med*. 2022;12(4):607. doi:10.3390/jpm12040607
24. Harrell J, Frank E, Harrell FE. Cox proportional hazards regression model. In: Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. 2nd ed. Springer; 2015. doi:10.1007/978-3-319-19425-7
25. Bertolaccini L, Pardolesi A, Davoli F, Solli P. Nanos gigantium humeris insidentes: the awarded Cox proportional hazards model. *J Thorac Dis*. 2016;8(11):3464-3470. doi:10.21037/jtd.2016.11.84
26. Jennings B, Zhang T. A systematic review of multimodal AI models for ovarian cancer prognosis using whole slide imaging and omics data. *arXiv preprint*. 2025;arXiv:2507.16876.