Am J Cancer Res 2025;15(2):769-780
www.ajcr.us /ISSN:2156-6976/ajcr0161731

Original Article

Random survival forest model in patients with epithelial ovarian cancer: a study based on SEER database and single center data

Luwei Wei[1], Guowei Chen[1], Huiying Liang[1], Li Li[2]

[1]*Department of Gynecology, Liuzhou Workers' Hospital, Liuzhou 545001, Guangxi, China;* [2]*Department of Gynecology, The Affiliated Tumor Hospital of Guangxi Medical University, Nanning 530021, Guangxi, China*

Received November 4, 2024; Accepted February 13, 2025; Epub February 15, 2025; Published February 28, 2025

Abstract: Clinical data of 1,780 patients with epithelial ovarian carcinoma (EOC) in the Surveillance, Epidemiology and End Results (SEER) database were retrospectively analyzed. A random survival forest model and a nomogram model were built based on the prognostic factors. The clinical data of 140 patients with EOC treated in Liuzhou Worker's Hospital were collected for the validation of the prognostic model. Age (≥75 years), histology grade (poor differentiation or undifferentiation), histologic types (clear cell carcinoma or carcinosarcoma), T stage (T2 or T3), M stage (M1), surgical conditions, and chemotherapy situation (without chemotherapy) were identified as independent risk factors. Based on these factors, a random forest survival prediction model was established. In the training set, the area under the curve (AUC) for the random forest survival prediction model in predicting 1-, 3- and 5-year survival were 0.848, 0.859 and 0.890, respectively. In the test set, the AUCs for 1-, 3- and 5-year survival were 0.992, 0.795 and 0.883, respectively. A nomogram prediction model was also established. In the training set, the AUCs for the nomogram prediction model for 1-, 3- and 5-year survival were 0.789, 0.803 and 0.838, respectively. In the test set, the AUCs for 1-, 3- and 5-year survival were 0.926, 0.748 and 0.836, respectively. The results indicated that the random forest survival model established in this study holds significant clinical value. Physicians can develop personalized follow-up strategies or treatment regimens for patients based on the predicted survival risk, potentially improving long-term outcomes.

Keywords: Ovarian neoplasms, prognosis, random survival forest, SEER database

Introduction

Ovarian carcinoma (OC) is one of the deadliest gynecologic malignancies [1], with epithelial ovarian carcinoma (EOC) being the most common subtype. EOC originates from the epithelium of the endometrium, ovary, or fallopian tube, and is characterized by an insidious onset and rapid progression. According to statistics, about 75% of patients are diagnosed at an advanced stage, with a five-year survival rate of only 29% [2]. The origin and pathogenesis of EOC remain unclear. Survival rates have shown little improvement over the past 5 to 30 years [3], making EOC one of the most challenging carcinomas. Although patients are highly sensitive to cisplatin in the early stage of therapy, multidrug resistance often develops in the later stage, leading to recurrence, metastasis, and even death in most patients, which seriously threatening the health and survival of patients [4, 5]. Therefore, tailoring treatment plans to individual patient's condition is crucial for improving their prognosis. An accurate prognostic model for EOC is vital for both clinicians and patients. Although previous studies have developed nomogram-based prediction models for EOC, their predictive efficiency and discrimination are insufficient [6].

Machine learning algorithms, such as random forests, have shown significant promise in various medical applications, including disease diagnosis, patient prognosis prediction, and drug discovery. The random survival forest model is particularly powerful for survival analysis, as it can handle complex, high-dimensional data sets and mitigate overfitting. Applying machine learning in EOC prognostic analysis represents an innovative approach that could

https://doi.org/10.62347/PLDH8547

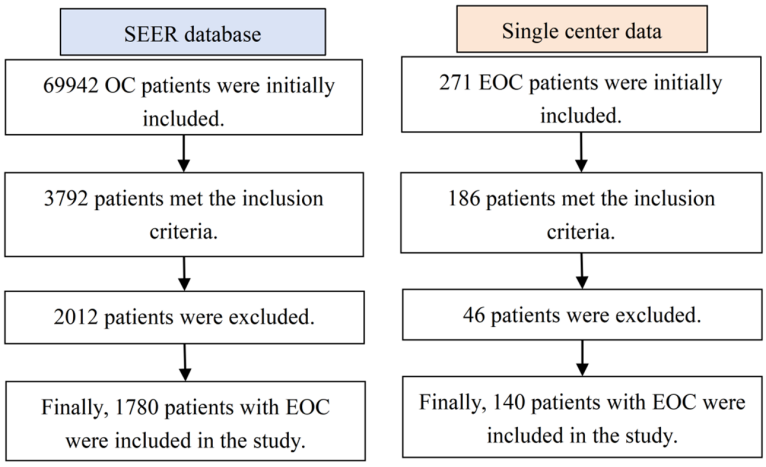

Figure 1. Sample screening process. EOC: Epithelial ovarian carcinoma; SEER: Surveillance, Epidemiology, and End Results.

potentially improve patient outcomes and optimize disease management [7].

Our study, based on the Surveillance, Epidemiology and End Results (SEER) database, funded by the National Carcinoma Institute (NCI) [8], aims to identify independent prognostic factors for EOC, develop a prognostic model, and externally validate this model using clinically collected data. This model is designed to assist clinicians in better understanding patient therapy and formulating more appropriate treatment plans.

Materials and methods

Participants

Patient data eligible for inclusion in this study were collected from the SEER database, which covers the clinical data of 18 cancer registries, representing 28% of the U.S. population. This database is characterized by large sample size and relatively complete follow-up information [9]. The data were collected using SEER*Stat 8.4.1, forming the training set. An additional 140 EOC patients, hospitalized at Liuzhou Worker's Hospital from December 2006 to December 2018, were selected as the test set. Our study was reviewed and approved by the Medical Ethics Committee of Liuzhou Worker's Hospital.

Inclusion and exclusion criteria

All 69,942 OC patients in the SEER database were collected. Inclusion criteria for training-set: (1) Patients with EOC coded as OVA-C56.9 according to ICD-O-3; (2) Patients with complete data, including age at diagnosis, race, marital status, histological grading, tumor size, 7th edition American Joint Committee on carcinoma (AJCC) TNM staging, surgery, radiotherapy and chemotherapy; (3) Patients with complete follow-up information who died exclusively from ovarian carcinoma. Exclusion criteria for training-set: (1) Patients with non-EOC; (2) Patients with incomplete data on age, race, marital status, histological grading, the seventh edition of AJCC TNM staging, tumor size, surgery, and radiotherapy and chemotherapy; (3) Patients with incomplete follow-up time, cause of death from other conditions, or unknown death status. Finally, a total of 1,780 cases met the above criteria for the training set.

Inclusion criteria for the test-set: (1) Patients diagnosed with EOC at Liuzhou Worker's Hospital; (2) Patients with complete data on age, race, marital status, histological grading, tumor size, TNM staging of EOC, surgery and chemotherapy; (3) Patients with complete follow-up information and those who died exclusively from ovarian carcinoma. Exclusion criteria for test-set: (1) Patients with non-EOC; (2) Patients with incomplete data on age, race, marital status, histological grading, TNM staging of EOC, tumor size, surgery and chemotherapy; (3) Patients who died of other causes or lost to follow-up from December 2006 to December 2018. Finally, a total of 140 cases met the above criteria. The case selection processes are shown in **Figure 1**.

Sample size estimation: The sample size was calculated according to the principle of events per variable (EPV), where the formula is: sample size = number of variables × 10/incidence. In this study, the number of variables in this study was 8, and the estimated five-year mortality rate was 70%. Thus, the required sample size was calculated as: = 8 * 10/0.7 = 114. The inclusion of 140 patients in this study as an external validation set met the minimum requirements for statistical analysis.

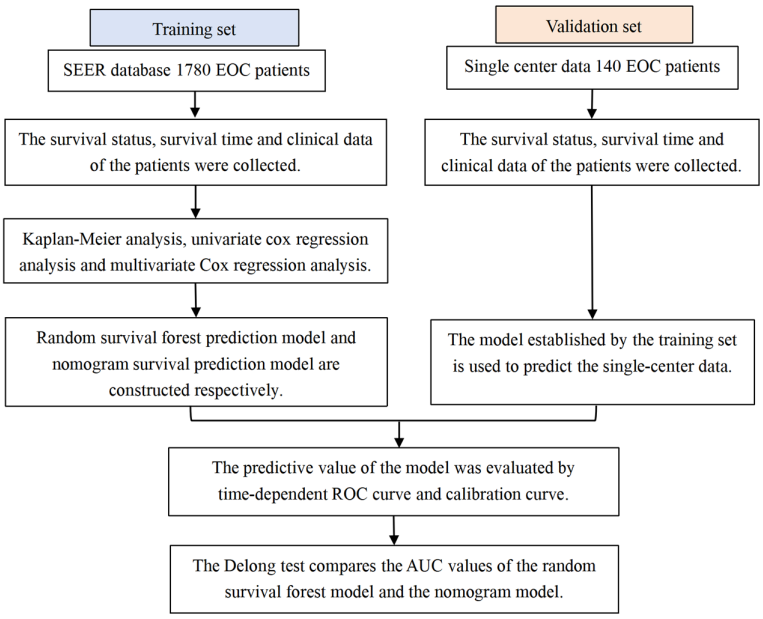

Figure 2. Research flow chart. EOC: Epithelial ovarian carcinoma; ROC: Receiver operating characteristic; AUC: area under ROC curve; SEER: Surveillance, Epidemiology, and End Results.

Clinical case characteristics

The clinical data collected included age, race, marriage, degree of differentiation, histological type, TNM staging, tumor size, surgery, radiotherapy and chemotherapy. Age was categorized into three groups: ≤54 years, 55 to 74 years, and ≥75 years, optimal cutoff points for age and tumor size were analyzed using X-tile. Racial groups, as defined by the SEER database, included American Indian, Asian, and Pacific Islander. According to the 2014 World Health Organization classification of tumors of female reproductive organs, patients were classified into the following histological types: serous carcinoma, mucinous carcinoma, endometrioid carcinoma, clear cell carcinoma, carcinosarcoma, and Brenner tumor.

Statistics process

Descriptive statistics were used to summarize the collected data on EOC patients. Univariate analysis was performed using Log-rank χ^2 test in SPSS 26.0. Variables with statistical significance in univariate analysis (P<0.05) were included in multivariate Cox regression analysis, and the Kaplan-Meier survival curves were plotted using SPSS 26.0. Independent prognostic factors identified by Cox multivariate regression analysis were integrated, and the "rfsrc" function of the "randomForestSRC" package in R-4.2.3 was used to build a random forest prediction model. In addition, a nomogram prediction model was developed using the "rms" package. The clinical data collected from hospital were used as the test set for external validation of the model. The area under the receiver operating characteristic (ROC) curve (AUC) was used to evaluate the discrimination of the model. An AUC value closer to 1 indicates higher discrimination and better prediction performance. The Delong test was used to compare two AUC values. Calibration curves were used to evaluate the model's calibration. A flow chart of this study is shown in **Figure 2**.

Results

Epidemiological features

A total of 1780 EOC patients were included in the training set. Among the samples, 40.28% had an age of ≤54 years and 43.48% were between 55 and 74. The majority of patients were White (80.06%), and most were married (57.75%). The undifferentiated and poorly differentiated tumors accounted for 28.76% and 33.93%, respectively. The most common histological types were serous carcinoma (38.67%) and endometrioid carcinoma (33.9%). The majority of patients had a tumor size of <1.5 cm. In terms of T stage, 46.12% were classified as T1, 15.79% as T2, and 38.09% as T3; In N and M stages, the majority of patients were N0 (81.91%) and M0 (89.66%). Over 99% of the patients underwent surgery, while 77.36% received chemotherapy. However, 98.31% of the patients did not receive radiotherapy. See **Table 1** for details.

Univariate analysis of prognosis

Univariate analysis using the Log-rank χ^2 test indicated that age, race, marital status, histo-

Table 1. Cox univariate and multivariate analyses of factors affecting patients' survival

Characteristics	*n (%)*	Survival time (months)	*Log-rank* χ^2 test		*COX* analysis	
			HR (95% CI)	*P*	*HR (95% CI)*	*P*
Age				<0.001**		
≤54	717 (40.28)	96.432 (93.525-99.338)	1 (Reference value)		1 (Reference value)	
55-74	774 (43.48)	83.719 (80.564-86.875)	1.705 (1.417-2.051)	<0.001**	1.157 (0.955-1.401)	0.127
≥75	289 (16.24)	63.192 (57.820-68.564)	3.176 (2.572-3.921)	<0.001**	1.592 (1.262-2.009)	<0.001**
Race				0.001**		
White	1425 (80.06)	84.995 (82.679-87.311)	1 (Reference value)		1 (Reference value)	
Black	95 (5.34)	73.946 (64.410-83.482)	1.414 (1.040-1.924)	0.027*	1.089 (0.790-1.500)	0.688
Other	260 (14.61)	92.831 (87.635-98.026)	0.718 (0.561-0.918)	0.008**	1.057 (0.816-1.396)	0.744
Marital status				<0.001**		
Unmarried	409 (22.98)	88.930 (84.862-92.997)	1 (Reference value)		1 (Reference value)	
Married	1028 (57.75)	87.104 (84.423-89.784)	1.137 (0.930-1.389)	0.210	0.840 (0.683-1.032)	0.124
Divorced or widowed	343 (19.27)	75.539 (70.502-80.576)	1.675 (1.329-2.112)	<0.001**	1.092 (0.855-1.396)	0.494
Grade of histology				<0.001**		
Well differentiated: grade I	296 (16.63)	113.983 (111.560-116.407)	1 (Reference value)		1 (Reference value)	
Moderate differentiation: grade II	368 (20.67)	103.351 (99.804-106.897)	3.334 (1.927-5.765)	<0.001**	2.256 (1.280-3.975)	0.002**
Poor differentiation: grade III	604 (33.93)	73.288 (69.653-76.923)	12.248 (7.408-20.248)	<0.001**	4.055 (2.341-7.023)	<0.001**
Undifferentiated: grade IV	512 (28.76)	71.318 (67.238-75.399)	12.862 (7.764-21.309)	<0.001**	3.946 (2.253-6.910)	<0.001**
Histologic Types				<0.001**		
Serous carcinoma	688 (38.67)	66.953 (63.643-70.264)	1 (Reference value)		1 (Reference value)	
Mucinous carcinoma	61 (3.43)	110.711 (103.712-117.710)	0.100 (0.041-0.242)	<0.001**	0.957 (0.378-2.423)	0.773
Endometrioid carcinoma	603 (33.9)	107.660 (105.224-110.095)	0.159 (0.124-0.203)	<0.001**	0.816 (0.608-1.095)	0.174
Clear cell carcinoma	345 (19.39)	90.650 (86.017-95.282)	0.436 (0.351-0.541)	<0.001**	1.430 (1.106-1.850)	0.009**
carcinosarcoma	82 (4.61)	46.440 (36.660-56.221)	1.704 (1.294-2.243)	<0.001**	1.883 (1.402-2.529)	<0.001**
Brenner	1 (0.0006)	43.000 (43.000-43.000)	2.013 (0.283-14.339)	0.485	2.555 (0.347-18.792)	0.454
Tumour size (cm)				<0.001**		
≤0.77	673 (37.81)	80.935 (77.512-84.358)	1 (Reference value)		1 (Reference value)	
0.78-1.5	768 (43.15)	85.869 (82.727-89.011)	0.839 (0.712-0.990)	0.038*	0.833 (0.703-0.987)	0.033*
≥1.51	339 (19.04)	94.221 (89.689-98.752)	0.568 (0.447-0.722)	<0.001**	0.698 (0.541-0.900)	0.007**
T stage				<0.001**		
T1	821 (46.12)	110.703 (108.927-112.478)	1 (Reference value)		1 (Reference value)	
T2	281 (15.79)	85.618 (80.272-90.964)	4.578 (3.404-6.156)	<0.001**	3.598 (2.604-4.971)	<0.001**
T3	678 (38.09)	55.947 (52.735-59.160)	11.682 (9.182-14.864)	<0.001**	4.970 (3.231-7.552)	<0.001**

N stage						
N0	1458 (81.91)	91.379 (89.193-93.566)	1 (Reference value)		1 (Reference value)	
N1	322 (18.09)	59.399 (54.666-64.132)	2.790 (2.366-3.289)	<0.001**	1.158 (0.968-1.386)	0.063
M stage						
M0	1596 (89.66)	89.977 (87.881-92.074)	1 (Reference value)		1 (Reference value)	
M1	184 (10.34)	47.452 (41.323-53.581)	3.649 (3.020-4.410)	<0.001**	1.406 (1.154-1.715)	<0.001**
Surgical conditions				<0.001**		
No surgery	11 (0.62)	11.364 (2.653-20.074)	1 (Reference value)		1 (Reference value)	
Ovariectomy + hysterectomy	783 (43.99)	100.554 (97.997-103.111)	0.038 (0.02-0.07)	<0.001**	0.116 (0.061-0.223)	<0.001**
Oophorectomy only	200 (11.23)	98.198 (92.865-103.530)	0.044 (0.022-0.085)	<0.001**	0.115 (0.058-0.230)	<0.001**
Cytoreductive, cytoreductive surgery	769 (43.20)	69.090 (65.804-72.377)	0.129 (0.071-0.237)	<0.001**	0.144 (0.077-0.272)	<0.001**
Pelvic exenteration	17 (0.95)	41.750 (24.547-58.953)	0.257 (0.116-0.569)	0.001**	0.187 (0.081-0.430)	<0.001**
Chemotherapy				<0.001**		
Early absence of chemotherapy	335 (18.82)	108.633 (105.327-111.938)	1 (Reference value)		1 (Reference value)	
Advanced stage without chemotherapy	68 (3.82)	39.248 (29.320-49.176)	16.448 (10.737-25.197)	<0.001**	0.755 (0.388-1.469)	0.408
Early Chemotherapy	767 (0.43)	102.384 (99.896-104.873)	1.810 (1.251-2.620)	0.002**	0.779 (0.529-1.149)	0.208
Late Chemotherapy	610 (34.27)	57.749 (54.390-61.108)	9.737 (6.890-13.758)	<0.001**	0.382 (0.205-0.710)	0.002**
Situation of radiotherapy						
No radiotherapy	1750 (98.31)	85.373 (83.277-87.469)	1 (Reference value)			
Radiation therapy	30 (1.69)	90.514 (78.179-102.850)	0.609 (0.289-1.282)	0.191		

**P<0.05, **P<0.01.*

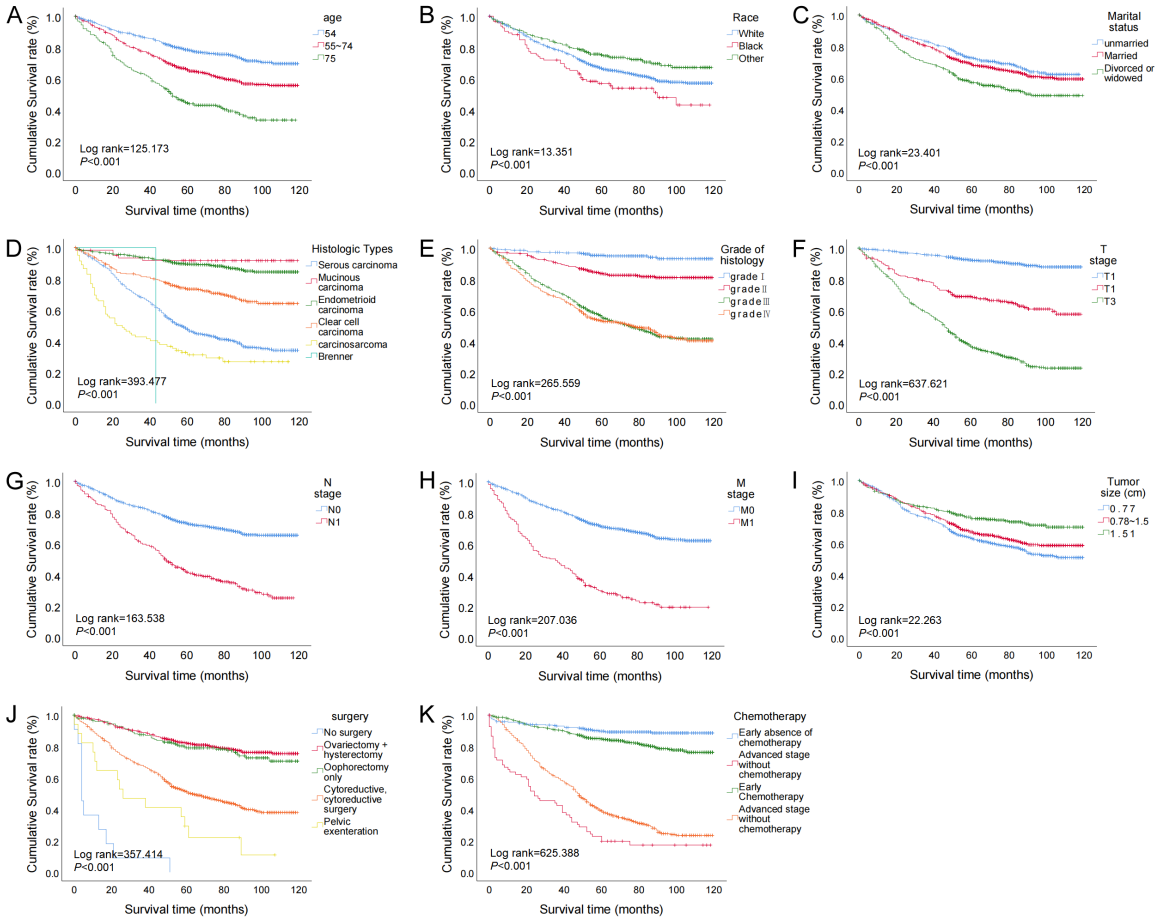

Figure 3. K-M survival curves for patients stratified by independent prognostic factors. A: Age; B: Race; C: Marital status; D: Histological type; E: Histological grading; F: T stage; G: N stage; H: M stage; I: Tumor size; J: Surgical situations; K: Chemotherapy.

logical grading, histological type, tumor size, TNM staging, surgery, and chemotherapy were significantly associated with the survival of EOC patients (all P<0.05), as shown in **Table 1** and **Figure 3**. Patients aged ≤54 years had better outcomes compared to older age groups. Patients from races other than White or Black had better outcomes. Married patients had better outcomes than those with other marital statuses. Patients with higher histological grades had better outcomes than those with lower grades. Mucinous carcinoma patients had better outcomes than those with other histological types. Patients with tumor size ≥1.51 cm had better outcomes than those with smaller tumors. Patients with T1 stage had better outcomes than those with T2 or T3 stages. Patients with N0 and M0 stages had better outcomes than those with other N or M stage classifications. Patients who underwent surgery had better outcomes than those who did not. Additionally, patients with advanced carcinoma treated with chemotherapy had better outcomes than those who did not receive chemotherapy. All of these results were statistically significant (P<0.01) (see **Table 1**).

Multivariate analysis of prognosis

Multivariate Cox regression analysis identified several independent risk factors for poor prognosis, including age ≥75 years, moderate differentiation (grade II), poor differentiation (grade III), undifferentiated carcinoma (grade IV), clear cell carcinoma, carcinosarcoma, T2 and T3 stages, and M1 stage (all P<0.001). In contrast, independent protective factors included tumor sizes of 0.78-1.5 cm and ≥1.51 cm, oophorectomy with hysterectomy, simple oophorectomy, cytoreductive surgery, pelvic

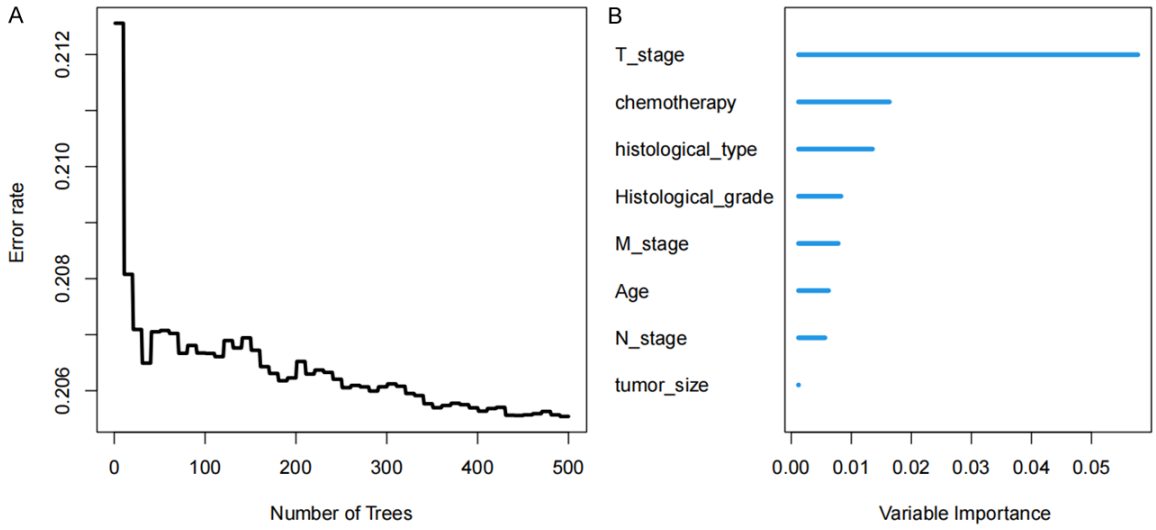

Figure 4. Analysis of model performance and variable importance in survival tree models. A: The relationship between out-of-pocket data error rate and the number of survival trees; B: Significance of variables (VIMP) plot.

exenteration, and late chemotherapy (all P< 0.01). See **Table 1**.

Model construction and validation

Eight independent prognostic factors (Age, histological grade, histological type, T stage, N stage, M stage, tumor size, chemotherapy) from Cox multivariate regression analysis were integrated, and the random forest prediction model was developed. The VIMP diagram, illustrating the relationship between out-of-bag data error rate and the number of survival trees, revealed that the forest stabilized when the number of survival trees reached 400 (**Figure 4A**). The variable importance map emphasized that the T stage was the most influential factor affecting survival (**Figure 4B**).

The AUCs for the constructed random forest prediction model for 1-, 3-, and 5-year survival in the training set were 0.848, 0.859 and 0.890, respectively; and the AUCs in the test set were 0.992, 0.795 and 0.883, respectively. The AUCs were all greater than 0.7, confirming that the prediction model had good discrimination (**Figure 5**). There was no significant difference in the AUCs between the training set and the validation set (all P>0.05). Additionally, the calibration curves of 1-, 3- and 5-year survival displayed a good agreement between the model's predictions and actual observations (**Figure 6**).

To compare the performance of the random forest prediction model with other prediction models, a nomogram prediction model was also constructed (**Figure 7**). In the training set, the AUCs for the nomogram in predicting 1-, 3- and 5-year survival were 0.789, 0.803 and 0.838, respectively (**Figure 8A**). In the test set, the AUCs were 0.926, 0.748 and 0.836, respectively (**Figure 8B**). The calibration curves for 1-, 3- and 5-year survival displayed a good agreement between the model's predictions and actual observations (**Figure 9**). Delong test showed that there was no significant difference in the AUC values between the random forest prediction model and the nomogram prediction model (P>0.05).

Discussion

Epithelial ovarian carcinoma (EOC) is one of the deadliest gynecologic carcinomas worldwide. The advanced stage at diagnosis is a major contributor to its high mortality rate. The 5-year relative survival rate for patients with advanced ovarian carcinoma is 29%, compared to 92% for those diagnosed at an early stage; and due to the absence of early symptoms, approximately 75% of patients are diagnosed at an advanced stage [10]. Survival outcomes for EOC largely depend on early diagnosis and access to appropriate surgical and systemic therapy [11]. Identifying the optimal treatment strategy for each patient is crucial [12]. To pre-

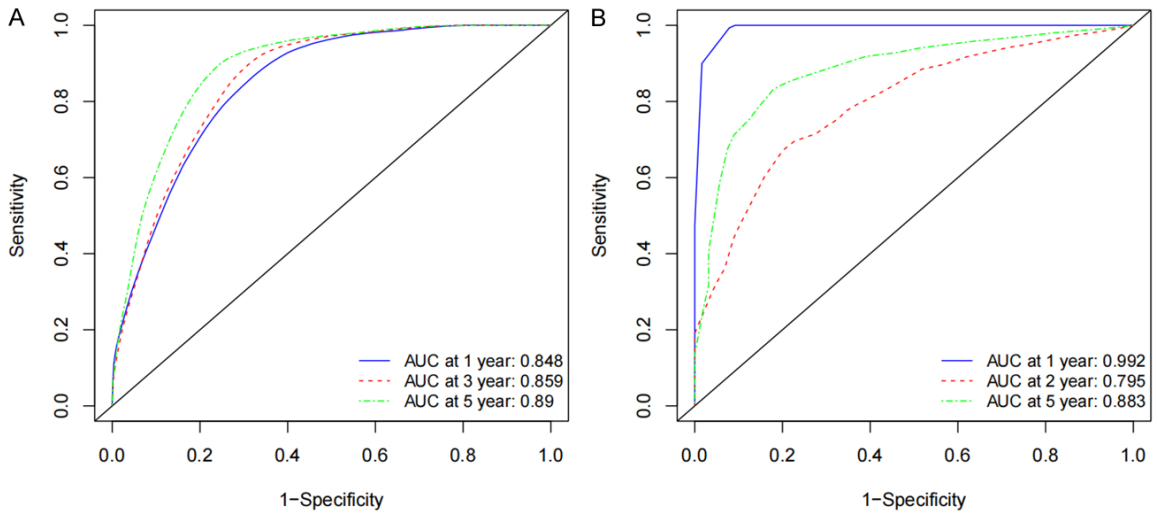

Figure 5. ROC curves for random survival forest-based model in predicting 1-, 3-, and 5-year survival. A: Training set; B: Test set.

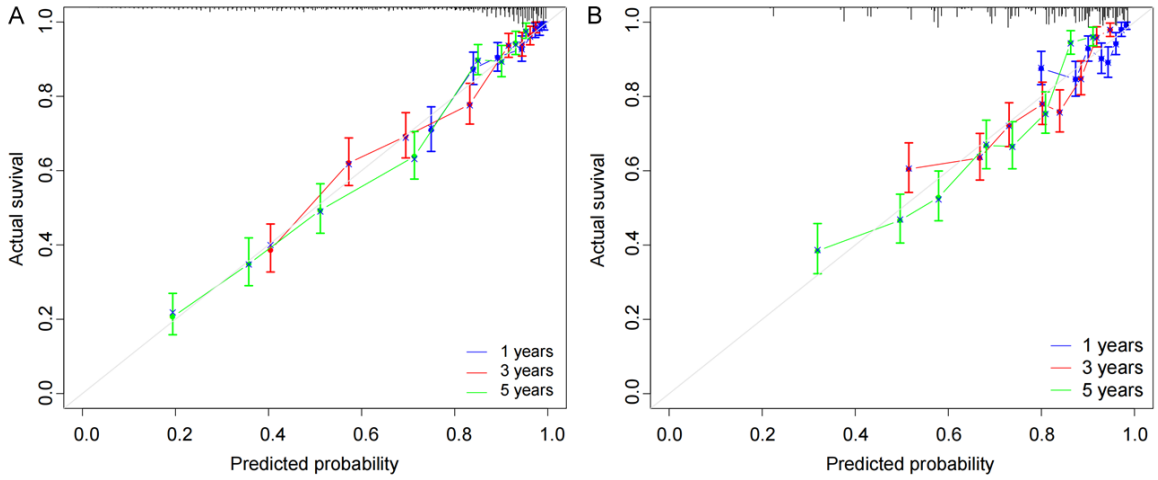

Figure 6. Calibration curves for random survival forest prediction model for 1-, 3-, and 5-year survival. A: Training set; B: Test set.

dict the patient outcome, clinicians must devise tailored treatment plans, especially for those in later stages. Although previous studies have developed nomogram prediction models for EOC outcomes, the prediction efficiency of these models remain suboptimal [13, 14]. Based on the large public carcinoma database established by the National Carcinoma Institute, this study analyzed patient data on EOC, identified independent prognostic factors, and developed a prognostic model with improved prediction efficiency. After external validation, the model demonstrated high AUC, sensitivity, and specificity.

Cox multivariate regression analysis based on clinical data from the SEER database identified several independent risk factors for poor prognosis in EOC, including age ≥75 years, moderate differentiation (grade II), poor differentiation (grade III), undifferentiated carcinoma (grade IV), clear cell carcinoma, carcinosarcoma, T2 and T3 stages, and M1 stage. Previous studies have also highlighted advanced age as an independent risk factor, with patients aged >73 years having the worst outcomes, consistent with our findings [15]. Regarding histological grading, poorer differentiation correlates with worse prognosis. In the histological classi-

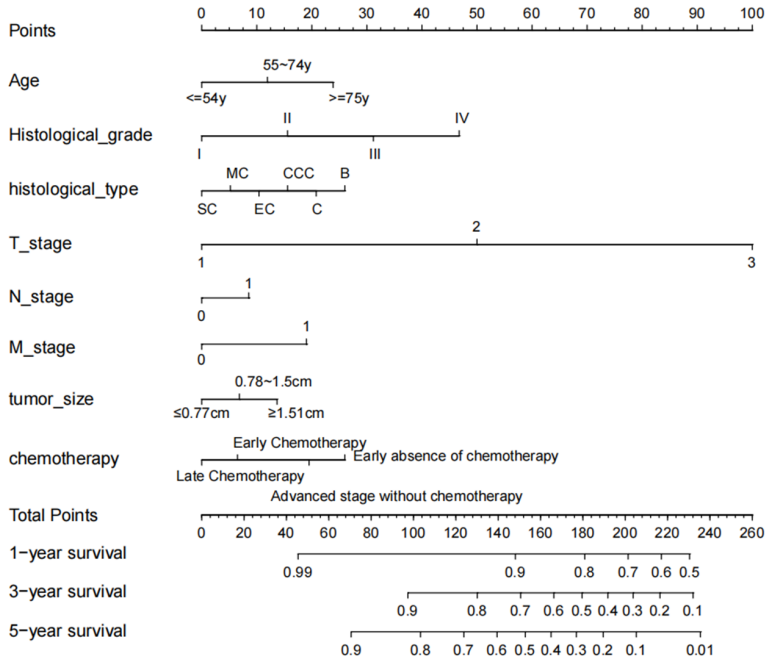

Figure 7. Nomogram prediction model.

fication of ovarian epithelial carcinoma, clear cell carcinoma and carcinosarcoma are associated with poor prognoses, consistent with findings from Peres et al. [16]. In TNM staging, the hazard ratio (HR) for T2 and T3 stages were 3.906 and 7.763, respectively, with 95% CIs of 2.843-5.366 and 5.786-10.415, respectively (both P<0.001). The M1 stage was also identified as an independent risk factor, aligning with prior research that indicates poor outcomes for ovarian carcinoma with distant metastasis [17].

Tumor size (≥0.78 cm), oophorectomy combined with hysterectomy, oophorectomy only, cytoreductive surgery, pelvic exenteration, and late chemotherapy were identified as independent protective factors. Although tumor size has a limited effect on the probability of metastasis, even small tumors can metastasize. Moreover, small tumors may lead some patients to forgo chemotherapy [18]. Our study found that patients who underwent any form of surgery had better outcomes than those who did not, which aligns with existing biological knowledge. Surgical treatment for recurrent ovarian carcinoma has been shown to improve survival, especially in the best surgical candidates, as it can reduce the tumor to the maximum extent possible [19]. The hazard ratio for chemotherapy was 0.382 (95% CI: 0.205-0.710), confirming that patients with advanced carcinoma who received chemotherapy had a better prognosis compared to those who did not. Li et al. [20] also emphasized that surgery combined with chemotherapy in patients with advanced ovarian carcinoma reduces surgery duration, intraoperative bleeding, and ascites. Therefore, surgery is recommended for patients with indications for therapy and no contraindications, and adjuvant chemotherapy is recommended when necessary.

Ovarian carcinoma is the second most common cause of gynecologic cancer-related deaths worldwide, with 90% of cases being epithelial ovarian carcinoma (EOC), the most aggressive form. Despite surgery combined with chemotherapy being the standard treatment, approximately 66% of EOC patients are diagnosed at an advanced stage, and within 16 months, half of them will experience a relapse [21]. Treatment options for EOC patients vary significantly, making it crucial for clinicians to have a reliable prediction model to better understand prognostic factors and potential outcomes. The random forest algorithm, a popular ensemble machine learning tool, has gained recognition for its effectiveness in clinical decision support and prognostic prediction tasks [22-24]. Random forest is particularly well-suited for these tasks due to its high accuracy, ability to handle nonlinear data, and low tendency to overfit [25-27]. In our study, we established a random forest-based predictive model, which enhances prediction accuracy and stability by constructing multiple decision trees. This model has high predictive value, providing clinicians with more accurate predictions to inform clinical decision-making.

However, there are some limitations to our study: (1) The data available in large public carcinoma databases are limited, meaning several potentially impactful clinical factors were not included in the model; (2) The test

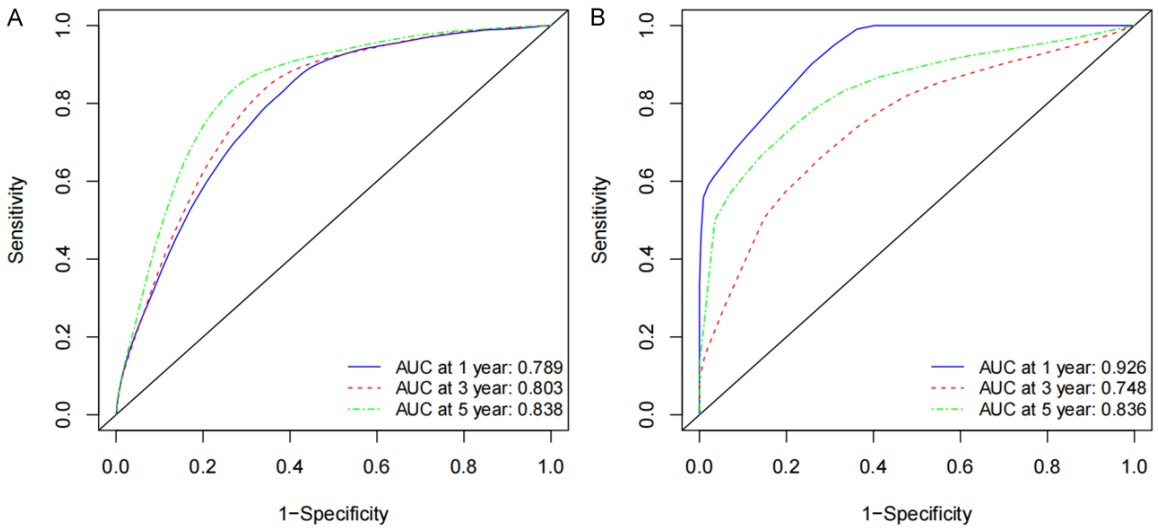

Figure 8. ROC curves for nomogram model in predicting 1-, 3-, and 5-year survival. A: Training set; B: Test set.

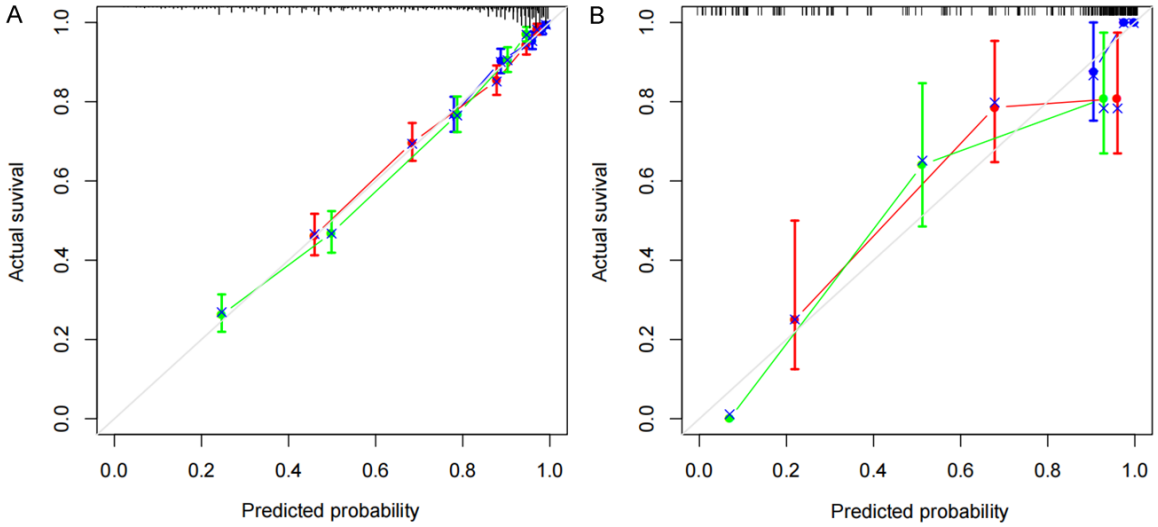

Figure 9. Calibration curves for nomogram prediction model for 1-, 3-, and 5-year survival. A: Training set; B: Test set.

set, which consists of patient data from 1975 to 2019, is retrospective and may introduce significant bias compared to randomized clinical trials. Additionally, the retrospective nature of the test set limits the model's prospective predictive capabilities; (3) Although the training set from the SEER database was large, the smaller sample size of the test set may have introduced errors in model performance testing; and (4) While the predictive performance of the model surpasses that of nomograms in other studies, its complexity requires clinicians to have a basic understanding of programming, which limits its practical accessibility.

Conclusion

In summary, our study identified independent prognostic factors for EOC using the SEER database and developed a prognostic prediction model.

This model provides valuable insights into patient prognosis and offers data-driven support to clinicians for making informed decisions regarding subsequent treatment options.

Acknowledgements

This study was supported by Guangxi Natural Science Foundation (2018GXNSFBA138052)

and National Natural Science Foundation of China (82060470).

Disclosure of conflict of interest

None.

Address correspondence to: Li Li, Department of Gynecology, The Affiliated Tumor Hospital of Guangxi Medical University, 71 Hedi Road, Qingxiu District, Nanning 530021, Guangxi, China. E-mail: lili@gxmu.edu.cn

References

[1] Veneziani AC, Gonzalez-Ochoa E, Alqaisi H, Madariaga A, Bhat G, Rouzbahman M, Sneha S and Oza AM. Heterogeneity and treatment landscape of ovarian carcinoma. Nat Rev Clin Oncol 2023; 20: 820-842.

[2] Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I and Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024; 74: 229-263.

[3] Kelliher L and Lengyel E. Understanding long-term survival of patients with ovarian cancer-the tumor microenvironment comes to the forefront. Cancer Res 2023; 83: 1383-1385.

[4] Vergote I, Pérez-Fidalgo JA, Hamilton EP, Valabrega G, Van Gorp T, Sehouli J, Cibula D, Levy T, Welch S, Richardson DL, Guerra EM, Scambia G, Henry S, Wimberger P, Miller DS, Klat J, Martínez-Garcia J, Raspagliesi F, Pothuri B, Romero I, Bergamini A, Slomovitz B, Schochter F, Høgdall E, Fariñas-Madrid L, Monk BJ, Michel D, Kauffman MG, Shacham S, Mirza MR and Makker V; ENGOT-EN5/GOG-3055/SIENDO Investigators. Oral Selinexor as Maintenance Therapy after first-line chemotherapy for advanced or recurrent endometrial cancer. J Clin Oncol 2023; 41: 5400-5410.

[5] Knisely A, Hinchcliff E, Fellman B, Mosley A, Lito K, Hull S, Westin SN, Sood AK, Schmeler KM, Taylor JS, Huang SY, Sheth RA, Lu KH and Jazaeri AA. Phase 1b study of intraperitoneal ipilimumab and nivolumab in patients with recurrent gynecologic malignancies with peritoneal carcinomatosis. Med 2024; 5: 311-320, e3.

[6] Wang J, Chen S, Zhong F, Zhu T and Zhao Y. A LASSO-derived prediction model for assessing the risk of lymph node metastasis in T1 and T2 epithelial ovarian cancer: an international retrospective cohort study. Int J Surg 2023; [Epub ahead of print].

[7] Zhang H, Chi M, Su D, Xiong Y, Wei H, Yu Y, Zuo Y and Yang L. A random forest-based metabolic risk model to assess the prognosis and metabolism-related drug targets in ovarian cancer. Comput Biol Med 2023; 153: 106432.

[8] Omofuma OO, Cook MB, Abnet CC and Camargo MC. Race and ethnicity, stage-specific mortality, and cancer treatment in esophageal and gastric cancers: surveillance, epidemiology, and end results (2000-2018). Gastroenterology 2023; 164: 473-475, e4.

[9] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019; 69: 7-34.

[10] Lheureux S, Gourley C, Vergote I and Oza AM. Epithelial ovarian cancer. Lancet 2019; 393: 1240-1253.

[11] Lheureux S, Braunstein M and Oza AM. Epithelial ovarian cancer: evolution of management in the era of precision medicine. CA Cancer J Clin 2019; 69: 280-304.

[12] Bommert M, Harter P, Heitz F and du Bois A. When should surgery be used for recurrent ovarian carcinoma? Clin Oncol (R Coll Radiol) 2018; 30: 493-497.

[13] Wang H, Liu J, Yang J, Wang Z, Zhang Z, Peng J, Wang Y and Hong L. A novel tumor mutational burden-based risk model predicts prognosis and correlates with immune infiltration in ovarian cancer. Front Immunol 2022; 13: 943389.

[14] Zhou Y, Wang A, Sun X, Zhang R and Zhao L. Survival prognosis model for elderly women with epithelial ovarian cancer based on the SEER database. Front Oncol 2023; 13: 1257615.

[15] Cheng H, Xu JH, Kang XH, Wu CC, Tang XN, Chen ML, Lian ZS, Li N and Xu XL. Nomograms for predicting overall survival and cancer-specific survival in elderly patients with epithelial ovarian cancer. J Ovarian Res 2023; 16: 75.

[16] Peres LC, Cushing-Haugen KL, Köbel M, Harris HR, Berchuck A, Rossing MA, Schildkraut JM and Doherty JA. Invasive epithelial ovarian cancer survival by histotype and disease stage. J Natl Cancer Inst 2019; 111: 60-68.

[17] Gardner AB, Charo LM, Mann AK, Kapp DS, Eskander RN and Chan JK. Ovarian, uterine, and cervical cancer patients with distant metastases at diagnosis: most common locations and outcomes. Clin Exp Metastasis 2020; 37: 107-113.

[18] Vargo JA, Gill BS, Balasubramani GK and Beriwal S. Treatment selection and survival outcomes in early-stage diffuse large B-cell lymphoma: do we still need consolidative radiotherapy? J Clin Oncol 2015; 33: 3710-7.

[19] Stone R, Sakran JV and Long Roche K. Salpingectomy in ovarian cancer prevention. JAMA 2023; 329: 2015-2016.

[20] Li X and Du X. Neoadjuvant chemotherapy combined with interval cytoreductive surgery in ovarian carcinoma. J BUON 2019; 24: 2035-2040.
[21] Yan T, Ma X, Hu H, Gong Z, Zheng H, Xie S, Guo L and Lu R. Serology-based model for personalized epithelial ovarian cancer risk evaluation. Curr Oncol 2022; 29: 2695-2705.
[22] Zhao J, Henriksson A, Kvist M, Asker L and Boström H. Handling temporality of clinical events for drug safety surveillance. AMIA Annu Symp Proc 2015; 2015: 1371-80.
[23] Arevalillo JM, Sztein MB, Kotloff KL, Levine MM and Simon JK. Identification of immune correlates of protection in Shigella infection by application of machine learning. J Biomed Inform 2017; 74: 1-9.
[24] Cox AP, Raluy-Callado M, Wang M, Bakheit AM, Moore AP and Dinet J. Predictive analysis for identifying potentially undiagnosed post-stroke spasticity patients in United Kingdom. J Biomed Inform 2016; 60: 328-33.
[25] Hayes T, Baraldi AN and Coxe S. Random forest analysis and lasso regression outperform traditional methods in identifying missing data auxiliary variables when the MAR mechanism is nonlinear (p.s. Stop using Little's MCAR test). Behav Res Methods 2024; 56: 8608-8639.
[26] Rahman SA, Maynard N, Trudgill N, Crosby T, Park M, Wahedally H, Underwood TJ and Cromwell DA; NOGCA Project Team and AUGIS. Prediction of long-term survival after gastrectomy using random survival forests. Br J Surg 2021; 108: 1341-1350.
[27] Wang J, Sun X, Cheng Q and Cui Q. An innovative random forest-based nonlinear ensemble paradigm of improved feature extraction and deep learning for carbon price forecasting. Sci Total Environ 2021; 762: 143099.