World Journal of
Clinical Oncology

World Journal of
Clinical Oncology

Contents

Contents

ABOUT COVER

Editorial Board Member of *World Journal of Clinical Oncology*, Guangzhen Wu, MD, PhD, Postgraduate Tutor, Professor. First Affiliated Hospital of Dalian Medical University, No.222 Zhongshan Road, Xigang District, Dalian, Liaoning, 116000, China. wuguang0613@hotmail.com

AIMS AND SCOPE

The primary aim of *World Journal of Clinical Oncology* (*WJCO*, *World J Clin Oncol*) is to provide scholars and readers from various fields of oncology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

WJCO mainly publishes articles reporting research results and findings obtained in the field of oncology and covering a wide range of topics including art of oncology, biology of neoplasia, breast cancer, cancer prevention and control, cancer-related complications, diagnosis in oncology, gastrointestinal cancer, genetic testing for cancer, gynecologic cancer, head and neck cancer, hematologic malignancy, lung cancer, melanoma, molecular oncology, neurooncology, palliative and supportive care, pediatric oncology, surgical oncology, translational oncology, and urologic oncology.

INDEXING/ABSTRACTING

The *WJCO* is now abstracted and indexed in PubMed, PubMed Central, Emerging Sources Citation Index (Web of Science), Reference Citation Analysis, China National Knowledge Infrastructure, China Science and Technology Journal Database, and Superstar Journals Database. The 2022 edition of Journal Citation Reports® cites the 2021 Journal Citation Indicator (JCI) for *WJCO* as 0.35.

RESPONSIBLE EDITORS FOR THIS ISSUE

ORIGINAL ARTICLE

Retrospective Cohort Study

Machine learning-assisted ensemble analysis for the prediction of urinary tract infection in elderly patients with ovarian cancer after cytoreductive surgery

Jiao Ai, Yao Hu, Fang-Fang Zhou, Yi-Xiang Liao, Tao Yang

Jiao Ai, Fang-Fang Zhou, Yi-Xiang Liao, Tao Yang, Department of Urology, Jingzhou Central Hospital, Jingzhou Hospital Affiliated to Yangtze University, Jingzhou 434020, Hubei Province, China

Yao Hu, Department of Obstetrics and Gynaecology, Jingzhou Central Hospital, Jingzhou Hospital Affiliated to Yangtze University, Jingzhou 434020, Hubei Province, China

Corresponding author: Tao Yang, MD, Surgical Oncologist, Department of Urology, Jingzhou Central Hospital, Jingzhou Hospital Affiliated to Yangtze University, No. 26 Chuyuan Road, Jingzhou 434020, Hubei Province, China. yangtaotao123321@163.com

Abstract

BACKGROUND
Urinary tract infection (UTI) is a common type of postoperative infection following cytoreductive surgery for ovarian cancer, which severely impacts the prognosis and quality of life of patients.

AIM
To develop a machine learning assistant model for the prevention and control of nosocomial infection.

METHODS
A total of 674 elderly patients with ovarian cancer who were treated at the Department of Gynaecology at Jingzhou Central Hospital between January 31, 2016 and January 31, 2022 and met the inclusion criteria of the study were selected as the research subjects. A retrospective analysis of the postoperative UTI and related factors was performed by reviewing the medical records. Five machine learning-assisted models were developed using two-step estimation methods from the candidate predictive variables. The robustness and clinical applicability of each model were assessed using the receiver operating characteristic curve, decision curve analysis and clinical impact curve.

RESULTS
A total of 12 candidate variables were eventually included in the UTI prediction model. Models constructed using the random forest classifier, support vector machine, extreme gradient boosting, and artificial neural network and decision

tree had areas under the receiver operating characteristic curve ranging from 0.776 to 0.925. The random forest classifier model, which incorporated factors such as age, body mass index, catheter, catheter intubation times, blood loss, diabetes and hypoproteinaemia, had the highest predictive accuracy.

CONCLUSION
These findings demonstrate that the machine learning-based prediction model developed using the random forest classifier can be used to identify elderly patients with ovarian cancer who may have postoperative UTI. This can help with treatment decisions and enhance clinical outcomes.

Key Words: Cytoreductive surgery; Machine learning; Ovarian cancer; Risk factors; Urinary tract infection

Core Tip: Using a machine learning-based algorithm, we developed a feasible and robust method to identify factors that are significant for predicting urinary tract infections. The random forest classifier was especially robust and can improve the prediction and early detection of urinary tract infections in patients with ovarian cancer. In addition, the five most crucial factors were age, body mass index, catheter, catheter intubation times, blood loss, diabetes and hypoproteinaemia. Clinicians may find it extremely helpful to assess the individualised risk of urinary tract infections in clinical practice by incorporating the presentation of simple clinical data.

INTRODUCTION

Ovarian cancer is a gynaecological malignant tumour with the highest degree of malignancy and mortality[1]. Approximately 70%–80% of patients have advanced to the middle and late stage at the initial diagnosis owing to the asymptomatic nature of ovarian cancer in the early stage and lack of sensitive screening methods. In addition, 80% of patients with ovarian cancer experience a relapse within 1-2 years after surgery[2,3]. According to the statistics of the International Union of Obstetrics and Gynaecology, patients with ovarian cancer have a 5-year overall survival rate of < 40%, and a 5-year clinical stage IV survival rate of < 5%[4,5]. Currently, the first-line treatment for ovarian cancer is carboplatin combined with paclitaxel platinum chemotherapy following surgery, with a clinical remission rate of 60%–80%[6-8].

Advanced ovarian cancer is surgically treated with tumour cytoreductive surgery, which is the most effective surgical procedure[8]. Tumour cytoreductive surgery with a satisfactory tumour reduction ratio can prolong the survival time of patients and improve their overall survival rate. However, the scope of surgical resection includes not only ovaries, uterus and omentum but also pelvic and abdominal metastases and affected lymph nodes with a diameter of > 2 cm[9]. The operation is challenging, the injury obtained from the procedure is significant, and there are numerous complications since the procedure often involves the intestinal tract, the urinary tract and pelvic vessels. Additionally, some patients must undergo 2-3 courses of neoadjuvant chemotherapy prior to the surgery to have sufficient operation conditions[10]. The postoperative rehabilitation process and the quality of life of patients will suffer significantly from the high incidence of postoperative complications.

Urinary tract infection (UTI) is a common type of postoperative infection following tumour cell reduction surgery for ovarian cancer. It is related to the surgical procedure and the unique physiological structure of the female urinary tract[11,12]. Elderly patients with ovarian cancer have a higher incidence of postoperative UTI owing to their weak immune system, poor organ reserve capacity and a high proportion of basic diseases[13]. The evaluation of related factors is crucial for the prevention and management of nosocomial infection. However, there is no specific study on the related factors of UTI in elderly patients with ovarian cancer who underwent cytoreductive surgery at home and abroad.

Nowadays, predictive models based on advanced algorithms have been gradually applied to the medical field, which also enables many diseases to be detected and diagnosed early[14,15]. Among them, the machine learning (ML) algorithm relies on repeated iterative operations to accurately output the results, so it can improve the accuracy and robustness of prediction. Given the superior ability of the

ML-based algorithm to improve the accuracy of muscular invasion prediction, we applied the ML-assisted decision-support model to assess the risk of UTI using clinical parameters and direct clinical decision-making prior to treatment decisions.

MATERIALS AND METHODS

Patient selection
As the research subjects, 674 elderly patients with ovarian cancer who received treatment at the Department of Gynaecology at Jingzhou Central Hospital between January 31, 2016 and January 31, 2022 and met the inclusion criteria were selected. A retrospective analysis of the postoperative UTI and related factors was performed using medical records. The inclusion criteria for patients were as follows: (1) All patients met the diagnostic criteria in the clinical practice guidelines for ovarian cancer developed by the National Comprehensive Cancer Network and were diagnosed by imaging examination and postoperative pathology[16]; (2) All the patients were older than 60-years-old. The clinical stages were stage III and above, and the pelvic and abdominal masses were fixed; and (3) All the patients were scheduled for cytoreductive surgery. The clinical data were complete, and the postoperative hospital stay exceeded 5 d. The exclusion criteria for patients were as follows: (1) Patients undergoing secondary cytoreductive surgery for recurrent ovarian cancer; (2) Patients with liver and kidney insufficiency, cardiovascular and cerebrovascular accidents, blood diseases, autoimmune diseases or immunodeficiency diseases and other malignant tumours; (3) Patients diagnosed with acute and chronic infection prior to surgery; and (4) Patients who had long-term usage of immunosuppressants or glucocorticoids. The guidelines of the Helsinki Declaration (2013 revision) were followed by the study protocol. It was approved by the Institutional Review Committee of Jingzhou Central Hospital (JZ-2022014). Owing to its traceability, patient information was managed with the utmost confidentiality, and informed consent was waived. The workflow for patient selection and model construction is summarized in Figure 1.

Diagnosis of postoperative urinary tract infection
The diagnostic criteria were as follows: The patient had urinary tract irritation symptoms such as frequent micturition, urgency and pain following the surgery. By microscopic examination of the urine sediment, the average number of leukocytes per high-power visual field was ≥ 5, and the urine pathogen was present. Based on a diagnosis of UTI, patients were divided into an infection group and a non-infection group.

Data collection and quality assessment
The following data were collected from all patients: age, body mass index (BMI), catheter retention time, catheter intubation times, operation time, intraoperative blood loss, length of hospital stay, diabetes, hypertension, prophylactic use of antibiotics and postoperative hypoproteinaemia. In most cases, the median was applied to variables with missing values. A variable was excluded from variable screening for the final model if ≥ 10% of its values were missing.

### Development and validation of ML-based models
The data were randomly divided into a training set (70%) and a verification set (30%) to verify the prediction model. The inclusion principle of variables reported in previous studies was followed to screen variables. The principle of 'OOB error' was employed to screen the model variables (*i.e.* characteristic variables)[17], as follows: Gini (D) = $1 - \sum_{i=1}^{m} P_i^2$. If the Gini index was small, the probability of selecting mixed samples in the set was low, that is, the higher the purity of the set was and vice versa. However, the Gini index approaches zero if every sample in the set was of the same class. Based on the above algorithm principles, we have included five commonly used machine algorithm prediction models in this study, namely random forest classifier (RFC), support vector machine, extreme gradient boosting, artificial neural network (ANN) and decision tree (DT). Among them, RFC and DT are based on the algorithm principle of "branching and pruning," while ANN is based on "hidden layer" iteration. Support vector machine and extreme gradient boosting are also based on their iterative algorithm principle.

Prediction efficiency evaluation of ML-based models
The optimal subset variables for the modelling were obtained based on the intersection of variable sets. The receiver operating characteristic curve was used to evaluate the prediction accuracy of the model in the training and validation set. The discrimination ability of each model was quantified by the area under the receiver operating characteristic curve, decision curve analysis and clinical impact curve.

Statistical analysis
For descriptive analysis, median (interquartile range) and frequencies (%) were assessed for continuous
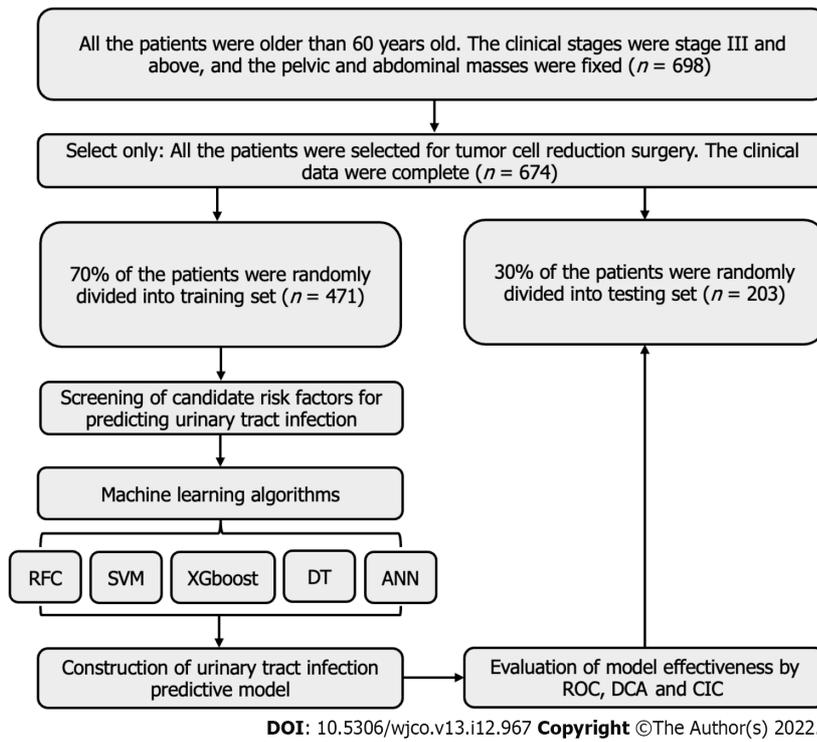
Figure 1 Flow chart of the patient selection and data process. ANN: Artificial neural network; CIC: Clinical impact curve; DCA: Decision curve analysis; DT: Decision tree; RFC: Random forest classifier; ROC: Receiver operating characteristic curve; SVM: Support vector machine; XGboost: Extreme gradient boosting.

and categorical variables, respectively. Bonferroni corrected probability values were used to compare the qualitative data[18]. Wilcoxon rank-sum test or χ^2 test was used to compare the differences between diverse groups. The best subset of randomly selected explanatory variables or features was used to further divide each node during the selecting process, and the class prediction values generated by each tree were collected. Finally, the candidate variables of the prediction model, namely the Gini index, were determined according to the weight. All analysis was performed using the Python programming language (version 3.9.2, Python Software Foundation, https://www.python.org/) and R project for statistical computing (version 4.0.4, http://www.r-project.org/). All P values were two-tailed, and $P <$ 0.05 was considered statistically significant.

RESULTS

Baseline characteristics of the study population

The comprehensive clinical features and baseline data of 674 elderly patients with ovarian cancer are presented in Table 1. Using the caret package, patients were randomly divided into a training set (70%, n = 471) and a validation set (30%, n = 203) for internal validation of the model. As presented in Table 1 and Supplementary Table 1, 96 patients had postoperative UTI, with an infection rate of 14.24%. The clinical symptoms and signs of the patients were primarily urinary tract irritation, fever, poor urination or urinary retention, renal percussion pain and urethral mouth itching. In addition, there were significant differences in catheter retention time, catheter intubation times, intraoperative bleeding, length of hospital stay, the proportion of patients with diabetes and the incidence of postoperative hypoproteinaemia ($P <$ 0.05) between the infection group and the non-infection group.

Selection of candidate variables

Feature selection is the aspect of ML that concentrates on selecting candidate variables[19]. The iterative analysis screened the candidate covariates of each algorithm. We executed 13 variables *via* Pearson correlation analysis. The correlation matrix revealed that UTIs significantly correlated with image factors and some clinical variables (Figure 2A). In addition, every significant candidate variable, such as age, BMI, catheter, catheter intubation times, blood loss, diabetes and hypoproteinaemia, contributed to the ML-based model (Figure 2B). These seven were the top predictors, which were consistent with the findings of the correlation analysis.

**Table 1 Baseline demographic and clinicopathological characteristics of patients, *n* (%)**

Variables	Training set				Testing set			
	Overall, *n* = 471	Yes, *n* = 70	No, *n* = 401	*P* value	Overall, *n* = 203	Yes, *n* = 26	No, *n* = 177	*P* value
Age [median (IQR)], yr	64.00 (63.00, 66.00)	69.00 (67.00, 71.00)	64.00 (62.00, 65.00)	< 0.001	64.00 (62.00, 65.50)	68.50 (65.25, 70.75)	63.00 (62.00, 65.00)	< 0.001
BMI [median (IQR)], kg/m²	23.00 (22.00, 24.00)	24.00 (23.00, 25.00)	23.00 (22.00, 24.00)	< 0.001	23.00 (22.00, 24.00)	25.00 (23.00, 26.00)	23.00 (22.00, 24.00)	< 0.001
Catheter [median (IQR)], d	8.00 (7.00, 10.00)	13.00 (10.00, 14.00)	8.00 (7.00, 9.00)	< 0.001	8.00 (7.00, 9.00)	13.00 (11.00, 13.00)	8.00 (6.00, 9.00)	< 0.001
Catheter intubation times								
≥ 3	148 (31.4)	49 (70.0)	99 (24.7)	< 0.001	59 (29.1)	23 (88.5)	36 (20.3)	< 0.001
< 3	323 (68.6)	21 (30.0)	302 (75.3)		144 (70.9)	3 (11.5)	141 (79.7)	
Operation time [median (IQR)], h	3.60 (2.90, 4.40)	3.80 (3.10, 4.50)	3.60 (2.80, 4.30)	0.061	3.70 (2.90, 4.60)	3.85 (2.93, 4.60)	3.70 (2.90, 4.50)	0.373
Blood loss [median (IQR)], mL	476.00 (434.50, 515.50)	627.00 (592.75, 658.75)	465.00 (429.00, 499.00)	< 0.001	470.00 (432.00, 504.00)	646.00 (616.25, 670.00)	461.00 (426.00, 494.00)	< 0.001
Hospitalization [median (IQR)], d	11.00 (9.00, 13.00)	15.00 (14.00, 17.00)	10.00 (8.00, 12.00)	< 0.001	10.00 (8.00, 12.00)	16.00 (15.00, 17.00)	10.00 (8.00, 12.00)	< 0.001
Diabetes								
Yes	162 (34.4)	57 (81.4)	105 (26.2)	< 0.001	59 (29.1)	22 (84.6)	37 (20.9)	< 0.001
No	309 (65.6)	13 (18.6)	296 (73.8)		144 (70.9)	4 (15.4)	140 (79.1)	
Hypertension								
Yes	283 (60.1)	45 (64.3)	238 (59.4)	0.519	132 (65.0)	19 (73.1)	113 (63.8)	0.483
No	188 (39.9)	25 (35.7)	163 (40.6)		71 (35.0)	7 (26.9)	64 (36.2)	
Antibiotics								
Yes	295 (62.6)	47 (67.1)	248 (61.8)	0.477	128 (63.1)	16 (61.5)	112 (63.3)	1
No	176 (37.4)	23 (32.9)	153 (38.2)		75 (36.9)	10 (38.5)	65 (36.7)	
Hypoproteinaemia								
Yes	122 (25.9)	53 (75.7)	69 (17.2)	< 0.001	54 (26.6)	22 (84.6)	32 (18.1)	< 0.001
No	349 (74.1)	17 (24.3)	332 (82.8)		149 (73.4)	4 (15.4)	145 (81.9)	
NACT								
Yes	293 (62.2)	18 (25.7)	275 (68.6)	< 0.001	133 (65.5)	12 (46.2)	121 (68.4)	0.045
No	178 (37.8)	52 (74.3)	126 (31.4)		70 (34.5)	14 (53.8)	56 (31.6)	

BMI: Body mass index; IQR: Interquartile range; NACT: Neoadjuvant chemotherapy.

Construction of ML-based UTI predictive model

Positive or negative training results for each patient were entered for training data, and the final judgment result was the output, as indicated in the following formula: Gini(D)= $1-\sum_{(i = 1)}^{m} P\_i^2$. The RFC algorithm represents a computational method for effectively navigating the free parameter space to obtain a robust model (Figure 3A). The variable Gini index in the RFC model is presented in Supplementary Table 2. The top seven candidate variables were age, BMI, catheter, catheter intubation times, blood loss, diabetes and hypoproteinaemia, which were consistent with the predicted results. In addition, data mining through the DT model, as demonstrated by impurity analysis: Gini (p) = $\sum_{(K = 1)}^{K} [Pk(1\text{-}Pk)]$, was advantageous. At the branch of DT, age and catheter functioned as the irreplaceable weight in addition to clinical factor indicators (Figure 3B). In contrast, the RFC model outperformed the ANN model, which outperformed other models, in terms of prediction efficiency (Figure 4 and Supplementary Table 3).

A

B
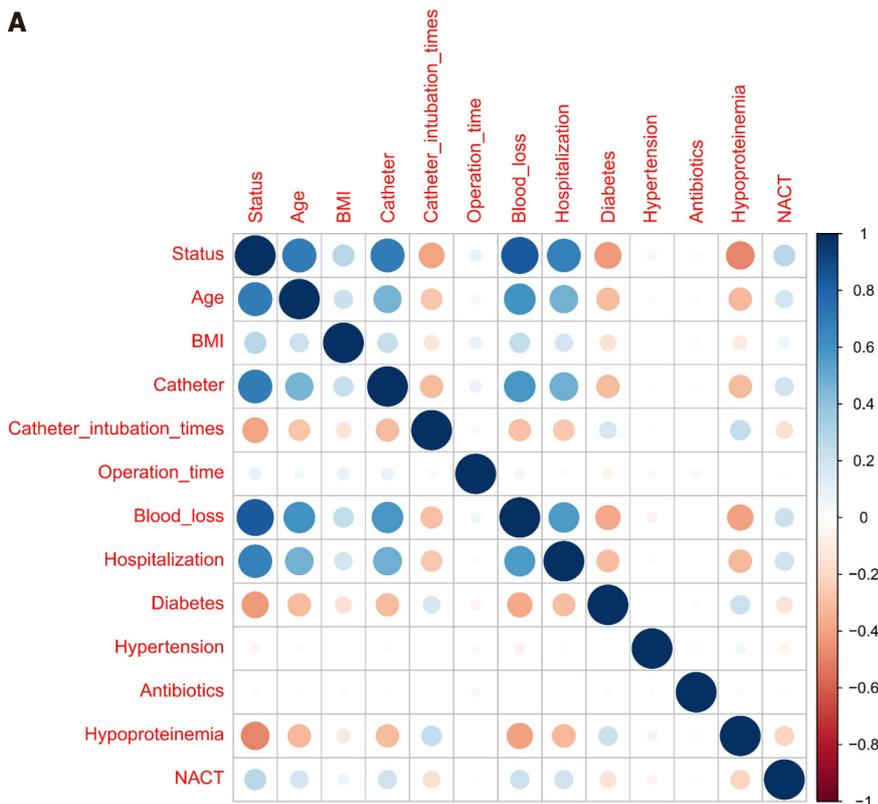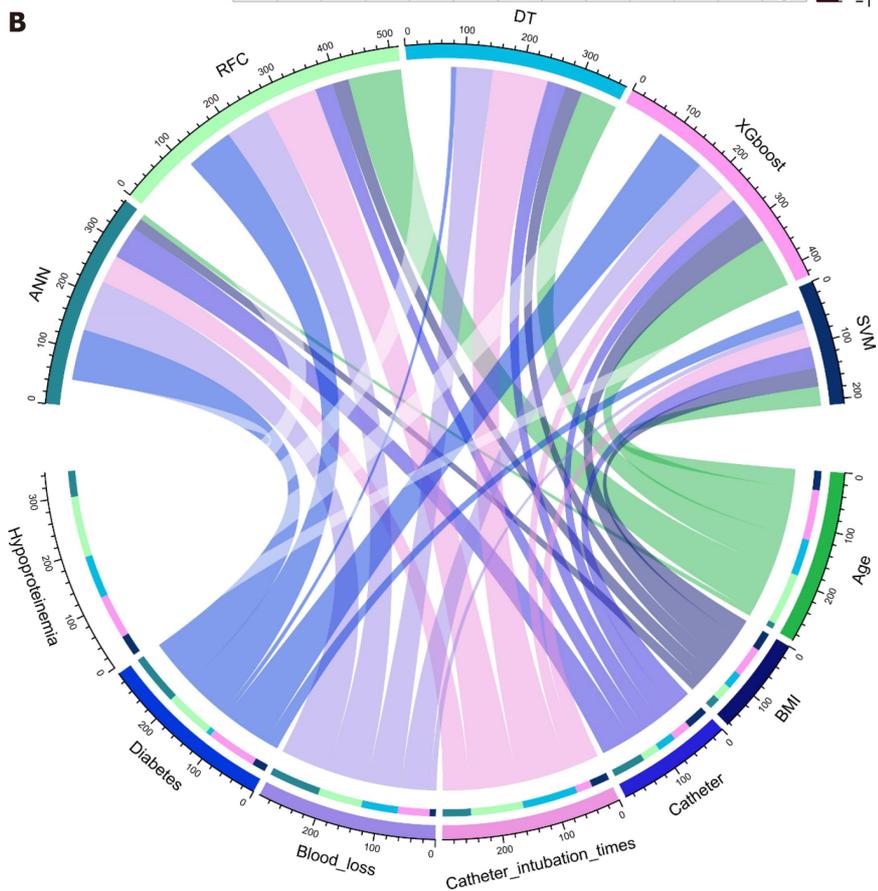
Figure 2 Variable screening and weight allocation. A: Variable screening; B: weight allocation. ANN: Artificial neural network; BMI: Body mass index; DT: Decision tree; NACT: Neoadjuvant chemotherapy; RFC: Random forest classifier; SVM: Support vector machine; XGboost: Extreme gradient boosting.

Comparison across ML-based models

We used five supervised learning models for UTI assessment to investigate whether ML-based models

A

B
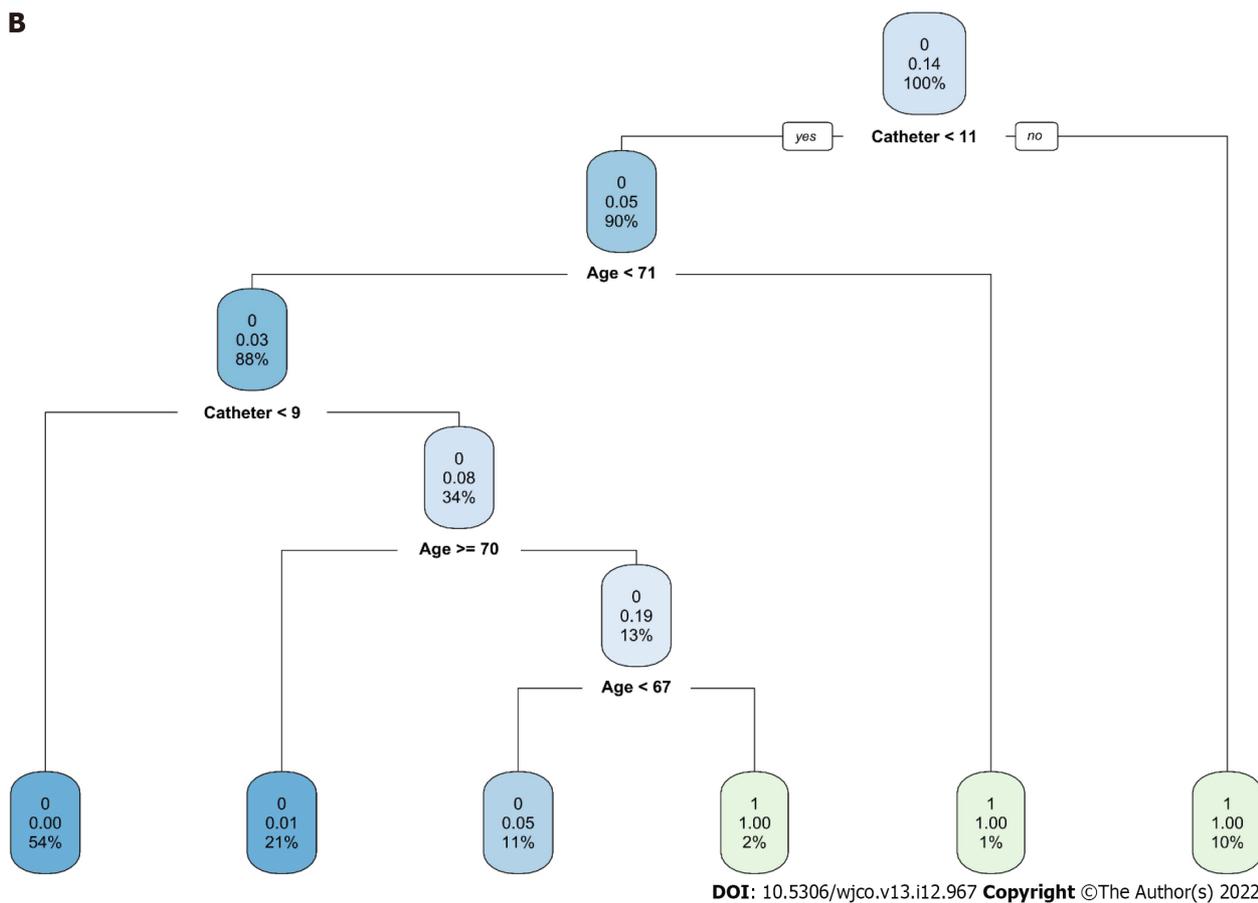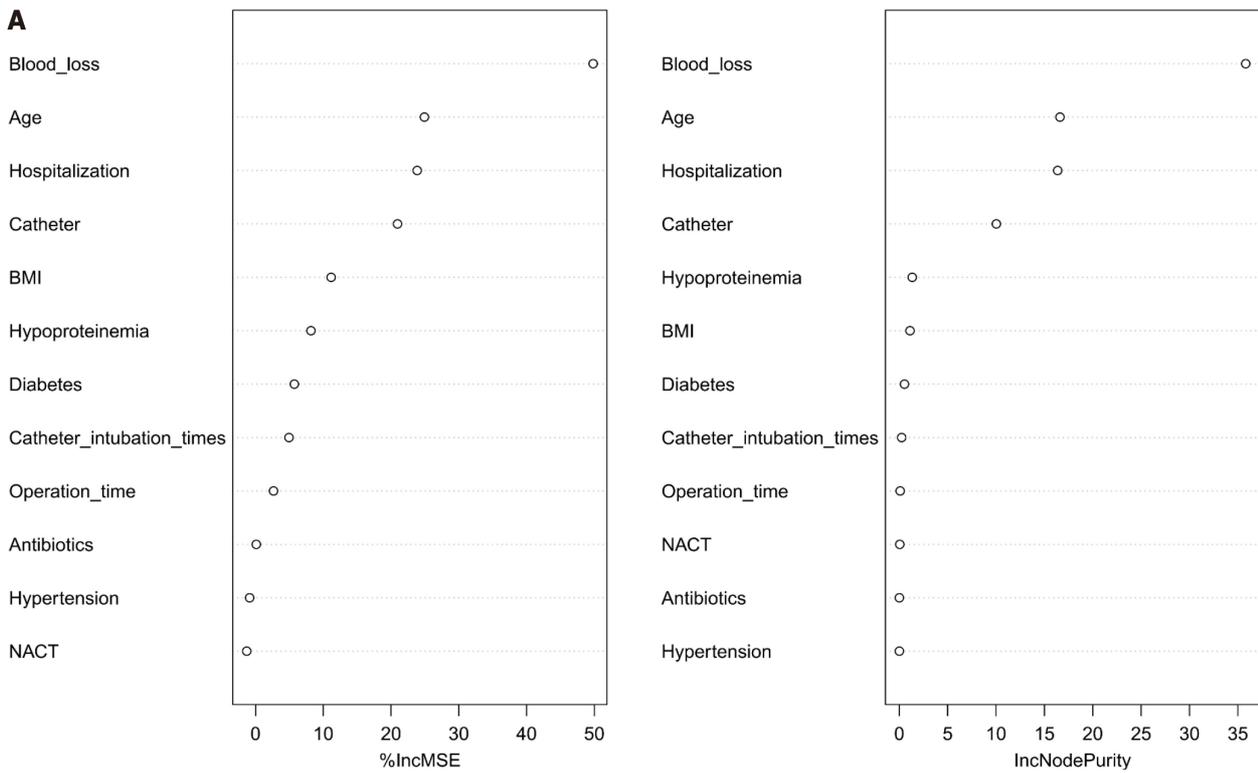
Figure 3 Predictive model visualization based on machine learning-based algorithm. A: The random forest classifier algorithm represents a computational method for effectively navigating the free parameter space to obtain a robust model; B: At the branch of decision tree, age and catheter functioned as the irreplaceable weight in addition to clinical factor indicators. BMI: Body mass index; NACT: Neoadjuvant chemotherapy.

can improve prediction performance. The RFC model demonstrated a strong prediction performance in the training and validation cohorts based on decision curve analysis (Figure 5). In addition, the area under the curve of the RFC models peaked when the seven variables were added, followed by those of
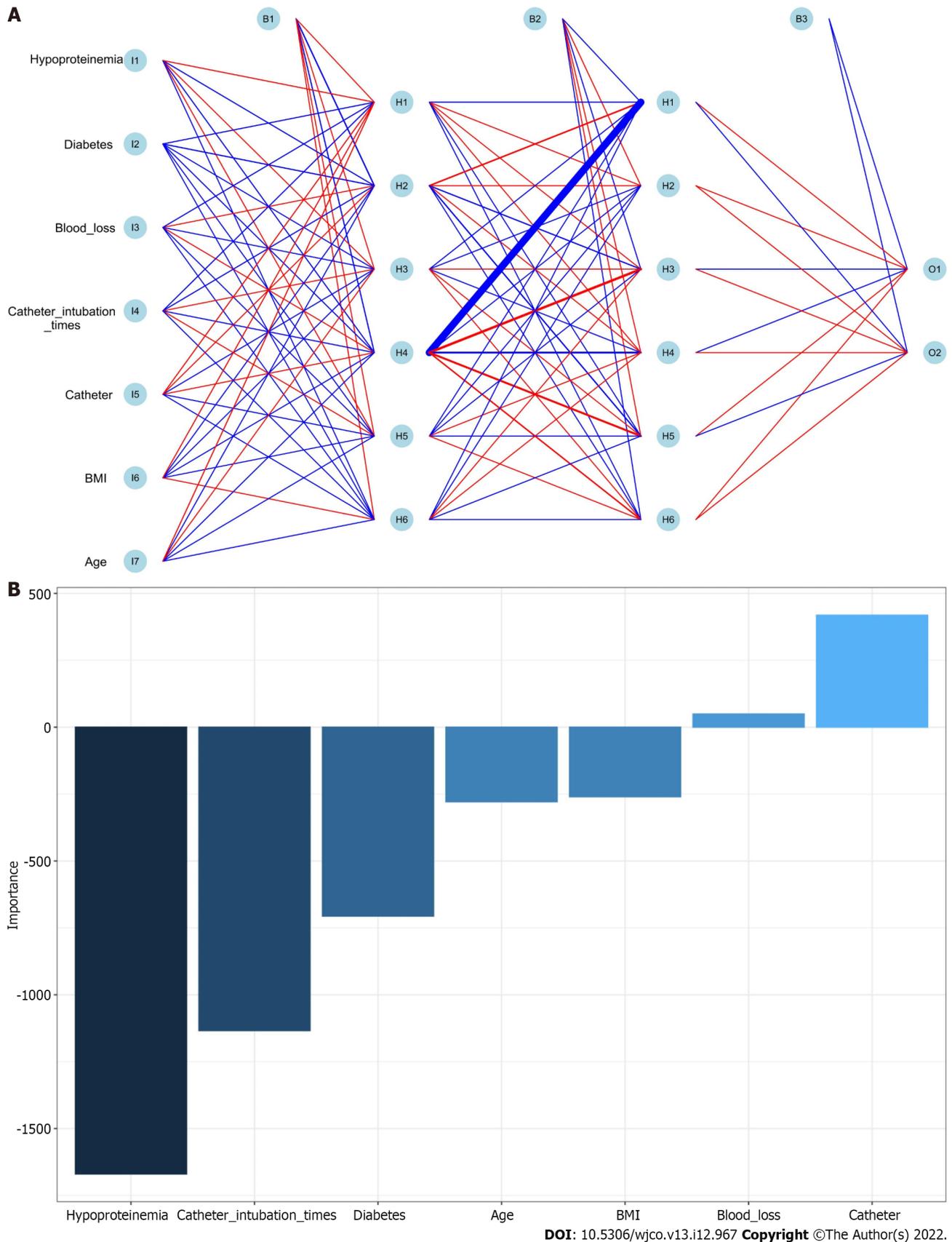
Figure 4 Predictive model visualization based on artificial neural network algorithm. BMI: Body mass index.

ANN, DT, support vector machine and extreme gradient boosting (Table 2 and Supplementary Table 4). Undoubtedly, RFC outperformed the generalised linear model in terms of prediction accuracy. Thus, both RFC and DT (ML-assisted decision-support) models were used to guide UTI prediction using the iterative algorithm analysis of supervised learning.

Table 2 Receiver operating characteristic curve analyses for predicting urinary tract infection in each machine learning-based model

Model	Training set			Testing set		
	AUC mean	AUC 95%CI	Variables[1]	AUC mean	AUC 95%CI	Variables[1]
RFC	0.925	0.868-0.982	7	0.918	0.861-0.975	7
SVM	0.787	0.730-0.844	7	0.779	0.722-0.836	7
DT	0.776	0.719-0.833	7	0.769	0.712-0.826	7
ANN	0.879	0.822-0.936	6	0.854	0.797-0.911	6
XGboost	0.797	0.740-0.854	7	0.788	0.731-0.845	7

[1]Variables included in the model. ANN: Artificial neutral network; AUC: Area under the curve; CI: Confidence interval; DT: Decision tree; RFC: Random forest classifier; SVM: Support vector machine; XGboost: Extreme gradient boosting.

### *Internal validation of the optimal predictive model*

We also used the clinical impact curve to assess the accuracy to further validate the RFC model's ability to predict outcomes. The clinical impact curve revealed that UTI stratification was achieved in the training cohorts (Supplementary Figure 1). These were consistent with the results of validation cohorts, indicating that RFC performed best in terms of discrimination, calibration and overall performance, in particular the candidate systemic inflammation markers that were highly relevant to UTIs.

DISCUSSION

This study's findings indicated that the factors influencing the risk of UTI in elderly patients with ovarian cancer after tumour cell reduction include not only the patients' basic diseases but also their indwelling catheter and postoperative nutritional level. There has been a lack of specialised research on UTIs after tumour cell reduction in recent years, and some researchers have examined the operation or resection of patients with ovarian cancer. These findings demonstrated that the most common clinical manifestations of patients with UTI are urinary tract irritation, urinary retention, urethral mouth itching and urine turbidity, which is consistent with the results of a previous study[20]. However, the incidence of UTIs reported in these studies ranges from < 10% to > 40%. The infection rate reported in this study was 14.24%, which is considered moderate; this may be due to the exclusion of patients with severe basic diseases such as liver and kidney dysfunction[21,22]. In addition, concerning infection-related factors, these reports have drawn similar conclusions as this study, albeit they also stated that age, intubation times, length of hospital stay, paraaortic lymph node dissection and intestinal resection, haemoglobin and other factors can all affect the risk of infection.

In general, cytoreductive surgery is a relatively traumatic procedure for advanced ovarian cancer, and the scope of the operation is likely to involve the urinary system, causing a significant increase in the risk of postoperative UTI. Simultaneously, the risk factors for postoperative UTI in elderly patients differ from those in young and middle-aged patients. Therefore, the research objects with high heterogeneity are selected for analysis, and the demonstrability and repeatability of the results are insufficient. Furthermore, a thematic analysis for the patient population with specific surgical procedures and similar conditions and susceptibility factors should be performed. Considering this situation, this study included elderly patients with ovarian cancer who had undergone cytoreductive surgery as the research. Our findings indicated that actively controlling catheter-related UTIs and correcting postoperative malnutrition were important links to preventing and controlling UTIs in the elderly after ovarian cancer cell reduction.

Through clinical observation, researchers have discovered that postoperative UTI caused by an indwelling catheter is one of the most common postoperative infections in clinics in recent years. The operation, catheter selection, bladder flushing and patient factors are the main causes of infection, and catheter placement time, difficulty in catheter intubation, multiple intubations, previous catheter retention history, long anaesthesia time, history of diabetes, age, consciousness disorder improper bladder flushing, gastrointestinal decompression, enema, long replacement time of urine collection bag and other factors are related to postoperative catheter-related UTI[23]. Numerous studies have demonstrated that the time of catheter placement is an independent risk factor for postoperative UTI. Thus, to effectively prevent catheter-related UTI, it is necessary to strictly control the indications of a long-term indwelling catheter, reduce urethral injury, improve the skills of operators, prevent retrograde infection, improve the tightness of the catheter system and ensure the patency of the catheter system in clinical work.
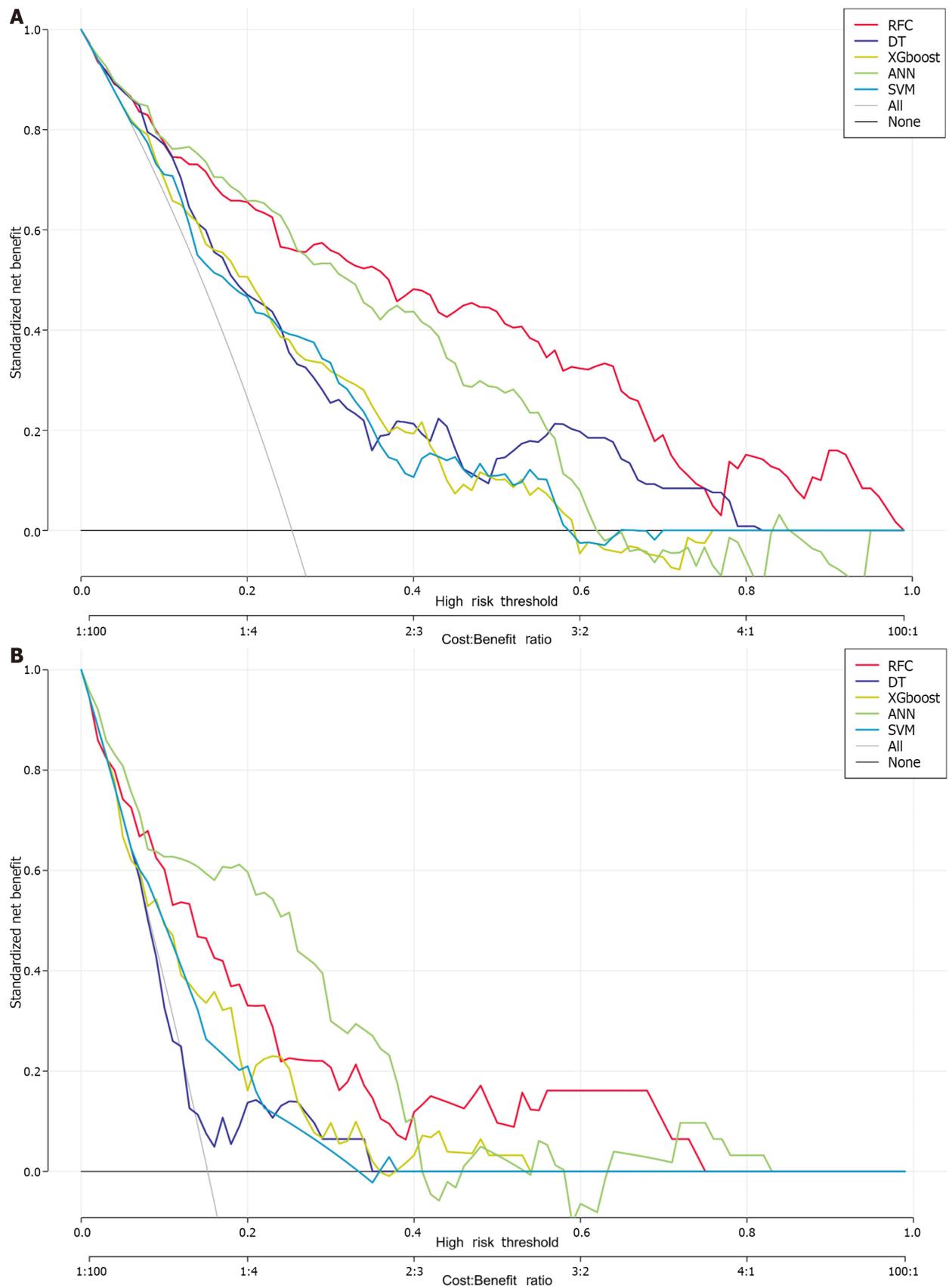
Figure 5 Prediction performance of candidate models based on machine learning-based algorithm. ANN: Artificial neural network; DT: Decision tree; RFC: Random forest classifier; SVM: Support vector machine; XGboost: Extreme gradient boosting.

One of the most common complications after major surgery is hypoproteinaemia. Its causes are complex, and it is closely linked to surgical trauma. Operation post-stress is related to infection and other factors[24,25]. Simultaneously, plasma albumin level in severe patients is correlated with the expression of serum inflammatory factors and peripheral blood T cell subsets. Hypoproteinaemia can improve the degree of inflammatory stress, cause immune dysfunction, significantly increase the risk of bacterial and fungal infection and have a serious adverse impact on the disease outcome[26,27]. In patients undergoing surgery for a malignant tumour or organ function decompensation, postoperative hypoproteinaemia can increase the incidence and mortality of complications, such as postoperative infection, and seriously impact the surgical efficacy. The incidence of postoperative complications significantly decreases as plasma albumin levels rise. Therefore, early postoperative nutritional support is an important link in the treatment and prevention of infection in elderly patients with ovarian cancer. Early intravenous nutrition support should be strengthened for patients who cannot use early enteral nutrition to correct the negative nitrogen balance caused by surgical stress and maintain normal nutrition levels in particular.

It is noteworthy that, owing to the limitations of clinical medical records, the risk factors associated with cytoreductive surgery for ovarian cancer in the elderly examined in this study are not comprehensive, which is a flaw in this study. In addition to the relevant factors analysed in this study, ascites volume, operation scope and other factors will also affect it. Laparoscopic cytoreductive surgery for advanced ovarian cancer is becoming more prevalent as laparoscopic technology advances, which reduces the risk of surgical trauma and postoperative infection to some extent. These factors must be researched and analysed further.

CONCLUSION

In conclusion, using an ML-based algorithm, we created a feasible and robust method for identifying factors important for predicting UTIs. The RFC in particular, which can improve the prediction and early detection of UTIs in patients with ovarian cancer, was robust. In addition, age, BMI, catheter, catheter intubation times, blood loss, diabetes and hypoproteinaemia were five crucial factors. In clinical practice, incorporating the presentation of simple clinical data may be helpful for clinicians to identify the individualised risk of UTI.

ARTICLE HIGHLIGHTS

Research background
Nowadays, predictive models based on advanced algorithms have been gradually applied to the medical field, which also enables many diseases to be detected and diagnosed early. Among them, the machine learning (ML) algorithm relies on repeated iterative operations to accurately output the results. Therefore, it can improve the accuracy and robustness of prediction.

Research motivation
Given the superior ability of the ML-based algorithm to improve the accuracy of muscular invasion prediction, we applied the ML-assisted decision-support model to assess the risk of urinary tract infection (UTI) using clinical parameters and direct clinical decision-making prior to treatment decisions.

Research objectives
We developed an ML assistant model for the prevention and control of nosocomial infection.

Research methods
A total of 674 elderly patients with ovarian cancer treated between January 31, 2016 and January 31, 2022 and met the inclusion criteria of the study were selected as the research subjects. A retrospective analysis of the postoperative UTI and related factors was performed by reviewing the medical records. Five ML-assisted models were developed using two-step estimation methods from the candidate predictive variables. The robustness and clinical applicability of each model were assessed using the receiver operating characteristic curve, decision curve analysis and clinical impact curve.

Research results
A total of 12 candidate variables were eventually included in the UTI prediction model. Models constructed using the random forest classifier (RFC), support vector machine, extreme gradient boosting, artificial neural network and decision tree had areas under the receiver operating characteristic curve ranging from 0.776 to 0.925. The RFC model, which incorporated factors such as age, body

mass index, catheter, catheter intubation times, blood loss, diabetes and hypoproteinaemia, had the highest predictive accuracy.

Research conclusions

These findings demonstrated that the ML-based prediction model developed using the RFC can be used to identify elderly patients with ovarian cancer who may have postoperative UTI. This can help with treatment decisions and enhance clinical outcomes.

Research perspectives

Using an ML-based algorithm, we developed a feasible and robust method to identify factors that are significant for predicting UTIs. The RFC, which can improve the prediction and early detection of UTIs in patients with ovarian cancer, was particularly robust. In addition, the five most crucial factors were age, body mass index, catheter, catheter intubation times, blood loss, diabetes and hypoproteinaemia. Clinicians may find it extremely helpful to assess the individualised risk of UTI in clinical practice by incorporating the presentation of simple clinical data.

ACKNOWLEDGEMENTS

FOOTNOTES

REFERENCES

1   **Webb PM**, Jordan SJ. Epidemiology of epithelial ovarian cancer. *Best Pract Res Clin Obstet Gynaecol* 2017; **41**: 3-14 [PMID: 27743768 DOI: 10.1016/j.bpobgyn.2016.08.006]
2   **Rooth C**. Ovarian cancer: risk factors, treatment and management. *Br J Nurs* 2013; **22**: S23-S30 [PMID: 24067270 DOI: 10.12968/bjon.2013.22.Sup17.S23]
3   **Nash Z**, Menon U. Ovarian cancer screening: Current status and future directions. *Best Pract Res Clin Obstet Gynaecol* 2020; **65**: 32-45 [PMID: 32273169 DOI: 10.1016/j.bpobgyn.2020.02.010]
4   **Dinkelspiel HE**, Champer M, Hou J, Tergas A, Burke WM, Huang Y, Neugut AI, Ananth CV, Hershman DL, Wright JD. Long-term mortality among women with epithelial ovarian cancer. *Gynecol Oncol* 2015; **138**: 421-428 [PMID: 26050923

DOI: 10.1016/j.ygyno.2015.06.005]

5   **Wright JD**, Ananth CV, Tsui J, Glied SA, Burke WM, Lu YS, Neugut AI, Herzog TJ, Hershman DL. Comparative effectiveness of upfront treatment strategies in elderly women with ovarian cancer. *Cancer* 2014; **120**: 1246-1254 [PMID: 24443159 DOI: 10.1002/cncr.28508]

6   **Orr B**, Edwards RP. Diagnosis and Treatment of Ovarian Cancer. *Hematol Oncol Clin North Am* 2018; **32**: 943-964 [PMID: 30390767 DOI: 10.1016/j.hoc.2018.07.010]

7   **Friedrich M**, Friedrich D, Kraft C, Rogmans C. Multimodal Treatment of Primary Advanced Ovarian Cancer. *Anticancer Res* 2021; **41**: 3253-3260 [PMID: 34230119 DOI: 10.21873/anticanres.15111]

8   **Lheureux S**, Braunstein M, Oza AM. Epithelial ovarian cancer: Evolution of management in the era of precision medicine. *CA Cancer J Clin* 2019; **69**: 280-304 [PMID: 31099893 DOI: 10.3322/caac.21559]

9   **Moran BJ**. Cytoreductive surgery and hyperthermic intraperitoneal chemotherapy for colorectal peritoneal metastases: from novelty to routine in selected cases. *Tech Coloproctol* 2017; **21**: 767-769 [PMID: 28936563 DOI: 10.1007/s10151-017-1679-8]

10   **Knutson KL**, Maurer MJ, Preston CC, Moysich KB, Goergen K, Hawthorne KM, Cunningham JM, Odunsi K, Hartmann LC, Kalli KR, Oberg AL, Goode EL. Regulatory T cells, inherited variation, and clinical outcome in epithelial ovarian cancer. *Cancer Immunol Immunother* 2015; **64**: 1495-1504 [PMID: 26298430 DOI: 10.1007/s00262-015-1753-x]

11   **Tew WP**. Ovarian cancer in the older woman. *J Geriatr Oncol* 2016; **7**: 354-361 [PMID: 27499341 DOI: 10.1016/j.jgo.2016.07.008]

12   **Matsuo K**, Prather CP, Ahn EH, Eno ML, Tierney KE, Yessaian AA, Im DD, Rosenshein NB, Roman LD. Significance of perioperative infection in survival of patients with ovarian cancer. *Int J Gynecol Cancer* 2012; **22**: 245-253 [PMID: 22228425 DOI: 10.1097/IGC.0b013e31823bd6db]

13   **Cracco A**, Roy M, Simpfendorfer CH. Cytoreductive surgery and hyperthermic intraperitoneal chemotherapy combined with two-stage hepatectomy for multiple and bilobar desmoplastic small round cell tumor liver metastases. *J Gastrointest Oncol* 2017; **8**: E60-E64 [PMID: 28890830 DOI: 10.21037/jgo.2017.07.07]

14   **Martins CB**, De Bels D, Honore PM, Redant S. Early Prediction of Acute Kidney Injury by Machine Learning: Should We Add the Urine Output Criterion to Improve this New Tool? *J Transl Int Med* 2020; **8**: 201-202 [PMID: 33511045 DOI: 10.2478/jtim-2020-0031]

15   **Tulsi R**, Ul Haque MM, Hanif FM, Devi A, Mubarak M, Hassan Luck N. Metastasis of Duodenal Adenocarcinoma to the Urinary Bladder Presenting as Hematuria. *J Transl Int Med* 2021; **9**: 143-145 [PMID: 34497753 DOI: 10.2478/jtim-2021-0010]

16   **Morgan RJ Jr**, Armstrong DK, Alvarez RD, Bakkum-Gamez JN, Behbakht K, Chen LM, Copeland L, Crispens MA, DeRosa M, Dorigo O, Gershenson DM, Gray HJ, Hakam A, Havrilesky LJ, Johnston C, Lele S, Martin L, Matulonis UA, O'Malley DM, Penson RT, Percac-Lima S, Pineda M, Plaxe SC, Powell MA, Ratner E, Remmenga SW, Rose PG, Sabbatini P, Santoso JT, Werner TL, Burns J, Hughes M. Ovarian Cancer, Version 1.2016, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2016; **14**: 1134-1163 [PMID: 27587625 DOI: 10.6004/jnccn.2016.0122]

17   **Fan J**, Lv J. A Selective Overview of Variable Selection in High Dimensional Feature Space. *Stat Sin* 2010; **20**: 101-148 [PMID: 21572976]

18   **Armstrong RA**. When to use the Bonferroni correction. *Ophthalmic Physiol Opt* 2014; **34**: 502-508 [PMID: 24697967 DOI: 10.1111/opo.12131]

19   **Deo RC**. Machine Learning in Medicine. *Circulation* 2015; **132**: 1920-1930 [PMID: 26572668 DOI: 10.1161/CIRCULATIONAHA.115.001593]

20   **Pavlov MJ**, Ceranic MS, Latincic SM, Sabljak PV, Kecmanovic DM, Sugarbaker PH. Cytoreductive surgery and hyperthermic intraperitoneal chemotherapy for the treatment of advanced epithelial and recurrent ovarian carcinoma: a single center experience. *Int J Hyperthermia* 2018; **34**: 564-569 [PMID: 28877609 DOI: 10.1080/02656736.2017.1371341]

21   **Filippova OT**, Kim SW, Cowan RA, Chi AJ, Iasonos A, Zhou QC, Broach V, Zivanovic O, Long Roche K, Sonoda Y, Gardner G, Chi DS. Hematologic changes after splenectomy for ovarian cancer debulking surgery, and association with infection and venous thromboembolism. *Int J Gynecol Cancer* 2020; **30**: 1183-1188 [PMID: 32665236 DOI: 10.1136/ijgc-2020-001368]

22   **Kirmusaoglu S**, Yurdugül S, Metin A, Vehid S. The Effect of Urinary Catheters on Microbial Biofilms and Catheter Associated Urinary Tract Infections. *Urol J* 2017; **14**: 3028-3034 [PMID: 28299769]

23   **Redant S**, Nehar-Stern N, Honoré PM, Attou R, Haggenmacher C, Tolwani A, De Bels D, Biarent D. Acute Bronchiolitis: Why Put an IV Line? *J Transl Int Med* 2021; **9**: 185-189 [PMID: 34900629 DOI: 10.2478/jtim-2021-0013]

24   **Premuzic V**, Hudolin T, Pasini J, Zimak Z, Hauptman D, Jelakovic B, Kastelan Z. Hypoproteinemia as a prognostic risk factor for arteriovenous fistula failure. *Hemodial Int* 2018; **22**: 37-44 [PMID: 28133903 DOI: 10.1111/hdi.12538]

25   **Lorente L**. Antimicrobial-impregnated catheters for the prevention of catheter-related bloodstream infections. *World J Crit Care Med* 2016; **5**: 137-142 [PMID: 27152256 DOI: 10.5492/wjccm.v5.i2.137]

26   **Li F**, Yuan MZ, Wang L, Wang XF, Liu GW. Characteristics and prognosis of pulmonary infection in patients with neurologic disease and hypoproteinemia. *Expert Rev Anti Infect Ther* 2015; **13**: 521-526 [PMID: 25724113 DOI: 10.1586/14787210.2015.1019471]

27   **Deng J**, Chen X, Sun H, Liu Y, Li W, Chen B, Zhao S, Jia K, Wang H, Guo H, Jiang M, Xu Y, He Y, Zhou C. Hypoproteinemia being a manifestation of immunotherapy-related liver dysfunction. *Ann Transl Med* 2020; **8**: 889 [PMID: 32793733 DOI: 10.21037/atm-20-4980]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com