

RESEARCH

Open Access



Development and validation of a prognostic prediction model for elderly gastric cancer patients based on oxidative stress biochemical markers

Xing-Qi Zhang^{1,2,3†}, Ze-Ning Huang^{1,2,3†}, Ju Wu^{1,4}, Chang-Yue Zheng^{1,5}, Xiao-Dong Liu⁶, Ying-Qi Huang^{1,2,3}, Qi-Yue Chen^{1,2,3}, Ping Li^{1,2,3}, Jian-Wei Xie^{1,2,3}, Chao-Hui Zheng^{1,2,3}, Jian-Xian Lin^{1,2,3*}, Yan-Bing Zhou^{6*} and Chang-Ming Huang^{1,2,3*}

Abstract

Background The potential of the application of artificial intelligence and biochemical markers of oxidative stress to predict the prognosis of older patients with gastric cancer (GC) remains unclear.

Methods This retrospective multicenter study included consecutive patients with GC aged ≥ 65 years treated between January 2012 and April 2018. The patients were allocated into three cohorts (training, internal, and external validation). The GC-Integrated Oxidative Stress Score (GIOSS) was developed using Cox regression to correlate biochemical markers with patient prognosis. Predictive models for five-year overall survival (OS) were constructed using random forest (RF), decision tree (DT), and support vector machine (SVM) methods, and validated using area under the curve (AUC) and calibration plots. The SHapley Additive exPlanations (SHAP) method was used for model interpretation.

Results This study included a total of 1,859 older patients. The results demonstrated that a low GIOSS was a predictor of poor prognosis. RF was the most efficient method, with AUCs of 0.999, 0.869, and 0.796 in the training, internal validation, and external validation sets, respectively. The DT and SVM models showed low AUC values. Calibration and decision curve analyses demonstrated the considerable clinical usefulness of the RF model. The SHAP results identified pN, pT, perineural invasion, tumor size, and GIOSS as key predictive features. An online web calculator was constructed based on the best model.

[†]Xing-Qi Zhang and Ze-Ning Huang contributed equally to this work and should be considered co-first authors.

*Correspondence:

Jian-Xian Lin

linjian379@163.com

Yan-Bing Zhou

zhouyanbing@qduhospital.cn

Chang-Ming Huang

hcmlr2002@163.com

Full list of author information is available at the end of the article



Conclusions Incorporating the GIOSS, the RF model effectively predicts postoperative OS in older patients with GC and is a robust prognostic tool. Our findings emphasize the importance of oxidative stress in cancer prognosis and provide a pathway for improved management of GC.

Trial registration Retrospectively registered at ClinicalTrials.gov (trial registration number: NCT06208046, date of registration: 2024-05-01).

Keywords Elderly, Gastric cancer, Oxidative stress, Machine learning, Overall survival

Background

Gastric cancer (GC), which is prevalent and lethal among older individuals, is a global health concern [1]. Various complex factors influence the prognosis of older patients with GC; thus, traditional clinical and pathological metrics such as the tumor-node-metastasis (TNM) staging system are less effective [2, 3]. Conventional methods have failed to address the intricacies of this disease, particularly the variations in individual biological characteristics. Systemic changes in the levels of biochemical markers including albumin (ALB), bilirubin, and uric acid (UA) are increasingly recognized for their significant roles in GC development, progression, and prognosis, as well as their correlation with degenerative aging processes [4–7]. These biochemical markers that indicate systemic oxidative stress, are hypothesized to have substantial value for predicting the prognosis of older patients with GC [8, 9].

Oxidative stress is caused by the negative effects of free radicals in the body and is an important cause of aging and disease [10, 11]. Impaired antioxidant systems frequently contribute to cancer development and elevated oxidative stress exacerbates cancer risk [9]. Conversely, mitigation of oxidative stress can help combat malignancies, highlighting the importance of oxidative regulation in cancer prevention and treatment. Decreased cellular antioxidant enzyme activity renders older patients with GC more susceptible to oxidative stress [12], leading to DNA, protein, and lipid damage. This decrease exacerbates cellular aging and disease progression, including cancer [4, 13]. Moreover, oncological treatments such as radiotherapy and chemotherapy can amplify oxidative stress, posing additional risks to older patients by intensifying oxidative damage [14].

Traditional modeling approaches often inadequately process and interpret extensive bioinformatic data, underscoring the need for innovative methods to enhance the prediction of GC prognosis by considering a wider array of biomarkers and patient health profiles. Supervised machine learning, a subset of artificial intelligence, is increasingly favored because of its ability to discern complex, nonlinear patterns in large, varied datasets [15–17].

Given the significance of oxidative stress markers in GC, in the present study, we identified key biochemical markers and developed the novel GC-Integrated Oxidative Stress Score (GIOSS). We examined the prognostic utility of the GIOSS and established three machine learning prognostic models based on the GIOSS to predict the five-year survival rate. Finally, we validated the model using an independent cohort study. The objective of this study was to develop a machine learning model to predict five-year postoperative survival and contribute to improved clinical decision-making. A model with strong predictive ability can provide individualized survival predictions for older patients with GC based on their clinicopathological data, enabling the implementation of targeted follow-up strategies.

Methods

Patient selection and research methodology

This multicenter, retrospective cohort study included four tertiary hospitals in China and 1,859 older patients with GC. These patients underwent radical gastrectomy and were registered in the gastric surgery databases across the participating centers between January 2012 and April 2018 (Fig. 1). The inclusion criteria were: (1) biopsy-proven gastric adenocarcinoma, (2) age at diagnosis ≥ 65 years, (3) radical surgical resection, and (4) availability of complete clinical and pathological data. The exclusion criteria were: (1) postoperative pathology confirming non-gastric primary tumors, (2) distant metastasis, (3) incomplete clinical data, and (4) the detection of other concurrent malignancies within five years. We employed a random sampling strategy to allocate patients from one of the clinical centers into training and internal validation sets at an 8:2 ratio ($n = 1,367$) [18, 19]. In particular, the 'sample' function was used to randomly select 80% of the samples as the training set for the development and enhancement of three machine learning models. The remaining 20% formed the internal validation cohort, which was used to evaluate the models and fine-tune the hyperparameters. To compare the performances of the machine learning models, we conducted an external validation using a validation dataset from three other clinical centers ($n = 492$). Event occurrence and censoring

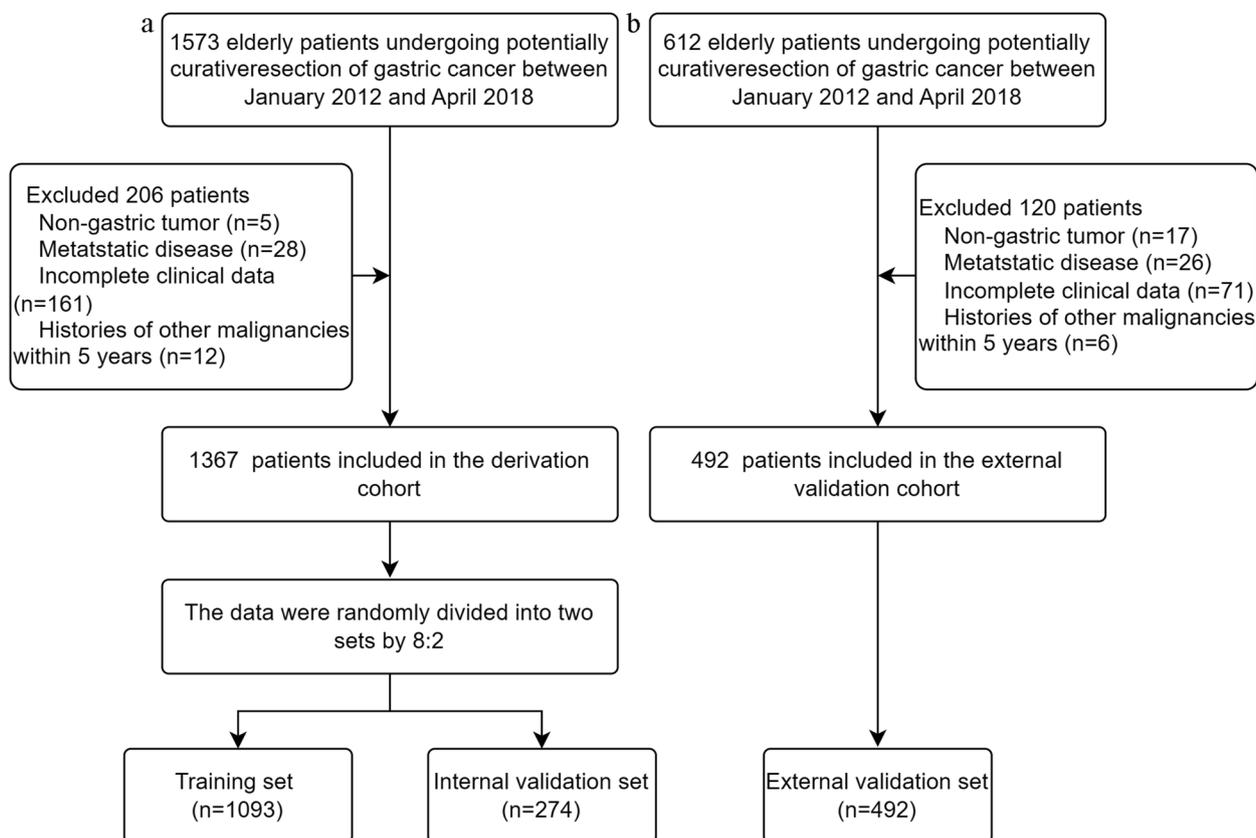


Fig. 1 Diagram representing the selection of the study population. **a** Derivation cohort; **b** External validation cohort

timelines were established from the date of surgery to the last visit (death or final follow-up).

Candidate predictive variables

Comprehensive data on preoperative assessments, intraoperative conditions, postoperative recovery, and pathological examinations were collected. Routine blood and biochemical analyses were performed on admission. The clinical TNM stage was determined according to the eighth edition of the American Joint Committee on Cancer and the Union for International Cancer Control (AJCC/UICC) TNM staging system [20]. Biochemical indices were categorized based on optimal cutoff values determined by the X-tile software as low (below the cutoff) or high (above the cutoff) for analytical precision. Oxidative stress markers included ALB, total bilirubin (TBIL), direct bilirubin (DBIL), blood urea nitrogen (BUN), and UA [8, 21–23]. Using the training set, we developed a GIOSS derived from β coefficients via multivariate Cox analysis. Patients were stratified into risk groups according to GIOSS tertiles [24, 25], followed

by validation in the training set and the entire study population.

The selection of additional clinically pertinent features for training machine learning models was determined through a consensus among researchers grounded in clinical reasoning, an extensive literature review, and their routine availability, ensuring the wide applicability of the models across diverse clinical contexts. The detailed operational definitions of each variable are provided in Additional File 1.

Construction and establishment of the machine learning models

To predict five-year postoperative survival, we used random forest (RF), decision tree (DT), and support vector machine (SVM) models, which were chosen for their proven effectiveness in analogous medical analyses [26]. For model building, we used the R package Tidymodels. To tune the hyperparameters of our classifiers, we used a grid search with cross-validation, which selected hyperparameters based on the model accuracy score. Cross-validation was integral to this process to optimize the complexity parameters and minimize errors. Model

performance was assessed using the area under the precision-recall curve (AUC). Calibration and decision curves were generated to test calibration performance and clinical utility. Finally, we employed the SHapley Additive exPlanations (SHAP) model to perform a post-hoc explainability analysis of the best-performing ML model.

Follow-up

The patients underwent structured follow-ups every three months during the first two years and subsequently every six months for up to five years postoperatively. The final assessment was conducted in April 2024. The routine follow-ups typically include physical examinations, laboratory assessments, chest radiography, abdominal ultrasound or computed tomography (CT) scans, and annual endoscopy. The endpoint was five-year postoperative survival status, distinguishing patients who were alive or deceased at this time point.

Ethics approval and consent to participate

This retrospective study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Institutional Review Board of Fujian Medical University Union Hospital (approval number: 2023KY237), which waived the requirement for informed consent owing to the retrospective design. The study protocol was registered at ClinicalTrials.gov (NCT06208046). All results were reported in accordance with the Strengthening The Reporting Of Cohort Studies in Surgery (STROCSS) Guidelines [27].

Statistical analysis

Data analysis and graphing were performed using R version 4.3.1. Continuous variables were compared using the independent samples t-test or Wilcoxon test depending on their normal distribution. Categorical variable distributions were analyzed using Pearson's χ^2 test and Fisher's exact test. Overall survival (OS) curves were generated using the Kaplan–Meier method, and differences were assessed using the log-rank test. X-tile software v3.6.1 (Yale University) was used to identify the optimal cutoff values for the selected indices [28]. The training datasets were used to refine the model parameters, whereas the independent validation sets were used to evaluate the performance of the trained model using the AUC and calibration curves. The DeLong test facilitated pairwise AUC comparisons among predictive models [29], with Holm–Bonferroni correction for multiple comparisons.

Results

Study cohorts

The derivation and external validation cohorts included 1367 and 492 older patients with GC, respectively

(Table 1). The derivation cohort included 1082 men (79.2%) and 285 women (20.8%), with an average age of 71.2 ± 5.0 years. Following an 8:2 ratio, the derivation cohort was randomly divided into a training set ($n=1093$) and an internal validation set ($n=274$). The training cohort included 861 (78.8%) men and 232 (21.2%) women, with an average age of 71.3 ± 5.1 years. The internal validation cohort comprised 221 men (80.7%) and 53 women (19.3%), with an average age of 70.9 ± 4.9 years.

Among the training and internal validation sets, only pathological T stage differed significantly ($p=0.007$), with no significant differences among the other variables ($p>0.05$). Compared with the training cohort, the external validation cohort exhibited differences in most variables, including age, American Society of Anesthesiologists (ASA) physical status, pathological T stage, and differentiation grade (all $p<0.05$); however, sex, body mass index (BMI), pathological N stage, perineural invasion, residual GC, and UA levels did not differ significantly between the groups (Table 1, $p>0.05$).

Regarding survival outcomes, approximately 40% of the patients in the derivation cohort died within five years postoperatively. The five-year OS rates in the training, internal validation, and external validation cohorts were 60.2%, 59.1%, and 57.5%, respectively (Fig. 2A). Table S1 (Additional File 2) presents the comparison of data between patients in the derivation and external validation cohorts who died and those who survived within five years postoperatively.

Creation of a novel oxidative stress score

In the training set, the optimal cutoff values for oxidative stress indicators determined using X-tile were as follows: ALB, 42.1 g/dL; TBIL, 8.3 $\mu\text{mol/L}$; DBIL, 2.8 $\mu\text{mol/L}$; BUN, 4 mg/dL; and UA, 395 $\mu\text{mol/L}$. These factors were then included in the Cox proportional hazards multivariate analysis, which revealed ALB, DBIL, and BUN as independent risk factors for five-year OS (Table 2). Based on the β coefficients, a prognostic model for the GIOSS was subsequently developed (GIOSS=0–68 ALB-20 TBIL-23 BUN) (Table 2). The patients were stratified into low ($n=427$), medium ($n=396$), and high- ($n=270$) GIOSS groups based on tertiles of GIOSS (–88 and –43). Detailed stratification across the training, internal validation, and external validation cohorts elucidated the distribution of critical prognostic factors such as age, tumor size, and pathological stage among patients with varying GIOSS levels (Table S2, Additional File 3). Kaplan–Meier survival analysis revealed significantly poorer survival rates in the low GIOSS group than in the medium and high GIOSS groups ($p<0.001$; Fig. 2B). Multivariate regression analysis demonstrated that a low GIOSS was an independent risk factor for

Table 1 Characteristics of the derivation cohort and external validation set of older patients undergoing GC surgery

| | Derivation set | | | External validation set (n = 492) | P-value ^b |
|------------------------------------|----------------------------|--------------------------------------|----------------------|--------------------------------------|----------------------|
| | Training set (n = 1093) | Internal validation set (n = 274) | P-value ^a | | |
| Age, mean (SD), years | 71.29 ± 5.07 | 70.93 ± 4.89 | 0.295 | 70.44 ± 4.39 | 0.001 |
| Sex, male | 861 (78.8) | 221 (80.7) | 0.547 | 382 (77.6) | 0.659 |
| BMI, mean (SD) | 22.42 (3.16) | 22.57 (3.29) | 0.502 | 22.57 (3.44) | 0.397 |
| Previous abdominal surgery | 200 (18.3) | 51 (18.6) | 0.974 | 39 (7.9) | < 0.001 |
| Age-adjusted comorbidity index ≤ 2 | 384 (35.1) | 90 (32.8) | 0.522 | 215 (43.7) | 0.001 |
| ASA physical status | | | 0.978 | | < 0.001 |
| 1 | 53 (4.8) | 13 (4.7) | | 14 (2.8) | |
| 2 | 985 (90.1) | 248 (90.5) | | 418 (85.0) | |
| 3 | 55 (5.0) | 13 (4.7) | | 60 (12.2) | |
| Differentiation | | | 0.502 | | < 0.001 |
| G1 | 62 (5.7) | 21 (7.7) | | 38 (7.7) | |
| G2 | 471 (43.1) | 120 (43.8) | | 165 (33.5) | |
| G3 | 501 (45.8) | 122 (44.5) | | 285 (57.9) | |
| G4 | 59 (5.4) | 11 (4.0) | | 4 (0.8) | |
| Tumor size ≤ 5 cm | 619 (56.6) | 162 (59.1) | 0.499 | 343 (69.7) | < 0.001 |
| Location | | | 0.511 | | < 0.001 |
| Low | 421 (38.5) | 110 (40.1) | | 200 (40.7) | |
| Middle | 184 (16.8) | 50 (18.2) | | 116 (23.6) | |
| Mixed | 116 (10.6) | 21 (7.7) | | 172 (35.0) | |
| Upper | 372 (34.0) | 93 (33.9) | | 4 (0.8) | |
| pT | | | 0.007 | | < 0.001 |
| 1 | 200 (18.3) | 75 (27.4) | | 97 (19.7) | |
| 2 | 121 (11.1) | 25 (9.1) | | 63 (12.8) | |
| 3 | 460 (42.1) | 97 (35.4) | | 90 (18.3) | |
| 4 | 312 (28.5) | 77 (28.1) | | 242 (49.2) | |
| pN | | | 0.144 | | 0.193 |
| 0 | 346 (31.7) | 97 (35.4) | | 182 (37.0) | |
| 1 | 194 (17.7) | 40 (14.6) | | 84 (17.1) | |
| 2 | 216 (19.8) | 65 (23.7) | | 84 (17.1) | |
| 3 | 337 (30.8) | 72 (26.3) | | 142 (28.9) | |
| pTNM | | | 0.116 | | < 0.001 |
| I | 235 (21.5) | 75 (27.4) | | 175 (35.6) | |
| II | 299 (27.4) | 69 (25.2) | | 136 (27.6) | |
| III | 559 (51.1) | 130 (47.4) | | 181 (36.8) | |
| Microvascular invasion | 411 (37.6) | 95 (34.7) | 0.407 | 102 (20.7) | < 0.001 |
| Perineural invasion | 331 (30.3) | 89 (32.5) | 0.527 | 145 (29.5) | 0.789 |
| Lymphatic invasion | 118 (10.8) | 27 (9.9) | 0.732 | 145 (29.5) | < 0.001 |
| Open approach | 101 (9.2) | 27 (9.9) | 0.845 | 196 (39.8) | < 0.001 |
| Reconstruction | | | 0.469 | | < 0.001 |
| Billroth-I | 204 (18.7) | 54 (19.7) | | 4 (0.8) | |
| Billroth-II | 232 (21.2) | 66 (24.1) | | 68 (13.8) | |
| Roux-en-Y | 657 (60.1) | 154 (56.2) | | 420 (85.4) | |
| Neoadjuvant therapy | 41 (3.8) | 11 (4.0) | 0.978 | 7 (1.4) | 0.019 |
| Adjuvant therapy | 286 (26.2) | 64 (23.4) | 0.381 | 253 (51.4) | < 0.001 |
| Emergency | 26 (2.4) | 4 (1.5) | 0.485 | 3 (0.6) | 0.026 |
| Surgery duration ≤ 200 min | 813 (74.4) | 209 (76.3) | 0.570 | 220 (44.7) | < 0.001 |
| Blood ≥ 100 mL | 329 (30.1) | 76 (27.7) | 0.489 | 467 (94.9) | < 0.001 |

Table 1 (continued)

| | Derivation set | | | External validation set (n = 492) | P-value ^b |
|---|----------------------------|--------------------------------------|----------------------|--------------------------------------|----------------------|
| | Training set (n = 1093) | Internal validation set (n = 274) | P-value ^a | | |
| Comprehensive Complication Index ≤ 26.2 | 222 (20.3) | 55 (20.1) | 0.997 | 459 (93.3) | < 0.001 |
| Remnant stomach | 12 (1.1) | 4 (1.5) | 0.854 | 6 (1.2) | 1 |
| ALB, mean (SD) | 38.48 (5.11) | 38.96 (4.73) | 0.157 | 39.09 (4.56) | 0.021 |
| TBIL, mean (SD) | 11.38 (6.44) | 11.17 (5.82) | 0.631 | 7.65 (4.55) | < 0.001 |
| DBIL, mean (SD) | 3.10 (1.41) | 3.11 (1.42) | 0.913 | 5.43 (4.56) | < 0.001 |
| BUN, mean (SD) | 5.26 (1.84) | 5.38 (1.64) | 0.335 | 5.12 (1.94) | 0.175 |
| UA, mean (SD) | 330.95 (90.0) | 341.79 (90.6) | 0.075 | 300.15 (88.7) | < 0.001 |

GC gastric cancer, SD standard deviation, BMI Body mass index, ASA American Society of Anesthesiologists physical status, ALB albumin, TBIL total bilirubin, DBIL direct bilirubin, BUN blood urea nitrogen, UA uric acid

^a training cohort vs. internal validation group

^b derivation set vs. external validation group. Data are expressed as numbers (percentages) of participants unless otherwise indicated

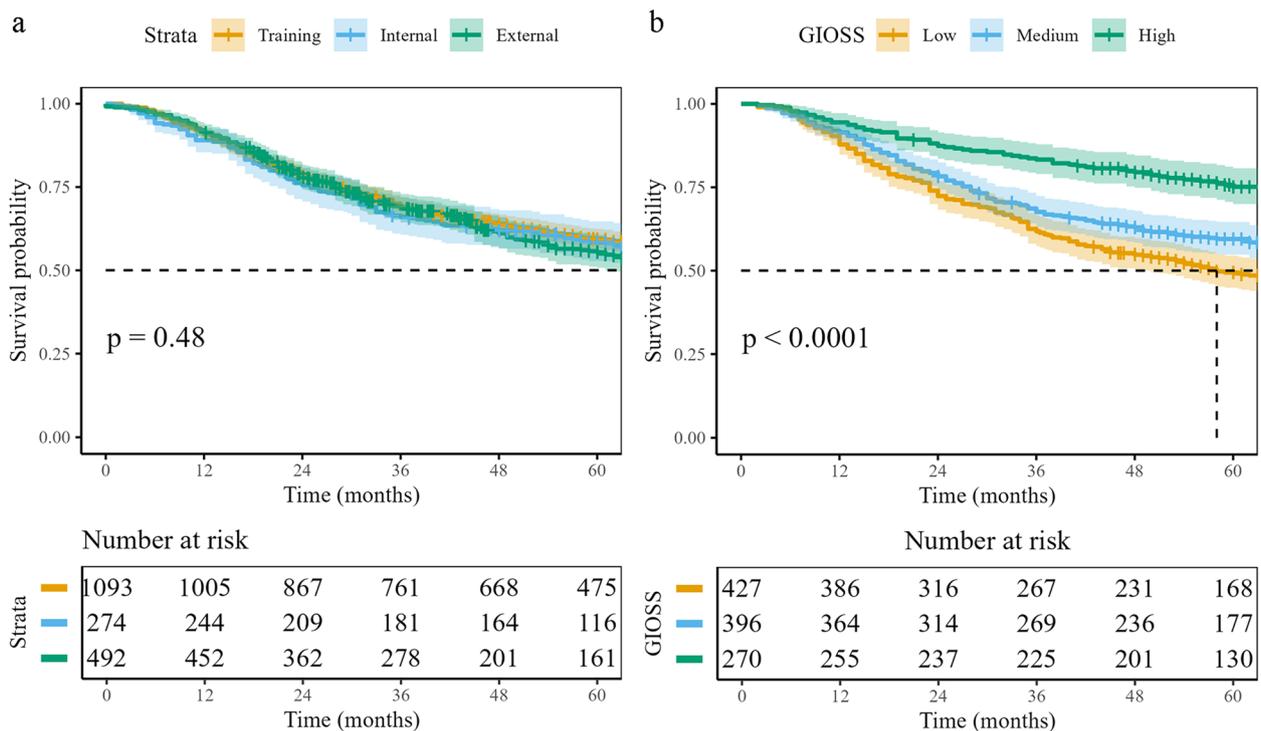


Fig. 2 Kaplan–Meier survival curves stratified by cohort and GIOSS. **a** Survival curves for training, internal validation, and external validation cohorts; **b** Survival curves stratified by GIOSS. Abbreviations: RF, random forest; DT, decision tree; SVM, support vector machine

five-year OS (Table S3, Additional File 4). Additionally, correlation analysis in both the derivation and external validation cohorts indicated that the GIOSS score was not strongly associated with other clinical or pathological factors (Figure S1, Additional File 5). The age distributions across the high, middle, and low GIOSS groups were comparable across the different groups, indicating

consistent categorization across the different cohorts (Figure S2, Additional File 6).

Model performance: calibration performance and clinical utility

We incorporated the GIOSS into machine learning and constructed three models. Table 3 displays the AUC

Table 2 Cox regression models of biochemical markers in the training set

| | Univariate Cox regression | | | Multivariate Cox regression | | |
|------|---------------------------|--------------------|---------|-----------------------------|--------------------|---------|
| | β | HR | P value | β | HR | P value |
| ALB | -0.754 | 0.470(0.363–0.610) | <0.001 | -0.680 | 0.506(0.388–0.662) | <0.001 |
| TBIL | -0.353 | 0.703(0.580–0.852) | <0.001 | -0.199 | 0.820(0.673–0.998) | 0.048 |
| DBIL | -0.165 | 0.848(0.703–1.023) | 0.086 | | | |
| BUN | -0.304 | 0.738(0.594–0.916) | 0.006 | -0.236 | 0.790(0.635–0.983) | 0.034 |
| UA | -0.213 | 0.808(0.637–1.025) | 0.079 | | | |

Gastric Cancer Integrated Oxidative Stress Score = $0-68*ALB-20*TBIL-23*BUN$

Abbreviations: HR Hazard Ratio, ALB Albumin, TBIL Total Bilirubin, DBIL Direct Bilirubin, BUN, Blood Urea Nitrogen, UA Uric Acid

Table 3 Areas under the receiver operating characteristic curves for five-year overall survival

| | RF | DT | SVM |
|-------------------------|---------------------|---------------------|---------------------|
| Training set | 0.999 (0.999–0.999) | 0.784 (0.756–0.811) | 0.832 (0.807–0.856) |
| Internal validation set | 0.869 (0.826–0.911) | 0.796 (0.741–0.851) | 0.839 (0.789–0.890) |
| External validation set | 0.796 (0.755–0.837) | 0.741 (0.699–0.783) | 0.789 (0.748–0.830) |

Values are presented as observed area under the receiver operating characteristic curve (95% confidence interval)

Abbreviations: RF Random Forest, DT Decision Tree, SVM Support Vector Machine

values of these models for predicting five-year OS in the training, internal validation, and external validation sets. The AUC values for RF were 0.999 (95% confidence interval [CI]: 0.999–1.000), 0.868 (95% CI: 0.826–0.911), and 0.796 (95% CI: 0.755–0.837) for the training, internal validation, and external validation sets, respectively. For DT, the AUC values were 0.784 (95% CI: 0.756–0.811), 0.796 (95% CI: 0.741–0.850), and 0.741 (95% CI: 0.699–0.783), whereas SVM presented values of 0.832 (95% CI: 0.807–0.856), 0.839 (95% CI: 0.789–0.890), and 0.789 (95% CI: 0.748–0.830), respectively. Compared with the DT and SVM models, the RF model consistently exhibited superior AUC values across all datasets. Analysis of the ROC curves revealed the performance of the RF, DT, and SVM models across the training, internal validation, and external validation sets (Fig. 3). Across all datasets, the RF model consistently demonstrated the highest AUC values, followed by the SVM model, whereas the DT model exhibited relatively lower performance.

Model performance: calibration and decision curves

The calibration curves revealed that the RF model performed robustly across all datasets, with its predicted probabilities closely aligned with the observed event frequencies, notably in the higher probability ranges (Figure S3a,d,g, Additional File 7). In contrast, the DT model typically aligned predictions with the actual outcomes; however, it exhibited slight deviations in external validation (Figure S3b/e/h, Additional File 7), possibly indicating overfitting to the training data and limited

generalizability. Calibration of the SVM model closely approximated the 45-degree line in both the training and validation sets, with minimal deviations at certain probability intervals (Figure S3c,f,i, Additional File 7), revealing a consistent calibration in probability prediction. Furthermore, the results of the decision curve analysis indicated that the RF model yielded the highest net benefit across most threshold probabilities in the training set, outperforming the similar but lower net benefits of the DT model and SVM (Figure S4a, Additional File 8). In the internal validation set, the RF model maintained greater net benefits across most thresholds, whereas the DT model and SVM showed similar but lower performances (Figure S4b, Additional File 8). Although all models exhibited reduced performance in the external validation set, the RF model outperformed the other models at most thresholds (Figure S4c, Additional File 8).

Model interpretation

To further examine the relationship between clinicopathological characteristics and prognosis, we analyzed the dependence plots of all features in mortality prediction, as assessed by the SHAP values. In the RF model, the SHAP results revealed that the five most significant variables were pN stage, pT stage, perineural invasion, tumor size, and GIOSS (Figure S5, Additional File 9).

To explore the contribution of various features to the predictive model, we employed partial dependence plots to visualize how the model's predictions were influenced by individual features while keeping all other variables

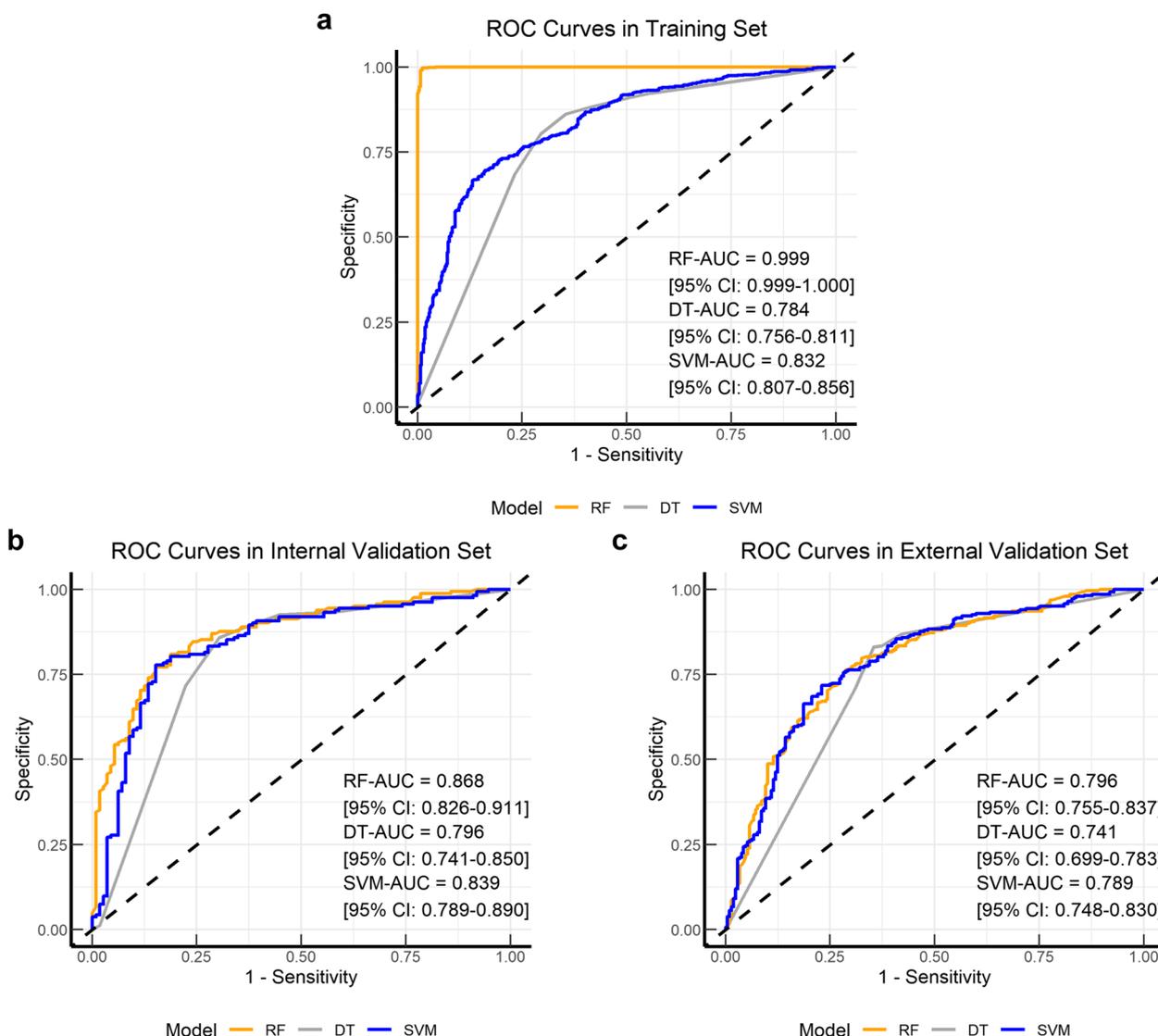


Fig. 3 Receiver operating characteristic curves for the three models. **a** Training set; **b** Internal validation set; **c** External validation set. Abbreviations: ROC, receiver operating characteristic; RF, random forest; DT, decision tree; SVM, support vector machine; AUC, area under the curve; CI, confidence interval

constant (Figure S6, Additional File 10). Lower GIOSS were associated with an increased average prediction of mortality risk.

We generated waterfall plots for all features to predict the mortality rates and examined the link between individual characteristics and patient prognosis. This machine-learning model allowed us to evaluate the five-year overall mortality of randomly selected patients while simultaneously revealing the key characteristics (Fig. 4).

SHAP analysis revealed that higher feature values were associated with an increased risk of mortality, and features with substantial weight were associated with life expectancy.

To support further validation and application, we developed a simplified interactive web tool for risk calculation using the top seven variables. This tool is available at https://fmuuh.shinyapps.io/GIOSS-based_Survival_Calculator_in_Aged_GC/. A snapshot of the online calculator is shown in Figure S7 (Additional File 11).

Discussion

The five-year survival rate after GC surgery is an essential parameter for assessing the efficacy of long-term oncological surgical care [30]. Adults >65 years of age face a unique set of biopsychosocial challenges [31, 32]. The escalation of comorbidities, increased frailty,

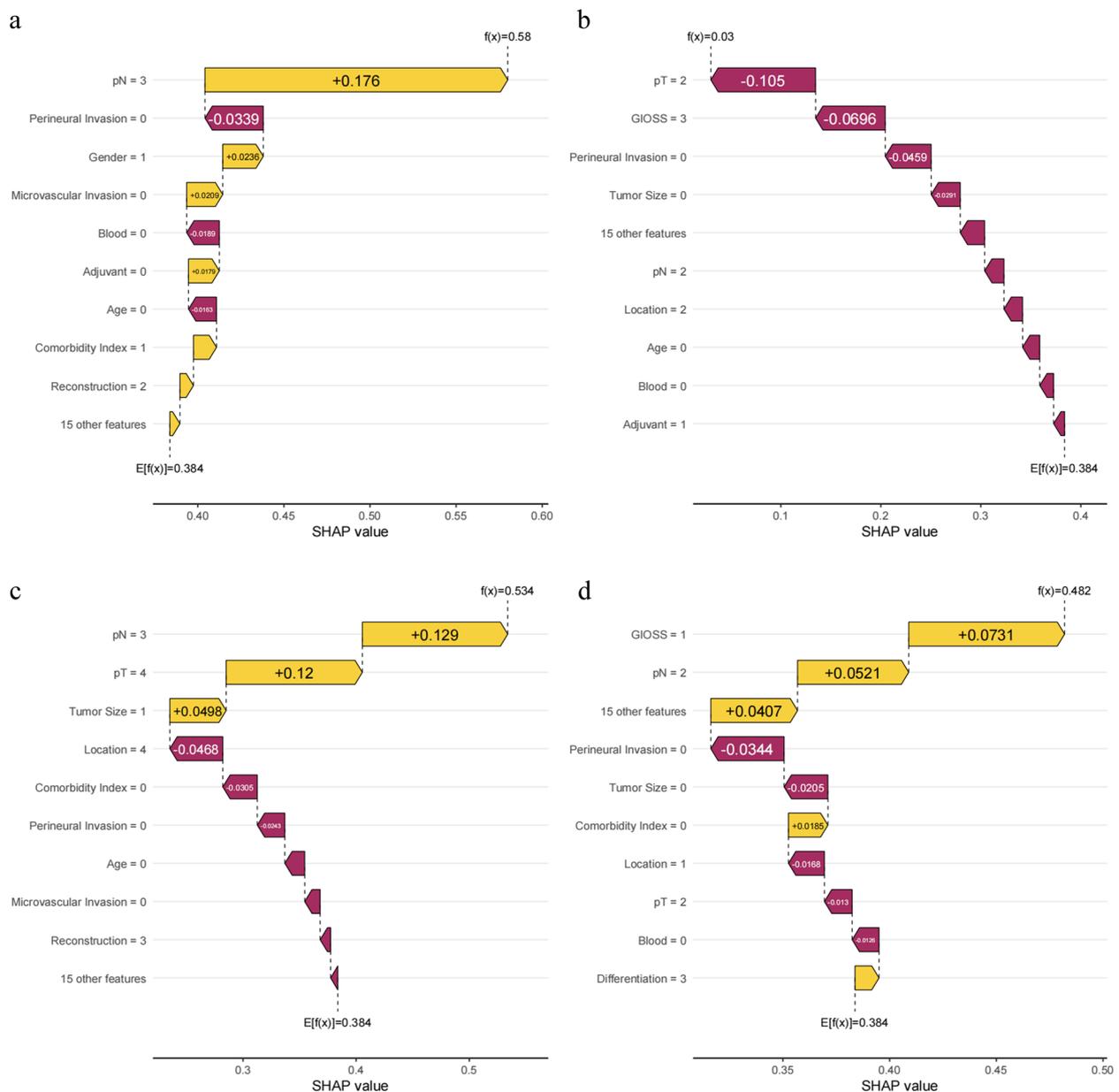


Fig. 4 Waterfall plots for all features to predict mortality and links between characteristics and patient prognosis (a) True-positive, (b) True-negative, (c) False-positive, and (d) False-negative Samples from the training dataset on a probability scale

diminished stress tolerance, and deterioration of physical and cognitive abilities significantly influence postoperative survival [33, 34]. This observation underscores the need for a sophisticated and individualized approach to post-surgical care and prognostic evaluation. To our knowledge, specific predictive models for five-year OS in older post-surgical patients with GC are currently lacking. Therefore, we investigated the five-year OS rate in this demographic population using machine-learning approaches (DT, RF, and SVM). Our results

demonstrated the superior efficacy of the RF model. Using this approach, clinicians can more accurately formulate personalized treatment plans and conduct targeted monitoring and follow-up, thereby enhancing the precision and effectiveness of patient management.

Animal model experiments have demonstrated that when mice are subjected to external stimuli, an oxidative stress response is triggered, which is characterized by increased levels of oxidative stress factors and significant increases in the levels of biochemical markers, including

TBIL, lactate dehydrogenase, creatinine, and BUN. Prospective studies have indicated that oxidative stress leads to alterations in TBIL, ALB, and UA levels [23]. Moreover, patients who receive antioxidant therapy show lower mortality rates than those who do not [35]. Although oxidative stress is associated with cancer onset and progression, the use of biochemical markers to predict GC prognosis remains underexplored. Particularly, their effectiveness as potential indicators for improving the accuracy of prognostic prediction in older patients with GC warrants further investigation. Therefore, we propose the GIOSS as an indicator of oxidative stress. Our findings revealed that patients with lower GIOSS scores had poorer prognoses than those with higher GIOSS. Correlation analysis revealed that the GIOSS weakly overlapped with the age-adjusted comorbidity index score, which represents different aspects of a patient's health status. The GIOSS was developed using a training cohort from one of our institutions, which facilitated the use of detailed clinical data and enabled long-term follow-up. Therefore, we hypothesized that a predictive model incorporating GIOSS would more effectively predict the prognosis of older patients with GC.

Previous models, such as the Yonsei Gastric Cancer Prognosis Prediction Model [36], TNM staging system [37], and gene signature prediction models [38], have been used to predict the OS and prognosis of patients with GC. Woo et al. developed a non-machine-learning predictive model for post-gastrectomy prognosis in patients with GC. This model demonstrated robust C-indices ranging from 0.798 to 0.868 across four external validation cohorts in three different countries [36]. However, it exhibited only moderate discriminative accuracy when applied to the Surveillance, Epidemiology, and End Results (SEER) database [39]. Li et al. developed an SVM model that integrates nine clinicopathological factors [2]. This model achieved an AUC of 0.773 (95% CI: 0.708–0.838) for predicting the five-year OS and an AUC of 0.852 (95% CI: 0.810–0.894) in the validation group, effectively forecasting the efficacy of postoperative chemotherapy in stage II/III patients [2]. Similarly, the machine learning model by Rahman et al., which incorporated 29 clinical and pathological variables, exhibited outstanding performance in the internal validation set, achieving a 5-year time-dependent AUC of 0.80 (95% CI: 0.78–0.82) [3].

The rigorous cross-validation performed in the present study confirmed the higher AUC of the newly developed RF model compared with previous models. This enhancement is attributable to the comprehensive incorporation of the physiological characteristics of older patients, including oxidative stress biochemical markers, age-adjusted comorbidity index [40], and a

comprehensive complication index [41]. The integration of biochemical markers indicative of oxidative stress into the model highlights patients who may benefit from continuous therapies aimed at mitigating oxidative damage, thereby leading to novel therapeutic strategies. These factors were combined with pathological characteristics such as pT, pN, and tumor location, offering a more detailed reflection of the traits specific to older patients with GC. Decision curve analysis is an invaluable tool for evaluating and comparing the efficacy of different models across various thresholds, thereby facilitating informed model selection and application. The RF model demonstrated superior performance across the training, testing, and validation datasets, indicating its potential to enhance generalizability. This model represents a significant advance in predicting the five-year postoperative OS of older patients with GC, potentially assisting clinicians in enhancing the precision and effectiveness of patient management.

This model represents a significant advancement in predicting the five-year postoperative OS in older patients with GC. Potential therapeutic strategies based on the results of this study include personalized adjustment of postoperative treatment plans, such as more tailored chemotherapy regimens for high-risk patients identified by the model, as well as closer monitoring and follow-up schedules to promptly address complications and recurrence. Additionally, integrating the RF model into clinical practice could aid in identifying patients who may benefit from adjunctive therapies targeting oxidative stress, thus improving overall patient management and outcomes. Similarly, earlier small-sample, single-center, retrospective studies introduced the concept of oxidative stress scores for patients with GC and used nomograms for their predictions [42, 43]. In contrast, our study collected multidimensional data from a large sample of older patients across multiple centers. By leveraging advanced machine learning techniques, our study provides a more precise prognostic tool specifically tailored to this patient group.

This study has some limitations, primarily its retrospective nature and consequent inherent risk of selection bias. Prospective studies should be conducted to corroborate our findings. Additionally, while the use of external data from three distinct hospitals as independent validation sets strengthened the generalizability of our models, these datasets may not fully capture the diversity of the older patient population with GC owing to geographic and institutional constraints. Moreover, our database lacked information on key factors known to influence GC outcomes, such as high-risk gene mutations, use of immunotherapy drugs, and socioeconomic status. The incorporation of these variables may enhance the model's

performance. We aim to expand the scope of this study by incorporating more representative centers, thereby enhancing the representativeness and significance of our research findings. Future prospective studies should also implement stricter standardization of surgical procedures to eliminate biases resulting from differences in surgeon skills. Furthermore, these studies should explore the association between tissue oxidative stress markers and serum biochemical markers and validate the clinical application of dynamic changes in GIOSS.

Conclusion

The machine learning clinical prediction model developed using the GIOSS effectively predicted the postoperative prognosis of older patients with GC following curative surgery.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12885-025-13545-x>.

Additional file 1: Variable definitions.

Additional file 2: Table S1. Associations between different potential risk factors for 5-year mortality in the derivation cohort.

Additional file 3: Table S2. Relationship between gastric cancer oxidative stress score and clinical characteristics.

Additional file 4: Table S3. Univariate and multivariate analyses of prognostic factors for 5-year overall survival.

Additional file 5: Figure S1. Correlation matrix heatmap. (a) Derivation cohort; (b) External validation cohort. Abbreviations: ASA, American Society of Anesthesiologists physical status; BMI, body mass index; CCI, comprehensive complication index; GIOSS, gastric cancer-integrated oxidative stress score.

Additional file 6: Figure S2. Age distribution by GIOSS groups. (a) Training set; (b) Internal validation set; (c) External validation set. Abbreviations: GIOSS, gastric cancer oxidative stress score.

Additional file 7: Figure S3. Calibration curves of the three models. (a–c) Calibration curves for the RF, DT, and SVM models, respectively, in the training set. (d–f) Calibration curves for the RF, DT, and SVM models, respectively, in the internal validation set. (g–i) Calibration curves for the RF, DT, and SVM models, respectively, in the external validation set. Abbreviations: RF, random forest; DT, decision tree; SVM, support vector machine.

Additional file 8: Figure S4. Decision curve analysis. (a) Training set; (b) Internal validation set; (c) External validation set. Abbreviations: RF, random forest; DT, decision tree; SVM, support vector machine.

Additional file 9: Figure S5. Feature importance of random forests based on SHAP. Abbreviations: RF, random forest; DT, decision tree; SVM, support vector machine; GIOSS, gastric cancer-integrated oxidative stress score; BMI, body mass index; ASA, American Society of Anesthesiologists physical status; SHAP, Shapley additive explanations.

Additional file 10: Figure S6. Partial dependence plots for 5-year overall survival using the random forest model. Abbreviations: ASA, American Society of Anesthesiologists physical status; BMI, body mass index; CCI, comprehensive complication index; GIOSS, gastric cancer-integrated oxidative stress score.

Additional file 11: Figure S7. Internet browser-based online calculator. Online tool: (https://fmuuh.shinyapps.io/GIOSS-based_Survival_Calculator_in_Aged_GC/).

Acknowledgements

We thank those who have devoted a lot to this study, including nurses, pathologists, further-study doctors, statisticians, reviewers, and editors. We would like to thank Editage (www.editage.cn) for English language editing. We also thank Figdraw (www.figdraw.com) for their assistance in creating diagrams.

Authors' contributions

X-Q.Z., Z-N.H., and C-M.H. conceptualized and designed the study. X-Q.Z., Z-N.H., and J.W. performed formal analysis and investigation. X-Q.Z., Z-N.H., and J-X.L. drafted the original manuscript. All authors contributed to manuscript review and editing. Funding was acquired by C-M.H. Resources were provided by all authors, and supervision was carried out by Z-N.H., Q-Y.C., P.L., J-W.X., C-H.Z., and Y-B.Z. All authors reviewed and approved the manuscript.

Funding

This study was supported by the Joint Funds for the Innovation of Science and Technology, Fujian Province (No. 2023Y9208), the Natural Science Foundation of Fujian Province (No. 2023J01674), and the Construction Funds for "High-level Hospitals and Clinical Specialties" of Fujian Province (No. [2021]76).

Data availability

The datasets supporting the conclusions of this article are available in the GIOSS_FOR_GC repository, https://github.com/Justin-Zhangxq/GIOSS_FOR_GC.

Declarations

Ethics approval and consent to participate

This study was conducted in accordance with the Declaration of Helsinki and approved by the institutional review board of Fujian Medical University Union Hospital (approval number: 2023KY237). The institutional review board provided ethics approval for this study and exemption from informed consent due to the retrospective nature of the study. All procedures performed in this study were in accordance with the ethical standards of the institutional and/or national research committee.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Gastric Surgery, Fujian Medical University Union Hospital, No.29 Xinquan Road, Fuzhou, Fujian Province 350001, China. ²Department of General Surgery, Fujian Medical University Union Hospital, Fuzhou, Fujian Province, China. ³Key Laboratory of Ministry of Education of Gastrointestinal Cancer, Fujian Medical University, Fuzhou, Fujian Province 350108, China. ⁴Department of General Surgery, Affiliated Zhongshan Hospital of Dalian University, Dalian, Liaoning Province, China. ⁵Department of Gastrointestinal Surgery, The Affiliated Hospital of Putian University, Putian, Fujian Province, China. ⁶Department of General Surgery, The Affiliated Hospital of Qingdao University, NO.16 Jiangsu Road, Qingdao, Shandong 266000, China.

Received: 29 October 2024 Accepted: 16 January 2025

Published online: 01 February 2025

References

1. Rebecca LS, Angela NG, Ahmedin J. Cancer statistics, 2024. *CA Cancer J Clin.* 2024;74(1):12–49.
2. Li X, Zhai Z, Ding W, Chen L, Zhao Y, Xiong W, Zhang Y, Lin D, Chen Z, Wang W, et al. An artificial intelligence model to predict survival and chemotherapy benefits for gastric cancer patients after gastrectomy development and validation in international multicenter cohorts. *Int J Surg (London, England).* 2022;105:106889.
3. Rahman S, Maynard N, Trudgill N, Crosby T, Park M, Wahedly H, Underwood T, Cromwell D. Prediction of long-term survival after gastrectomy using random survival forests. *Br J Surg.* 2021;108(11):1341–50.

4. Hayes J, Dinkova-Kostova A, Tew K. Oxidative stress in cancer. *Cancer Cell*. 2020;38(2):167–97.
5. Arfin S, Jha N, Jha S, Kesari K, Ruokolainen J, Roychoudhury S, Rathi B, Kumar D. Oxidative stress in cancer cell metabolism. *Antioxidants* (Basel, Switzerland). 2021;10(5):642.
6. Abdelhamid R, Nagano S. Crosstalk between oxidative stress and aging in neurodegeneration disorders. *Cells*. 2023;12(5):753.
7. Wang J, Sun Y, Zhang X, Cai H, Zhang C, Qu H, Liu L, Zhang M, Fu J, Zhang J, et al. Oxidative stress activates NORAD expression by H3K27ac and promotes oxaliplatin resistance in gastric cancer by enhancing autophagy flux via targeting the miR-433-3p. *Cell Death Dis*. 2021;12(1):90.
8. Kim Y, Choi C, Lee Y, Choi S, Kim H, Shin M, Kweon S. Association between albumin, total bilirubin, and uric acid serum levels and the risk of cancer: a prospective study in a Korean population. *Yonsei Med J*. 2021;62(9):792–8.
9. Yang Y, Dai C, Chen X, Feng J. The relationship between serum trace elements and oxidative stress of patients with different types of cancer. *Oxid Med Cell Longev*. 2021;2021:4846951.
10. Piao M, Tu Y, Zhang N, Diao Q, Bi Y. Advances in the application of phyto-genic extracts as antioxidants and their potential mechanisms in ruminants. *Antioxidants* (Basel, Switzerland). 2023;12(4):879.
11. Chu Q, Jia R, Chen W, Liu Y, Li Y, Ye X, Jiang Y, Zheng X. Purified Tetragymna hemsleyanum vines polysaccharide attenuates EC-induced toxicity in Caco-2 cells and *Caenorhabditis elegans* via DAF-16/FOXO pathway. *Int J Biol Macromol*. 2020;150:1192–202.
12. Toh D, Lee W, Zhou H, Sutanto C, Lee D, Tan D, Kim J. Lycium barbarum Wolfberry () consumption with a healthy dietary pattern lowers oxidative stress in middle-aged and older adults: a randomized controlled trial. *Antioxidants* (Basel, Switzerland). 2021;10(4):567.
13. Sanidad K, Sukamtoh E, Xiao H, McClements D, Zhang G. Curcumin: recent advances in the development of strategies to improve oral bioavailability. *Annu Rev Food Sci Technol*. 2019;10:597–617.
14. Liu M, Rao H, Liu J, Li X, Feng W, Gui L, Tang H, Xu J, Gao W, Li L. The histone methyltransferase SETD2 modulates oxidative stress to attenuate experimental colitis. *Redox Biol*. 2021;43:102004.
15. Van Calster B, Wynants L. Machine learning in medicine. *N Engl J Med*. 2019;380(26):2588.
16. Deo R. Machine learning in medicine. *Circulation*. 2015;132(20):1920–30.
17. Jiang Y, Zhang Z, Yuan Q, Wang W, Wang H, Li T, Huang W, Xie J, Chen C, Sun Z, et al. Predicting peritoneal recurrence and disease-free survival from CT images in gastric cancer with multitask deep learning: a retrospective study. *Lancet Digit Health*. 2022;4(5):e340–50.
18. D'Ascenzo F, De Filippo O, Gallone G, Mittone G, Deriu M, Iannaccone M, Ariza-Solé A, Liebetrau C, Manzano-Fernández S, Quadri G, et al. Machine learning-based prediction of adverse events following an acute coronary syndrome (PRAISE): a modelling study of pooled datasets. *Lancet* (London, England). 2021;397(10270):199–207.
19. Gao Q, Zhou H, Wang G, Ma Z, Li J, Wang H, Sun G, Xue X. A model for predicting clinical prognosis in patients with WHO grade 2 glioma. *J Oncol*. 2022;2022:2795939.
20. Mahul BA, Frederick LG, Stephen BE, Carolyn CC, Jeffrey EG, Robert KB, Laura M, Donna MG, David RB, David PW. The Eighth Edition AJCC cancer staging manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin*. 2017;67(2):93–9.
21. Claps G, Faouzi S, Quidville V, Chehade F, Shen S, Vagner S, Robert C. The multiple roles of LDH in cancer. *Nat Rev Clin Oncol*. 2022;19(12):749–62.
22. Kühn T, Sookthai D, Graf M, Schübel R, Freisling H, Johnson T, Katzke V, Kaaks R. Albumin, bilirubin, uric acid and cancer risk: results from a prospective population-based study. *Br J Cancer*. 2017;117(10):1572–9.
23. Sandesc M, Rogobete A, Bedreag O, Dinu A, Papurica M, Cradigati C, Sarandau M, Popovici S, Bratu L, Bratu T, et al. Analysis of oxidative stress-related markers in critically ill polytrauma patients: an observational prospective single-center study. *Bosn J Basic Med Sci*. 2018;18(2):191–7.
24. Fazel S, Långström N, Hjern A, Grann M, Lichtenstein P. Schizophrenia, substance abuse, and violent crime. *JAMA*. 2009;301(19):2016–23.
25. Armstrong A, Lin P, Higano C, Sternberg C, Sonpavde G, Tombal B, Templeton A, Fizazi K, Phung D, Wong E, et al. Development and validation of a prognostic model for overall survival in chemotherapy-naïve men with metastatic castration-resistant prostate cancer. *Ann Oncol*. 2018;29(11):2200–7.
26. Noble W. What is a support vector machine? *Nat Biotechnol*. 2006;24(12):1565–7.
27. Mathew G, Agha R, Albrecht J, Goel P, Mukherjee I, Pai P, D'Cruz A, Nixon I, Roberto K, Enam S, et al. STROCSS 2021: strengthening the reporting of cohort, cross-sectional and case-control studies in surgery. *Int J Surg* (London, England). 2021;96:106165.
28. Camp R, Dolled-Filhart M, Rimm D. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res*. 2004;10(21):7252–9.
29. DeLong E, DeLong D, Clarke-Pearson D. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–45.
30. Haga Y, Ikejiri K, Wada Y, Ikenaga M, Takeuchi H. Preliminary study of surgical audit for overall survival following gastric cancer resection. *Gastric Cancer*. 2015;18(1):138–46.
31. Zu H, Wang H, Li C, Kang Y, Xue Y. Clinico-pathological features and prognostic analysis of gastric cancer patients in different age groups. *Hepatogastroenterology*. 2015;62(137):225–30.
32. Joharatnam-Hogan N, Shiu K, Khan K. Challenges in the treatment of gastric cancer in the older patient. *Cancer Treat Rev*. 2020;85:101980.
33. Hikage M, Tokunaga M, Makuuchi R, Irino T, Tanizawa Y, Bando E, Kawamura T, Terashima M. Surgical outcomes after gastrectomy in very elderly patients with gastric cancer. *Surg Today*. 2018;48(8):773–82.
34. Sakurai K, Muguruma K, Nagahara H, Kimura K, Toyokawa T, Amano R, Kubo N, Tanaka H, Ohtani H, Yashiro M, et al. The outcome of surgical treatment for elderly patients with gastric carcinoma. *J Surg Oncol*. 2015;111(7):848–54.
35. Lima W, Martins-Santos M, Chaves V. Uric acid as a modulator of glucose and lipid metabolism. *Biochimie*. 2015;116:17–23.
36. Woo Y, Son T, Song K, Okumura N, Hu Y, Cho G, Kim J, Choi S, Noh S, Hyung W. A novel prediction model of prognosis after gastrectomy for gastric carcinoma: development and validation using asian databases. *Ann Surg*. 2016;264(1):114–20.
37. Park K, Jun K, Song K, Chin H, Lee H. Development of a staging system and survival prediction model for advanced gastric cancer patients without adjuvant treatment after curative gastrectomy: a retrospective multicenter cohort study. *Int J Surg* (London, England). 2022;101:106629.
38. Cheong J, Wang S, Park S, Porembka M, Christie A, Kim H, Kim H, Zhu H, Hyung W, Noh S, et al. Development and validation of a prognostic and predictive 32-gene signature for gastric cancer. *Nat Commun*. 2022;13(1):774.
39. Woo Y, Goldner B, Son T, Song K, Noh S, Fong Y, Hyung W. Western validation of a novel gastric cancer prognosis prediction model in US gastric cancer patients. *J Am Coll Surg*. 2018;226(3):252–8.
40. Koseki Y, Hikage M, Fujiya K, Kamiya S, Tanizawa Y, Bando E, Terashima M. Utility of a modified age-adjusted Charlson Comorbidity Index in predicting cause-specific survival among patients with gastric cancer. *Eur J Surg Oncol*. 2021;47(8):2010–5.
41. Artiles-Armas M, Roque-Castellano C, Conde-Martel A, Marchena-Gómez J. The Comprehensive complication index is related to frailty in elderly surgical patients. *J Surg Res*. 2019;244:218–24.
42. Xinyu W, Limin Z. The systemic oxidative stress score has a prognostic value on gastric cancer patients undergoing surgery. *Front Oncol*. 2024;14:1307662.
43. Yu-Hang L, Rui M, Bing Z, Qi-Qi Z, Xin Y, Guan-Yi D, Chun-Liang J, Qian-Yu L, Wei-Guo X. Integrated oxidative stress score for predicting prognosis in stage III gastric cancer undergoing surgery. *Pathol Oncol Res*. 2023;29:1610897.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.