



Machine learning models and nomogram based on clinical, laboratory profiles and skeletal muscle index to predict pancreatic fistula after pancreatoduodenectomy

Yile Dai^{1,2}, Wenqian Huang², Liming Xu³, Qiyu Zhang², Xiaming Huang²

¹The First School of Clinical Medicine, Wenzhou Medical University, Wenzhou, China; ²Department of Hepatobiliary Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China; ³Department of Hepatobiliary Surgery, Wenzhou Central Hospital, Wenzhou, China

Contributions: (I) Conception and design: X Huang, Q Zhang, Y Dai; (II) Administrative support: None; (III) Provision of study materials or patients: X Huang, Q Zhang; (IV) Collection and assembly of data: Y Dai, W Huang, L Xu; (V) Data analysis and interpretation: Y Dai, W Huang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Xiaming Huang, MD; Qiyu Zhang, MD. Department of Hepatobiliary Surgery, The First Affiliated Hospital of Wenzhou Medical University, Nanbaixiang Street, Wenzhou 325000, China. Email: huangxiam9@163.com; qiyu_zhang@wmu.edu.cn.

Background: Postoperative pancreatic fistula (POPF) is a perilous complication that may arise subsequent to pancreaticoduodenectomy (PD). In recent times, there has been an escalating interest in employing machine learning (ML) techniques to aid in treatment decision-making. The purpose of this research is to assess the effectiveness of ML in comparison to conventional models, while also conducting an initial evaluation of the predictive capability of skeletal muscle index (SMI) concerning POPF.

Methods: This retrospective observational study was carried out at The First Affiliated Hospital of Wenzhou Medical University from January 2012 to January 2021, encompassing data from 269 patients who underwent PD. After identifying independent factors associated with the condition, a logistic regression model was employed to construct a nomogram, alongside the establishment of five ML models. To assess their effectiveness, the best-performing ML model and nomogram were evaluated on a separate test group comprising 77 additional patients. The evaluation involved comparing the area under the curve (AUC) and Brier score.

Results: Among the 269 patients studied, the incidence of POPF was found to be 56.9%, with 106 patients (69.3%) experiencing clinically-relevant POPF. We identified six independent factors associated with POPF, including body mass index (BMI), SMI, pancreatic duct dilatation, tumor size, triglyceride levels, and the ratio of aspartate aminotransferase to alanine aminotransferase (AST/ALT) on the first postoperative day. When evaluated on the test set, the Gaussian Naive Bayes (GNB) model, which was the best-performing ML model, achieved an AUC of 0.824 and a Brier score of 0.175. The corresponding performance indicators for the nomogram were 0.844 for AUC and 0.165 for the Brier score.

Conclusions: This study found that there is minimal difference between ML and the nomogram based on logistic regression in predicting POPF. Additionally, SMI shows promise as a potential and practical tool for assessing the risk of POPF.

Keywords: Machine learning (ML); nomogram; skeletal muscle index (SMI); pancreaticoduodenectomy (PD); postoperative pancreatic fistula (POPF)

Submitted Sep 30, 2023. Accepted for publication Dec 17, 2023. Published online Feb 27, 2024.

doi: 10.21037/gs-23-410

View this article at: <https://dx.doi.org/10.21037/gs-23-410>

Introduction

Pancreaticoduodenectomy (PD) remains the primary curative treatment for pancreaticobiliary tract cancer, despite its inherent complexity and multi-step surgical procedure that contribute to a high postoperative complication rate of approximately 50%. Of particular concern is postoperative pancreatic fistula (POPF) (1), which has been reported in 10% to 30% of pancreatic cancer patients undergoing surgery and is associated with increased complication rates and prolonged hospital stays (2,3). Consequently, it is crucial to identify high-risk patients to optimize clinical decision-making.

The utilization of skeletal muscle index (SMI) as a diagnostic tool for identifying sarcopenic patients is a well-established practice. SMI quantifies the proportion of skeletal muscle at the third lumbar vertebra (L3). Studies have demonstrated a relationship between sarcopenia and prognosis in patients with cirrhosis and cancer (4,5). Furthermore, Pecorelli *et al.* have identified sarcopenia as a reliable predictor of postoperative mortality (6). However, the association between SMI and complications in pancreatic surgery remains inconclusive, necessitating further exploration of its clinical significance.

The field of medicine has shown a growing interest in machine learning (ML) due to its ability to uncover complex non-linear relationships. ML's capacity for continual learning enables models to autonomously learn from dynamic input data and adapt their behavior (7). Consequently, there is considerable interest in applying

ML techniques to medical diagnosis, treatment decision-making, and gene analysis (8-10). Although traditional logistic regression has been widely used in recent years due to its simplicity and interpretability, ML offers promising alternatives. Therefore, this study aims to develop a novel prediction model for POPF using ML and a nomogram based on multivariate logistic regression analysis and to compare the performance of these two approaches. We present this article in accordance with the TRIPOD reporting checklist (available at <https://gs.amegroups.com/article/view/10.21037/gS-23-410/rc>).

Methods

Patient cohort

Data were retrospectively collected from a cohort of 290 patients who underwent the Whipple operation at The First Affiliated Hospital of Wenzhou Medical University between January 2012 and January 2021. Exclusion criteria included patients with tumors displaying distant metastasis (n=11) and cases with significant data loss (n=10). Ultimately, a total of 269 patients (158 men and 111 women) were included in the analysis. Seven surgeons performed all operations during the study period. Additionally, to assess the performance of the models, prospectively collected data from 77 patients who underwent the same surgery between January 2021 and January 2023 at The First Affiliated Hospital of Wenzhou Medical University were used following the same procedure.

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by Ethics Board of The First Affiliated Hospital of Wenzhou Medical University (No. KY2023-R135). Informed consent was obtained from the patients.

Date collection

A total of 36 variables potentially related to POPF were initially collected for analysis. Any variables with missing values exceeding 20% were excluded from the study. These variables encompassed various aspects, including patient characteristics, preoperative and intraoperative data, and laboratory results obtained on the first postoperative day (refer to *Table 1*). The size of the pancreatic duct and SMI were measured using preoperative imaging, predominantly computed tomography). In this study, SMI was calculated as the area of skeletal muscle at the level of the L3 vertebra

Highlight box

Key findings

- The skeletal muscle index (SMI) value may relate to pancreatic fistula after pancreaticoduodenectomy. Moreover, machine learning (ML) and nomogram based on logistic regression have similar predictive capabilities in small sample studies.

What is known and what is new?

- Pancreatic fistula is a main complication after pancreaticoduodenectomy and it was common to use simple multivariate analysis to build predictive models in the past.
- SMI can be a potential factor for assessing the risk of pancreatic fistula and ML was used to build models.

What is the implication, and what should change now?

- Further studies are needed to establish a pancreatic fistula evaluation system based on clinical characteristics and SMI for helping surgeons to implement timely medical intervention.

Table 1 Comparison of perioperative variables between POPF group and non-POPF group

Variables	All (N=269)	Non-POPF group (n=116)	POPF group (n=153)	P value
Age (years)	63 [56, 69]	61 [57, 68]	63 [56, 69]	0.281
Sex				
Female	111 (41.26)	50 (43.10)	61 (39.87)	0.594
Male	158 (58.74)	66 (56.90)	92 (60.13)	
BMI (kg/m ²)	22.43 (20.08, 24.46)	20.92 (19.53, 23.14)	23.44 (20.83, 25.15)	<0.001*
Hypertension	83 (30.86)	32 (27.59)	51 (33.33)	0.314
Diabetes	62 (23.05)	29 (25.00)	33 (21.57)	0.510
History of alcohol	42 (15.61)	15 (12.93)	27 (17.65)	0.293
Hepatitis	12 (4.46)	5 (4.31)	7 (4.58)	0.919
History of previous surgery	52 (19.33)	21 (18.10)	31 (20.26)	0.657
Pre-and intraoperative indicators				
Neutrophils (×10 ⁹ /L)	4.00 (2.95, 5.30)	3.65 (2.72, 5.05)	4.19 (3.20, 5.34)	0.023*
Hemoglobin (g/L)	117.87 (18.43)	116.20 (17.64)	119.14 (18.91)	0.196
Platelets (×10 ⁹ /L)	253.00 (203.00, 328.00)	252.00 (197.00, 312.00)	255.00 (213.00, 345.00)	0.138
Total bilirubin (μmol/L)	66.00 (15.00, 187.00)	76.00 (14.00, 187.00)	61.00 (15.00, 183.00)	0.728
Direct bilirubin (μmol/L)	56.00 (8.00, 155.00)	64.00 (9.00, 155.00)	53.00 (8.00, 150.00)	0.561
AST/ALT	0.88 (0.65, 1.21)	0.83 (0.60, 1.21)	0.89 (0.67, 1.22)	0.181
Alkaline phosphatase (U/L)	305.00 (135.00, 548.00)	318.00 (157.00, 607.00)	284.00 (121.00, 525.00)	0.171
Total cholesterol (mmol/L)	5.22 (4.13, 6.57)	5.05 (4.02, 6.43)	5.32 (4.23, 6.61)	0.296
Triglyceride (mmol/L)	1.78 (1.07, 2.78)	1.54 (0.96, 2.25)	1.96 (1.19, 3.07)	0.002*
Creatinine (μmol/L)	58.00 (50.00, 69.00)	55.00 (49.00, 65.00)	59.00 (52.00, 71.00)	0.031*
Albumin (g/L)	37.23 (4.33)	37.20 (4.34)	37.25 (4.32)	0.926
SMI (cm ² /m ²)				
≤30 (female), ≤40 (male)	213 (79.18)	78 (67.24)	135 (88.24)	<0.001*
>30 (female), >40 (male)	56 (20.82)	38 (32.76)	18 (11.77)	
CA199 (U/mL)	82.50 (27.60, 255.70)	92.30 (38.40, 536.90)	71.90 (20.00, 204.60)	0.014*
Pancreatic duct dilatation				
No (≤4 mm)	130 (48.33)	36 (31.03)	94 (61.44)	<0.001*
Yes (>4 mm)	139 (51.67)	80 (68.97)	59 (38.56)	
Tumor size (cm)				
≤4	213 (79.18)	78 (67.24)	135 (88.24)	<0.001*
>4	56 (20.82)	38 (32.76)	18 (11.77)	
Surgical modality				
Laparotomy	254 (94.42)	111 (95.69)	143 (93.46)	0.431
Laparoscopy	15 (5.58)	5 (4.31)	10 (6.54)	

Table 1 (continued)

Table 1 (continued)

Variables	All (N=269)	Non-POPF group (n=116)	POPF group (n=153)	P value
Duration of surgery (h)	5.08 (4.18, 6.02)	5.17 (4.00, 6.13)	5.08 (4.38, 5.98)	0.707
Intraoperative transfusion				
Yes	127 (47.39)	56 (48.28)	71 (46.71)	0.799
No	141 (52.61)	60 (51.72)	81 (53.29)	
Pathology				
Malignant	238 (88.48)	103 (88.79)	135 (88.24)	0.880
Benign	31 (11.52)	13 (11.21)	18 (11.77)	
Postoperative indicators (postoperative day 1)				
Neutrophils ($\times 10^9/L$)	10.60 (8.23, 13.40)	9.50 (7.67, 12.60)	11.40 (8.92, 14.00)	0.002*
Hemoglobin (g/L)	103.00 (92.00, 115.00)	100.00 (91.00, 114.00)	104.00 (92.00, 115.00)	0.220
Platelets ($\times 10^9/L$)	219.00 (167.00, 287.00)	211.00 (158.00, 268.00)	229.00 (176.00, 303.00)	0.120
Total bilirubin ($\mu\text{mol/L}$)	54.00 (23.00, 132.00)	59.00 (20.00, 124.00)	51.00 (26.00, 154.00)	0.716
Direct bilirubin ($\mu\text{mol/L}$)	43.00 (12.00, 108.00)	46.00 (12.00, 100.00)	42.00 (13.00, 130.00)	0.863
AST/ALT	0.98 (0.74, 1.37)	0.88 (0.61, 1.21)	1.07 (0.84, 1.44)	<0.001*
Alkaline phosphatase (U/L)	86.75 (54.00, 173.00)	84.00 (53.75, 184.00)	87.00 (57.00, 152.50)	0.892
Creatinine ($\mu\text{mol/L}$)	55.50 (46.00, 66.00)	53.00 (46.00, 63.00)	57.00 (47.00, 69.00)	0.037*
Albumin (g/L)	32.20 (5.61)	31.90 (5.45)	32.43 (5.71)	0.446

Data are presented as median (Q1, Q3), mean (SD) or n (%). *, significance was set at $P < 0.100$. POPF, postoperative pancreatic fistula; BMI, body mass index; AST/ALT, the ratio of aspartate aminotransferase to alanine aminotransferase; SMI, skeletal muscle index; SD, standard deviation.

divided by the square of the patient's height. Postoperative outcomes were assessed, with a specific focus on the presence of pancreatic fistulas according to the diagnostic guidelines outlined by the International Study Group of Pancreatic Fistulas (ISGPS) (11). Other complications were evaluated based on the latest guidelines and consensus (12-15).

SMI acquisition

The images were processed using the GE Healthcare Advantage workstation (version 4.6) and the Infinity PACS system. The entire procedure was conducted by a radiologist with more than 5 years of experience who was blinded to any clinical information regarding the patients.

Once the image capturing the longest distance between the bilateral transverse processes at the L3 level was obtained, the Hounsfield unit (HU) scale was used to differentiate muscle tissue. The radiologist manually delineated the regions of interest in all images to calculate the skeletal muscle area, as depicted in *Figure 1*. To

determine the SMI, the following formula was applied: $\text{SMI} = \text{skeletal muscle area (cm}^2\text{)}/\text{height (m}^2\text{)}$. The cutoff value for SMI was determined based on a receiver operating characteristic (ROC) curve, with $40 \text{ cm}^2/\text{m}^2$ for males and $30 \text{ cm}^2/\text{m}^2$ for females.

Surgical technique and perioperative management

In cases where patients presented with evident jaundice before surgery, we followed the conventional approach of endoscopic retrograde cholangiopancreatography (ERCP) or percutaneous transhepatic cholangial drainage (PTCD) to alleviate liver dysfunction. To prevent infection, antibiotic prophylaxis was administered approximately half an hour to an hour before surgery, typically using second-generation cephalosporins. Perioperative administration of somatostatin and its analogs was routinely implemented. In the absence of specific indications, a standardized Whipple procedure was performed using either laparoscopy or laparotomy, with pancreatic duct stenting included in



Figure 1 Region of interest delineation. A 44-year-old male patient was undergoing a preoperative imaging examination. A plain CT scan of the abdomen was performed with the region of interest delineation of the skeletal muscle at the transverse process of the third lumbar vertebra. CT, computed tomography.

the surgical procedures. Following the procedure, two drains were placed, with one positioned anterior to the pancreaticoenteric anastomosis and the other placed posteriorly. The drainage fluid was collected for measuring amylase content levels on postoperative days one, three, and five, with further follow-up examinations conducted as deemed necessary.

Method of missing values

For continuous variables with missing values constituting less than 20% of the data, we employed the k-nearest neighbors (KNN) approach with $N=3$ for imputation. This method introduces minimal bias compared to using complete observations.

Statistical analysis and ML

The normality of the data was assessed using Shapiro-Wilk tests. Continuous variables were reported as mean \pm standard deviation (SD) or median and interquartile range (IQR) based on the results of the Shapiro-Wilk tests. Independent-sample t -tests or Mann-Whitney U tests were employed, depending on the nature of the variables. Categorical variables were presented as percentages and compared using the χ^2 test or Fisher's exact test. Variables with a P value <0.1 were included in the LASSO regression to address issues of collinearity and prevent overfitting. Independent factors were identified through binary logistic regression analysis, and a predictive nomogram was

developed, evaluated using the area under the curve (AUC) and calibration curve.

The study population ($n=269$) was randomly divided into a training set and a validation set in a ratio of 4:1. To achieve the highest prediction accuracy, we employed ten-fold cross-validation on the training datasets with various models: eXtreme Gradient Boosting (XGBoost), random forest (RF), AdaBoost, Gaussian Naive Bayes (GNB), support vector machine (SVM), and KNN. The best-performing model was selected based on the maximum AUC value and the minimum Brier score.

The performance of the nomogram and the best ML model was evaluated using a test set of 77 patients. Furthermore, the SHAP value was used to visually represent the importance and contribution of each factor to the occurrence of POPF. All data analysis was conducted using SPSS V26.0 (IBM Corp., Armonk, NY, USA), R version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria), and Python version 3.7 (Google, Mountain View, CA, USA). The flowchart illustrating the data analysis process is depicted in *Figure 2*.

Results

General characteristics

A total of 269 patients participated in this study, with an average hospital stay of 24 days. The study population consisted of 158 males and 111 females, with a median age of 63 years. *Table 2* displays the composition of surgically removed samples, with malignant neoplasms comprising a substantial portion (up to 88.48%). Among the histological findings, 136 patients (50.56%) were diagnosed with pancreatic tumors, of which 101 were pancreatic adenocarcinomas (37.55%).

The incidence of POPF was 56.9%, with 106 patients (69.3%) experiencing clinically-relevant POPF. The occurrence of postoperative complications differed significantly between the two groups (*Table 3*). No significant differences were observed among the seven participating surgeons in terms of the incidence rate of POPF ($P=0.27$). Furthermore, there was no significant difference in the occurrence of POPF between the testing group consisting of 77 patients and the previous group of 269 patients ($P=0.443$).

Predictors of POPF

Table 1 presents a comparison of perioperative variables

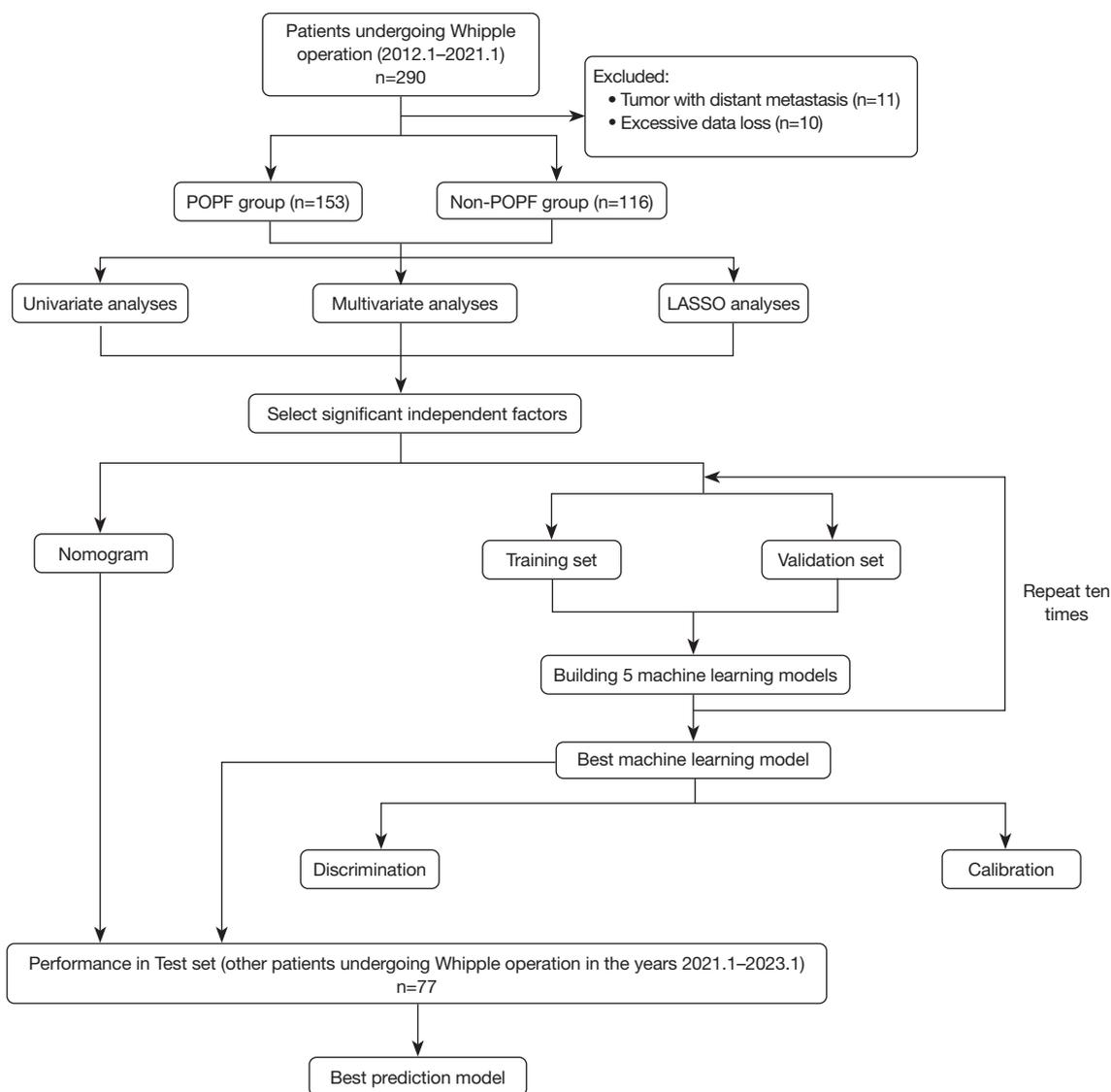


Figure 2 Technology roadmap for this study. Patients who were hospitalized for Whipple operation at The First Affiliated Hospital of Wenzhou Medical University from January 2012 to January 2021 were employed to construct various machine learning models and a nomogram model. Patients enrolled in the same hospital from January 2021 to January 2023 were used to evaluate and compare the two separate models. POPF, postoperative pancreatic fistula; LASSO, least absolute shrinkage and selection operator.

between the non-POPF group and the POPF group. Significant differences were observed in preoperative laboratory test indicators, including TG, CA199, neutrophils, and creatinine, as well as postoperative laboratory test indicators, including neutrophils, AST/ALT, and creatinine. The POPF group had a lower proportion of patients with pancreatic duct dilation and a higher proportion of patients with tumor sizes less than 4 cm in diameter. Other variables, such as SMI and BMI, also

showed P values less than 0.1.

Through LASSO regression with ten-fold cross-validation, the minimum distance standard error ($\lambda=0.059$) was selected, resulting in six features with nonzero coefficients (Figure 3). The variables included in the model building were BMI, SMI, pancreatic duct dilation, tumor size, TG, and AST/ALT (postoperative day 1).

In the multivariate analysis, BMI [OR =1.137, 95% confidence interval (CI): 1.036–1.254, P=0.008], SMI

Table 2 The demographic and histopathological details of 269 patients undergoing pancreaticoduodenectomy

Characteristics	Values
Age (years)	63 [56, 69]
Sex	
Female	111 (41.26)
Male	158 (58.74)
Length of postoperative stay (days)	24 [18, 34]
Malignant tumors	238 (88.48)
Benign tumors	31 (11.52)
Pancreatic abnormalities	136 (50.56)
Pancreatic adenocarcinoma	101 (37.55)
Pancreatic neuroendocrine neoplasms	5 (1.86)
Pancreatic cystadenoma	11 (4.09)
Intraductal papillary mucinous neoplasms	3 (1.12)
Other pancreatic abnormalities	16 (5.95)
Bile abnormalities	51 (18.96)
Bile duct adenocarcinoma	49 (18.22)
Bile duct neuroendocrine carcinoma	1 (0.37)
Chronic inflammation of bile duct mucosa	1 (0.37)
Duodenal abnormalities	81 (30.11)
Duodenal adenocarcinoma	73 (27.14)
Other duodenal abnormalities	8 (2.97)
Gastric adenocarcinoma	1 (0.37)

Data are presented as median [Q1, Q3] or n (%).

(female >30, male >40) (OR =2.534, 95% CI: 1.191–5.631, P=0.018), TG (OR =1.326, 95% CI: 1.047–1.716, P=0.025), and AST/ALT on postoperative day 1 (OR =2.668, 95% CI: 1.502–4.961, P=0.001) were identified as independent risk factors for POPF. Pancreatic duct dilatation (OR =0.340, 95% CI: 0.190–0.601, P<0.001) and tumor size >4 cm (OR =0.287, 95% CI: 0.137–0.579, P=0.001) were identified as protective factors (Table 4). Furthermore, these indicators did not differ between patients with a biochemical leak and those who developed POPF grades B or C, except for SMI (Table S1).

Predictive models

A nomogram was constructed based on the significant

variables identified through multivariate logistic regression analysis (Figure 4). The nomogram exhibited an AUC of 0.804 in the training set. The calibration curve, displayed in, indicated a mean absolute error of 0.014 and a Brier score of 0.179, suggesting a good fit between the predicted and actual test values.

Furthermore, we assessed the performance of the models by incorporating the variables into five ML models and visualizing their features through a forest plot of the AUC score and a calibration plot. In Figure 5, the GNB model demonstrated the highest AUC score (0.764; 95% CI: 0.584–0.937). Additionally, when comparing the Brier scores of the models (Figure 6), the GNB model achieved a score of 0.102, which was lower than that of the other four models. Moreover, the GNB model exhibited the highest accuracy, specificity, and Youden index, as presented in Table 5. Based on this analysis, we can conclude that the GNB model was suitable for predicting POPF and outperformed other algorithms. Table 5 provides an overview of the performance values for all models in the validation set.

Model test

In the test set, the performance of two models was as follows: the GNB model achieved an AUC of 0.824 (95% CI: 0.782–0.921), a Brier score of 0.175, and exhibited an accuracy of 0.779, sensitivity of 0.725, specificity of 0.838, Youden index of 0.563, and F1 score of 0.773. On the other hand, the nomogram demonstrated an AUC of 0.844 (95% CI: 0.745–0.934), a Brier score of 0.165, an accuracy of 0.753, sensitivity of 0.825, specificity of 0.676, Youden index of 0.501, and F1 score of 0.776 (Table 6 and Figure 7). The nomogram showed a slight advantage in predicting POPF, displaying better AUC, Brier score, and sensitivity compared to the GNB model. However, when considering the F1 score, the GNB model and the nomogram showed similar performance. SHAP summary plots were utilized to rank the features in descending order: pancreatic duct dilatation, tumor size, BMI, SMI, TG, and AST/ALT (postoperative day 1) (Figure 8).

Discussion

Since Alessandro Codivilla performed the first “en block” resection of the head of the pancreas and duodenum in Italy, advancements and refinements in pancreatic surgical techniques have been ongoing (16,17). The focus on POPF

Table 3 Complications of 269 patients undergoing pancreaticoduodenectomy

Postoperative condition	All (N=269)	Non-POPF group (n=116)	POPF group (n=153)	P value
Pancreatic leak	153 (56.9)	–	–	–
Biochemical leak	47 (17.7)	–	–	–
Grade B	88 (32.7)	–	–	–
Grade C	18 (6.7)	–	–	–
Bile leakage	24 (8.9)	5 (1.9)	19 (7.1)	0.021*
Delayed gastric emptying [†]	70 (26.0)	10 (3.7)	60 (22.3)	<0.001*
Hemorrhage	33 (12.3)	3 (1.1)	30 (11.2)	<0.001*
Intraperitoneal infection with drainage therapy	17 (6.3)	2 (0.7)	15 (5.6)	0.007*
Pulmonary complications	30 (11.2)	5 (1.9)	25 (9.3)	0.002*
Additional surgery or interventional therapy	27 (10.0)	1 (0.4)	26 (9.7)	<0.001*
Clavien-Dindo complications				
I–II	112 (41.6)	35 (13.0)	77 (28.6)	0.001*
III–V	42 (15.6)	3 (1.1)	39 (14.5)	<0.001*

Data are presented as n (%). [†], indwelling gastric tube >14 days or indwelling catheterization again. *, significance was set at P<0.05. POPF, postoperative pancreatic fistula.

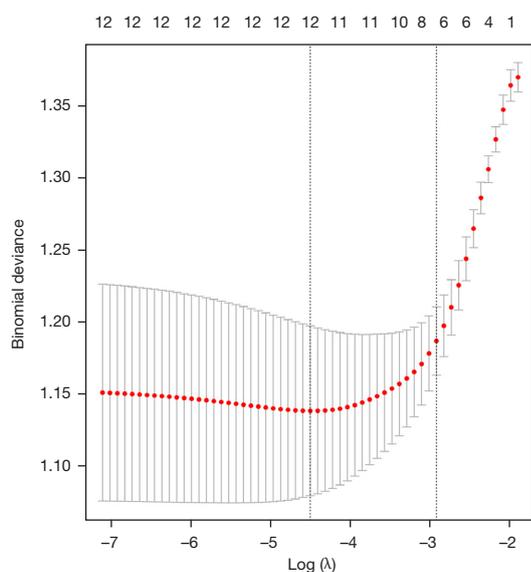


Figure 3 Coefficients of the LASSO model: LASSO coefficient profiles of the 6 features. The LASSO analyses contained variables with a P value <0.1 in univariate analyses. When the standard errors of the minimum distance for λ were 0.059, the variables corresponding to the model were BMI, skeletal muscle index, pancreatic duct dilatation, tumor size, triglyceride, and AST/ALT (postoperative day 1). LASSO, least absolute shrinkage and selection operator; BMI, body mass index; AST/ALT, the ratio of aspartate aminotransferase to alanine aminotransferase.

remains constant due to its association with an increased risk of complications. Early intervention is crucial in preventing the occurrence of POPF, as it is a complex phenomenon. Recent studies by Nakata *et al.* and Smits *et al.* consistently demonstrate that catheter drainage is a vital intervention for POPF, offering safer outcomes compared to surgical lavage, drainage, or pancreatectomy (18,19). A prospective, multicenter, randomized controlled trial has also shown that continuous infusion of somatostatin post-operation may reduce the incidence of grade B and C pancreatic fistulae within 30 days (1). Although these postoperative preventive measures have played a corresponding role in some studies, they have not yet achieved consistent and complete prevention of pancreatic fistula in clinical practice. However, regardless of the circumstances, it can be inferred that early identification and prediction are essential in the management of POPF, as conservative measures can effectively reverse the condition.

It is important to note that this study examines pancreatic fistulae, encompassing both biochemical leaks and clinically significant manifestations. Among the 153 patients with pancreatic fistulae, the majority (69.3%) experienced clinically relevant pancreatic fistulae. No significant differences were observed between patients with biochemical leaks and those with POPF grades B or C, except for SMI, in terms of five independent correlated

Table 4 Multivariate logistic regression analysis of risk factors and protective factor correlated with pancreatic fistula after pancreaticoduodenectomy

Characteristics	OR	95% CI	P value
BMI (kg/m ²)	1.137	1.036–1.254	0.008*
SMI, >30 (female), >40 (male)	2.534	1.191–5.631	0.018*
Pancreatic duct dilatation	0.340	0.190–0.601	<0.001*
Tumor size >4 cm	0.287	0.137–0.579	0.001*
TG (mmol/L)	1.326	1.047–1.716	0.025*
Postoperative day 1 AST/ALT	2.668	1.502–4.961	0.001*

*, significance was set at $P < 0.05$. OR, odds ratio; CI, confidence interval; BMI, body mass index; SMI, skeletal muscle index; TG, triglyceride; AST/ALT, the ratio of aspartate aminotransferase to alanine aminotransferase.

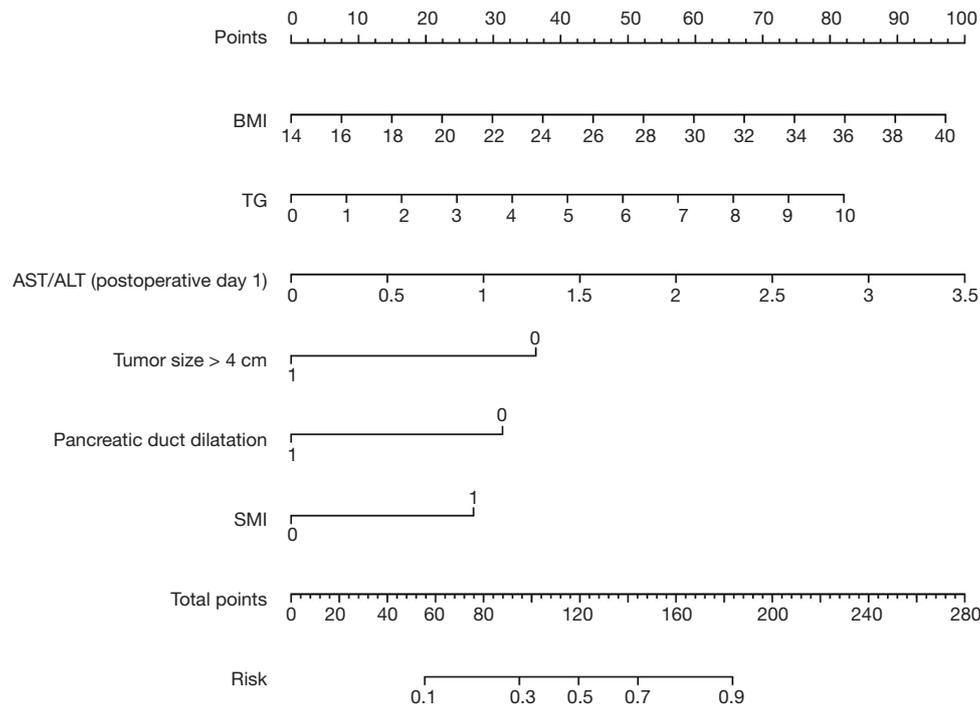


Figure 4 Nomogram for predicting pancreatic fistula after pancreatoduodenectomy. BMI, body mass index; TG, triglyceride; AST/ALT, the ratio of aspartate aminotransferase to alanine aminotransferase; SMI, skeletal muscle index.

factors (Table S1). Furthermore, it is noteworthy that biochemical leaks can contribute to an increase in the Clavien-Dindo classification (Table S2). Considering that the Clavien-Dindo classification reflects not only morbidity but also serves as a significant predictor of costs, the inclusion of biochemical leaks in this study is justified.

The study ultimately encompassed 269 patients who underwent the Whipple operation (child reconstruction) without any distant metastases. Following this, predictive

models were developed with the incorporation of six independent correlative factors as predictive variables. These factors include SMI, BMI, pancreatic duct dilatation, TG, tumor size, and postoperative AST/ALT levels. In the realm of ML models that learn via training sets, the GNB model illustrated outstanding performance in the validation set. However, discerning the advantages and disadvantages of the nomogram and GNB models within the external validation queue poses a significant challenge. Interestingly,

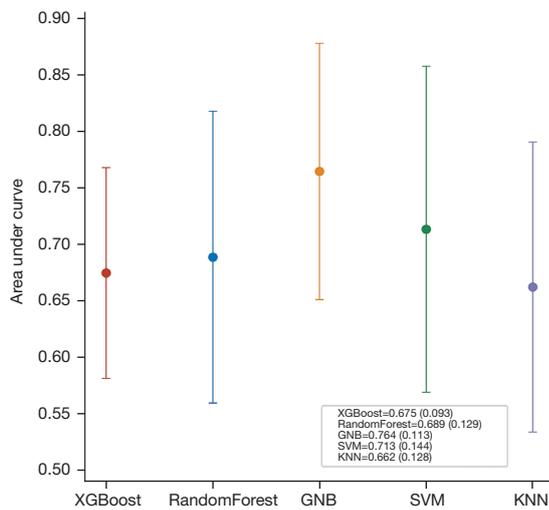


Figure 5 Forest plot of the AUC score of five machine learning models. The GNB achieved a better AUC value (0.764±0.113) in contrast to the other models. AUC, area under curve; XGBoost, eXtreme Gradient Boosting; GNB, Gaussian Naive Bayes; SVM, support vector machine; KNN, K-nearest neighbor.

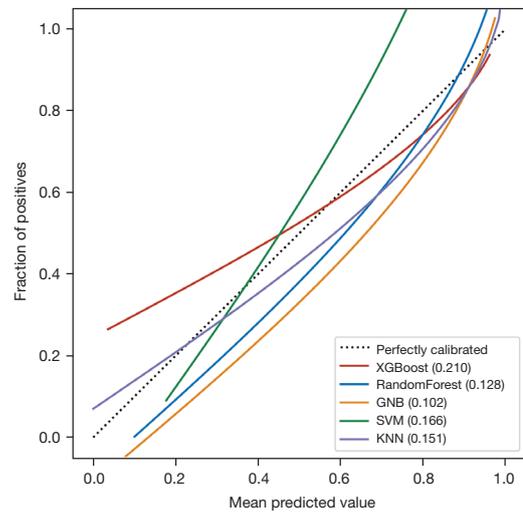


Figure 6 Calibration plots of five machine learning models. The GNB model achieved a lower Brier score (0.102) in contrast to the other models. XGBoost, eXtreme Gradient Boosting; GNB, Gaussian Naive Bayes; SVM, support vector machine; KNN, K-nearest neighbor.

Table 5 Performance metrics of five models in the validation dataset

Model	AUC (95% CI)	Brier score	Accuracy	Sensitivity	Specificity	Youden index	F1 score
XGBoost	0.675 (0.464–0.883)	0.210	0.587	0.836	0.545	0.381	0.741
RandomForest	0.689 (0.468–0.877)	0.128	0.624	0.810	0.570	0.380	0.731
GNB	0.764 (0.584–0.937)	0.102	0.646	0.776	0.799	0.575	0.726
SVM	0.713 (0.524–0.898)	0.166	0.639	0.739	0.733	0.472	0.717
KNN	0.662 (0.468–0.858)	0.151	0.590	0.569	0.733	0.302	0.594

AUC, area under curve; CI, confidence interval; XGBoost, eXtreme Gradient Boosting; GNB, Gaussian Naive Bayes; SVM, support vector machine; KNN, K-nearest neighbor.

Table 6 Comparison of two prediction models on external dataset

Model	AUC (95% CI)	Brier score	Accuracy	Sensitivity	Specificity	Youden index	F1 score
Nomogram	0.844 (0.745–0.934)	0.165	0.753	0.825	0.676	0.501	0.776
GNB	0.824 (0.782–0.921)	0.175	0.779	0.725	0.838	0.563	0.773

AUC, area under curve; CI, confidence interval; GNB, Gaussian Naive Bayes.

both models surpassed the predictive model platform established by Mungroop *et al.* in terms of AUC levels (20). Unlike other scoring systems that depend on subjective indicators, for instance, pancreatic texture, the employment of numerical indicators in this study enhances the scoring

process’s objectivity (20,21).

Incorporating radiographic techniques, such as extracellular volume fraction and the pancreatic parenchymal-to-portal venous phase IC ratio, has been recognized as a significant area of research in recent years. These techniques offer

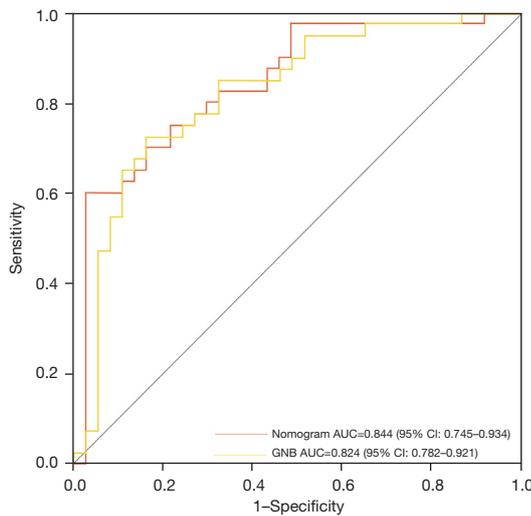


Figure 7 Receiver-operating characteristic curve of the nomogram model and Gaussian Naive Bayes model in the test set. AUC, area under curve; CI, confidence interval; GNB, Gaussian Naive Bayes.

potential advantages for the early detection of POPF (22,23). Jin *et al.*'s recent meta-analysis, which included 1,814 patients, indicated that the existence of sarcopenia, as characterized by SMI, was correlated with unfavorable clinical outcomes in ovarian cancer patients (5). Earlier research had explored the prognostic value of body composition in forecasting the incidence of POPF (6,24). Our study generated several insightful findings, including a significant difference in the incidence of POPF between patients with high and low SMI. Patients with a higher SMI were more likely to develop POPF, with odds ratios reaching 7.213. It is crucial to emphasize that diagnosing sarcopenia based solely on SMI is inadequate from a stringent perspective. Nonetheless, in the ongoing debate concerning whether sarcopenia is a protective or risk factor for POPF (25-27), our study contributes new evidence about SMI to this discussion. Further investigations are warranted to elucidate potential

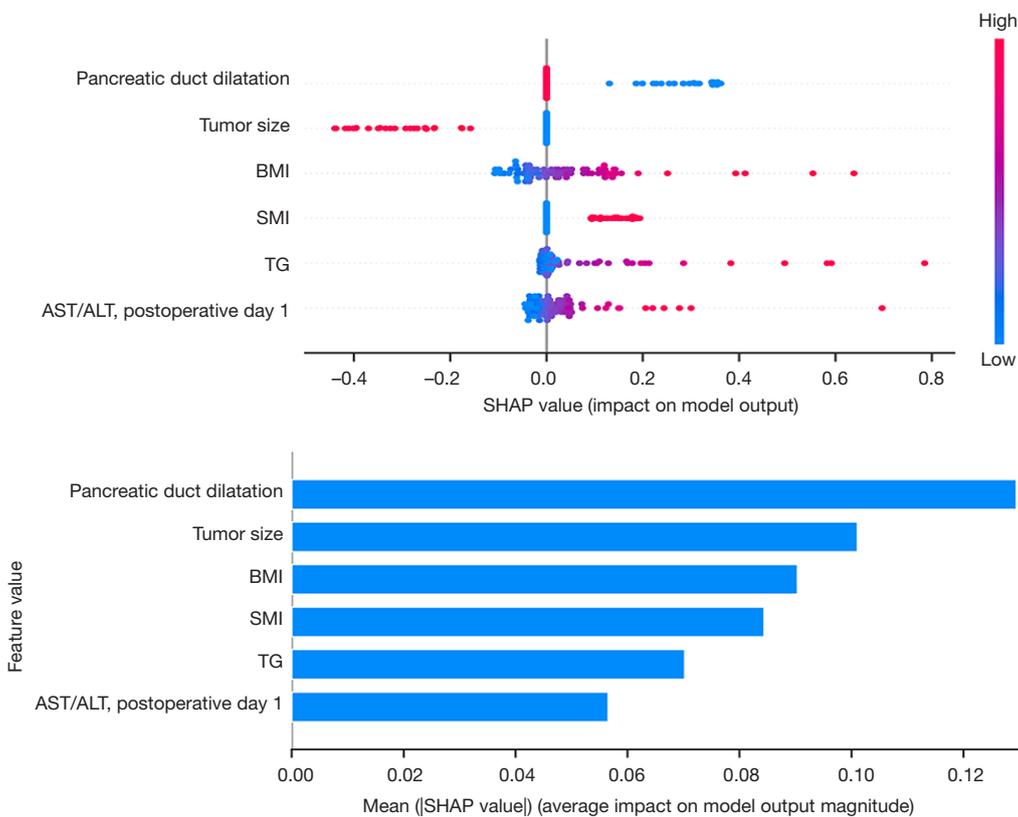


Figure 8 SHAP analysis of the Gaussian Naive Bayes model. A visualization of each sample, where each point represents a sample feature, with red representing higher values and blue representing lower values. These features were ranked according to the mean absolute value of SHAP. AST/ALT, the ratio of aspartate aminotransferase to alanine aminotransferase; BMI, body mass index; SMI, skeletal muscle index; TG, triglyceride.

mechanisms, which might include the relationship between muscle, fat, and inflammation (28,29). Moreover, supplementary *Table 1* suggests that SMI may influence the development of POPF.

The dominant viewpoint holds that both BMI and the diameter of the pancreatic duct are reliable predictors for POPF, a conclusion that is supported by our research findings (20,30). A SHAP analysis of our GNB model revealed that BMI, pancreatic duct dilation, and TG bear considerable influence, emphasizing their essential role in predicting POPF. This could be due to the technical challenges encountered while operating on smaller pancreatic ducts and the anatomical and physiological changes that occur in patients with obesity, potentially affecting their susceptibility to severe illness (21,31). The rising prevalence of obesity has brought pancreatic steatosis to the forefront as a significant clinical concern. Ramkissoon *et al.* proposed that obesity impacts pancreatic steatosis, leading to enhanced progression and lethality of pancreatic cancer (32). Consequently, it is unequivocal that obesity represents a significant risk to patient recovery post-surgery.

Tumor size was found to be significantly associated with POPF in both univariate and multivariate analysis. To our knowledge, this factor has not been reported previously. Surgeons with substantial experience suggest that large tumor volume may amplify surgical difficulty and duration. We postulate that a connection might exist between tumor size and volume of the remnant pancreas. In distal pancreatectomy, the volume of the remaining pancreas has been implicated in POPF (33). However, due to the data being gathered through imaging reports and surgical documentation, there are limitations to this study that warrant further exploration with a larger sample size.

Plasma levels of AST and ALT are commonly used indicators of liver function, signaling hepatocyte injury. Generally, AST is present in various tissues, while ALT is more specific and enriched in the liver. AST has been identified as an essential component in glycolysis within the mitochondria (34). Notably, cancer pathophysiology frequently correlates with increased glycolytic activity (35). Thus, there is evidence to suggest that the AST/ALT ratio may serve as a promising biomarker for predictive purposes. Stocken *et al.* observed a positive association between elevated AST levels and poor prognosis in patients with pancreatic cancer (36). In a recent study, an AST/ALT ratio greater than 0.89 resulted in a significant increase in postoperative mortality following PD (37).

Correspondingly, our studies also revealed that as the AST/ALT ratio in the nomogram exceeded 0.5, the incidence of POPF significantly increased. We hypothesize that POPF might induce oxidative stress, a multifaceted pathological mechanism, contributing to the onset and progression of hepatocyte impairment (38).

The retrospective nature of this study inherently imposes limitations. Potential loss of unknown information may arise due to the absence of direct observation and real-time recording. Nonetheless, this study, conducted solely at one institution, ensures a degree of standardization in perioperative management. While a single-center external validation of the model was performed, prospective external validation involving multiple centers is still necessary. Given the 'black-box' characteristic of ML, future efforts should aim to enhance model visualization, thereby facilitating its clinical utility via web-based platforms.

Conclusions

Both the GNB model and the nomogram exhibit predictive potential in the onset of POPF following pancreatoduodenectomy, with the nomogram possibly possessing a slight edge. Their respective high F1 scores and sensitivity lend credence to their potential utility in clinical practice, particularly for early detection and the selection of the most effective therapeutic strategy. SMI, serving as a straightforward indicator of sarcopenia, enables the objective and accurate assessment of POPF risk.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://gs.amegroups.com/article/view/10.21037/gc-23-410/rc>

Data Sharing Statement: Available at <https://gs.amegroups.com/article/view/10.21037/gc-23-410/dss>

Peer Review File: Available at <https://gs.amegroups.com/article/view/10.21037/gc-23-410/prf>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://gs.amegroups.com>)

[com/article/view/10.21037/gs-23-410/coif](https://doi.org/10.21037/gs-23-410/coif)). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by Ethics Board of First Affiliated Hospital of Wenzhou Medical University (No. KY2023-R135). Informed consent was obtained from the patients.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Cao Z, Qiu J, Guo J, et al. A randomised, multicentre trial of somatostatin to prevent clinically relevant postoperative pancreatic fistula in intermediate-risk patients after pancreaticoduodenectomy. *J Gastroenterol* 2021;56:938-48.
2. Shen Z, Chen H, Wang W, et al. Machine learning algorithms as early diagnostic tools for pancreatic fistula following pancreaticoduodenectomy and guide drain removal: A retrospective cohort study. *Int J Surg* 2022;102:106638.
3. Schouten TJ, Henry AC, Smits FJ, et al. Risk Models for Developing Pancreatic Fistula After Pancreatoduodenectomy: Validation in a Nationwide Prospective Cohort. *Ann Surg* 2023;278:1001-8.
4. Montano-Loza AJ, Duarte-Rojo A, Meza-Junco J, et al. Inclusion of Sarcopenia Within MELD (MELD-Sarcopenia) and the Prediction of Mortality in Patients With Cirrhosis. *Clin Transl Gastroenterol* 2015;6:e102.
5. Jin Y, Ma X, Yang Z, et al. Low L3 skeletal muscle index associated with the clinicopathological characteristics and prognosis of ovarian cancer: a meta-analysis. *J Cachexia Sarcopenia Muscle* 2023;14:697-705.
6. Pecorelli N, Carrara G, De Cobelli F, et al. Effect of sarcopenia and visceral obesity on mortality and pancreatic fistula following pancreatic cancer surgery. *Br J Surg* 2016;103:434-42.
7. Lee CS, Lee AY. Clinical applications of continual learning machine learning. *Lancet Digit Health* 2020;2:e279-81.
8. Machine learning in cancer diagnostics. *EBioMedicine* 2019;45:1-2.
9. Howard FM, Kochanny S, Koshy M, et al. Machine Learning-Guided Adjuvant Treatment of Head and Neck Cancer. *JAMA Netw Open* 2020;3:e2025881.
10. Du AX, Ali Z, Ajgeiy KK, et al. Machine learning model for predicting outcomes of biologic therapy in psoriasis. *J Am Acad Dermatol* 2023;88:1364-7.
11. Bassi C, Marchegiani G, Dervenis C, et al. The 2016 update of the International Study Group (ISGPS) definition and grading of postoperative pancreatic fistula: 11 Years After. *Surgery* 2017;161:584-91.
12. Wente MN, Bassi C, Dervenis C, et al. Delayed gastric emptying (DGE) after pancreatic surgery: a suggested definition by the International Study Group of Pancreatic Surgery (ISGPS). *Surgery* 2007;142:761-8.
13. Wente MN, Veit JA, Bassi C, et al. Postpancreatectomy hemorrhage (PPH): an International Study Group of Pancreatic Surgery (ISGPS) definition. *Surgery* 2007;142:20-5.
14. Koch M, Garden OJ, Padbury R, et al. Bile leakage after hepatobiliary and pancreatic surgery: a definition and grading of severity by the International Study Group of Liver Surgery. *Surgery* 2011;149:680-8.
15. Mao CC, Chen XD, Lin J, et al. A Novel Nomogram for Predicting Postsurgical Intra-abdominal Infection in Gastric Cancer Patients: a Prospective Study. *J Gastrointest Surg* 2018;22:421-9.
16. WHIPPLE AO. Observations on radical surgery for lesions of the pancreas. *Surg Gynecol Obstet* 1946;82:623-31.
17. Fuks D. Minimally invasive pancreatoduodenectomy. *Br J Surg* 2020;107:630-1.
18. Smits FJ, van Santvoort HC, Besselink MG, et al. Management of Severe Pancreatic Fistula After Pancreatoduodenectomy. *JAMA Surg* 2017;152:540-8.
19. Nakata K, Mori Y, Ikenaga N, et al. Management of postoperative pancreatic fistula after pancreatoduodenectomy: Analysis of 600 cases of pancreatoduodenectomy patients over a 10-year period at a single institution. *Surgery* 2021;169:1446-53.
20. Mungroop TH, van Rijssen LB, van Klaveren D, et al. Alternative Fistula Risk Score for Pancreatoduodenectomy (a-FRS): Design and International External Validation.

- Ann Surg 2019;269:937-43.
21. Schuh F, Mihaljevic AL, Probst P, et al. A Simple Classification of Pancreatic Duct Size and Texture Predicts Postoperative Pancreatic Fistula: A classification of the International Study Group of Pancreatic Surgery. *Ann Surg* 2023;277:e597-608.
 22. Sofue K, Ueshima E, Masuda A, et al. Estimation of pancreatic fibrosis and prediction of postoperative pancreatic fistula using extracellular volume fraction in multiphase contrast-enhanced CT. *Eur Radiol* 2022;32:1770-80.
 23. Shi HY, Lu ZP, Li MN, et al. Dual-Energy CT Iodine Concentration to Evaluate Postoperative Pancreatic Fistula after Pancreatoduodenectomy. *Radiology* 2022;304:65-72.
 24. Tanaka K, Yamada S, Sonohara F, et al. Pancreatic Fat and Body Composition Measurements by Computed Tomography are Associated with Pancreatic Fistula After Pancreatectomy. *Ann Surg Oncol* 2021;28:530-8.
 25. Sui K, Okabayashi T, Iwata J, et al. Correlation between the skeletal muscle index and surgical outcomes of pancreaticoduodenectomy. *Surg Today* 2018;48:545-51.
 26. Tsukagoshi M, Harimoto N, Araki K, et al. Impact of preoperative nutritional support and rehabilitation therapy in patients undergoing pancreaticoduodenectomy. *Int J Clin Oncol* 2021;26:1698-706.
 27. Perra T, Sotgiu G, Porcu A. Sarcopenia and Risk of Pancreatic Fistula after Pancreatic Surgery: A Systematic Review. *J Clin Med* 2022;11:4144.
 28. Reilly SM, Saltiel AR. Adapting to obesity with adipose tissue inflammation. *Nat Rev Endocrinol* 2017;13:633-43.
 29. Kirk B, Feehan J, Lombardi G, et al. Muscle, Bone, and Fat Crosstalk: the Biological Role of Myokines, Osteokines, and Adipokines. *Curr Osteoporos Rep* 2020;18:388-400.
 30. Roberts KJ, Sutcliffe RP, Marudanayagam R, et al. Scoring System to Predict Pancreatic Fistula After Pancreaticoduodenectomy: A UK Multicenter Study. *Ann Surg* 2015;261:1191-7.
 31. Shashaty MG, Stapleton RD. Physiological and management implications of obesity in critical illness. *Ann Am Thorac Soc* 2014;11:1286-97.
 32. Ramkissoon R, Gardner TB. Pancreatic Steatosis: An Emerging Clinical Entity. *Am J Gastroenterol* 2019;114:1726-34.
 33. Frozanpor F, Albiin N, Linder S, et al. Impact of pancreatic gland volume on fistula formation after pancreatic tail resection. *JOP* 2010;11:439-43.
 34. Greenhouse WV, Lehninger AL. Occurrence of the malate-aspartate shuttle in various tumor types. *Cancer Res* 1976;36:1392-6.
 35. Hsu PP, Sabatini DM. Cancer cell metabolism: Warburg and beyond. *Cell* 2008;134:703-7.
 36. Stocken DD, Hassan AB, Altman DG, et al. Modelling prognostic factors in advanced pancreatic cancer. *Br J Cancer* 2008;99:883-93.
 37. Kapoor D, Perwaiz A, Singh A, et al. Factors predicting 30-day mortality after pancreaticoduodenectomy—the impact of elevated aspartate aminotransferase. *Langenbecks Arch Surg* 2023;408:130.
 38. Li S, Tan HY, Wang N, et al. The Role of Oxidative Stress and Antioxidants in Liver Diseases. *Int J Mol Sci* 2015;16:26087-124.

Cite this article as: Dai Y, Huang W, Xu L, Zhang Q, Huang X. Machine learning models and nomogram based on clinical, laboratory profiles and skeletal muscle index to predict pancreatic fistula after pancreaticoduodenectomy. *Gland Surg* 2024;13(2):164-177. doi: 10.21037/gs-23-410