Experimental Research

INTERNATIONAL JOURNAL OF SURGERY

OPEN

AutoFRS: an externally validated, annotation-free approach to computational preoperative complication risk stratification in pancreatic surgery – an experimental study

Fiona R. Kolbinger, MD[a,b,c,d,*], Nithya Bhasker[e], Felix Schön, MD[f], Daniel Cser, BSc[e], Alex Zwanenburg, PhD[b,g], Steffen Löck, PhD[b,g], Sebastian Hempel, MD[a,b], André Schulze, MD[a,b,h], Nadiia Skorobohach, MD[f], Hanna M. Schmeiser[a], Rosa Klotz, MD[h], Ralf-Thorsten Hoffmann, MD[f], Pascal Probst, MD[i], Beat Müller, MD[h,j], Sebastian Bodenstedt, PhD[e], Martin Wagner, MD[a,b,h,k], Jürgen Weitz, MD[a,b,k], Jens-Peter Kühn, MD[f], Marius Distler, MD[a,b], Stefanie Speidel, PhD[e,k,*]

Background: The risk of postoperative pancreatic fistula (POPF), one of the most dreaded complications after pancreatic surgery, can be predicted from preoperative imaging and tabular clinical routine data. However, existing studies suffer from limited clinical applicability due to a need for manual data annotation and a lack of external validation. We propose AutoFRS (automated fistula risk score software), an externally validated end-to-end prediction tool for POPF risk stratification based on multimodal pre-operative data.

Materials and methods: We trained AutoFRS on preoperative contrast-enhanced computed tomography imaging and clinical data from 108 patients undergoing pancreatic head resection and validated it on an external cohort of 61 patients. Prediction performance was assessed using the area under the receiver operating characteristic curve (AUC) and balanced accuracy. In addition, model performance was compared to the updated alternative fistula risk score (ua-FRS), the current clinical gold standard method for intraoperative POPF risk stratification.

Results: AutoFRS achieved an AUC of 0.81 and a balanced accuracy of 0.72 in internal validation and an AUC of 0.79 and a balanced accuracy of 0.70 in external validation. In a patient subset with documented intraoperative POPF risk factors, AutoFRS (AUC: 0.84 ± 0.05) performed on par with the uaFRS (AUC: 0.85 ± 0.06). The AutoFRS web application facilitates annotation-free prediction of POPF from preoperative imaging and clinical data based on the AutoFRS prediction model.

Conclusion: POPF can be predicted from multimodal clinical routine data without human data annotation, automating the risk prediction process. We provide additional evidence of the clinical feasibility of preoperative POPF risk stratification and introduce a software pipeline for future prospective evaluation.

Keywords: artificial intelligence, pancreaticoduodenectomy, pancreatic surgery, radiomics, surgical complications

Introduction

Pancreatic surgery is associated with severe postoperative complications, causing considerable morbidity and mortality. Postoperative pancreatic fistula (POPF) affects about 10-30% of patients undergoing pancreatic head resection and is the

[a]Department of Visceral, Thoracic and Vascular Surgery, University Hospital and Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany, [b]National Center for Tumor Diseases (NCT/UCC), Dresden, Germany; German Cancer Research Center (DKFZ), Heidelberg, Germany; Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany; Helmholtz-Zentrum Dresden-Rossendorf (HZDR), Dresden, Germany, [c]Weldon School of Biomedical Engineering, Purdue University, West Lafayette, Indiana, USA, [d]Regenstrief Center for Healthcare Engineering, Purdue University, West Lafayette, Indiana, USA, [e]Department of Translational Surgical Oncology, National Center for Tumor Diseases (NCT), NCT/UCC Dresden, a partnership between DKFZ, Faculty of Medicine and University Hospital Carl Gustav Carus, TUD Dresden University of Technology, and Helmholtz-Zentrum Dresden-Rossendorf (HZDR), Germany, [f]Faculty of Medicine and University Hospital Carl Gustav Carus Dresden, Technische Universität Dresden, Institute and Polyclinic for Diagnostic and Interventional Radiology, Dresden, Germany, [g]Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, OncoRay – National Center for Radiation Research in Oncology, Helmholtz-Zentrum Dresden-Rossendorf, Dresden, Germany, [h]Department of General, Visceral and Transplantation Surgery, Heidelberg University Hospital, Heidelberg, Germany, [i]Cantonal Hospital Thurgau, Frauenfeld, Switzerland, [j]Clarunis, University Digestive Health Care Center Basel, Switzerland and [k]Technische Universität Dresden, Centre for Tactile Internet with Human-in-the-Loop (CeTI), Dresden, Germany

F.R.K. and N.B. contributed equally to this workundefined

*Corresponding Author. Address: Department of Visceral, Thoracic and Vascular Surgery, University Hospital and Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Fetscherstraße 74, 01307 Dresden, Germany.

E-mail: fiona.kolbinger@uniklinikum-dresden.de (F. Kolbinger);
Department of Translational Surgical Oncology, National Center for Tumor Diseases (NCT/UCC), Partner Site Dresden, Fetscherstrasse 74, 01307 Dresden, Germany.
E-mail: stefanie.speidel@nct-dresden.de (S. Speidel).

International Journal of Surgery (2025) 111:3212–3223

Received 6 November 2024; Accepted 3 March 2025

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.lww.com/international-journal-of-surgery.

Published online 28 March 2025

http://dx.doi.org/10.1097/JS9.0000000000002327

leading root cause of in-hospital death after pancreatic resection[1-4]. Affected patients often experience a significant prolongation of their postoperative course and a decline in performance status[5,6]. In consequence, patients with pancreatic cancer who suffer from POPF receive adjuvant chemotherapy less often and have a shorter overall survival than patients with an uncomplicated postoperative course, especially after neoadjuvant chemotherapy[7-9].

Accurate preoperative complication risk stratification, therefore, has the potential to augment both patient informed consent and surgical decision-making. In clinical routine, a preoperative estimation of POPF risk could facilitate a more informed discussion with patients and their relatives as well as optimizations of the surgical process and postoperative follow-up. For example, depending on the surgery indication, it would allow for consideration of alternative conservative or surgical treatment options or personalized adjustment of intra- and postoperative care decisions adapted to complication risk[10-12]. The current clinical gold standard method for POPF risk stratification in patients undergoing both minimally invasive and open pancreatic head resection is the updated alternative fistula risk score (ua-FRS)[13,14], which includes small pancreatic duct diameter, soft texture of the pancreatic remnant, high body mass index (BMI), and male sex as risk factors for POPF. An inherent limitation of this risk stratification method is that it is based on intraoperative risk factors: surgeon-determined remnant texture and intraoperatively measured pancreatic duct width.

Manual measurements of certain regions on preoperative computed tomography (CT) images have been demonstrated to be prognostic of POPF risk, for example, the pancreatic duct diameter at the location of anticipated dissection[15] and pancreatic neck thickness[16]. Advancements in computer vision have facilitated the automatic analysis of medical image features, capturing characteristics that are difficult to assess visually. Computational analysis of routine CT imaging can reliably predict POPF before surgery in patients undergoing pancreatic head resection[17-19]. However, of the published models, only one has been externally validated[18]. All approaches published to date rely on manual annotation by radiologists, which limits the potential for clinical translation. In addition, few published works combine different data types: Despite BMI and sex being known risk factors for POPF[13], most computational approaches for preoperative POPF risk prediction exclusively consider preoperative imaging and do not incorporate other abundant clinical routine data.

To overcome these limitations, we propose AutoFRS, a radiomics-based multimodal model predicting POPF from preoperative routine data (Fig. 1). Our approach requires no manual segmentation and has been validated on an external patient dataset, showcasing its generalizability. An end-to-end data processing pipeline qualifies our approach for prospective clinical evaluation.

HIGHLIGHTS

- AutoFRS is the first segmentation-free, externally validated computational model that predicts the risk of postoperative pancreatic fistula (POPF) in patients undergoing pancreatic head resection based on routinely available preoperative imaging and clinical data.
- Model risk stratification performance based on preoperative data is on par with the updated alternative fistula risk score (ua-FRS), the clinical gold standard for intraoperative complication risk stratification.
- The web-based AutoFRS tool facilitates POPF risk stratification from patient data within seconds without any manual annotation and could improve preoperative patient information for shared decision-making.

Materials and methods

Patient cohort

This study investigates a retrospective patient cohort treated at the University Hospital Carl Gustav Carus Dresden between 2011 and 2019 (Dresden cohort) and the prospective patient cohort of the NURIMAS trial[20] at the University Hospital Heidelberg between 2014 and 2015 (Heidelberg cohort). All patients had a clinical indication for pancreatic head resection and all patients underwent open surgery. The inclusion criteria were (i) history of pancreatic head resection for a malignant tumor or cystic neoplasia, (ii) availability of a preoperative contrast-enhanced venous-phase CT, and (iii) documentation of basic patient characteristics and POPF-related outcomes. The main exclusion criteria were calcifying chronic or acute pancreatitis, peripancreatic fluid collections, pancreatic pseudocysts, or a stent in the common bile duct. We provide an extended description of the selection criteria in Supplementary Material S1 (available at: http://links.lww.com/JS9/E69). Patient characteristics are summarized in Table 1.

Frequency and severity of POPF were assessed for at least 30 days after the pancreatic head resection or until discharge from the hospital, whichever occurred last. Following ISGPS standards[21], the POPF endpoint included Grade B POPF (defined as a clinically evident fistula with mild restriction of patient condition, typically managed using antibiotic medication, transfusion, parenteral nutrition, prolonged drain maintenance, or additional interventional drain placement) and Grade C POPF (a severe complication typically necessitating intensive care or an operative intervention). All included patients were treated based on institutional standard procedures. In Dresden, the POPF risk was assessed intraoperatively using the FRS[22]. In patients with an FRS of 3 or above, two drains were routinely inserted at the biliodigestive anastomosis and the pancreatic duct anastomosis. Somatostatin analogs were not routinely administered. In patients with an abdominal drain, drain amylase was assessed on postoperative days (POD) 1 and 3. In Heidelberg, drains were routinely inserted. Drain amylase was assessed on POD 1 and 3, and drains were removed on POD 3 when possible[20]. Somatostatin analogs were not routinely administered.

Age, BMI, sex, smoking history, and preoperative diabetes status were recorded for each patient. Patients were categorized as smokers (active tobacco use at the time of surgery or smoking cessation within six weeks before surgery) or non-smokers (smoking cessation over six weeks before surgery or never smokers), as described previously[23]. Six weeks was selected to separate smokers from non-smokers on the basis of studies

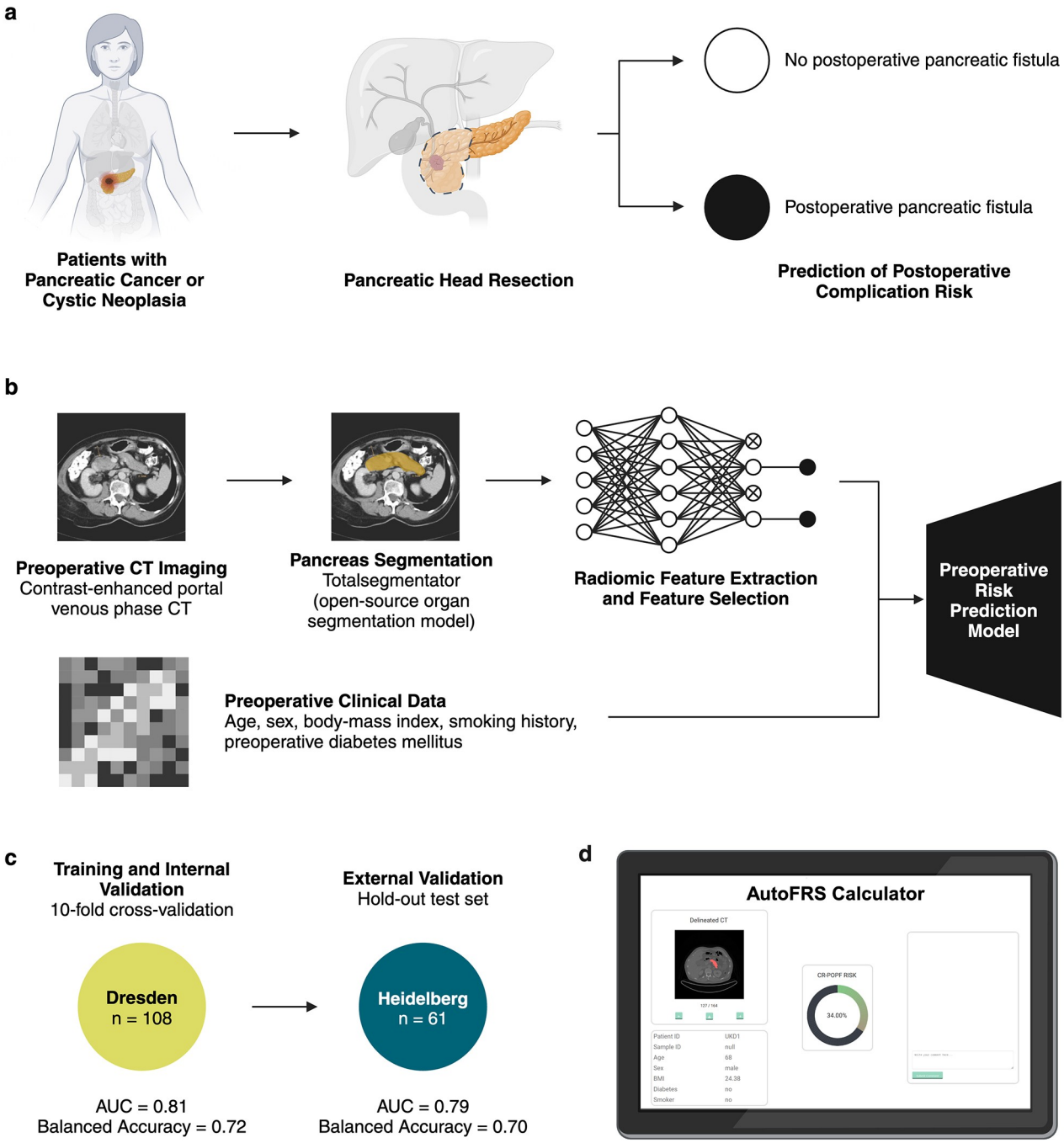

Figure 1. Annotation-free preoperative prediction of POPF in patients undergoing pancreatic head resection. (a) Prediction models were implemented on data from a total of 169 patients with pancreatic cancer or cystic neoplasia undergoing pancreatic head resection at two German centers. (b) Radiomic features were extracted from preoperative CT imaging after automated segmentation of the pancreas. Selected radiomic features were combined with clinical data points into a preoperative risk prediction model. (c) Following training on retrospective data from Dresden, model performance was evaluated in an independent prospectively collected patient cohort from Heidelberg. (d) User interface of the AutoFRS Calculator prototype displaying the results of automated POPF risk prediction for future integration into clinical workflows. Abbreviations: area under the receiver operating characteristic curve (AUC). Figure created with BioRender.com.

suggesting improved postoperative outcomes with smoking cessation at least 6 weeks prior to elective surgical procedures[24,25].

Preoperative diabetes mellitus was defined according to American Diabetes Association criteria for the diagnosis of diabetes[26], categorizing patients with a documented diagnosis of diabetes mellitus (type 1 or type 2) before surgery, either based on fasting plasma glucose ≥126 mg/dL (7.0 mmol/L), HbA1c ≥6.5%, or 2 h plasma glucose ≥200 mg/dl (11.1 mmol/l) during an oral glucose tolerance test, as diabetic.

Table 1

Patient characteristics

	Statistic	Dresden (development cohort)	Heidelberg (external test cohort)	*p*-value
Patients	n	108	61	
Sex:	n (%)			1.00
Female		48 (44.4)	27 (44.3)	
Male		60 (55.6)	34 (55.7)	
Age (years)	Mean ± standard deviation	68.5 ± 9.4	62.4 ± 12.8	0.002
BMI (kg/m2)	Mean ± standard deviation	25.2 ± 4.3	24.8 ± 4.4	0.60
Time between CT and surgery (days)	Mean ± standard deviation	62.6 ± 90.4	43.8 ± 100.8	0.002
Tumor classification:	n (%)			<0.001
Benign		3 (2.8)	21 (34.4)	
Malignant		105 (97.2)	40 (65.6)	
POPF:	n (%)			0.002
No		75 (69.4)	55 (90.2)	
Yes		33 (30.6)	6 (9.8)	

Abbreviations: body mass index (BMI), postoperative pancreatic fistula (POPF), computed tomography (CT), pylorus-preserving pancreaticoduodenectomy (PPPD), pancreatic ductal adenocarcinoma (PDAC). The *p*-values were calculated using Fisher's exact test for categorical variables and Mann–Whitney U-test for continuous variables.

Extraction of radiomic features from CT scans without manual annotation

After patient selection, the Dresden and Heidelberg cohorts comprised 108 and 61 patients, respectively. The characteristics of the corresponding CT scans are provided in Supplementary Material S2 (available at: http://links.lww.com/JS9/E69). To obtain annotations of the entire pancreas from CT scans, we used a pre-trained publicly available segmentation model, Totalsegmentator v.1.5.6[27], for all the patients (Supplementary Material S3, available at: http://links.lww.com/JS9/E69). Totalsegmentator segmentations of the pancreas were qualitatively and quantitatively compared with manual segmentations of the pancreas, which had been annotated by a radiologist with five years of experience in abdominal CT imaging and independently reviewed by a second radiologist with 20 years of experience in pancreatic imaging, as described previously[19]. We used the DICE score and Surface DICE score (https://github.com/google-deepmind/surface-distance) with a tolerance of 3 mm as the performance metrics for the quantitative evaluation. Subsequently, CT scans and pancreatic segmentations (Totalsegmentator) were fed to the MIRP framework (v2.0.0, https://github.com/oncoray/mirp)[28,29] for the extraction of radiomic features. A total of n = 654 features were extracted. These included morphological (n = 25), intensity-based (n = 57), and texture features (n = 136) extracted from both base image and response maps of spatial filters (wavelet coiflet-1 and Laplacian of Gaussian). Detailed configurations for feature extraction are provided in Supplementary Material S4 (available at: http://links.lww.com/JS9/E69)[30,31].

Development of AutoFRS

The AutoFRS was developed on the Dresden cohort in a two-step process: First, two separate preliminary models were developed based on radiomic and clinical features, respectively. The most relevant features from both the radiomic and the clinical preliminary models were subsequently combined into the multimodal AutoFRS using simple logistic regression (Fig. 1).

The preliminary radiomic model was developed using the FAMILIAR framework (version 0.0.0.54)[32] as described previously[19,33]. Pre-processing steps included z-score standardization with Yeo-Johnson transformation, hierarchical agglomerative clustering, and feature selection. More details about feature clustering are provided in Supplementary Material S5 (available at: http://links.lww.com/JS9/E69). We used a 10-fold cross-validation scheme repeated three times to develop preliminary risk models. Feature selection and hyperparameter optimization were performed on 10× bootstrapped samples from training folds. To assess performance variability across different methods, we compared three machine learning algorithms, (i) logistic regression, logistic regression with (ii) lasso penalty, and (iii) elastic-net penalty, and four feature selection methods, (i) mutual information maximization (MIM), (ii) minimum redundancy maximum relevance (MRMR), and logistic regression models with (iii) lasso penalty and (iv) elastic-net penalty.

The clinical variables used to develop the preliminary clinical risk model were age, BMI, sex, smoking history, and preoperative diabetes status. The selection of a subset of the input features enhances the performance of traditional machine learning methods when the number of input features is greater than the number of samples in the training set. As this is the case with our study, we selected a subset of the most predictive radiomic features (radiomic signature) for training the machine learning models. Supplementary Material S6 (available at: http://links.lww.com/JS9/E69) provides details about the development of the radiomic signature. The clinical signature was developed in the same manner. Area under the receiver operating characteristic curve (AUC) and balanced accuracy were used to evaluate model performance. Additionally, we conducted a univariate analysis with Benjamini-Hochberg correction to establish the relationship of individual features with the outcome. Kendall's tau was used to compare relevant radiomic features with intraoperative texture.

External validation of AutoFRS

Data from the Heidelberg cohort was pre-processed using the same transformation steps applied to the Dresden cohort. Specifically, features were processed using the parameters obtained in the development cohort (shift and scale parameters for normalization and λ for Yeo-Johnson power transformation) and grouped into the same clusters used in the development cohort. Subsequently, the AutoFRS model was validated using the Heidelberg cohort. We report model performance using AUC and balanced accuracy.

Data availability

Due to the participating centers' data privacy regulations, the datasets generated and analyzed in this study cannot be made available. Extracted features are available at https://gitlab.com/nct_tso_public/autofrs.git. Data requestors must sign a data use agreement to gain access.

Code availability

Totalsegmentator is publicly available at https://github.com/wasserth/TotalSegmentator. The code and the processed features for the AutoFRS model are available at https://gitlab.com/nct_tso_public/autofrs.git.

Results

Quantitative and qualitative evaluation of automatic segmentations of the pancreas

As the first image processing step, we used Totalsegmentator, an organ segmentation model for CT imaging, to create segmentations of the pancreas. To evaluate the accuracy of automatic pancreas segmentation, we quantitatively and qualitatively compared automatic segmentations with annotations encompassing the healthy pancreas, the pancreatic duct, and the pancreatic pathology curated by radiologists[19]. There was substantial overlap between manual and automatic segmentations, with an average DICE score of 0.77 ± 0.09 and an average surface DICE score of 0.90 ± 0.07 over the Dresden cohort (Supplementary Material S7, available at: http://links.lww.com/JS9/E69). Figure 2 illustrates the qualitative evaluation of automatic segmentations.

Development and validation of AutoFRS for POPF risk prediction

The results of all the risk models developed on the Dresden cohort are provided in Supplementary Material S8 (available at: http://links.lww.com/JS9/E69). Based on the results of the preliminary radiomic and clinical risk models, two radiomic features and BMI had the highest predictive value for the POPF outcome. Among the two radiomic features, one belonged to the texture feature family ($F1_T$) and the other to the morphological feature family ($F2_M$). Individual regressors built using $F1_T$, $F2_M$, and BMI had a significant ($p < 0.05$) impact on the POPF outcome upon univariate analysis (Table 2). These features were further used to train the AutoFRS model on the Dresden cohort. The formula for the AutoFRS model is given below, with $F1_T$, $F2_M$, and BMI indicating values after feature pre-processing. Pre-processing parameters are provided in Supplementary Material S9 (available at: http://links.lww.com/JS9/E69).

$$Risk\ of\ CR-POPF\ (\%) = 100 \times \frac{1}{1+\exp\left(1.1087\ -\ 0.7916\,F1_T\ +\ 0.5847\,F2_M\ -\ 0.8027\,BMI\right)}$$

AutoFRS achieved an AUC of 0.81 and a balanced accuracy of 0.72 on the Dresden cohort. Figure 3a shows the expression of AutoFRS features with respect to the actual outcome for the Dresden cohort. We externally validated AutoFRS on unseen data from a different medical center (Heidelberg cohort). On external validation, AutoFRS achieved an AUC of 0.79 and a balanced accuracy of 0.70. Given the extracted features, the runtime of AutoFRS in inference mode was 2.80 ± 0.07 seconds for the Heidelberg cohort on a computer equipped with Intel i7-7700 K processor with 32GB RAM. For a use case where a CT scan was provided as input along with the clinical data, the AutoFRS model took 6.73 ± 0.05 minutes per patient in inference mode.

We observed significant associations between the intraoperative texture of the pancreatic remnant, a known POPF risk factor, with both the radiomic $F1_T$ feature ($p = 0.017$; Fig. 3b) and the radiomic $F2_M$ feature ($p < 0.001$; Fig. 3c).

Comparison with existing risk prediction models

We compared AutoFRS with the recently published RAD-FRS model for preoperative POPF prediction, which, in contrast to AutoFRS, requires manual annotations of the pancreas[18]. RAD-FRS outlines two morphological and one texture features as important for the prediction of POPF. We used the MIRP version of these features, z-score standardized and extracted from Totalsegmentator pancreatic segmentations, as input to the model. AutoFRS (AUC: 0.81 ± 0.05 [0.74, 0.88], balanced accuracy: 0.72 ± 0.05 [0.65, 0.80]) performed similar to the RAD-FRS (AUC: 0.71 ± 0.05 [0.62, 0.80], balanced accuracy: 0.71 ± 0.05 [0.63, 0.78]) on the Dresden cohort (Fig. 4a). Similarly, in the Heidelberg cohort, AutoFRS and RAD-FRS performed on par (AutoFRS: AUC: 0.79 ± 0.08 [0.65, 0.92], balanced accuracy: 0.70 ± 0.10 [0.52, 0.87], RAD-FRS: AUC: 0.71 ± 0.10 [0.54, 0.86], balanced accuracy: 0.62 ± 0.10 [0.46, 0.79]; Fig. 4b). When compared using the DeLong test[34], AutoFRS and RAD-FRS yielded AUCs with no significant differences for both the Dresden ($p = 0.10$) and Heidelberg cohorts ($p = 0.58$).

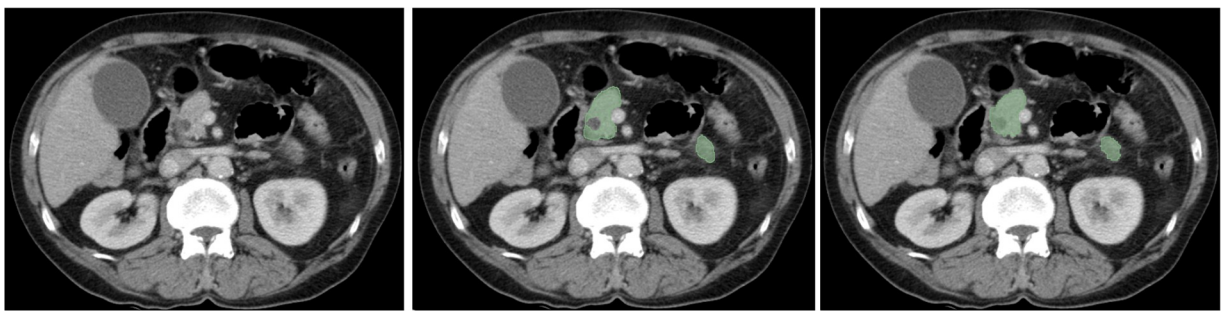

Figure 2. Qualitative evaluation of pancreatic segmentations. The figure visualizes an example from the Dresden cohort, comparing the expert radiologist segmentation, encompassing the healthy pancreas, the pancreatic duct, and the pancreatic pathology, with the pancreatic segmentation by Totalsegmentator (DICE score: 0.87).

Table 2

Radiomic and clinical signature

Feature	Corrected *P* value	AUC
BMI	0.00039	0.76
$F1_T$	0.0085	0.68
$F2_M$	0.00039	0.74

Results from univariate logistic regression of the radiomic and clinical features. The *P* value was corrected using the Benjamini-Hochberg method. The AUC indicates the performance of individual features for predicting POPF outcome. Abbreviations: area under the receiver operating characteristic curve (AUC), body mass index (BMI), postoperative pancreatic fistula (POPF).

Last, we compared AutoFRS and the RAD-FRS to the ua-FRS[13], the clinical gold standard for intraoperative POPF risk stratification, on a subset of the Dresden cohort with available intraoperative parameters. These intraoperative parameters were not recorded for the Heidelberg cohort and, therefore, were not included in the comparison. The annotation-free, preoperative AutoFRS (AUC: 0.84 ± 0.05, [0.75, 0.92], balanced accuracy: 0.73 ± 0.05, [0.64, 0.82]) performed on par with the preoperative RAD-FRS (AUC: 0.73 ± 0.07, [0.61, 0.83], balanced accuracy: 0.71 ± 0.06, [0.61, 0.80]) and the intraoperative ua-FRS (AUC: 0.85 ± 0.06, [0.75, 0.94], balanced accuracy: 0.79 ± 0.05, [0.71, 0.87]; Fig. 5). Predicted labels were calculated at a 20% threshold for balanced accuracy score calculation for ua-FRS[13]. When compared using the DeLong test[34], ua-FRS yielded AUC with no significant differences to AUCs of both AutoFRS (p = 0.82) and RAD-FRS (p = 0.18).

Similarly, the AutoFRS risk predictions aligned well with the risk predictions based on the original FRS, the first published intraoperative POPF risk score based on pancreatic texture, pancreatic duct width, blood loss, and tumor classification (malignant vs. benign)[22], which is still commonly used for POPF risk stratification at surgical centers worldwide (Supplementary Material S10, available at: http://links.lww.com/JS9/E69).

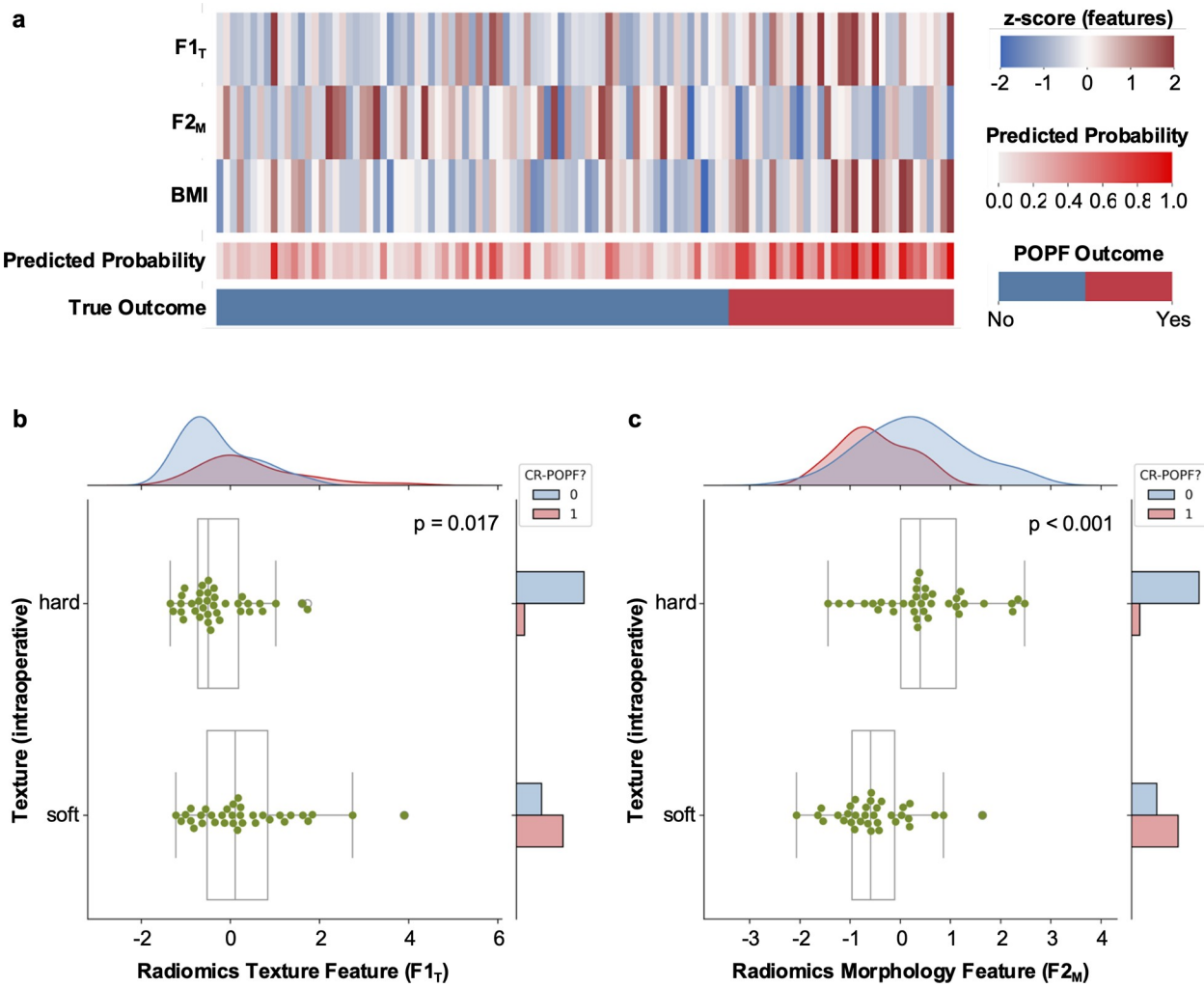

Figure 3. AutoFRS model development and prediction performance. (a) Expression of relevant radiomic and clinical features in relation to AutoFRS model predictions and true outcomes in the development cohort. (b) Correlation between radiomics texture feature ($F1_T$) and intraoperative pancreatic texture in relation to the POPF outcome. (c) Correlation between radiomics morphology feature ($F2_M$) and intraoperative pancreatic texture in relation to the POPF outcome. The relations between intraoperative and image-based radiomics features (b, c) are displayed for a subset of the Dresden cohort (n = 70) with available intraoperative features. Abbreviations: body mass index (BMI), postoperative pancreatic fistula (POPF).

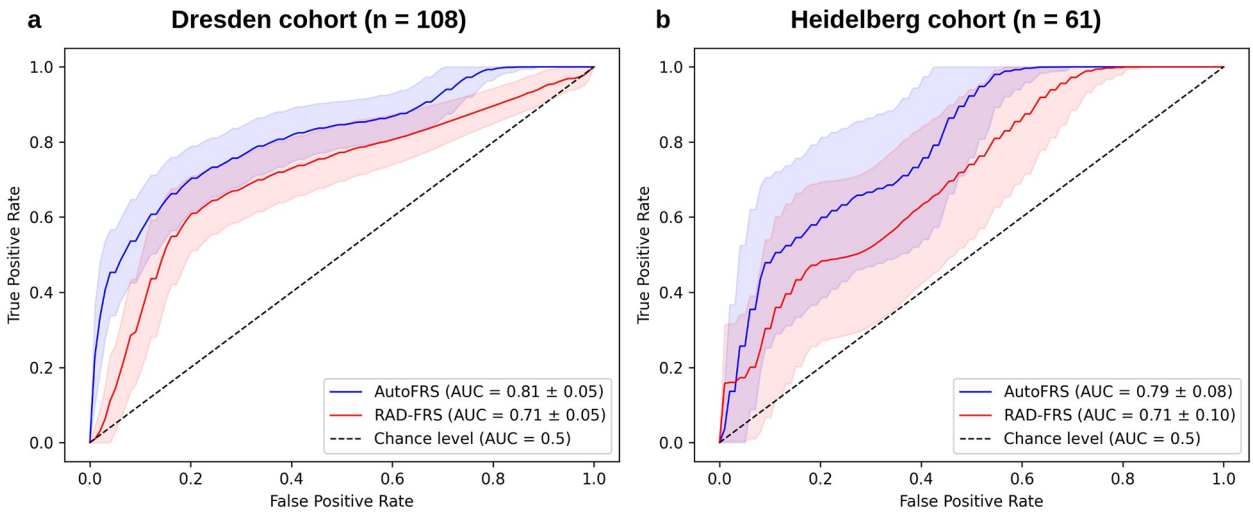

Figure 4. Comparison of the AutoFRS and the RAD-FRS. The figure displays receiver operating characteristic (ROC) curves and AUCs for AutoFRS and RAD-FRS models for the Dresden cohort (a) and Heidelberg cohort (b). The positive label is the occurrence of POPF. The shaded area represents ± one standard deviation. Abbreviations: area under the receiver operating characteristic curve (AUC), postoperative pancreatic fistula (POPF).

Web application for POPF prediction

We implemented the AutoFRS Calculator, a prototype web application for predicting POPF based on the AutoFRS model for future prospective evaluation (Fig. 6). After providing a contrast-enhanced CT image and the required clinical information, AutoFRS starts the evaluation pipeline, including segmentation of the CT image, feature extraction and preprocessing, and risk prediction. The users can also compare a given patient to another from a local database and leave comments. Further details regarding the application are provided in Supplementary Material S11 (available at: http://links.lww.com/JS9/E69).

Discussion

Postoperative deaths represent the third largest contributor to mortality worldwide and pose a considerable financial burden to

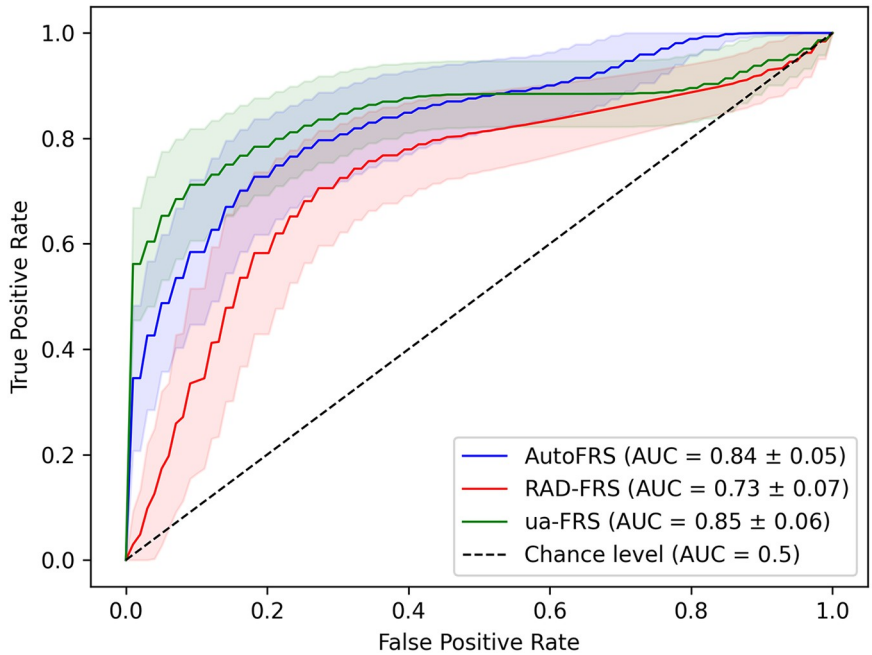

Figure 5. Performance comparison with the clinical gold standard for intraoperative POPF risk stratification. The figure displays receiver operating characteristic (ROC) curves for AutoFRS, RAD-FRS, and ua-FRS on a subset of the Dresden cohort (n = 70) with documented intraoperative risk factors. The shaded area represents ± one standard deviation. Abbreviations: area under the receiver operating characteristic curve (AUC), postoperative pancreatic fistula (POPF).

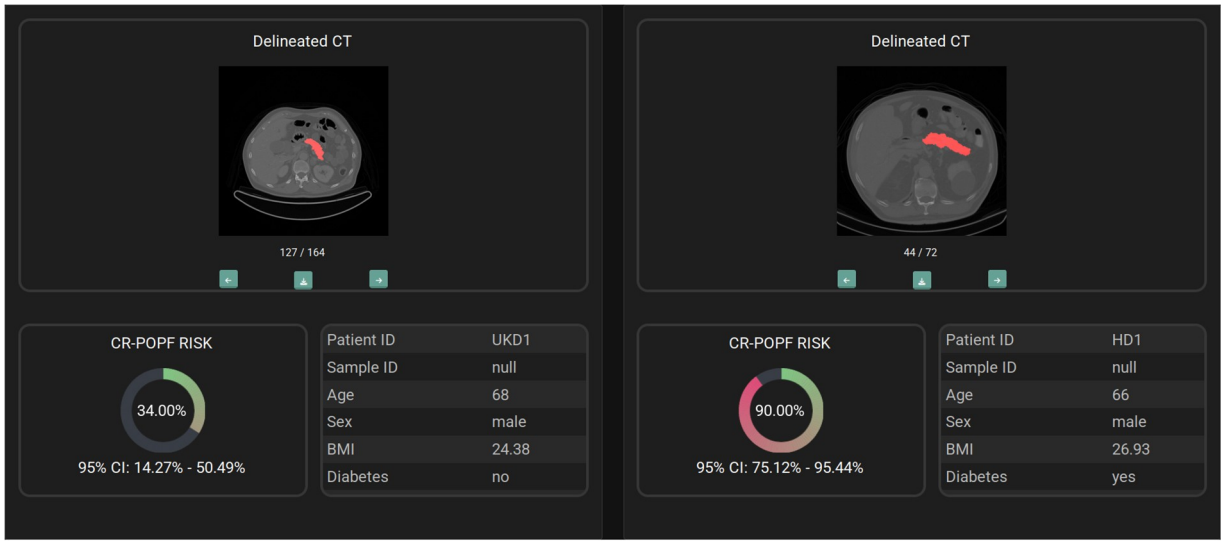

Figure 6. Prototype of the web application for POPF prediction. The figure illustrates the patient comparison feature of the application. The risk for POPF (including confidence intervals), delineated CT scans, and selected patient information are visualized to facilitate the comparison.

healthcare systems[35]. The incidence and severity of postoperative complications vary depending on the type of surgery[5,36,37]. While the relationship between preoperative variables, some of which can be derived from imaging, is well-established in pancreatic surgery[5,13], clinically available risk stratification systems do not utilize the breadth of available preoperative data to estimate complication risks preoperatively but rather build on intraoperative parameters that are, in part, subject to inter-observer variation. The objective of this work is to establish and validate a clinically applicable tool for computational preoperative risk stratification of POPF.

In summary, AutoFRS addresses three current limitations in POPF risk stratification: First, it builds on preoperative information, facilitating consideration of POPF risk when planning a pancreatic head resection. Second, it includes not only clinical data but also an automated computational and, thus, objective evaluation of standard CT imaging in POPF risk predictions. Third, unlike all existing approaches for image-based POPF risk estimation, it requires no manual measurements or segmentations, facilitating implementation in the clinical routine. Furthermore, AutoFRS is validated on an independent external cohort, and our findings from internal validation suggest that stratification performance is on par with the intraoperative ua-FRS.

Across surgical fields, general[38,39] or procedure-specific[40,41] complication risk stratification models have been developed to facilitate preoperative patient and caregiver information in clinical routine and adapt intra- and postoperative treatment to account for the likelihood of postoperative sequelae[42]. Most published risk stratification tools for POPF after pancreatic head resection rely on preoperatively available clinical parameters such as age or sex as well as intraoperatively determined factors like pancreatic duct diameter or pancreatic texture[13,22,43,44]. While some risk stratification tools include parameters derived from imaging (i.e., pancreatic duct or pancreatic neck diameter)[15,16], these measurements, like intraoperatively determined texture or pancreatic duct diameter, are subject to measurement errors and interrater variation. Additionally, in minimally invasive pancreatic surgery, intraoperative evaluation of pancreatic duct width and remnant texture, which are key components of traditional POPF risk stratification models[13,22], may not be feasible due to the surgical setup[45,46].

Existing studies for radiomics-based prediction of POPF typically use expert annotations to delineate pancreatic structures[18,19,47]. Creating these annotations is time-consuming, limiting their clinical applicability. In contrast, our results indicate that pre-trained segmentation models can automatically obtain annotations of the pancreas from CT scans. This approach considerably reduces the time and expertise required for risk assessment, making it more feasible for clinical implementation. Moreover, our model's runtime on affordable hardware is clinically feasible, enhancing its potential for integration into routine clinical workflows.

Many studies have shown the benefit of combining radiomic features with clinical variables derived from electronic health records in various clinical contexts[19,48,49]. The AutoFRS combines radiomic and clinical signatures. We found the two most relevant radiomic features, the $F1_T$ feature and the $F2_M$ feature, to modestly relate with the intraoperative texture of the remnant pancreas, which is a known indicator of POPF risk. This potentially provides a link between preoperative imaging and established clinical knowledge. In addition, the $F1_T$ feature, which belongs to the texture feature family, may be similar to the zone size variance feature that was found to be relevant in the RAD-FRS model by Ingwersen *et al.*[18]. Both features are related to gray-level size zone matrix calculations, suggesting a consistent pattern of texture heterogeneity that correlates with POPF risk.

While deep learning approaches have shown promise in medical image analysis[50,51], they typically require large datasets for training. In the context of POPF prediction, no public datasets are currently available. In this setting, feature-based approaches can effectively capture relevant image characteristics even with smaller datasets, facilitating computational complication prediction within the constraints of limited data availability. By making both our code and extracted features publicly available, we

facilitate reproducibility and encourage other researchers to build upon our work. This transparency also allows for model refinement and future integration of deep learning techniques as more data from different centers may become available in the future.

Our results suggest that AutoFRS can help identify patients at high risk for POPF. Compared to existing methods, which allow for estimation of the POPF risk at the end of the surgical procedure, the preoperative availability of a patient's individual estimated POPF risk as well as a measure of the reliability of this risk prediction (i.e., a confidence interval) could add a new dimension to shared preoperative decision-making. Specifically, AutoFRS could enable more comprehensive patient counseling and provide an avenue toward risk-adapted surgical approaches. Practically, such risk-adapted management options for patients with a known high risk of POPF could include alternative anastomotic techniques[52,53], prophylactic application of somatostatin analogs[54,55], individualized drain management[56,57], intensified postoperative care[58], or even consideration of total pancreatectomy with islet cell autotransplantation as an option[10,59,60], particularly in patients with preoperatively impaired glucose tolerance[61-63].

Future work could benefit from advancements in pancreatic imaging analysis, such as the PANORAMA study, which aims to improve pancreatic cancer detection and segmentation[64]. Integration of lesion-specific features into complication prediction approaches could further refine and enhance the predictive power of models like AutoFRS. Similarly, future refinements of automatic pancreas segmentation models like Totalsegmentator[27], for example, for segmentation of sub-structures like the pancreatic duct or adjacent blood vessels, could offer model improvements over current segmentation models that generate one segment for the entire pancreas.

The limitations of our work are mainly related to its retrospective nature, the small number of two centers as data sources for the development (one center) and validation (independent cohort from a second center) of the AutoFRS, and the use of automated segmentation models as part of the pipeline. First, as intraoperative data was only available for the Dresden cohort, we could only use these patients to compare the preoperative AutoFRS and the intraoperative gold standard for POPF risk stratification. Second, some clinical parameters differed significantly between the Dresden and the Heidelberg cohorts: POPF rates in the Dresden cohort and the Heidelberg cohort were significantly different at 30.6% (33 out of 108 patients) and 9.8% (6 out of 61 patients), respectively. These figures represent the upper and lower boundaries of POPF rates described in the literature[21,65,66]. The high POPF rate in the Dresden cohort may relate to the relatively early interventional treatment of patients with clinical symptoms based on institutional standards, which may inflate the number of patients with POPF. In addition, the two cohorts were significantly different in composition of malignant and benign tumors, with 97.2% of the Dresden cohort and 65.6% of the Heidelberg cohort presenting with malignant disease, respectively. While malignant and benign lesions may influence postoperative outcomes, recent studies suggest that the primary determinants of POPF risk are pancreatic duct size and the texture of the remnant pancreas rather than tumor characteristics themselves[13,44]. Both the Dresden and the Heidelberg cohorts had substantial variation in the time interval between preoperative CT imaging and pancreatic head resection, including nine patients with imaging obtained over 180 days before surgery. We observed negligible influence (Kendall's tau correlation measure, $\tau = -0.03$ for the Dresden cohort and $\tau = 0.13$ for the Heidelberg cohort) of this delay on the POPF risk predicted by the proposed preoperative signature. This implies that inherent parameters of the healthy pancreas that may be present before disease onset likely have a substantial impact on POPF risk. Third, some imaging metadata including contrast agent characteristics and dosage, were not available for the external validation cohort. This was due to protocol differences between the centers, which is a likely scenario in multicenter studies. The consequences of possible variations in contrast agent types are likely not decisive, as AutoFRS was trained and validated on contrast-enhanced venous-phase CT scans, ensuring consistency in image acquisition protocols. In addition, robust radiomic models can generalize across such variations in contrast agent protocols[67,68]. Fourth, we observed varying performance of the automated pancreas segmentation model when comparing automated organ segmentations to expert annotations. Poor segmentation performance of pancreatic parenchyma was often associated with variability in including or excluding the tumor in the generated pancreatic segment. Nevertheless, the automation of the prediction process offers immense improvements in clinical applicability. Our findings of comparable POPF prediction performance of the AutoFRS to the intraoperative ua-FRS indicate the high clinical potential of our proposed preoperative prediction model. An inherent limitation of retrospective artificial intelligence-based studies is that technical performance metrics, such as AUC and accuracy, do not necessarily correlate with clinical utility. This is particularly relevant when predictive models influence treatment decisions, as their impact on patient outcomes cannot be fully assessed using retrospective data alone. While our study demonstrates the feasibility of preoperative POPF risk stratification, its results should be considered preliminary due to the retrospective design and the limited sample size. Prospective interventional trials will be necessary to evaluate the effect of AutoFRS on surgical decision-making, perioperative management, and patient outcomes in real-world clinical practice.

Two aspects specifically qualify AutoFRS for clinical evaluation: First, no manual annotation of preoperative imaging is needed during data input. This reduces the expert effort needed to infer POPF risk predictions from routine data to a tolerable level in clinical routine. Second, the presented AutoFRS application couples model functionalities with a graphical user interface. Building on our results, we intend to prospectively validate AutoFRS in a multicenter clinical trial, analyzing both technical feasibility and the potential impact of preoperative POPF risk prediction on patient care.

Conclusion

AutoFRS is an externally validated, manual annotation-free approach to preoperative computational estimation of POPF risk in patients undergoing pancreatic head resection that integrates clinical parameters and imaging-derived features. Our findings suggest stratification performance on par with the ua-FRS, the clinical gold standard for intraoperative estimation of POPF risk. A clinically applicable software pipeline facilitates future prospective evaluation of AutoFRS.

Ethical approval

This study was performed following the ethical standards of the Declaration of Helsinki and its later amendments. The local Institutional Review Boards (ethics committees at the Technical University Dresden, approval number BO-EK-177032021, and University of Heidelberg, approval number S-248/2021) reviewed and approved this study.

Consent

Per the relevant local legislature, no individual patient consent was required for the analysis of previously collected pseudonymized data.

Sources of funding

This work was supported by the German Federal Ministry of Health (Surgomics, Grant BMG 2520DAT82) and by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) as part of Germany's Excellence Strategy (EXC 2050/1, Project ID 390696704) within the Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (CeTI) of the Dresden University of Technology. In addition, FRK receives support from the Joachim Herz Foundation (Add-On Fellowship for Interdisciplinary Life Science), the German Cancer Research Center (CoBot 2.0), and the Indiana Clinical and Translational Sciences Institute, funded, in part, by Grant Number UM1TR004402 from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

F.R.K., N.B., J.P.K., M.D., S.S.: conceptualization; F.R.K., N.B., F.S., A.Z., S.L.: methodology; N.B., D.C., A.Z., S.L.: software; F.R.K., N.B., F.S., A.Z., S.L., H.M.S., J.P.K., M.D.: validation; . F.R.K, N.B., F.S., A.Z., S.L.: formal analysis; F.R.K., N.B.: investigation; A.Z., S.L., S.H., R.T.H., B.M., M.W., J.W., J.P.K., M.D., S.S.: resources; F.R.K., N.B., F.S., S.H., A.S., N.S., H.M.S., R.K., R.T.H., P.P., B.M., M.W., J.W., J.P.K., M.D.: data curation; F.R.K., N.B.: writing – original draft; F.R.K., N.B., F.S., D.C., A.Z., S.L., S.H., A.S., N.S., H.M.S., R.K., R.T.H., P.P., B.M., S.B., M.W., J.W., J.P.K., M.D., S.S.: writing – review & editing; F.R.K., N.B., D.C.: visualization; A.Z., S.L., J.P.K., M.D., S.S.: supervision; R.T.H., B.M., S.B., M.W., J.W., J.P.K., M.D., S.S.: project administration; B.M., S.B., M.W., J.W., M.D., S.S.: funding acquisition.

Conflicts of interest disclosure

F.R.K. declares advisory roles for Radical Healthcare, Inc., San Francisco, CA; and the Surgical Data Science Collective (SDSC), Washington, DC. All other authors declare no conflicts of interest.

Guarantor

Fiona R. Kolbinger and Stefanie Speidel.

Research registration unique identifying number (UIN)

The study was prospectively registered at OSF on July 3rd, 2024 (https://osf.io/vf5ad) and at the German Registry of Clinical Trials (Deutsches Register Klinischer Studien, DRKS) on October 8th, 2024 (DRKS00035187). The protocol was not amended or changed.

Provenance and peer review

Not commissioned, externally peer reviewed.

Data availability statement

Data Availability: Due to the participating centers' data privacy regulations, the datasets generated and analyzed in this study cannot be made available. Extracted features are available at https://gitlab.com/nct_tso_public/autofrs.git. Data requestors must sign a data use agreement to gain access.

Code Availability: Totalsegmentator is publicly available at https://github.com/wasserth/TotalSegmentator. The code and the processed features for the AutoFRS model are available at https://gitlab.com/nct_tso_public/autofrs.git.

References

1. Hempel S, Wolk S, Kahlert C, *et al*. Outpatient drain management of patients with clinically relevant postoperative pancreatic fistula (POPF). Langenbecks Arch Surg 2017;402:821–29.
2. Nimptsch U, Krautz C, Weber GF, *et al*. Nationwide in-hospital mortality following pancreatic surgery in Germany is higher than anticipated. Ann Surg 2016;264:1082–90.
3. Pulvirenti A, Marchegiani G, Pea A, *et al*. Clinical implications of the 2016 international study group on pancreatic surgery definition and grading of postoperative pancreatic fistula on 775 consecutive pancreatic resections. Ann Surg 2018;268:1069.
4. Henry AC, Smits FJ, Daamen LA, *et al*. Root-cause analysis of mortality after pancreatic resection in a nationwide cohort. HPB (Oxford) 2025;27:461–469.
5. Davenport DL, Henderson WG, Khuri SF, *et al*. Preoperative risk factors and surgical complexity are more predictive of costs than postoperative complications: a case study using the National Surgical Quality Improvement Program (NSQIP) database. Ann Surg 2005;242:463–68.
6. Ho CK, Kleeff J, Friess H, *et al*. Complications of pancreatic surgery. HPB 2005;7:99–108.
7. Bonaroti JW, Zenati MS, Al-Abbas AI, *et al*. Impact of postoperative pancreatic fistula on long-term oncologic outcomes after pancreatic resection. HPB 2021;23:1269–76.
8. Uchida Y, Masui T, Nagai K, *et al*. Postoperative pancreatic fistulas decrease the survival of pancreatic cancer patients treated with surgery after neoadjuvant chemoradiotherapy: a retrospective analysis. Surg Oncol 2020;35:527–32.
9. Hank T, Sandini M, Ferrone CR, *et al*. Association between pancreatic fistula and long-term survival in the era of neoadjuvant chemotherapy. JAMA Surg 2019;154:943–51.
10. Hempel S, Oehme F, Tahirukaj E, *et al*. More is more? Total pancreatectomy for periampullary cancer as an alternative in patients with high-risk pancreatic anastomosis: a propensity score-matched analysis. Ann Surg Oncol 2021;28:8309–17.
11. Callery MP, Pratt WB, Vollmer CM. Prevention and management of pancreatic fistula. J Gastrointest Surg 2009;13:163–73.
12. Marchegiani G, Bassi C. Prevention, prediction, and mitigation of postoperative pancreatic fistula. Br J Surg 2021;108:602–04.
13. Mungroop TH, Klompmaker S, Wellner UF, *et al*. Updated alternative fistula risk score (ua-FRS) to include minimally invasive pancreatoduodenectomy. Ann Surg 2021;273:334–40.

14. 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.
15. Kolbinger FR, Lambrecht J, Leger S, *et al*. The image-based preoperative fistula risk score (preFRS) predicts postoperative pancreatic fistula in patients undergoing pancreatic head resection. Sci Rep 2022;12:1–11.
16. Sok C, Sandhu S, Shah H, *et al*. Simple preoperative imaging measurements predict postoperative pancreatic fistula after pancreatoduodenectomy. Ann Surg Oncol 2024;31:1898–905.
17. Ingwersen EW, Rijssenbeek PMW, Marquering HA, *et al*. Radiomics for the prediction of a postoperative pancreatic fistula following a pancreatoduodenectomy: a systematic review and radiomic score quality assessment. Pancreatology 2024;24:306–13.
18. Ingwersen EW, Bereska JI, Balduzzi A, *et al*. Pancreatobiliary and Hepatic Artificial Intelligence Research (PHAIR) consortium, Radiomics preoperative-fistula risk score (RAD-FRS) for pancreatoduodenectomy: development and external validation. BJS Open 2023;7:zrad100.
19. Bhasker N, Kolbinger FR, Skorobohach N, *et al*. Prediction of clinically relevant postoperative pancreatic fistula using radiomic features and preoperative data. Sci Rep 2023;13:7506.
20. Probst P, Haller S, Bruckner T, *et al*. Prospective trial to evaluate the prognostic value of different nutritional assessment scores in pancreatic surgery (NURIMAS pancreas). Br J Surg 2017;104:1053–62.
21. 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 (United States) 2017;161:584–91.
22. Callery MP, Pratt WB, Kent TS, *et al*. A prospectively validated clinical risk score accurately predicts pancreatic fistula after pancreatoduodenectomy. J Am Coll Surg 2013;216:1–14.
23. Graybill WS, Frumovitz M, Nick AM, *et al*. Impact of smoking on perioperative pulmonary and upper respiratory complications after laparoscopic gynecologic surgery. Gynecol Oncol 2012;125:556–60.
24. Møller AM, Villebro N, Pedersen T, *et al*. Effect of preoperative smoking intervention on postoperative complications: a randomised clinical trial. Lancet 2002;359:114–17.
25. Rodrigo C. The effects of cigarette smoking on anesthesia. Anesth Prog 2000;47:143–50.
26. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2011;34:S62–9.
27. Wasserthal J, Breit H-C, Meyer MT, *et al*. TotalSegmentator: robust segmentation of 104 anatomic structures in CT images. Radiol 2023;5:e230024.
28. Zwanenburg A, Leger S, Agolli L, *et al*. Assessing robustness of radiomic features by image perturbation. Sci Rep 2019;9:614.
29. Zwanenburg A, Löck S. MIRP: a Python package for standardised radiomics. J Open Source Softw 2024;9:6413.
30. Whybra P, Zwanenburg A, Andrearczyk V, *et al*. The image biomarker standardization initiative: standardized convolutional filters for reproducible radiomics and enhanced clinical insights. Radiology 2024;310:e231319.
31. Zwanenburg A, Vallières M, Abdalah MA, *et al*. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 2020;295:328–38.
32. Familiar: End-to-End Automated Machine Learning and Model Evaluation (2024). https://rdrr.io/cran/familiar/ (accessed July 1, 2024).
33. Leger S, Zwanenburg A, Leger K, *et al*. Comprehensive analysis of tumour sub-volumes for radiomic risk modelling in locally advanced HNSCC. Cancers 2020;12:1–27.
34. Sun X, Xu W. Fast implementation of DeLong's algorithm for comparing the areas under correlated receiver operating characteristic curves. IEEE Signal Process Lett 2014;21:1389–93.
35. Nepogodiev D, Martin J, Biccard B, *et al*. National institute for health research global health research unit on global surgery, global burden of postoperative death. Lancet 2019;393:401.
36. Dharap SB, Barbaniya P, Navgale S. Incidence and risk factors of postoperative complications in general surgery patients. Cureus 2022;14:e30975.
37. Eappen S, Lane BH, Rosenberg B, *et al*. Relationship between occurrence of surgical complications and hospital finances. JAMA 2013;309:1599–606.
38. Saklad M. Grading of patients for surgical procedures. Anesthesiology 1941;2:281–84.
39. Lee TH, Marcantonio ER, Mangione CM, *et al*. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. Circulation 1999;100:1043–49.
40. Gawande AA, Kwaan MR, Regenbogen SE, *et al*. An Apgar score for surgery. J Am Coll Surg 2007;204:201–08.
41. Regenbogen SE, Ehrenfeld JM, Lipsitz SR, *et al*. Utility of the surgical Apgar score: validation in 4119 patients. Arch Surg 2009;144:30–36.
42. Moonesinghe SR, Mythen MG, Das P, *et al*. Risk stratification tools for predicting morbidity and mortality in adult patients undergoing major surgery: qualitative systematic review. Anesthesiology 2013;119:959–81.
43. Verma A, Balian J, Hadaya J, *et al*. Machine learning-based prediction of postoperative pancreatic fistula following pancreaticoduodenectomy. Ann Surg 2024;280:325–31.
44. Schuh F, Mihaljevic AL, Probst P, *et al*. A simple classification of pancreatic duct size and texture predicts postoperative pancreatic fistula. Ann Surg 2023;277:e597–e608.
45. Klotz R, Mihaljevic AL, Kulu Y, *et al*. Robotic versus open partial pancreatoduodenectomy (EUROPA): a randomised controlled stage 2b trial. Lancet Reg Health Eur 2024;39:100864.
46. Da Dong X, Felsenreich DM, Gogna S, *et al*. Robotic pancreaticoduodenectomy provides better histopathological outcomes as compared to its open counterpart: a meta-analysis. Sci Rep 2021;11:3774.
47. Zhang W, Cai W, He B, *et al*. A radiomics-based formula for the preoperative prediction of postoperative pancreatic fistula in patients with pancreaticoduodenectomy. Cancer Manag Res 2018;10:6469–78.
48. Ligero M, Garcia-Ruiz A, Viaplana C, *et al*. Artificial intelligence combining radiomics and clinical data for predicting response to immunotherapy. Ann Oncol 2019;30:v476.
49. Dang LH, Hung SH, Le NTN, *et al*. Enhancing nasopharyngeal carcinoma survival prediction: integrating pre- and post-treatment MRI radiomics with clinical data. J Imaging Inform Med 2024;37:2474–89.
50. Litjens G, Kooi T, Bejnordi BE, *et al*. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88.
51. Liu X, Faes L, Kale AU, *et al*. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digital Health 2019;1:e271–e297.
52. Ratnayake CBB, Wells CI, Kamarajah SK, *et al*. Critical appraisal of the techniques of pancreatic anastomosis following pancreaticoduodenectomy: a network meta-analysis. Int J Surg 2020;73:72–77.
53. Kawaida H, Kono H, Amemiya H, *et al*. Anastomosis technique for pancreatojejunostomy and early removal of drainage tubes may reduce postoperative pancreatic fistula. World J Surg Oncol 2020;18:295.
54. Bootsma BT, Plat VD, van de Brug T, *et al*. Somatostatin analogues for the prevention of pancreatic fistula after open pancreatoduodenectomy: a nationwide analysis. Pancreatology 2022;22:421–26.
55. Li T, D'Cruz RT, Lim SY, *et al*. Somatostatin analogues and the risk of post-operative pancreatic fistulas after pancreatic resection – a systematic review & meta-analysis. Pancreatology 2020;20:158–68.
56. Takeda Y, Saiura A, Takahashi Y, *et al*. Conservative drain management increases the incidence of grade B postoperative pancreatic fistula without increasing serious complications: does persistent drainage reflect the quality of pancreatic surgery or institutional policy? J Hepatobiliary Pancreat Sci 2020;27:1011–18.
57. Seykora TF, Liu JB, Maggino L, *et al*. Drain management following distal pancreatectomy. Ann Surg 2020;272:1110–17.
58. Smits FJ, Henry AC, Besselink MG, *et al*. Algorithm-based care versus usual care for the early recognition and management of complications after pancreatic resection in the Netherlands: an open-label, nationwide, stepped-wedge cluster-randomised trial. Lancet 2022;399:1867–75.
59. Bellin MD, Eaton A, Ramanathan K, *et al*. Long-term islet graft functional assessments in more than 500 patients undergoing total pancreatectomy with intraportal islet autotransplantation. J Am Coll Surg 2025;240:405–413.
60. Barthold L, Smith KD, Chaidarun SS, *et al*. Quality of life following total pancreatectomy with islet autotransplantation: a patient experience survey. Pancreas 2024;53:e652–e656.
61. Balzano G, Zerbi A, Aleotti F, *et al*. Total pancreatectomy with islet autotransplantation as an alternative to high-risk pancreatojejunostomy after pancreaticoduodenectomy: a prospective randomized trial. Ann Surg 2023;277:894–903.

62. Ludwig S, Distler M, Schubert U, *et al*. Quality of life and metabolic outcomes after total pancreatectomy and simultaneous islet autotransplantation. Commun Med (Lond) 2022;2:24.
63. Hempel S, Oehme F, Ehehalt F, *et al*. The impact of pancreatic head resection on blood glucose homeostasis in patients with chronic pancreatitis. J Clin Med Res 2022;11:663–663.
64. Alves N, Schuurmans M, Rutkowski D, *et al*. The PANORAMA study protocol: pancreatic cancer diagnosis – radiologists meet AI. Zenodo 2024.
65. Zhang H, Zhu F, Shen M, *et al*. Systematic review and meta-analysis comparing three techniques for pancreatic remnant closure following distal pancreatectomy. Br J Surg 2015;102:4–15.
66. Xiong JJ, Tan CL, Szatmary P, *et al*. Meta-analysis of pancreaticogastrostomy versus pancreaticojejunostomy after pancreaticoduodenectomy. Br J Surg 2014;101:1196– 208.
67. Schön F, Kieslich A, Nebelung H, *et al*. Comparative analysis of radiomics and deep-learning algorithms for survival prediction in hepatocellular carcinoma. Sci Rep 2024;14:590.
68. Rinaldi L, De Angelis SP, Raimondi S, *et al*. Reproducibility of radiomic features in CT images of NSCLC patients: an integrative analysis on the impact of acquisition and reconstruction parameters. Eur Radiol Exp 2022;6:2.