fortune Journals

Research Article

JOURNAL OF RADIOLOGY AND CLINICAL IMAGING

ISSN: 2644-2809

# Preoperative Prediction of Postoperative Pancreatic Fistula After Pancreatic Head Resection Using Radiomics and Machine Learning Based on Computed Tomographic Diagnostics

**Johannes D Kaiser[1*], Matthias Benndorf[2], Esther A Biesel[1], Claudia Neubauer[2], Stefan Fichtner-Feigl[1], Fabian Bamberg[2], Uwe A Wittel[1], Jakob Neubauer[2]**

## Abstract

**Background:** Postoperative pancreatic fistula is one of the major complications after pancreatic head resection and can be life-threatening for patients. This study employed machine learning and radiomics to determine whether postoperative pancreatic fistulas (POPF) and perioperative drain amylase dynamics can be predicted prior to pancreaticoduodenectomy by evaluating the radiologic appearance of the pancreatic tissue.

**Methods:** In this retrospective trial 68 patients were included. For POPF prediction model (PPM) Radiomic features of the pancreas were extracted from the arterial phase of computed tomography (CT) at a 1 mm slice thickness for each patient. The radiomic features with highest correlation with POPF for our models, controlling for autocorrelation and applying Bonferroni correction for P-values were selected. For amylase prediction model (APM), radiomic features were correlated with postoperative maximum drain amylase levels at a cut-off of 1000U/l. ROC analysis was performed for evaluation of the resulting prediction models. The project was approved by the Ethics Committee of the University of Freiburg (246/20) in accordance with the Helsinki Declaration.

**Results:** POPF prediction model showed an area under the curve (AUC) of 0.897 (confidence interval (CI) =82.3-97.1%) in the cohort. The AUC of PPM was higher than that for the Roberts score, but the difference was not statistically significant. An attempt to predict postoperative amylase dynamics in the drainage fluid achieved an AUC of 0.936 (CI=88%-99.1%).

**Conclusions:** Preoperative prediction of POPF and drain amylase dynamics using radiomics and machine learning showed promising results. Both models offer new approaches to the clinical management of POPF.

**Keywords:** Radiomics; Machine Learning; Prediction Modeling; POPF; Pancreaticoduodenectomy.

## Background

The incidence and mortality rates of pancreatic cancer have increased over the past few decades [4, 6, 25, 31]. Not only the developments of the past, but also the prognoses of increasing case numbers for the future pose new challenges for pancreatic surgery.

**Affiliation:**

[1]Department of General and Visceral Surgery, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Germany

[2]Department of Radiology, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Germany

**Corresponding author:**
Johannes D Kaiser, Department of General and Visceral Surgery, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Germany Hugstetter Str. 55, 79106 Freiburg, Germany.

**Citation:** Johannes D Kaiser, Matthias Benndorf, Esther A Biesel, Claudia Neubauer, Stefan Fichtner-Feigl, Fabian Bamberg, Uwe A Wittel, Jakob Neubauer. Preoperative Prediction of Postoperative Pancreatic Fistula After Pancreatic Head Resection Using Radiomics and Machine Learning Based on Computed Tomographic Diagnostics. Journal of Radiology and Clinical Imaging. 7 (2024): 88-98.

**Received:** October 18, 2024
**Accepted:** October 28, 2024
**Published:** November 07, 2024

Kaiser JD, et al., J Radiol Clin Imaging 2024
DOI:10.26502/jrci.2809101

Significant improvements have been achieved by understanding premalignant lesions, preventing pancreatic cancer and adjuvant chemotherapy, which significantly prolongs patient survival after complete tumor resection [4, 6, 8, 21, 25, 31]. Perioperative complications after pancreatic surgery severely diminish both achievements [14, 28].

Perioperative morbidity is reported to be 30-50% and the driver of postoperative complications after pancreas resections are postoperative pancreatic fistulas (POPF) accounting for 10-36% of complications [9, 28]. POPF summarize the observation of digestive enzymes leaking from the pancreatic anastomosis and are classified according to the International Study Group of Pancreatic Fistula (ISGPF) into biochemical leak (BL) and clinically relevant grade B and C fistulas [6, 7]. Based on these definitions, the classification can only be performed retrospectively; thus, in clinical practice, less-standardized observations such as cut of drain amylase values are regularly used for clinical decision-making [11, 17, 30].

Despite detailed knowledge of the adverse effects of pancreatic fistulas, strategies for fistula prevention have not been found in routine clinical practice. In this context, preoperative prediction of POPF is essential for an individualized approach during the pre- and early perioperative course of pancreatic resections [4]. Although the scores differ in their architecture, they reflect the scientific consensus regarding the risk factors for the development of POPF [2, 24, 29] such as pancreatic duct width, texture of the pancreatic parenchyma and body mass index (BMI) [29, 31].

Purely preoperative scores have been developed by Yamamoto, Wellner and Roberts. The Roberts score is the only preoperative score that has been validated to date and serves as the reference in this study [2, 24, 29, 32]. This simple score includes preoperative pancreatic duct width and BMI, which are always available without additional costs [29].

In addition to the clinically assessable body mass index, the pancreatic duct width is always available because CT has become the reference standard for preoperative staging [23]. CT is required to assess resectability, vascular infiltration, lymph node involvement and potential metastases [23]. In terms of qualitative characteristics, CT provides information that is not directly visible to humans. Technological advances have made it possible to quantify and harness these imaging features. The respective method, called radiomics, is based on being able to predict a defined endpoint reliant on quantitative imaging features. Furthermore, these features can be correlated with clinical variables according to machine learning principles and statistically analyzed [12, 15, 19]. Several studies have shown that radiomics has enormous potential for the prediction, diagnosis, classification, progression, planning, treatment and treatment response in oncology [19].

## Methods

### Study Design

This study employed a retrospective, monocentric clinical design. The data pertaining to patients undergoing pancreatic surgery were extracted from a prospectively maintained database. The CT scans were obtained from the hospital radiology database. The project was approved by the Ethics Committee of the University of Freiburg (246/20) and was conducted in accordance with the Helsinki Declaration.

### Study population

The cohort consisted of 977 patients diagnosed with pancreatic head carcinoma who underwent partial pancreaticoduodenectomy at the Department of General and Visceral Surgery at the University Hospital Center of the University of Freiburg between 1998 and 2019. Only participants who underwent contrast-enhanced abdominal CT in the late arterial phase were included in the trial. Following the selection process, 68 patients were included in the study. All imaging was obtained at our radiology center to ensure the generation of homogeneous data collection in accordance with the principles of basic research. Patients who did not meet the inclusion criteria or whose examinations exhibited qualitative deficiencies, such as artifacts and incomplete visualization of the pancreas, were excluded (Figure 1). The Ethics Committee of the Albert-Ludwigs-University Freiburg waived the requirement for a written declaration of consent.

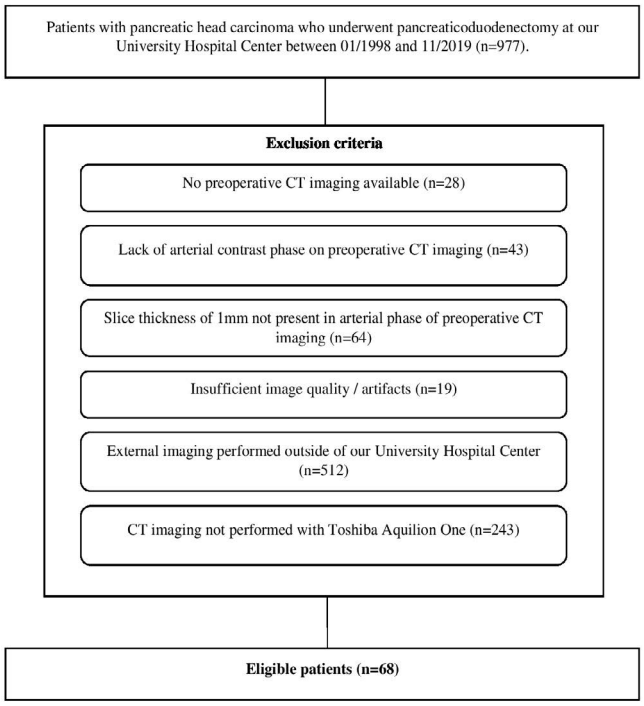


**Figure 1:** Flowchart of the recruitment of the total collective.

**Abbreviation:** CT, computed tomography.

**Citation:** Johannes D Kaiser, Matthias Benndorf, Esther A Biesel, Claudia Neubauer, Stefan Fichtner-Feigl, Fabian Bamberg, Uwe A Wittel, Jakob Neubauer. Preoperative Prediction of Postoperative Pancreatic Fistula After Pancreatic Head Resection Using Radiomics and Machine Learning Based on Computed Tomographic Diagnostics. Journal of Radiology and Clinical Imaging. 7 (2024): 88-98.

Kaiser JD, et al., J Radiol Clin Imaging 2024
DOI:10.26502/jrci.2809101

## CT examination

All patients underwent multiphase contrast-enhanced CT, performed using a Toshiba Aquilion ONE (320-line detector). The pitch factor was 0.813, the current voltage was 120 kV, and the current intensity was 55 mAs. The reconstructions were performed using a matrix of 512 × 512 and an axial reconstruction window of 440 × 440 mm. The layer thickness and spacing were both 1 mm. A total of 110 milliliters of Imeron 400 was administered to the patients at a 90% dilution, with isotonic saline serving as the contrast agent. The injection rate was 4 ml/s. The region of interest was placed in the aorta of the upper abdomen for bolus tracking. The mean delay was 20 seconds, which corresponds to the late arterial phase. This approach is recommended by the American College of Radiology for the diagnosis of pancreatic cancer, as it allows for the optimal contrast enhancement of the pancreatic parenchyma, resulting in the strongest image contrast between tumor tissue and healthy parenchyma [23]. This is also important for segmentation.

## Segmentation and Radiomics Analysis

3D Slicer version 4.10.2 (Figure 2) was used for Segmentation of the pancreas. 3D Slicer is an open source software developed by the Surgical Planning Laboratory at Brigham and Women's Hospital and the MIT Artificial Intelligence Laboratory [10]. The segmentation was performed manually, and the results were then confirmed by a Board-approved consulting radiologist. The segmentation process involved the delineation of the pancreas on each slice as a region of interest (ROI), with a cutline running from anterior to posterior at the level of the superior mesenteric vein at a 90° angle to the coronary slice plane (Figure 3). This imaginary line corresponds to the surgical incision pattern of the Whipple procedure typically performed in this area. Cysts and dilated pancreatic ducts were included in the region of interest (ROI) regardless of their extent. The pancreatic head, including the tumor tissue, feeding vessels, and calcifications, was excluded from the region of interest (ROI). Radiomics analysis was conducted using PyRadiomics in Python 2.4, an open-source code that facilitates the extraction of features from medical imaging [13]. No image normalization or preprocessing was employed.

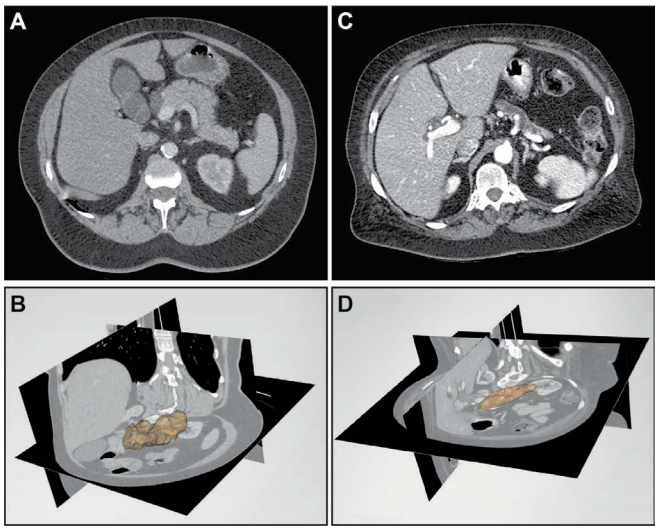


**Figure 2:** Segmentation of different p ancreas m orphologies with 3D-slicer.

**Notes:** Two-dimensional (A+C) and three-dimensional (B+D) representations of segmentations performed in this study from a patient whose pancreas is normal in appearance (A+B) and from a patient whose pancreas is atrophic and whose pancreatic duct is markedly dilated (C+D).

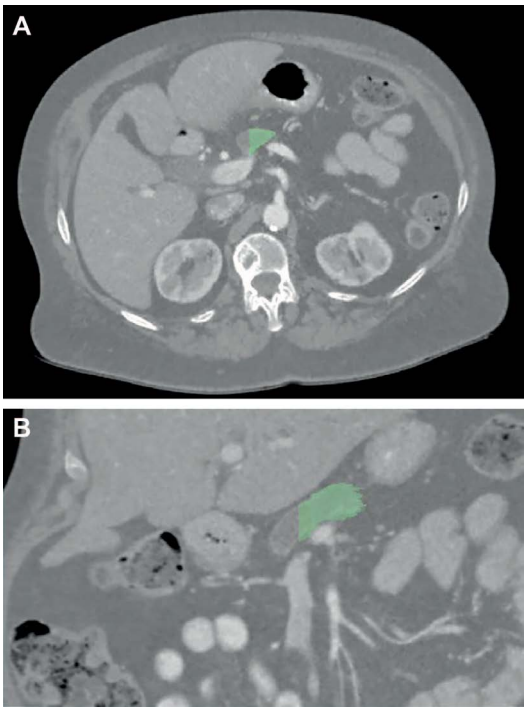


**Figure 3:** Segmentation of the ROI using 3D slicer.

**Notes:** Axial (A) and coronal (B) CT reconstruction with segmentation (highlighted in green) to demonstrate the segmentation margin at the border with the pancreatic head, equivalent to the site of surgical transection.

**Abbreviations:** ROI, region of interest.

For each segmentation, the following feature classes were extracted: Shape, First-order and Second-order consisting of Gray Level Co-Occurence (GLCM), Gray Level Run Length (GLRLM) and Gray Level Size Zone (GLSZM) matrices. The term "shape features" refers to characteristics that result from the boundary of the region of interest, or the shape of the segmentation and describes two- and three-dimensional shape characteristics. First-order features represent the distribution of gray values or their intensity. The spatial location is not considered. In contrast, the frequency of the gray values is depicted as a histogram. From this, a variety of information, including mean values, medians, intensities, entropy, uniformity, and others, can be calculated. Second-

order features represent the texture of the region of interest (ROI) by creating different matrices. GLCM describes the frequency with which a specific pixel or voxel of a given grey value is related to another specific grey value. Additionally, the distance between the grey values within the region of interest (ROI) is also fixed. Angular changes can be employed to capture multi-dimensional spaces and expressed mathematically. GLRLM describes the frequency with which a specific gray value recurs in a single direction until it is interrupted. This sequence is referred to as a "run." GLSZM indicates areas of neighboring pixels with identical gray values. It examines the relationships between gray value groups in all directions, as well as the group size. Given that the number of features to be extracted is theoretically unlimited, it is crucial to exercise caution when selecting the features to be included in the model. Including an excess of or an insufficient number of features may compromise the validity of the data obtained, potentially leading to over- or underfitting [19].

Ultimately, 16 shape features and 18 intensity features were extracted. In addition, 54 texture features were generated. The extracted features included 22 GLCM, 16 GLRLM, and 16 GLSZM features. A total of 86 features were extracted. Given that the collective size of the cohort was 68 patients, there was a risk of overfitting with the 86 extracted features. Therefore, it was necessary to restrict the analysis to the named classes in order to keep the dimensionality of the feature space as low as possible from the outset. For further information on pyradiomics and the radiomics workflow, we direct the reader to the work of van Griehuysen and Liu et al. [13, 19].

## Statistical analysis

All extracted features underwent an autocorrelation analysis, employing a cluster correlation matrix and visualized through a heat map. Features that exhibited a statistically significant correlation with the development of POPF or an amylase threshold of >1000 U/l in the drain were identified and quantified using Spearman's rank correlation coefficient. The significance level of the feature correlation following the Bonferroni correction was adjusted according to the number of features ($p^* < 0.05/86 = 0.0005814$). Regularization entails the selection of the most significantly correlated features. The selected features were subsequently subjected to autocorrelation testing via a cluster correlation matrix, with the results presented in the form of heat maps. The heat maps afforded the opportunity to graphically read significant correlations using color spectra (Figure 4). Subsequently, the identified features were employed in the generation of the model. Multivariate analysis also retained the radiomic features obtained. To assess the models for their predictive accuracy, we applied a 10-fold cross-validation with a 90:10 ratio. Subsequently, the models were extended to include clinical markers using the same system. The predictive ability of the constructed radiomics models was depicted graphically using receiver operating characteristic (ROC) curves. To enhance the accuracy of the results, we employed bootstrapping as an additional resampling method. The confidence intervals (CI) for the area under the curve (AUC) were calculated based on the analysis. All statistical analyses were conducted using R, version 4.3.2, with the graphical interface RStudio, version 2022.02.0, a free software environment for statistical computing and graphics developed by the R Foundation for Statistical Computing.

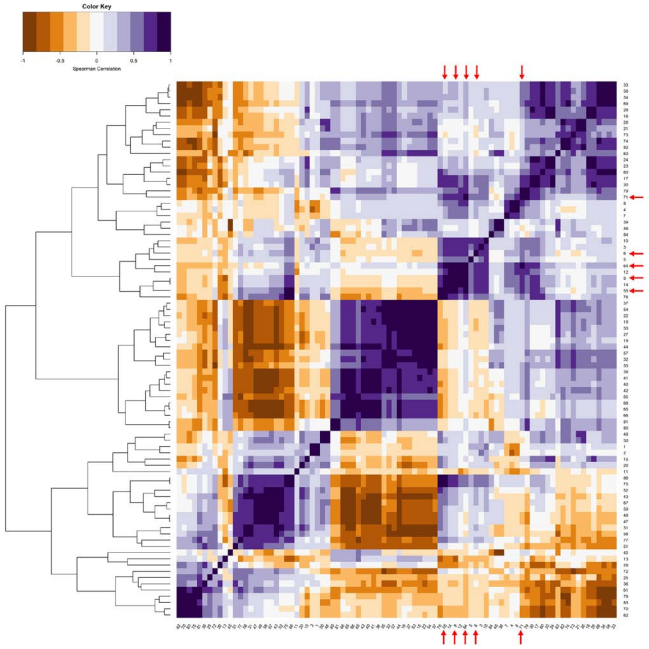


**Figure 4:** Heat map of the 86 Pyradiomics features.

**Notes:** Red arrows as markers of significantly correlated features for the development of POPF.

**Abbreviation:** POPF, postoperative pancreatic fistula.

## Surgery and Postoperative treatment

The surgical procedures were performed by experienced pancreatic surgeons. The pancreatic head was resected by dividing the left superior mesenteric vein. Subsequently, an end-to-side pancreaticojejunostomy (PJ) were created according to the description of Cattell and Warren. All patients receive a minimum of 2 drains which are placed dorsal and ventral of the pancreatoenteric anastomosis and passed through the abdominal wall in one insertion. Additional drains located distant to the panreatoenteric anastomosis (BDA or lower abdomen) were placed according to the surgeons preference. Following surgery, serum and peripancreatic drainage fluid amylase levels were measured in all patients on postoperative days 1 and 3. CT scans or further amylase measurements were not routinely performed, but were conducted depending on the clinical course. The peripancreatic drains were removed according to the following criteria: (1) the flow rate, (2) amylase levels in drain fluid

(removal if below 1000 U/l on the third postoperative day), (3) clinical findings, and (4) no evidence of fistula formation. In the event that infection was uncontrollable or instability persisted despite adequate clinical treatment, reoperation was performed. In the event of leakage or suspected infectious complications, the peripancreatic drains were retained and antibiotics were administered at the discretion of the treating physician. Percutaneous or endoscopic drainage was performed, depending on the location of the fluid collection. In cases of biochemical leakage, no additional treatment was administered.

## Results

The total cohort (n=68) consisted of 27 women and 41 men. The median age was 70.7 years (IQR=17.2). Detailed information on the clinical parameters of the cohort is shown in Table 1. The median body mass index (BMI) was 24.7 kg/m2 (IQR=4.4). The median width of the pancreatic duct measured on the preoperative computed tomography was 4 mm (IQR= 3.1). The median preoperative serum amylase level of the entire cohort was 21 U/l (IQR= 21). On the first postoperative day (POD1), the median serum amylase level was 43 U/l (IQR=118.5). The maximum postoperative drain amylase activity yielded a median of 374.5 U/l (IQR=3361). According to the ISGPF definition 39.7% (n = 27) of all patients in our cohort developed postoperative pancreatic fistula [3,11]. Of these, nine patients had a biochemical leak, 14 had grade B, and four had grade C biochemical leak. In this context, the architecture of our prediction models only allows a binary distinction between no fistula and fistula according to the ISGPF classification, which aligns our outcome stratification with the results of comparable studies.

**Table 1:** Characteristics of the total collective.

| Characteristics | Total collective (n=68) |
|---|---|
| Demographic characteristics | |
| Age, years | 70.7 (59.2-76.4) |
| Sex, male | 41 (60.3) |
| Histopathological characteristics | 51 (75) |
| Adenocarcinoma | |
| Acinar cell carcinoma | 1 (1.47) |
| Cystadenoma | 1 (1.47) |
| IPMN | 3 (4.41) |
| PanIN | 1 (1.47) |
| Chronic pancreatitis | 6 (8.82) |
| Inflammatory alteration | 1 (1.47) |
| Metastasis | 1 (1.47) |
| Neuroendocrine tumor | 1 (1.47) |
| Other | 2 (2.94) |
| Clinical characteristics | |
| BMI, kg/m$^2$ | 24.7 (23.1-27.4) |
| DM, yes | 12 (8.2) |
| Exocrine insufficiency, yes | 4 (2.7) |
| Preoperative creatinine, mg/dl | 0.8 (0.7-1) |
| Preoperative bilirubin, µmol/l | 0.8 (0.4-2.1) |
| CA 19-9, U/ml | 23.2 (0-136.5) |
| Preoperative serumamylase, U/l | 21 (13.5-34.5) |
| Serumamylase POD1, U/l | 43 (9-127.5) |
| Max. amylase in drainage, U/l | 374.5 (19.5-3380.5) |
| CT-predictors | |
| Duct width, mm | 4 (2.6-5.7) |
| POPF | 27 (39.7) |
| Total | |
| BL | 9 (33) |
| B | 14 (52) |
| C | 4 (15) |

**Notes:** unless otherwise indicated n(%) or (IQR).

**Abbreviations:** IQR, interquartile range; IPMN, intraductal papillary mucinous neoplasia; PanIN, intraepithelial neoplasia; BMI, body mass index; DW, duct width; DM, diabetes mellitus types 1 and 2; CA19-9, carbohydrate antigen 19-9; POD1, postoperative day 1.

**Citation:** Johannes D Kaiser, Matthias Benndorf, Esther A Biesel, Claudia Neubauer, Stefan Fichtner-Feigl, Fabian Bamberg, Uwe A Wittel, Jakob Neubauer. Preoperative Prediction of Postoperative Pancreatic Fistula After Pancreatic Head Resection Using Radiomics and Machine Learning Based on Computed Tomographic Diagnostics. Journal of Radiology and Clinical Imaging. 7 (2024): 88-98.

Kaiser JD, et al., J Radiol Clin Imaging 2024
DOI:10.26502/jrci.2809101

The heat map of all the extracted features revealed several autocorrelations. Regarding PPM, 6 of the 86 features from Pyradiomics, correlated significantly with a postoperative fistula. Of these, 2 were GLRLM features, 1 was GLSZM feature and 3 were shape features.

The ROC curve of Roberts' score for the total cohort showed an AUC of 0.813 (CI=70.6%–92%). In comparison, PPM without clinical markers was initially inferior to the Roberts score, with an AUC of 0.779 (CI=66.1- 89.6%). After the correlation of BMI, duct width and radiomics, the predictive power increased to an AUC of 0.897 (CI=82.3-97.1%), exceeding the level of the clinical Roberts' score, but without reaching statistical significance (p=0.144) (Figure 5).

Regarding APM, 11 radiomic features correlated significantly with the postoperative elevation of amylase enzyme activity in the drain fluid above a threshold of 1000 U/l on the third postoperative day, which is used as the cut-off level in clinical practice in our hospital. This threshold is used to determine whether a drain should be removed or left in place. After definitive selection, 5 of the 11 features corresponding to the most significant correlation were chosen. The final feature group is composed of one shape feature, one GLRLM feature and two first-order features. Evaluation of the ROC curve of APM resulted in a predictive strength with an AUC of 0.902 (CI=70.6-92%). BMI and duct width resulted in an increased AUC of 0.936 (CI=88-99.1%) (Figure 6).

## Discussion

In this study we successfully employed preoperative CT imaging, using radiomics, to predict postoperative pancreatic fistula (POPF), which was then compared with a preoperative multicenter validated scoring system. These results demonstrate the potential of radiomics for advancing the perioperative management of patients with pancreatic head carcinoma.

A purely preoperative gain of information about possible postoperative fistula formation offers the advantage of being able to reduce the risk of fistula formation preoperatively by means of medical therapy. For example, preoperative treatment with pasireotide was shown to significantly reduce the risk of fistula formation by Allen et al. [3].

PPM is useful for predicting the development of clinically relevant fistulas. Using APM, we were able to demonstrate that the amylase activity in the drainage on the third postoperative day is a suitable biomarker for predicting fistula development. This approach allowed us to provide a transfer possibility of radiological knowledge gain to clinical-surgical relevance.

Pancreatic fistulas, particularly postoperative pancreatic fistulas (POPF), are significant complications following pancreaticoduodenectomy. POPF result from the leakage of pancreatic fluid and can lead to severe morbidity and mortality. The incidence of POPF varies widely, influenced primarily by patient-specific factors such as gland texture, duct size, and underlying pathology. Current diagnostic and management strategies rely heavily on intraoperative findings

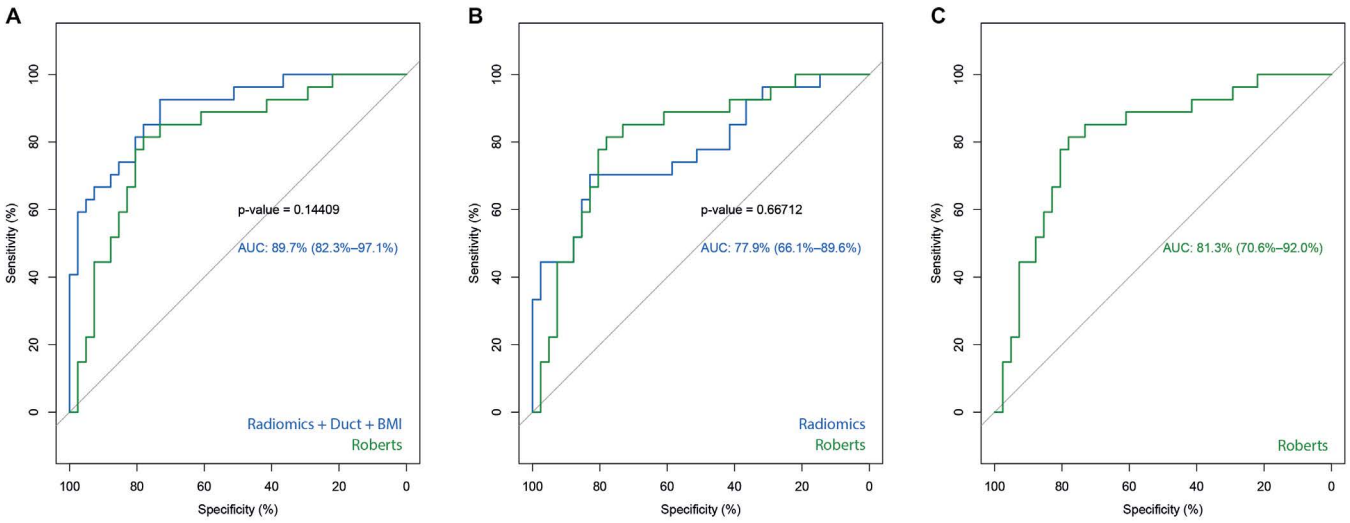


**Figure 5:** ROC curves of PPM and Roberts' score.

**Notes:** (A) Statistical comparison of ROC curves of Roberts' score (green) and Radiomics + Duct + BMI (blue) with an AUC of 0.897 (CI=82.3-97.1%) and p = 0.14409. (B) Statistical comparison of ROC curves of Roberts' score (green) and Radiomics (blue) with an AUC of 0.779 (CI=66.1-89.6%) and p= 0.66712. (C) ROC curve of Roberts' score for the total collective (green) with an AUC 0.813 (CI=70.6-92%).

**Abbreviations:** PPM, POPF prediction model; ROC, receiver operating characteristic; AUC(CI), area under the curve (confidence interval).

**Citation:** Johannes D Kaiser, Matthias Benndorf, Esther A Biesel, Claudia Neubauer, Stefan Fichtner-Feigl, Fabian Bamberg, Uwe A Wittel, Jakob Neubauer. Preoperative Prediction of Postoperative Pancreatic Fistula After Pancreatic Head Resection Using Radiomics and Machine Learning Based on Computed Tomographic Diagnostics. Journal of Radiology and Clinical Imaging. 7 (2024): 88-98.

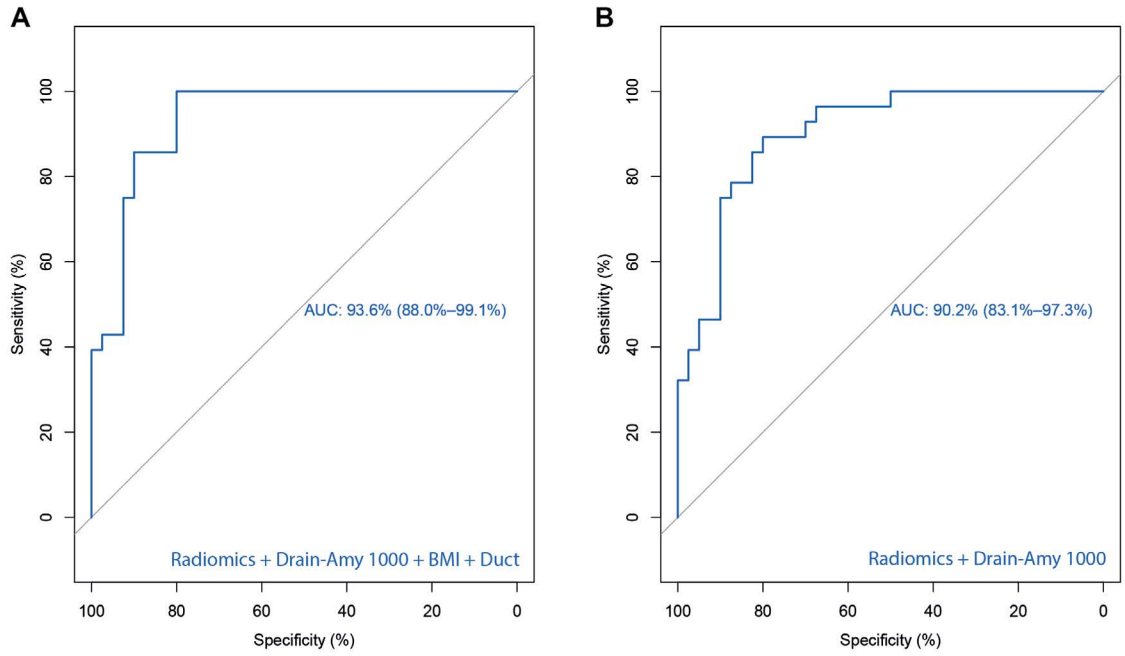


**Figure 6:** ROC curves of APM.

**Notes:** (A) ROC curve of results of Radiomics + Drain-Amy 1000 + BMI + Duct with an AUC of 0.936 (CI=88.0-99.1%). (B) ROC curve of the results of Radiomics + Drain-Amy 1000 with an AUC of 0.902 (CI=83.1-97.3%).

**Abbreviations:** APM, amylase prediction model; ROC, receiver operating characteristic; AUC (CI), area under the curve (confidence interval).

and postoperative biochemical analysis, underscoring the need for reliable preoperative predictive models. Following the current literature, the incidence ranges from 3% to 30% [6, 14, 28, 31]. The relatively elevated incidence observed within our cohort may potentially limit the generalizability of our findings to other patient groups. It could be attributed to selection bias, possibly stemming from patients with grade BL transitioning into B because they are ambulatory during clinical follow-up; therefore, the drain cannot be pulled until 21 days.

Most of the significantly correlated features in our models were Second-order features, which primarily reflected the relationships among voxels within the ROI using matrices. A possible explanation for this result is the complexity of the pancreatic structure and its large cell variance in a comparatively small organ volume [1]. Thus, the complexity of pancreatic parenchyma is reflected in the matrices of this class. This is a clue for radiomics as a relevant tool in pancreas-related research because the depth of Second-order matrices exceeds human image interpretation. This assumption is supported by the fact that in previous pancreas-related radiomics studies, the most prevalent significant radiomics features observed were Second-order as well [1].

Isolated radiological analysis of pancreatic parenchyma using radiomics (AUC=0.779) was inferior to prediction using the Roberts score (AUC=0.831), whose comparatively high value was consistent with the results of recent studies. Therefore, we investigated whether the combination of radiomics and clinical parameters used for the clinical score would alter the value of POPF prediction.

Although BMI did not have the expected effect on predictive strength, the implementation of duct width significantly improved the predictive performance of our models. We observed an increase in the AUC of PPM (+0,113) after combining it with the ductal size determined on preoperative CT scans. On the one hand, this could be explained by the influence of duct width on the possibility of surgical re-anastomosis. On the other hand, it could indicate a limitation of pure radiomics analysis to adequately distinguish the parenchyma and ductal system within the ROI. This finding supports our approach of combining radiomics with clinical parameters.

In contrast, the lack of a positive effect of BMI on the power of our prediction model may be explained by the distribution of BMI in our cohort, in which BMI was relatively low overall. The median BMI of 24.7 kg/m2 (IQR=4.4) was within the normal range. Nevertheless, we decided to include BMI to ensure the best possible comparability with the Roberts score [15, 24, 27].

Both models achieved a high level of predictive accuracy in the context of comparable literature [18, 19, 26, 34]. The potential of combining radiomics with clinical parameters is reflected in the increase in AUC due to the correlation with duct width and in the outcome of both models. The

predictive power of PPM compared to Roberts' score is critical for realistic consideration of the results. The lack of significance of our results can be explained by the strong clinical score and small cohort size [1, 24, 33]. In the future, our model should be applied to larger cohorts to further verify the significance level and transferability to other cohorts. Constructing the APM, we succeeded in detecting a total of 11 radiomics features that significantly correlated with an increase in enzyme activity in the drainage above 1000U/l. Our result, with an AUC of 0.936 (CI=88.0-99.1%) suggests a strong possibility to predict the healing process after PD perioperatively using radiomics [1, 11, 19].

Thus, to the best of our knowledge, this approach represents an innovation because there have been no comparable approaches in perioperative drainage management after PD. Postoperative management is controversial in the current literature regarding prophylactic drainage, timing of cutoff determination, cutoff level and timing of removal [11, 30].

This shows a gap between the theoretical retrospective classification according to ISGPF and its value for the clinical perioperative management of POPF. This fact is reflected in all prediction systems developed so far, which cannot distinguish between BL and clinically relevant grades B and C in the interpretation of results [19, 26, 29, 34]. This may explain the clinically practical importance of a cutoff level as a maxim for action.

The potential limitations of our study were its retrospective design and small patient population size. The results are based exclusively on data from a single center and require further external validation. Similarly, the small cohort size could equally influence the results. The large reduction in the number of patients could be explained by strict inclusion criteria. Furthermore, we decided not to use so-called "preprocessing". This was done to obtain as unbiased findings as possible [16, 19, 27, 33]. These considerations refer to a conflict in research with AI: On the one hand, the implemented data should be made comparable. However, any change in the datasets entails a change in the algorithm processes, and thus, the results. Furthermore, most radiomics studies to date have used manual segmentation, which could be a potential methodological limitation in terms of time and measurement error [1, 19, 27]. An increasingly explored solution could be automated segmentation based on deep learning. The requirements for segmentation models are complex due to the non-rigid nature of the pancreatic parenchyma and constitutional factors also influence deep learning segmentation results [5, 20]. Further testing and standardization are needed before they can be used on a large scale [19]. When extracting features, the feature space should be deliberately limited in order to ensure the best possible interpretation of the results. In their 2020 review of radiomics-based pancreas research, Abunahel et al. were able to show that there is a considerable variance (between 4 and 2041, median 166) in the number of extracted features [1]. This reinforced our approach to keep the feature space small right from the start of the extraction, to implement the selection methodology only for features that showed the most significant correlation and to use predefined feature classes whose interpretation should be comprehensible. This was done especially because the utility of radiometric features can vary depending on the application [1]. In our view, a higher number of features and undefined feature classes would have increased the scope for interpretation of the prediction results, which would have weakened the interpretation of results and reproducibility [15, 16, 27]. Nevertheless, this approach limited the detection of features of the pancreatic parenchyma that could improve the model results.

The current literature also suggests that deep learning will increasingly be used not only for segmentation, but also to develop predictive models for POPF itself. For example, Mu et al. were able to develop a preoperative model based on deep learning to identify high-risk patients with high predictive power [22]. Furthermore, Lee et al. were able to develop a POPF prediction model based on a combination of two machine learning and one deep learning model [18]. They also compared this with the Roberts score. The combined prediction model was significantly superior to the Roberts score in predicting preoperative POPF. The models performed equally well in predicting CR-POPF. However, these must be interpreted in the context of a predictive value of the Roberts score in this cohort that is significantly lower than in the literature described [24]. Despite this, the combination of machine and deep learning may be a promising approach for the future, but further research is needed.

Regarding APM, a cornerstone of radiomics research is the evaluation of one's model using the clinical gold standard, which in this case is a limitation [19, 27]. We were not aware of any standards against which validation would have been possible. The same applies to the upper limit of amylase in the drainage that was established. Recent advances in deep learning are revolutionizing medical imaging by enabling accurate and efficient analysis of datasets. Convolutional Neural Networks (CNNs) have been particularly successful in image classification, segmentation, and anomaly detection and are likely to be beneficial in the scenario of this study. However, CNNs were not the objective of this study and were therefore not tested on our dataset. Further research in this area would be important.

## Conclusion

In summary, PPM demonstrated that preoperative prediction of POPF is possible and achieved a higher value in the ROC analysis than the Roberts score. With the APM, we were able to provide a promising approach for postoperative drainage management after PD. Both models offer a new

contribution to informed preoperative decision making for the prediction and management of POPF. The results suggest that the combination of radiological and clinical parameters developed using only preoperatively available routine clinical and imaging data can compete with the potent Robert score for preoperative POPF risk stratification. In the future, the increased use of machine and deep learning methods for preoperative prediction of POPF seems likely. There is still a long way to go, and patience is needed before its implementation in clinical practice. On one hand, there is a need for standardization of clinical practice and, on the other hand, of radiomics workflows. This will require comprehensive studies with a prospective design, especially larger patient cohorts, tests against strong conventional scores and external multicenter validation.

## Abbreviations

POPF = Postoperative pancreatic fistula

PPM = POPF prediction model

APM = Amylase prediction model

CI = Confidence interval

AUC = Area under the curve

ISGPF = International study group on pancreatic fistula

BL = Biochemical leak

BMI = Body mass index

CT = Computer tomography

ROI = Region of interest

GLCM = Gray level co-occurrence matrix

GLRLM = Gray level run length matrix

GLSZM = Gray level size zone matrix

ROC = Receiver operating characteristic

PJ = Pancreaticojejunostomy

IQR = Interquartile range

POD = Postoperative day

CNN = Convolutional neural network

## Declarations

### Ethics approval and consent to participate

The project was approved by the Ethics Committee of the University of Freiburg (246/20) and was conducted in accordance with the Helsinki Declaration. The Ethics Committee of the Albert-Ludwigs-University Freiburg waived the requirement for a written declaration of consent.

## Consent for publication

Not applicable.

## Availability of Data and Materials

The data sets generated and analyzed as part of the current study are not publicly available as they are the subject of ongoing research. They are available as well as the packages used in R on reasonable request from the corresponding author.

## Competing interests

The authors declare that they have no competing interests.

## Funding

This work was funded entirely from internal resources. There was no external funding.

## Authors' contributions

J.K. was a major contributor in writing the manuscript, was instrumental in execution and design and statistical analysis of the experiment; M.B. was involved in the statistical analysis and support regarding radiomics; E.B. was a major contributor to the collection and analysis of clinical patient data; C.N. was responsible for the graphical representation of the results; S.F-F. and F.B. approved the manuscript and promoted the project; U.W. was the initiator of the project and was decisively involved in the conception on the surgical side, in the supervision and in the implementation; J.N. was radiologically involved in the conception, supervision, execution and statistical evaluation of the experiment. All authors read and approved the final manuscript.

## Acknowledgements

We thank Prof. Dr. Elmar Kotter of the Department of Radiology, University Medical Center Freiburg for scientific suggestions.

An early version of this work has been published as a preprint by Research Square. (https://doi.org/10.21203/rs.3.rs-3349227/v1; licensed under a CC BY 4.0 License).

## References

1. Abunahel BM, Pontre B, Kumar H, et al., Pancreas image mining: a systematic review of radiomics. Eur Radiol 31 (2021): 3447-3467.
2. Adamu M, Plodeck V, Adam C et al., Predicting postoperative pancreatic fistula in pancreatic head resections: which score fits all? Langenbecks Arch Surg (2021).
3. Allen PJ, Gönen M, Brennan MF, et al., Pasireotide for postoperative pancreatic fistula. N Engl J Med 370 (2014): 2014-2022.
4. Ansari D, Tingstedt B, Andersson B, et al., Pancreatic cancer: yesterday, today and tomorrow. Future Oncol Lond Engl 12 (2016): 1929-1946.

5. Bagheri MH, Roth H, Kovacs W, et al., Technical and Clinical Factors Affecting Success Rate of a Deep Learning Method for Pancreas Segmentation on CT. Acad Radiol 27 (2020): 689-695.

6. Bassi C, Dervenis C, Butturini G et al., Postoperative pancreatic fistula: an international study group (ISGPF) definition. Surgery 138 (2005): 8-13.

7. 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 161 (2017): 584-591.

8. Diener MK, Knaebel H-P, Heukaufer C, et al., A systematic review and meta-analysis of pylorus-preserving versus classical pancreaticoduodenectomy for surgical treatment of periampullary and pancreatic carcinoma. Ann Surg 245 (2007): 187-200.

9. Doi R, Imamura M, Hosotani R, et al., Surgery versus radiochemotherapy for resectable locally invasive pancreatic cancer: final results of a randomized multi-institutional trial. Surg Today 38 (2008): 1021-1028.

10. Fedorov A, Beichel R, Kalpathy-Cramer J, et al., 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 30 (2012): 1323-1341.

11. Fisher WE, Hodges SE, Silberfein EJ, et al., Pancreatic resection without routine intraperitoneal drainage. HPB 13 (2011): 503–510.

12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 278 (2016): 563–577.

13. van Griethuysen JJM, Fedorov A, Parmar C, et al., Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Res 77 (2017): e104-e107.

14. Kawaida H, Kono H, Hosomura N, et al., Surgical techniques and postoperative management to prevent postoperative pancreatic fistula after pancreatic surgery. World J Gastroenterol 25 (2019): 3722-3737.

15. Lambin P, Leijenaar RTH, Deist TM, et al., Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 14 (2017): 749-762.

16. Larue RTHM, Defraene G, De Ruysscher D, et al., Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. Br J Radiol 90 (2017): 20160665.

17. Lee SR, Kim HO, Shin JH, Significance of drain fluid amylase check on day 3 after pancreatectomy. ANZ J Surg 89 (2019): 497-502.

18. Lee W, Park HJ, Lee H-J, et al., Deep learning-based prediction of post-pancreaticoduodenectomy pancreatic fistula. Sci Rep 14 (2024): 5089.

19. Liu Z, Wang S, Dong D, et al., The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges. Theranostics 9 (2019): 1303-1322.

20. Man Y, Huang Y, Feng J, et al., Deep Q Learning Driven CT Pancreas Segmentation With Geometry-Aware U-Net. IEEE Trans Med Imaging 38 (2019): 1971–1980.

21. McGuigan A, Kelly P, Turkington RC, et al., Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes. World J Gastroenterol 24 (2018): 4846-4861.

22. Mu W, Liu C, Gao F, et al., Prediction of clinically relevant Pancreatico-enteric Anastomotic Fistulas after Pancreatoduodenectomy using deep learning of Preoperative Computed Tomography. Theranostics 10 (2020): 9779-9788.

23. Qayyum A, Tamm EP, Kamel IR, et al., ACR Appropriateness Criteria® Staging of Pancreatic Ductal Adenocarcinoma. J Am Coll Radiol JACR 14 (2017): S560-S569.

24. Roberts KJ, Hodson J, Mehrzad H et al (2014) A preoperative predictive score of pancreatic fistula following pancreatoduodenectomy. HPB 16:620-628.

25. Saif MW. Advancements in the management of pancreatic cancer: 2013. JOP J Pancreas 14 (2013): 112-118.

26. Skawran SM, Kambakamba P, Baessler B, et al., Can magnetic resonance imaging radiomics of the pancreas predict postoperative pancreatic fistula? Eur J Radiol 140 (2021): 109733.

27. van Timmeren JE, Cester D, Tanadini-Lang S, et al., Radiomics in medical imaging- “how-to” guide and critical reflection. Insights Imaging 11 (2020): 91.

28. Utzolino S, Kayser C, Keck T, et al., Perioperatives Management bei Pankreasresektionen. Viszeralmedizin 27 (2011): 19-27.

29. Vallance AE, Young AL, Macutkiewicz C, et al., Calculating the risk of a pancreatic fistula after a pancreaticoduodenectomy: a systematic review. HPB 17 (2015):1040-1048.

30. Van Buren G, Bloomston M, Hughes SJ, et al., A randomized prospective multicenter trial of pancreaticoduodenectomy with and without routine intraperitoneal drainage. Ann Surg 259 (2014): 605-612.

31. Vincent A, Herman J, Schulick R et al., Pancreatic cancer. Lancet Lond Engl 378 (2011): 607-620.

32. Yamamoto Y, Sakamoto Y, Nara S, et al., A preoperative predictive scoring system for postoperative pancreatic fistula after pancreaticoduodenectomy. World J Surg 35 (2011): 2747-2755.

33. Yip SSF, Aerts HJWL. Applications and limitations of radiomics. Phys Med Biol 61 (2016): R150-166.

34. 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 10 (2018): 6469-6478.

**Citation:** Johannes D Kaiser, Matthias Benndorf, Esther A Biesel, Claudia Neubauer, Stefan Fichtner-Feigl, Fabian Bamberg, Uwe A Wittel, Jakob Neubauer. Preoperative Prediction of Postoperative Pancreatic Fistula After Pancreatic Head Resection Using Radiomics and Machine Learning Based on Computed Tomographic Diagnostics. Journal of Radiology and Clinical Imaging. 7 (2024): 88-98.