



OPEN

Development of an interpretable machine learning model based on CT radiomics for the prediction of post acute pancreatitis diabetes mellitus

Xiyao Wan¹, Yuan Wang¹, Ziyi Liu¹, Ziyang Liu¹, Shuting Zhong² & Xiaohua Huang¹✉

This study sought to establish and validate an interpretable CT radiomics-based machine learning model capable of predicting post-acute pancreatitis diabetes mellitus (PPDM-A), providing clinicians with an effective predictive tool to aid patient management in a timely fashion. Clinical and imaging data from 271 patients who had undergone enhanced CT scans after first-episode acute pancreatitis from March 2017–June 2023 were retrospectively analyzed. Patients were classified into PPDM-A ($n=109$) and non-PPDM-A groups ($n=162$), and split into training ($n=189$) and testing ($n=82$) cohorts at a 7:3 ratio. 1223 radiomic features were extracted from CT images in the plain, arterial and venous phases, respectively. The radiomics model was developed based on the optimal features retained after dimensionality reduction, utilizing the extreme gradient boosting (XGBoost) algorithm. Five-fold cross-validation of the model was used to assess the performance of the model in the training and testing cohorts. The clinical performance of the model was assessed through a decision curve analysis, while insight into the predictions derived from this model was derived from Shapley additive explanations (SHAP). The final model incorporated five key radiomic features, and achieved area under the curve values in the training and testing cohorts of 0.947 (95% CI 0.915–0.979) and 0.901 (95% CI 0.838–0.964), respectively. SHAP analyses indicated that textural features were key features relevant to the prediction of PPDM-A incidence. The interpretable CT radiomics-based model developed in this study demonstrated good performance, enabling timely and effective interventions with the potential to improve patient outcomes.

Keywords Acute pancreatitis, Diabetes, CT scan, Radiomics, Interpretability

Post-acute pancreatitis diabetes mellitus (PPDM-A) is a type of secondary subtype of diabetes and is the second most commonly diagnosed type of newly diabetes in adults¹. Patients with PPDM-A exhibit more pronounced fluctuations in their blood sugar levels and face greater difficulty managing their condition relative to type 2 diabetes patients^{2,3}, exposing them to a higher risk of death. There are also differences in the medication approaches used to treat these two forms of diabetes. According to the American Diabetes Association (ADA), the progression of PPDM-A is associated with a 33-fold higher risk of pancreatic cancer development⁴. The identification of individuals facing a higher PPDM-A risk is thus vital to effectively prevent and manage this disease and sequelae. In prior studies, imaging data radiomics-based models have shown promise as tools for PPDM-A prediction⁵. As these machine learning models tend to exhibit a “black box”-like nature, however, determining the basis for particular predictions in specific patient populations can often be difficult. Radiomics approaches entail the extraction of a wide array of quantitative features from computed tomography (CT) or other imaging datasets, whereupon predictive models can be developed by machine learning or deep learning algorithms that process these datasets⁶. The inherent complexity of the resultant models typically precludes any efforts to intuitively understand their internal operations. As a result, this lack of decision-making transparency tends to limit the clinical implementation of these models despite their predictive accuracy. Model interpretation

¹Department of Radiology, Affiliated Hospital of North Sichuan Medical College, No.63 Wenhua Road, Shunqing District, Nanchong 637000, China. ²Department of Radiology, Chongqing University Cancer Hospital, No.181 Hanyu Road, Shapingba District, Chongqing 400030, China. ✉email: 15082797553@163.com

is also very challenging, as the medical data used for modeling tends to be complex and ambiguous⁷. The SHAP (SHapley additive explanations) technique has been developed in an effort to clarify the basis for machine learning model-based predictions. SHAP is based on the game theory concept of Shapley values, enabling the quantification of individual feature contributions to the prediction of a given model through the establishment of a visual, interpretable explanation of model outcomes⁸. In this study, CT-derived radiomics features were thus used for the development and validation of a machine learning model for the prediction of PDM-A to provide a tool that can aid the accurate detection of this disease at an early time point. To ensure the interpretability of the developed radiomics model, SHAP analysis was also performed.

Materials and methods

Study participants

The present study was approved by the Medical Ethics Committee of the Affiliated Hospital of North Sichuan Medical College and the Medical Ethics Committee of the Affiliated Hospital of North Sichuan Medical College waived the requirement of informed consent owing to its retrospective design (2022ER320-1). All the procedures performed in this study were in accordance with the Declaration of Helsinki (as revised in 2013).

This study ultimately enrolled 271 patients who had undergone contrast-enhanced CT scans within 7 days of an initial bout of acute pancreatitis (AP) between March 2017 and June 2023 (Fig. 1). These patients were separated into two groups based on whether or not they developed diabetes following AP, including a PPDM-A group ($n=109$) who underwent long-term follow-up for diabetes through June 2024, and a non-

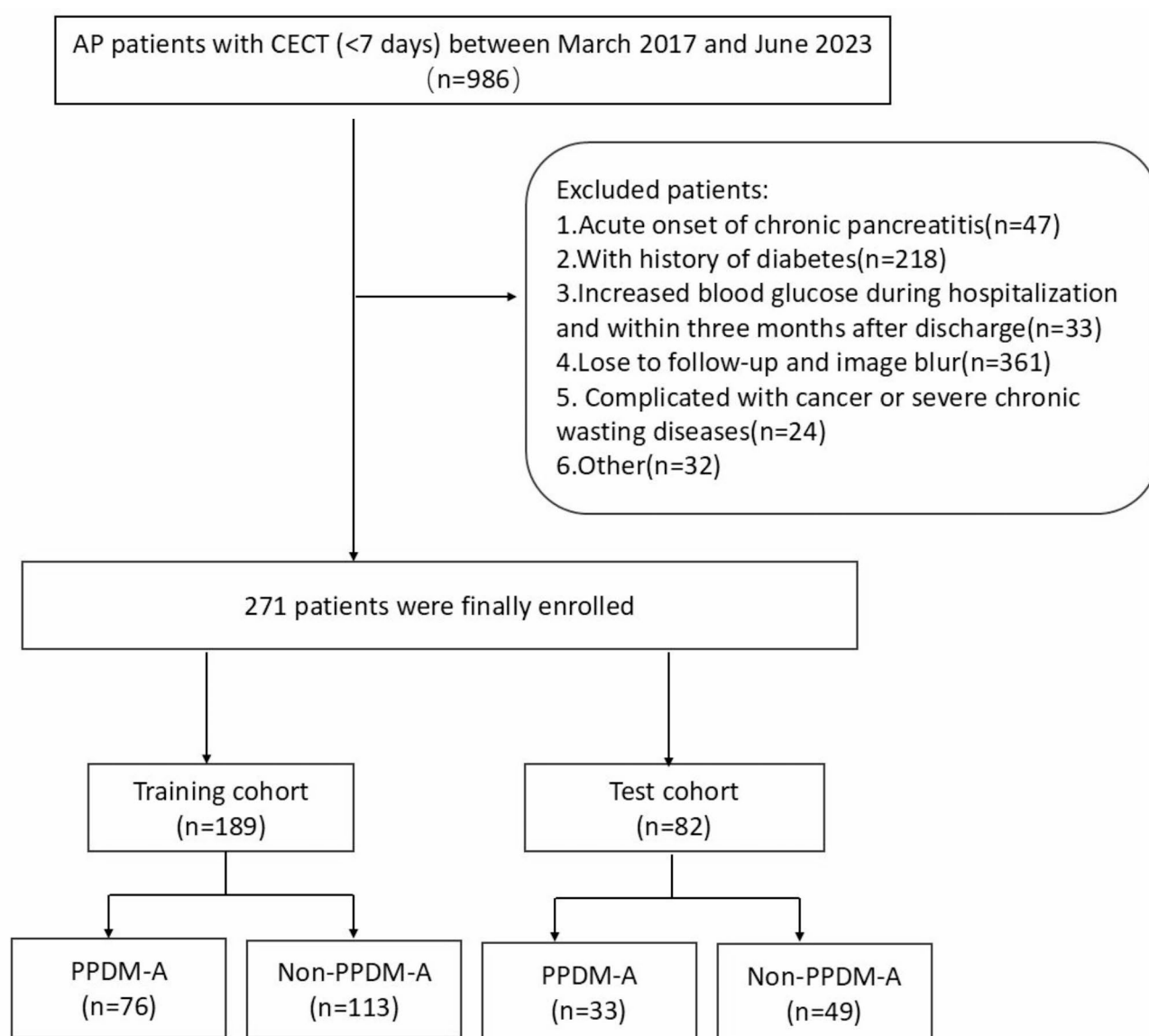


Fig. 1. Flowchart of patient recruitment in this study. AP acute pancreatitis, CECT contrast-enhanced computed tomography, PPDM-A post-acute pancreatitis diabetes mellitus.

PPDM-A group ($n = 162$) for whom follow-up was performed through medical record review and by telephone through June 2024 to confirm that they had not developed diabetes during this period⁹. These patients were further randomized into training ($n = 189$) and testing ($n = 82$) cohorts at a 7:3 ratio. Clinical and demographic features collected from these patients included age, sex, alcohol intake, smoking history, hypertension, fatty liver, disease characteristics (modified computed tomography severity index, severity, presence/absence of necrosis), laboratory indices (hypersensitive C-reactive protein, lipase, pancreatic amylase, serum amylase), and local complications (walled-off necrosis, pancreatic pseudocyst, acute peripancreatic fluid collection¹⁰).

PPDM-A group inclusion criteria (1) Diagnosed AP¹⁰; (2) No prior history of diabetes; (3) Diabetes diagnosed 3+ months after AP onset, as defined by one or more of the following: ① Fasting blood glucose ≥ 7.0 mmol/L, ② 2-h postprandial blood glucose ≥ 11.1 mmol/L (200 mg/dL), ③ HbA1c $\geq 6.5\%$, ④ typical symptoms of diabetes together with random blood glucose levels ≥ 11.1 mmol/L (200 mg/dL)^{4,11}.

Exclusion criteria (1) Acutely exacerbated chronic pancreatitis; (2) Any history of prediabetes or diabetes; (3) Blood glucose levels that were elevated while hospitalized or within 3 months following discharge; (4) Other comorbid conditions impacting pancreatic function (such as pancreatic resection, pancreatic cystic fibrosis, pancreatic tumors, pancreatic trauma, fibrous pancreatic disease, or hemochromatosis¹²); (5) Women who were pregnant or lactating; (6) Unclear imaging data or scanning parameters that were inconsistent; (7) < 18 or > 80 years of age; or (8) Incomplete clinical data or loss to follow-up.

Non-PPDM-A group inclusion criteria (1) First-episode AP; (2) No evidence of abnormal blood glucose levels over the course of follow-up.

Exclusion criteria (1) Acutely exacerbated chronic pancreatitis; (2) Any history of malignant tumors; (3) Unclear imaging data or scanning parameters that were inconsistent; (4) Incomplete clinical data or loss to follow-up. (5) Women who were pregnant or lactating; (6) < 18 or > 80 years of age.

Image acquisition

A 64-slice Siemens Somatom Force CT scanner and a 64-slice Philips Brilliance iCT scanner were used for CT scanning in helical mode, covering the area from the diaphragm to the lower edge of the pancreas. The scanning parameters included: tube voltage = 120 kV, automated mAs settings, pitch factor = 1, slice thickness = $0.625 \text{ mm} \times 64$, reconstruction slice thickness = 5 mm, slice spacing = 5 mm, tube rotation time = 0.50 s, and matrix size = 334×334 . A high-pressure injector was used for the injection of the non-ionic contrast agent iopamidol via the antecubital vein at 3.0–3.5 mL/s with an overall dose of 1.5 mL/kg, acquiring images at 25–30 s (arterial phase) and 45–50 s (venous phase) following contrast injection.

Radiomics feature selection

Two radiologists (5+ years experience) blinded to patient clinical details manually outlined the pancreas on plain, arterial, and venous phase CT images with 3D Slicer (v5.0.3, <https://www.slicer.org>), avoiding major structures including the common bile duct, splenic artery/vein, and pancreaticoduodenal artery/vein when delineating the region of interest (ROI) consisting of the entirety of the pancreas (Fig. 2). Laplacian Gaussian filtering and wavelet transform filtering were used to process these images in 3D Slicer for smoothing, edge enhancement, and noise reduction. The “PyRadiomics” package (v3.0.1; <https://www.slicer.org/>) was used to extract 1223 features from each ROI. Before these features were extracted, the resampling of all images at a $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$ voxel resolution was performed to ensure consistency.

To limit the effects of any subjective biases, the ROIs for 97 patients selected at random were independently delineated by the two radiologists participating in this study, and interclass correlation coefficients (ICCs) were then used to assess agreement between these observers.

Z-score normalization was performed for all radiomics features exhibiting an ICC > 0.75 , thereby ensuring that each had a mean value of 0 and a standard deviation of 1 as a means of eliminating the effects of scale-related differences among the different features. Feature selection was then performed through the variance threshold, K-best, and Least Absolute Shrinkage and Selection Operator (LASSO) methods. The detailed radiomics process is presented in Fig. 3.

Model construction and evaluation

The XGBoost algorithm was used to construct a radiomics model, the performance of which was evaluated via five-fold cross-validation to minimize overfitting and ensure model generalizability. The predictive performance of the model was assessed based on the accuracy, sensitivity, specificity, F1 score, and area under the curve (AUC) values. The clinical net benefit afforded by the model was assessed through a decision curve analysis (DCA).

Model interpretability

The influence of individual features on the predictions made by the developed model was assessed using the R shapviz package (v 4.3.2, <http://www.r-project.org/>) to calculate SHAP values, which enable the quantification of how these features affect model outputs, thereby improving interpretability¹³. SHAP values were visualized to support efforts to clarify the effects of particular features on model predictions.

Statistical analysis

SPSS 26.0 (IBM Corp, NY, USA) was used for all statistical analyses. After assessing data normality for continuous data with the Kolmogorov–Smirnov test, Skewed data were reported as medians with interquartile ranges and compared using the Mann–Whitney U test, while normally distributed data were reported as means \pm standard deviations and compared using independent sample t-tests. Categorical data were reported as numbers or

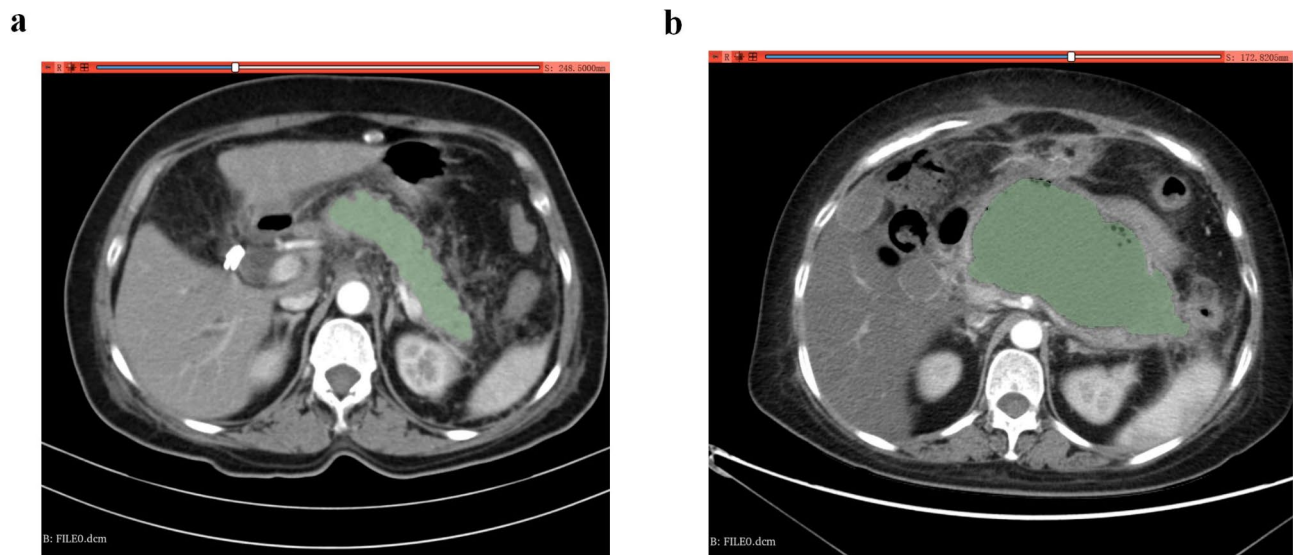


Fig. 2. Illustration of the delineation of the pancreatic ROI in acute pancreatitis patients. A patient with mild acute pancreatitis and slight pancreatic enlargement (a). Blurred peripancreatic fat spaces with strip-like exudates and bilateral perirenal fascia thickening are visible. No dilation of the main pancreatic duct is observed. The region of interest (ROI) is outlined layer by layer along the pancreatic edge, avoiding blood vessels and the common bile duct. A case of acute necrotizing pancreatitis with infection and walled-off necrosis (b). A large cystic low-density lesion (approximately 14.7×7.2 cm) with gas density is observed, with significant enhancement of the cyst wall on contrast-enhanced imaging. Residual normal pancreatic tissue is seen at the pancreatic head and tail, with no abnormal density or enhancement. The peripancreatic fat space is poorly defined. The ROI is outlined along the cyst boundary (including walled-off necrosis area), while avoiding critical structures such as blood vessels, the common bile duct, gastrointestinal tract, peritoneum, and mesentery. ROI region of interest.

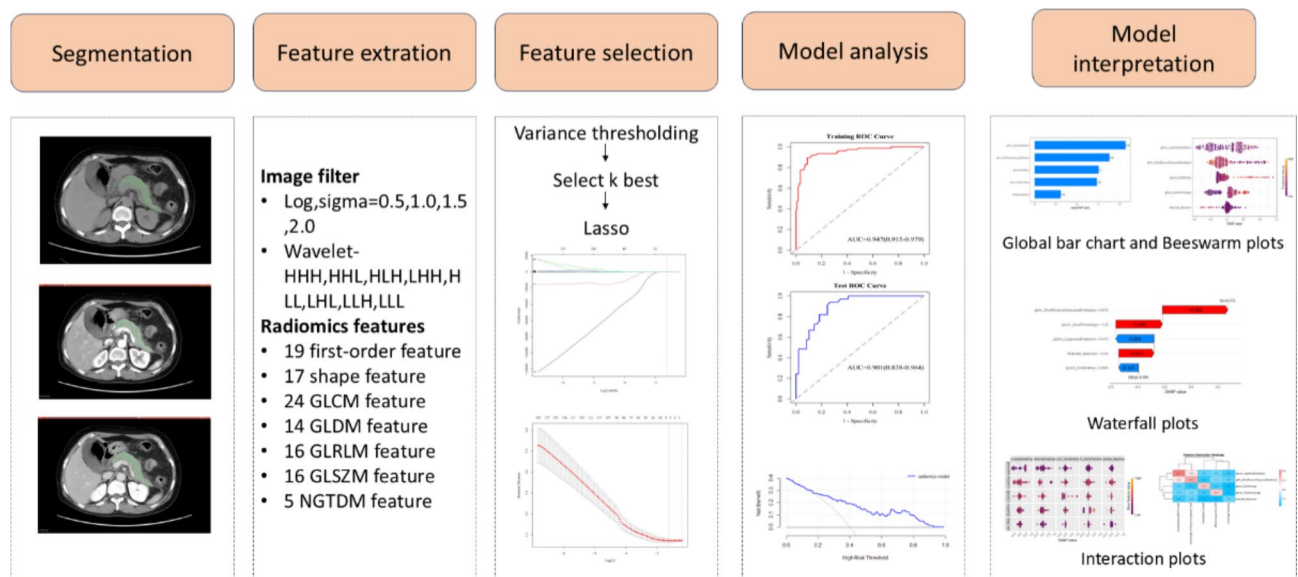


Fig. 3. Workflow of radiomics process in this study. GLCM gray level co-occurrence matrix, GLDM gray level dependence matrix, GLRLM gray level run-length matrix, GLSZM gray level size zone matrix, LASSO least absolute shrinkage and selection operator, NGTDM neighboring graytone difference matrix.

percentages and compared with chi-square tests. The analysis of radiomics features and associated modeling was performed with R (v 4.3.2, <http://www.r-project.org/>), and $P < 0.05$ was regarded as being significant.

Results

Patient information

In total, this study enrolled 271 patients with AP (177 males, 94 females). The clinical characteristics are summarized in Table 1. Of the 22 clinical characteristics measured, only the infection status significantly differed between the PPDM-A group and the non-PPDM-A group ($P < 0.001$).

Radiomics model establishment and evaluation

For each patient, 3669 radiomics features were extracted, of which 3078 were retained following ICC analysis, including 881, 1142, and 1055 features from the plain, arterial, and portal venous phase scans, respectively. The variance threshold and K-best means methods respectively reduced the numbers of features by 117 and 235, yielding a list of 2726 features. Using the LASSO algorithm with 10-fold cross-validation, the λ parameter was adjusted to an optimal list of 5 features retained for model construction, including 1 each from the plain and arterial phase scans, as well as 3 from venous phase scans (Supplementary Table S1). The model developed with the XGBoost algorithm yielded respective AUC values of 0.947 (95% CI 0.915–0.979) and 0.901 (95% CI 0.838–0.964) in the training and test cohorts (Fig. 4A,B, Supplementary Table S2). DCA analysis (Fig. 5A,B) demonstrated that the radiomics model provides a high clinical net benefit across a range of threshold probabilities. Specifically, the model outperforms the “Treat All” strategy in terms of net benefit, suggesting its potential for improving clinical decision-making.

Model interpretation

The five most influential features in the model according to the SHAP global bar chart (Fig. 6A) were glszm-LargeAreaEmphasis, glrlm-ShortRunLowGrayLevelEmphasis, glszm-ZoneEntropy, glszm-ZonePercentage, and firstorder-Maximum. The global importance of these features was assessed based on their average SHAP values across all samples. The waterfall plots (Fig. 6C) generated for the third patient reveal the contributions of each feature to the baseline prediction value of the model [$E[f(x)] = 0.395$], yielding the adjusted final prediction value of [$f(x) = 0.272$]. For instance, the glrlm-ShortRunLowGrayLevelEmphasis feature increased the prediction score for the patient by 0.488 relative to baseline. Beeswarm plots (Fig. 6B) further highlight the effects of each feature on prediction probability values, with positive and negative effects represented in orange and purple, respectively. Positive and negative SHAP values respectively suggest increased and decreased risk of PPDM-A¹⁴. Features including glszm-LargeAreaEmphasis, glszm-ZoneEntropy, and glszm-ZonePercentage had a positive influence on the prediction of PPDM-A incidence, whereas glrlm-ShortRunLowGrayLevelEmphasis and firstorder-Maximum had a negative impact. These SHAP interactions are summarized in Fig. 6D with features along the diagonal highlighting their individual effects on predictions, while areas away from the diagonal highlight the effects of combinations of features. Wider subplots are indicative of stronger interactions among features¹⁵. The established heatmap (Fig. 6E) highlights the significant interactions that exist between glrlm-ShortRunLowGrayLevelEmphasis and glszm-LargeAreaEmphasis.

Discussion

In this study, PPDM-A was predicted using a newly developed CT radiomics-based model. SHAP analyses were also implemented to improve the interpretability of this model, yielding a high degree of AUC. Radiomics approaches are thus an effective means of predicting PPDM-A risk, while SHAP can enhance the applicability of these methods by improving the overall transparency of the model decision-making process.

PPDM-A incidence has reportedly almost doubled over the past decade, and the annual incidence rate is reportedly around 2.8%¹⁶. Zhang et al.¹⁷ used an interpretable clinical index-based machine learning model in a prior study, but the AUC of this model (0.819) indicating that simple clinical models have limited value in predicting PPDM-A. By extracting radiomics features from only either enhanced or non-contrast-enhanced CT images, Zhong et al.¹⁸ and Hu et al.⁵ potentially overlooked key pieces of data. In the present study, data from both image types was combined, thereby capturing a wider range of anatomical and hemodynamic features, yielding greater predictive accuracy while overcoming the limitations imposed by reliance on a single imaging strategy. The resultant model, which was based on five highly reliable texture features, exhibits excellent accuracy and generalizability. Additionally, we found a significant difference in infection status between the PPDM-A and non-PPDM-A groups. Infection can affect changes in pancreatic texture and tissue characteristics, including local blood flow, potentially introducing bias. To mitigate this, we excluded major blood vessels during segmentation and used careful feature selection methods, such as variance threshold, K-best, and LASSO, to minimize the impact of irrelevant features. Repeated cross-validation analyses showed that infection did not significantly affect the performance of the model. While infection could be a confounding factor, our model remained stable and reliable. Future studies should better control for this variable to further validate our findings.

SHAP was herein leveraged to interpret the radiomics-based predictions of PPDM-A risk for the first time. In AP, damage to the pancreas results in the destruction of islet cells, reduced insulin secretion, and higher levels of insulin resistance^{19–21}. These shifts can be reflected by altered pancreatic morphology, density, and texture visible on CT images. Among the texture features incorporated into the model in this study, glszm (Gray-Level Size Zone Matrix) was particularly important. GLSZM is a texture feature that emphasizes the continuity of pixel values within two-dimensional regions. By computing characteristic parameters such as SZE (Size Zone Emphasis) and LZE (Long Zone Emphasis), it analyzes the heterogeneity of image textures and the distribution of region sizes⁶. The most significant predictor of PPDM-A in this study, according to SHAP analysis, was glszm LargeAreaEmphasis, likely owing to the ability of this feature to detect extensive pancreatic injury and

	PPDM-A (n = 109)	Non-PPDM-A (n = 162)	p value
Age (years)	47.10 ± 12.14	45.12 ± 11.58	0.178
Gender (n, %)			0.756
Male	70/64.2	107/66	
Female	39/35.8	55/34	
Smoking (n, %)	45/41.3	67/41.4	0.990
Drinking (n, %)	36/33	67/41.4	0.166
Severity (n, %)			0.076
Mild	26/23.9	38/23.5	
Moderate	53/48.6	97/59.9	
Severe	30/27.5	27/16.7	
Infection (n, %)	13/11.9	1/0.6	< 0.001
Hypertension (n, %)	19/17.4	20/12.3	0.242
Biliary stones (n, %)			
PPDM-A	20/18.3	34/21.0	0.594
Local complications (n, %)	80/73.4	124/76.5	0.556
Pancreatic necrosis (n, %)	32/29.4	40/24.7	0.394
Recurrence (n, %)	42/37.8	48/29.6	0.156
Fatty liver disease (n, %)	57/52.3	85/52.5	0.977
Hospital stay (days)	10.00 (8.00–14.00)	11.00 (8.00–15.00)	0.651
Ca (mmol/L)	2.25 (2.13–2.36)	2.27 (2.16–2.37)	0.641
Serum amylase (U/L)	119.50 (68.50–377.00)	154.30 (82.00–572.50)	0.167
Lipase (U/L)	176.00 (68.00–502.00)	188.60 (67.00–815.50)	0.208
Pancreatic amylase (U/L)	90.70 (46.45–313.00)	106.00 (49.00–477.00)	0.192
Hs-CRP (mg/L)	43.76 (10.98–106.06)	31.63 (6.96–83.29)	0.083
WBC (× 10 ⁹ l ⁻¹)	11.46 ± 4.44	12.46 ± 4.73	0.081
Triglyceride (mmol/L)	4.06 (1.71–8.24)	3.04 (1.295–7.77)	0.161
Cholesterol (mmol/L)	4.95 (4.04–6.52)	4.99 (4.06–6.43)	0.719
MCTSI	4 (4–6)	4 (4–6)	0.164

Table 1. Patient demographic and clinical characteristics stratified according to PPDM-A status. *Hs-CRP* hypersensitive C-reactive protein, *WBC* blood cell count, *MCTSI* modified computed tomography severity index.

inflammation²². Indeed, this feature reflects the overall pathological state of the pancreas owing to the fact that it captures extensive gray-level variation. As a measure of textural complexity and gray-level distribution, ZoneEntropy suggests that PPDM-A patients exhibit greater pancreatic heterogeneity consistent with the damage in these patients being more severe than that for patients without PPDM-A²³.

In prior studies, both disease severity and pancreatic necrosis have been identified as being independently associated with the risk of endocrine dysfunction after AP^{19,24,25}, with higher incidence following episodes of severe AP relative to mild AP (39% vs. 14%). This aligns well with the findings from the prior study, emphasizing the reliability of the established radiomics model. In the present model, morphological changes including diameter and volume were not significant, indicating that AP-related morphological shifts may not be good predictors of PPDM-A incidence. The established SHAP waterfall plots offer a detailed overview of the contributions of specific features to PPDM-A risk predictions, affording a higher degree of insight and specificity as compared to many such models, which suffer from a lack of interpretability.

The predictive model developed herein was generated using the XGBoost classifier, which enhanced predictive accuracy by employing ensemble learning to combine multiple weak models (individual decision trees) into a single model with greater power²⁶. Relative to other classifiers, XGBoost is capable of automatically capturing interactions among features through the examination of multiple features during decision tree splits, thereby providing a higher degree of insight relative to reliance on any single feature. The SHAP interaction plots generated herein highlight strong interactions between the LargeAreaEmphasis and glrlm-ShortRunLowGrayLevelEmphasis features, the latter of which serves as a measure of small region uniformity within CT images. Lesion areas may thus harbor both low-gray-level small regions and extensive areas of uniformity. These detailed insights can provide a more thorough understanding of how this predictive model functions, leading to greater overall accuracy.

Limitations

This study has some limitations. For one, the retrospective nature of this study contributed to a relatively limited sample size, highlighting a need for future validation in larger cohorts of patients. Secondly, these data were all derived from a single center and may be susceptible to selection bias such that further multicenter validation

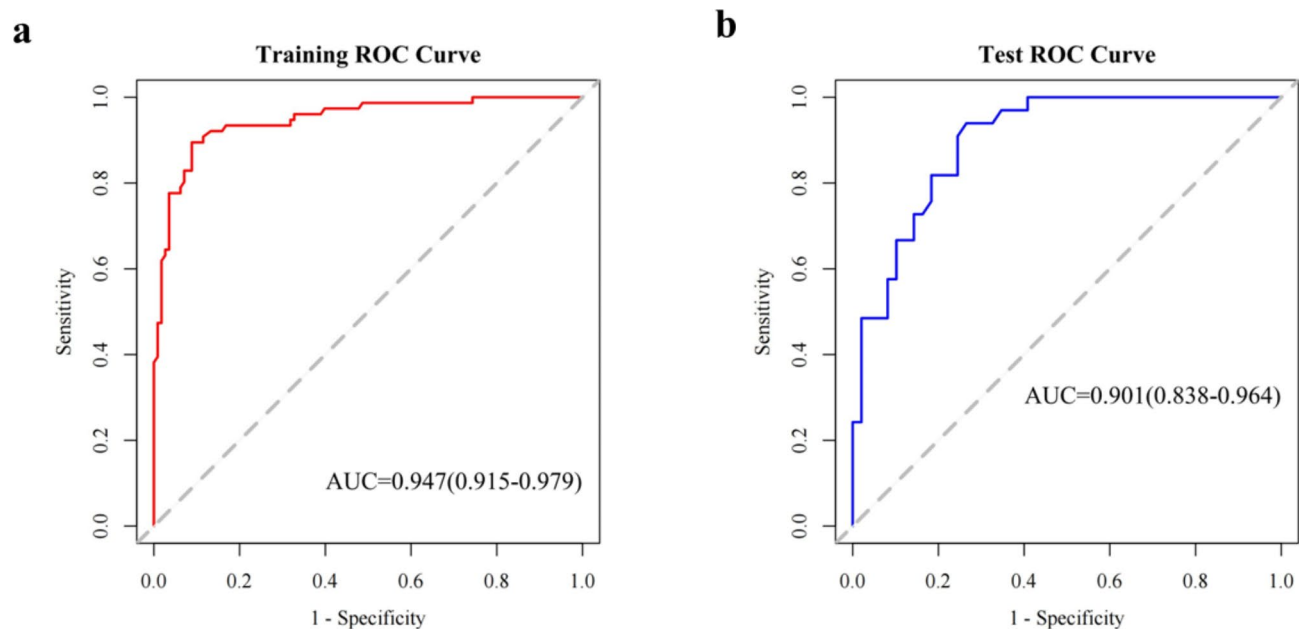


Fig. 4. ROC curves in (a) training and (b) test cohorts. ROC receiver operating characteristic.

of the robustness and generalizability of this model is essential. Third, this model did not incorporate clinical parameters, and certain key pieces of clinical data may thus have been overlooked. In the future, researchers should seek to combine radiomics features and clinical variables to establish more comprehensive machine learning-based models capable of predicting PPDM-A risk, thereby informing the understanding of this disease and aiding its diagnosis.

Conclusion

In conclusion, a robust, interpretable machine learning model based on CT radiomics features was herein developed as a tool to predict PPDM-A incidence. These results highlight the value of pancreatic radiomics features as biomarkers that can be leveraged to objectively predict PPDM-A, allowing for individualized clinical decision-making such that high-risk patients can receive appropriate interventions at an early time point. The generated SHAP values also ensure that this model is interpretable by clinicians, thereby ensuring the greater overall reliability of this radiomics model.

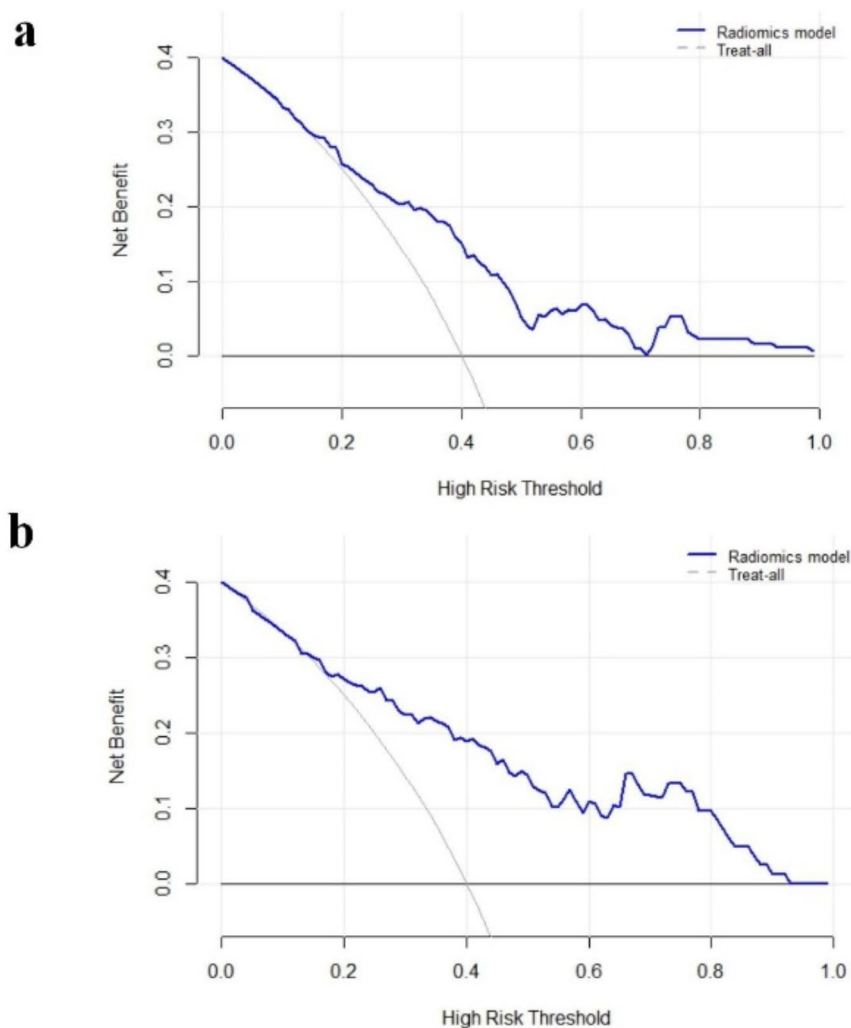


Fig. 5. DCA of the established models in the (a) training and (b) test cohorts. The grey curve represents the “Treat All Patients” strategy, where all patients are treated regardless of their predicted risk. Net Benefit on the y-axis reflects the model’s clinical utility, with higher values indicating greater benefit from the predictive model. DCA decision curve analysis.

DCA, decision curve analysis.

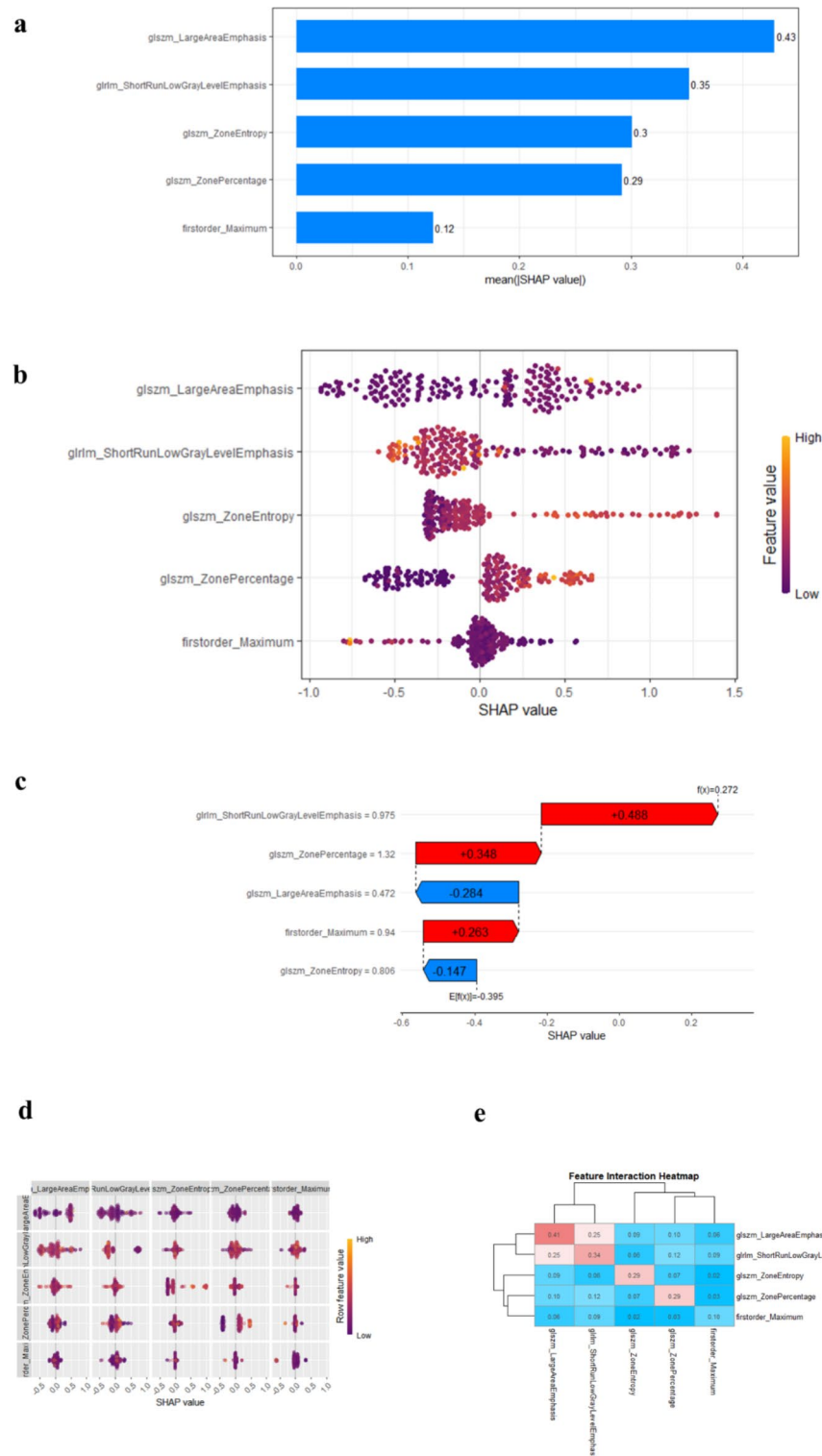


Fig. 6. SHAP analysis results for the radiomics model for PPDM-A prediction using XGboost. In the global bar chart (a), top variables contributed more significantly to the model, exhibiting higher predictive power compared to the ones positioned at the bottom. The beeswarm plot (b) illustrates the contribution of each feature to the model's prediction, with each point representing the SHAP value of a sample in the dataset. The SHAP waterfall plots (c) showed the individual interpretability of radiomics models. Red bar indicates increased predictive value and blue bar indicates decreased predictive value. Interaction summary plot (d) and heatmap (e) display both individual and interaction SHAP values, with diagonal points showing the individual SHAP values for each feature, and off-diagonal points revealing the interaction effects between feature pairs.

Data availability

The datasets generated and analysed during the current study are not publicly available due to our need to expand the sample size and conduct further research based on these datasets, but are available from the corresponding author on reasonable request.

Received: 6 October 2024; Accepted: 9 January 2025

Published online: 15 January 2025

References

1. Bharmal, S. H. et al. Trajectories of glycaemia following acute pancreatitis: A prospective longitudinal cohort study with 24 months follow-up. *J. Gastroenterol.* **55**, 775–788 (2020).
2. Cho, J., Scragg, R., Pandol, S. J., Goodarzi, M. O. & Petrov, M. S. Antidiabetic medications and mortality risk in individuals with pancreatic cancer-related diabetes and postpancreatitis diabetes: A nationwide cohort study. *Diabetes Care* **42**, 1675–1683 (2019).
3. Woodmansey, C. et al. Incidence, demographics, and clinical characteristics of diabetes of the exocrine pancreas (type 3c): A retrospective cohort study. *Diabetes Care* **40**, 1486–1493 (2017).
4. Classification and Diagnosis of Diabetes. Standards of medical care in diabetes-2022. *Diabetes Care* **45**, S17–S38 (2022).
5. Hu, R. et al. A radiomics model of contrast-enhanced computed tomography for predicting post-acute pancreatitis diabetes mellitus. *Quant. Imaging Med. Surg.* **14**, 2267–2279 (2024).
6. Mayerhoefer, M. E. et al. Introduction to radiomics. *J. Nucl. Med.* **61**, 488–495 (2020).
7. Coppola, F. et al. Human, all too human? An all-around appraisal of the “artificial intelligence revolution” in medical imaging. *Front. Psychol.* **12**, 710982 (2021).
8. Liu, Z. et al. Noninvasive prediction of perineural invasion in intrahepatic cholangiocarcinoma by clinicoradiological features and computed tomography radiomics based on interpretable machine learning: A multicenter cohort study. *Int. J. Surg.* **110**, 1039–1051 (2024).
9. Petrov, M. S. Diagnosis of endocrine disease: Post-pancreatitis diabetes mellitus: Prime time for secondary disease. *Eur. J. Endocrinol.* **184**, R137–R149 (2021).
10. Banks, P. A. et al. Classification of acute pancreatitis–2012: Revision of the Atlanta classification and definitions by international consensus. *Gut* **62**, 102–111 (2013).
11. Diagnosis and Classification of Diabetes. Standards of care in diabetes-2024. *Diabetes Care* **47**, S20–S42 (2024).
12. Petrov, M. S. Diabetes of the exocrine pancreas: American Diabetes Association-compliant lexicon. *Pancreatol.* **17**, 523–526 (2017).
13. Nohara, Y., Matsumoto, K., Soejima, H. & Nakashima, N. Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Comput. Methods Programs Biomed.* **214**, 106584 (2022).
14. Bang, M. et al. An interpretable radiomics model for the diagnosis of panic disorder with or without agoraphobia using magnetic resonance imaging. *J. Affect. Disord.* **305**, 47–54 (2022).
15. Cheng, X. et al. Predicting response to CCRT for esophageal squamous carcinoma by a radiomics-clinical SHAP model. *BMC Med. Imaging* **23**, 145 (2023).
16. Pendharkar, S. A., Mathew, J. & Petrov, M. S. Age- and sex-specific prevalence of diabetes associated with diseases of the exocrine pancreas: A population-based study. *Dig. Liver Dis.* **49**, 540–544 (2017).
17. Zhang, J. et al. Machine learning for post-acute pancreatitis diabetes mellitus prediction and personalized treatment recommendations. *Sci. Rep.* **13**, 4857 (2023).
18. Zhong, S. et al. Developing a CT-based radiomics nomogram for predicting post-acute pancreatitis diabetes mellitus incidence. *Br. J. Radiol.* **96**, 20230382 (2023).
19. Zhi, M. et al. Incidence of new onset diabetes mellitus secondary to acute pancreatitis: A systematic review and meta-analysis. *Front. Physiol.* **10**, 637 (2019).
20. Manrai, M. et al. Diabetes mellitus as a consequence of acute severe pancreatitis: Unraveling the mystery. *World J. Diabetes* **14**, 1212–1225 (2023).
21. Boxhoorn, L. et al. Acute pancreatitis. *Lancet* **396**, 726–734 (2020).
22. Zwanenburg, A. et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* **295**, 328–338 (2020).
23. Rundo, L. et al. clinically interpretable radiomics-based prediction of histopathologic response to neoadjuvant chemotherapy in high-grade serous ovarian carcinoma. *Front. Oncol.* **12**, 868265 (2022).
24. Hart, P. A. et al. Diabetes following acute pancreatitis. *Lancet Gastroenterol. Hepatol.* **6**, 668–675 (2021).
25. Yu, B. J. et al. Pancreatic necrosis and severity are independent risk factors for pancreatic endocrine insufficiency after acute pancreatitis: A long-term follow-up study. *World J. Gastroenterol.* **26**, 3260–3270 (2020).
26. Hou, N. et al. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: A machine learning approach using XGboost. *J. Transl. Med.* **18**, 462 (2020).

Author contributions

Conception and design: Xiyao Wan. Development of methodology: Xiyao Wan. Acquisition of data: Yuan Wang, Shuting Zhong. Analysis and interpretation of data: Ziyang Liu, Ziyi Liu. Writing, review, and/or revision of the manuscript: Xiyao Wan. Study supervision: Xiaohua Huang. All authors reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-86290-7>.

Correspondence and requests for materials should be addressed to X.H.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025