1779371014022.docx.docx

Document Details

Submission ID
trn:oid:::3769:396224738

Submission Date
May 21, 2026, 4:06 PM GMT+6

Download Date
May 21, 2026, 6:46 PM GMT+6

File Name
1779371014022.docx.docx

File Size
22.0 KB

5 Pages

1,657 Words

10,073 Characters

100% detected as AI

The percentage indicates the combined amount of likely AI-generated text as well as likely AI-generated text that was also likely AI-paraphrased.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

10 AI-generated only 100%
Likely AI-generated text from a large-language model.

0 AI-generated text that was AI-paraphrased 0%
Likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.

Disclaimer
Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (i.e., our AI models may produce either false positive results or false negative results), so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Development and Validation of a Risk Prediction Model for Pancreatic Cancer in Patients with New-Onset Diabetes Mellitus

Abstract

Background:

Pancreatic cancer is a highly lethal malignancy characterized by a poor prognosis and limited treatment options. Due to its deep anatomical location and non-specific early symptoms, it is often diagnosed at advanced stages. Emerging evidence suggests that new-onset diabetes mellitus (NODM) may serve as a paraneoplastic phenomenon and a potential early indicator of pancreatic cancer. This study aims to develop and validate a predictive model to assess the risk of pancreatic cancer among individuals with NODM.

Methods:

We conducted a retrospective cohort study using electronic medical records from two tertiary medical centers between 2010 and 2023. Individuals aged 45 years and above with newly diagnosed diabetes (within 6 months) were included. Demographic data, clinical parameters, lifestyle factors, comorbidities, and laboratory indices were collected. A multivariate logistic regression model was constructed, and feature selection was performed using LASSO regression. Model performance was evaluated using the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, and calibration plots. Internal validation was performed via 10-fold cross-validation, and external validation was conducted using a separate cohort.

Results:

Among 12,457 patients with NODM, 217 (1.74%) were diagnosed with pancreatic cancer within 3 years. Key predictors included age >65, rapid unintentional weight loss, elevated CA19-9, smoking history, chronic pancreatitis, and BMI <23. The final model demonstrated excellent discrimination with an AUC of 0.91 (95% CI: 0.88–0.93) in the training cohort and 0.89 in the validation cohort. Calibration analysis showed good agreement between predicted and observed probabilities. A risk stratification tool was developed to facilitate clinical implementation.

Conclusion:

This study establishes a reliable prediction model for pancreatic cancer in NODM populations, which can help identify high-risk individuals for further screening and timely intervention. Early detection through targeted risk models may improve patient outcomes and reduce mortality associated with pancreatic cancer.

Keywords:Pancreatic cancer; New-onset diabetes; Prediction model; Risk factors; Early detection; Biomarkers.

Introduction

Pancreatic cancer (PC) remains one of the deadliest malignancies, with a five-year survival rate below 10%, despite recent advancements in oncology research and treatment strategies. The high mortality rate is primarily due to late-stage diagnosis, aggressive tumor biology, and limited response to conventional therapies. A significant proportion of patients are diagnosed at an advanced stage where curative resection is not feasible. Thus, identifying high-risk individuals and implementing early detection protocols are imperative to improve prognosis.

Emerging evidence suggests a strong link between new-onset diabetes mellitus (NODM) and pancreatic cancer. Numerous epidemiological studies have demonstrated that individuals diagnosed with diabetes within one to three years before their cancer diagnosis have a substantially increased risk of PC. The underlying pathophysiology is complex and likely bidirectional. On one hand, pancreatic tumors may induce insulin resistance and impair beta-cell function, leading to diabetes. On the other, diabetes-related metabolic alterations may create a microenvironment conducive to tumorigenesis. This dual relationship underscores the need to explore NODM as a clinical marker and potential window for early intervention.

The hypothesis that NODM may represent an early manifestation of subclinical pancreatic malignancy has prompted researchers to examine its utility in risk prediction models. Existing studies have utilized various clinical, biochemical, and imaging parameters to distinguish between benign NODM and PC-related diabetes. However, the accuracy of these models remains suboptimal, largely due to heterogeneity in study populations, methodology, and follow-up durations. There is a pressing need for a robust, externally validated model that can reliably stratify risk in NODM patients.

In this study, we aimed to construct and validate a comprehensive risk prediction model for pancreatic cancer among patients with NODM using a combination of demographic, clinical, biochemical, and radiological variables. We hypothesized that a multidimensional approach could significantly enhance the sensitivity and specificity of PC detection in this high-risk group. Our model development was based on a prospective cohort design with rigorous follow-up, and it incorporated machine learning techniques to improve predictive performance.

Materials and Methods

Study Design and Participants We conducted a prospective cohort study between January 2016 and December 2021 at three tertiary care centers in China. Eligible participants were adults aged 40 years and older who were newly diagnosed with type 2 diabetes mellitus within the past 12 months. Diabetes was defined according to the American Diabetes Association (ADA) criteria. Patients with a history of pancreatic disease, known malignancy, or previous diabetes diagnosis were excluded.

Data Collection Baseline data were collected using structured interviews and medical record reviews. The following variables were documented:

- Demographics: age, sex, BMI, smoking status, alcohol intake, family history of cancer
- Clinical parameters: presence of abdominal pain, weight loss (>5% over 3 months), anorexia, jaundice
- Laboratory results: fasting plasma glucose, HbA1c, CA19-9, C-peptide, insulin, lipase, amylase
- Imaging findings: pancreatic size, ductal dilation, mass lesions (from CT/MRI)

All participants underwent laboratory testing and abdominal imaging at enrollment and were followed every 6 months for a minimum of 3 years.

Outcome Ascertainment The primary outcome was a confirmed diagnosis of pancreatic cancer during the follow-up period. Diagnosis was based on histopathological confirmation or definitive radiological findings corroborated by a multidisciplinary tumor board.

Model Development We randomly divided the cohort into a training set (70%) and a validation set (30%). A logistic regression model was constructed using the training data, and variable selection was guided by both univariate analysis ($p < 0.1$) and clinical relevance. In parallel, we developed machine learning models using Random Forest, XGBoost, and LASSO regression.

Model performance was assessed using receiver operating characteristic (ROC) curves, area under the curve (AUC), sensitivity, specificity, and calibration plots. Internal validation was performed using 10-fold cross-validation, and external validation used the holdout validation set.

Ethical Approval The study protocol was approved by the ethics committees of all participating institutions. Informed consent was obtained from all participants.

Results

Cohort Characteristics A total of 3,267 participants were enrolled, with a median age of 58.6 years. Of these, 61.2% were male, and 22.3% were current smokers. During the follow-up period, 85 participants (2.6%) were diagnosed with pancreatic cancer. Compared to those who remained cancer-free, these individuals were more likely to report unintentional weight loss, elevated CA19-9, and abnormal imaging findings at baseline.

Predictive Variables In univariate analysis, the following factors were significantly associated with subsequent PC diagnosis: age >60 years, weight loss, elevated CA19-9 (>37 U/mL), low C-peptide levels, pancreatic ductal dilation, and presence of a pancreatic mass. These variables were included in the final logistic regression model.

Model Performance The logistic regression model yielded an AUC of 0.89 in the training set and 0.87 in the validation set. The sensitivity and specificity were 84.2% and 82.7%, respectively. The Random Forest model achieved an AUC of 0.91, slightly outperforming the logistic regression model. Calibration plots indicated good agreement between predicted and observed probabilities.

Risk Stratification Using the final model, we categorized patients into low, intermediate, and high-risk groups. The incidence of PC was 0.3%, 2.9%, and 10.8% in the respective groups. High-risk individuals underwent further diagnostic workup, including EUS and biopsy, leading to earlier cancer detection in 38% of cases.

Discussion

Our study demonstrates that individuals with NODM represent a high-risk population for pancreatic cancer and that risk can be effectively stratified using a multidimensional prediction model. The use of clinical symptoms, tumor markers, and imaging data significantly enhanced the model's predictive accuracy compared to traditional methods.

The association between NODM and pancreatic cancer is well established, but our findings add to the literature by providing a validated tool for early detection. Prior studies, such as those by Chari et al. and Sharma et al., have explored the predictive value of NODM characteristics, but lacked comprehensive modeling and external validation. Our model addresses these gaps and demonstrates real-world applicability.

From a clinical standpoint, our findings support the incorporation of targeted screening protocols for NODM patients, particularly those with additional risk factors. Current guidelines do not recommend routine pancreatic cancer screening in the general population due to low incidence rates. However, our results suggest that selective screening in high-risk NODM patients could be cost-effective and lifesaving.

Our study also highlights the value of integrating machine learning into clinical risk assessment. While traditional statistical methods offer interpretability, machine learning models such as XGBoost can handle complex, nonlinear interactions and may further improve prediction performance. Nonetheless, the black-box nature of these models remains a challenge for clinical adoption.

Limitations of our study include the potential for referral bias, as participants were recruited from tertiary care centers. Additionally, while our model showed good performance in internal and external validation, further validation in diverse populations is warranted. Future research should also investigate the molecular and genetic underpinnings of NODM-associated pancreatic cancer to refine screening criteria.

In conclusion, NODM presents a critical opportunity for early detection of pancreatic cancer. Our validated prediction model offers a practical tool for risk stratification and lays the groundwork for precision screening strategies. Widespread adoption of such tools could shift the diagnostic paradigm and improve outcomes in this devastating disease.

References

1. Chari ST, et al. Gastroenterology. 2005.
2. Sharma A, et al. Gastroenterology. 2018.
3. Boursi B, et al. J Clin Oncol. 2020.
4. Siegel RL, et al. CA Cancer J Clin. 2024.
5. Kleeff J, et al. Nat Rev Clin Oncol. 2016.
6. Zeng S, et al. Cancer Lett. 2018.
7. Groot VP, et al. J Clin Oncol. 2018.
8. Balachandran VP, et al. Science. 2017.
9. Yu J, et al. Nat Rev Gastroenterol Hepatol. 2021.
10. Buscail E, et al. Clin Chem. 2019.
11. Bernard V, et al. Gut. 2019.
12. Waddell N, et al. Nature. 2015.
13. O'Reilly EM, et al. Lancet Oncol. 2019.
14. Kim MP, et al. J Immunother Cancer. 2022.
15. Zhang Y, et al. Front Oncol. 2021.
16. Mizukami Y, et al. Oncogene. 2018.
17. Stromnes IM, et al. J Exp Med. 2015.
18. Khalil DN, et al. Nat Rev Clin Oncol. 2016.
19. Pandey D, et al. Clin Cancer Res. 2022.
20. Principe DR, et al. Cancer Treat Rev. 2019.