

# 从传统评分到人工智能：胰十二指肠切除术后胰瘘预测模型的演进与展望

许志仁, 杨仕凡, 姚博, 朱盟, 李福宏, 王宇骁, 苏琨, 王连敏, 王滔, 杨夏威, 吴涛\*

昆明医科大学第二附属医院肝胆胰外科四病区, 云南 昆明

收稿日期: 2026年2月11日; 录用日期: 2026年3月4日; 发布日期: 2026年3月13日

## 摘要

胰十二指肠切除术(PD)是治疗胰头及壶腹周围肿瘤的主要手段, 尽管围术期死亡率有所下降, 但术后胰瘘(POPF)仍然是一个严重的并发症, 临床相关性胰瘘(CR-POPF)可能导致高病死率以及医疗负担增加。本文详细回顾了PD术后胰瘘预测模型的发展历程, 包括各个模型的具体改进和应用实例。从早期的传统评分系统, 比如原始FRS、a-FRS和ua-FRS, 到近年来应用的机器学习与深度学习模型, 预测准确性逐渐提升。传统模型虽然简单实用, 但在外部验证中常出现和报告结果不符的性能下降。随机森林、XGBoost、CatBoost、神经网络和深度学习等人工智能方法, 结合了多方面临床变量、术后引流数据以及放射组学特征, 内部验证的AUROC可超过0.80。部分联合模型在术前预测以及围术期管理中表现出优于传统FRS的潜力。然而, 现在的研究仍以回顾性、单/少中心为主, 外部验证数据不足, 报告透明度以及泛化能力还较低。未来需开展大规模多中心前瞻性研究, 推动模型融合、发展可解释AI、以及开发区域性模型。以实现从静态评分向动态、个性化精准风险分层工具的转变, 最终优化围术期管理、降低CR-POPF相关严重结局。

## 关键词

胰十二指肠切除术, 预测模型, 术后胰瘘, 机器学习, 深度学习, 放射组学

## From Traditional Scoring to Artificial Intelligence: Evolution and Prospects of Prediction Models for Postoperative Pancreatic Fistula after Pancreaticoduodenectomy

\*通讯作者。

文章引用: 许志仁, 杨仕凡, 姚博, 朱盟, 李福宏, 王宇骁, 苏琨, 王连敏, 王滔, 杨夏威, 吴涛. 从传统评分到人工智能: 胰十二指肠切除术后胰瘘预测模型的演进与展望[J]. 临床医学进展, 2026, 16(3): 2153-2162.

DOI: 10.12677/acm.2026.1631008

Zhiren Xu, Shifan Yang, Bo Yao, Meng Zhu, Fuhong Li, Yuxiao Wang, Kun Su, Lianmin Wang, Tao Wang, Xiawei Yang, Tao Wu\*

The Fourth Ward of Hepatobiliary and Pancreatic Surgery Department, The Second Affiliated Hospital of Kunming Medical University, Kunming Yunnan

Received: February 11, 2026; accepted: March 4, 2026; published: March 13, 2026

## Abstract

Pancreaticoduodenectomy (PD) is the primary therapeutic modality for periampullary and pancreatic head neoplasms. Although the perioperative mortality rate has declined, postoperative pancreatic fistula (POPF) remains a severe complication, and clinically relevant pancreatic fistula (CR-POPF) may lead to a high mortality rate and increased medical burden. This paper systematically reviews the developmental history of predictive models for POPF after PD. From early traditional scoring systems (e.g., the original Fistula Risk Score [FRS], adjusted FRS [a-FRS], and unadjusted FRS [ua-FRS]) to the machine learning and deep learning models widely applied in recent years, the predictive accuracy has been gradually improved. Despite their simplicity and practicability, traditional models often exhibit performance degradation in external validation. After integrating multi-dimensional clinical variables, postoperative dynamic drainage data and radiomic features, artificial intelligence methods (including random forest, XGBoost, CatBoost, neural networks and deep learning) have achieved an area under the receiver operating characteristic curve (AUROC) of over 0.80 in internal validation; some combined models outperform the traditional FRS and have demonstrated potential in preoperative prediction and perioperative management. However, current studies are still dominated by retrospective, single or small-sample multicenter research, with insufficient external validation data, low reporting transparency and poor generalization ability. In the future, large-scale multicenter prospective studies need to be conducted to promote model integration, develop explainable artificial intelligence, and establish regional predictive models. The ultimate goal is to realize the transformation from static scoring systems to dynamic and individualized precise risk stratification tools, thereby optimizing perioperative management and reducing CR-POPF-related severe outcomes.

## Keywords

Pancreaticoduodenectomy, Prediction Model, Postoperative Pancreatic Fistula, Machine Learning, Deep Learning, Radiomics

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

胰十二指肠切除术(pancreaticoduodenectomy, PD)是普外科里最经典的手术之一,目前仍是治疗胰头癌以及壶腹周围肿瘤唯一可能根治的办法。该手术的核心切除范围包括胰头、壶腹部、壶腹周围十二指肠以及胆总管远端[1]。近年来,随着外科手术技术的不断精进、围手术期综合管理策略的持续优化,以及加速康复外科(enhanced recovery after surgery, ERAS)理念的广泛推广,PD的临床适应证范围明显拓展,基本成为胰腺以及壶腹周围疾病治疗的核心术式。虽然在高容量中心,胰腺切除术后的围手术期死亡率

已经降至 5% 以下，但术后总体并发症发生率仍较高，通常在 30%~65% 之间[2]。其中，术后胰瘘 (postoperative pancreatic fistula, POPF) 是胰腺外科中最严重的并发症之一，其特征为胰液从手术吻合口或切除创面漏出。POPF 常导致患者住院时间延长、相关并发症发生率升高(如胃排空延迟、腹腔脓肿、术后胰腺出血、败血症等)，增加介入或再次手术风险，并推高医疗成本[3]。尤其在临床相关性胰瘘 (clinically relevant POPF, ISGPS B/C 级) 病例中，患者病死率和严重并发症风险进一步上升。这些持续存在的难题表明在现代胰腺外科中，医生需要开发更精准的风险分层工具、早期预测模型以及有效的预防策略，以减轻 POPF 对患者预后的负面影响。

因此，本文旨在梳理 PD 后胰瘘预测模型的发展历程，并总结机器学习在该领域临床应用中的最新进展。在此基础上，进一步分析当前面临的关键挑战与主要局限，展望未来的发展方向，希望能给胰腺外科医生提供一些实用的思路，帮助他们在围术期更精准、更安全地管理风险，实现真正个性化的临床决策。

## 2. 概述

### 2.1. 临床风险预测模型

临床预测模型 (Clinical Prediction Model) 是一种用于风险分层与医疗资源优化配置的实用工具，其核心价值在于能够识别并优先服务那些最能从中获益的患者[4]。临床风险预测模型常可以分为：诊断模型 (Diagnostic) 和预后/风险模型 (Prognostic)。POP 预测模型通过准确地筛选出 POPF 高风险的患者，从而指导临床团队根据患者的风险特征，制定针对性更强的围手术期管理策略。对高风险患者，外科医生应实施更强的监测、更积极的预防干预。而对低风险患者则可安全推进加速康复路径，比如尽早拔除腹腔引流管、缩短住院时间、减少不必要的侵入性操作。这样做既能省下医疗资源，又能让患者恢复得更舒服，并且体验更好。

### 2.2. 机器学习

机器学习 (machine learning, ML) 是人工智能 (Artificial Intelligence, AI) 的一个核心分支，其核心在于通过从大规模、高维度的训练数据中自动学习隐含模式和规律，能够构建高效的分类、回归、预测或估计模型，而无需依赖人工显式编程规则或先验假设[5]。机器学习通常分为三大类型：监督学习 (利用带标签数据训练模型，如逻辑回归、随机森林和深度神经网络，用于临床风险预测等任务)；无监督学习 (从无标签数据中发现内在结构，如聚类和降维)；以及强化学习 (通过与环境的交互和奖励/惩罚信号优化决策序列)[6]。近年来，机器学习技术已在医学领域中展现出很大潜力。它被广泛应用于放射影像的自动诊断、组织病理学特征筛选、临床疾病诊断、良恶性鉴别以及预后评估等相关领域[7]-[9]。在临床预测模型领域中，虽然其黑箱性质和泛化挑战仍需通过可解释性技术和严格验证来解决，但是机器学习已超越传统统计方法，成为提升预测准确性的关键工具[10]。

### 2.3. 十二指肠切除术后胰漏的定义及诊断标准

根据 2016 年国际胰腺外科研究组 (the International Study Group of Pancreatic Fistula, ISGPF) 更新的胰漏共识指南[11]，术后胰漏具体分级定义如下：A 级胰瘘 (生化漏, biochemical leak) 其诊断标准为：术后第 3 天以及以后，腹腔引流液淀粉酶浓度 > 正常血清淀粉酶上限的 3 倍，但无临床症状、无需改变原定治疗方案以及未引发不良结局，故不归为临床相关性胰瘘。临床相关性胰瘘 (clinically relevant postoperative pancreatic fistula, CR-POPF) 定义为：腹腔引流液淀粉酶浓度高于本机构血清淀粉酶正常上限 3 倍及以上 (无论引流量多少)，且伴随与胰瘘直接相关的临床表现或并发症。CR-POPF 进一步细分为 B 级与 C 级：

**B 级胰瘘：**需同时满足 CR-POPF 诊断标准，并具备以下至少 1 项特征：(1) 术后经皮引流管持续留置超过 21 天；(2) 需在影像引导下重新调整或放置引流管，以引流未充分引流的腹腔积液；(3) 发生与胰瘘相关的出血或假性动脉瘤，需接受输血和(或)血管造影干预；(4) 出现与胰瘘相关的器官间隙轻度感染，仅需抗生素治疗且未发生器官衰竭。**C 级胰瘘：**为最严重类型，指胰瘘进展至需再次手术干预，或引发单器官/多器官功能衰竭，甚至导致患者死亡的严重情形。

### 3. 传统胰漏预测评分系统

过去十年间，已有超过六十种针对胰十二指肠切除术后胰瘘(POPF)的预测模型相继产生。这些模型在临床应用中均展现出一定的参考价值，为术后胰瘘的风险评估提供了实用工具。其中最早的预测模型源于 Gaujoux S 等人开展的一项单中心前瞻性研究，这项研究纳入了 100 例接受胰十二指肠切除术的患者，旨在识别术后发生 PF 的危险因素[12]。结果显示，BMI  $\geq 25$  kg/m<sup>2</sup>、胰腺脂肪浸润以及无胰腺纤维化是 PF 的独立预测因素。综合这三个因素构建的评分系统能够对患者发生 PF 风险预测范围包含从最低的 7% (无危险因素)到最高的 78% (具备全部三个危险因素)。尤其对于有临床意义的 B 级或 C 级胰瘘，这个评分系统的预测准确率更高(风险范围 0%~81%)，相较于仅依赖胰腺质地(软/硬)的传统评估方法更为客观，有助于临床制定个体化的预防策略。然而，这项研究仅纳入了 100 例连续患者，虽采用前瞻性设计，但是样本量仍显不足。特别是在按危险因素进行分层后，有些亚组(如同时具备三个危险因素的亚组仅含 18 例患者)的样本量较少，统计效力有限。

#### 3.1. 原始瘘管风险评分(FRS)

尽管早期预测模型存在一定局限性，但随着相关研究的不断深入，新一代预测模型陆续被提出。这些模型纳入了更全面的变量体系，并且提升了预测的精确性与临床实用性。例如，2013 年 Callery MP 等人基于美国三家医学中心 445 例接受胰十二指肠切除术患者的术前和术中数据，通过相关分析，筛选出四项独立危险因素：胰腺质地软、特定病理类型、主胰管直径  $\leq 3$  mm，以及术中出血量  $> 1000$  mL。他们根据这四项危险因素构建了 10 分制的胰瘘风险评分(FRS) [13]。该评分将患者划分为四个风险等级：极低风险(0 分)、低风险(1~2 分)、中风险(3~6 分)和高风险(7~10 分)。与之对应的临床相关术后胰瘘(CR-POPF)的预测发生率分别为 0%、6%、22%和 88%，受试者工作特征曲线下面积(AUROC)达到 0.942。

#### 3.2. 替代瘘管风险评分(a-FRS)

由于 FRS 将术中出血量作为预测因子之一，但近年两项外部验证研究显示该因子并不具有明显的预测意义，且存在记录不全、估计偏差、易受手术质量影响等实际问题。此外，随着 2016 年国际胰腺外科研究组(ISGPS)更新了术后胰瘘(POPF)的定义，临床需要一种既能兼容 2005 年与 2016 年定义、又具备易获取变量和稳定预测效能的 POPF 预测模型。为此，Mungroop TH 等人于 2019 年开发并验证了替代瘘管风险评分(a-FRS) [14]。该评分仅基于胰腺质地(软/非软)、主胰管直径( $\leq 5$  mm)和体重指数(BMI)这三个易于获取的变量构建。在包含 1924 例患者的内部验证队列中，其 AUC 为 0.75；在涵盖 926 例患者的国际外部验证队列中，针对 ISGPS 2005 版定义的 POPF，AUC 达 0.78，针对 2016 版定义则为 0.72。该模型的预测效能不逊于原始 FRS (2005 版定义 AUC 0.75，2016 版定义 AUC 0.70)，且能够成功将患者划分为低(0%~5%)、中( $> 5\% \sim 20\%$ )、高( $\geq 20\%$ )三个风险层级，能够有效指导胰十二指肠切除术后 POPF 的风险分层与临床决策。

#### 3.3. 更新的替代瘘管风险评分(ua-FRS)

随着时间推移以及腹腔镜技术的推广，微创胰十二指肠切除术(MIPD)因为创伤小，在术后恢复较快

等方面的优势,越来越多外科医生选择该术式。**a-FRS** 基于胰腺质地、主胰管直径和体重指数这三个变量,虽在开放胰十二指肠切除术(OPD)中得到验证,但未在 MIPD 中进行专门评估。并且原版 **a-FRS** 在 MIPD 患者中的预测效能较差。为此, Mungroop TH 等人对 **a-FRS** 进行了优化,开发了适用于 MIPD 的更新版评分(**ua-FRS**) [15]。该评分在保留了原 **a-FRS** 三项变量的基础上,新增男性性别这一风险因素。他们通过涵盖 7 个国家、26 个中心的 952 例 MIPD 患者(包括腹腔镜、机器人辅助以及混合式三种术式)进行了泛欧洲验证。结果显示, **ua-FRS** 的预测效能有所提升, AUC 达到 0.75(原 **a-FRS** 为 0.68),且校准度良好;同时,该模型在开放胰十二指肠切除术(OPD)中也验证有效。

在临床应用中, **FRS**、替代-**FRS** 以及更新替代-**FRS** 均为广泛使用的预测工具。然而, Pande 等人的系统综述表明,许多模型在外部验证中的表现与其原始开发研究存在较大差异[16]。该研究对 2020 年以前发表的、针对胰腺术后胰瘘(POPF)的所有预测模型进行了外部验证汇总分析。值得注意的是,虽然 **FRS** 在原始研究中的开发中显示 AUROC 高达 0.94,但在纳入的 19 项外部验证研究中,其综合 AUROC 仅为 0.71。这些模型中表现最优的 **ua-FRS** 也仅为 0.72,且两者差异无统计学意义( $p > 0.05$ )。这一差异反映了预测模型在开发阶段可能存在的“过拟合”现象。简单说,由于选择的人群有较高同质性或变量选择偏倚等原因,原始研究可能高估了模型的预测效能。而外部验证的异质性人群更加符合真实临床环境,因而更能客观的评估模型的预测能力。

#### 4. 人工智能、机器学习在胰漏预测中的应用

由于传统预测模型通常仅纳入 3~6 个变量,且依赖线性假设,难以充分考虑风险因素间的复杂非线性影响。近年来,人工智能(AI)和机器学习(ML)方法逐渐成为解决这一问题的关键工具。他们通过自动学习大规模临床数据的相关规律,在提升预测准确性能力方面展现出较大优势。Han IW 等人首次采用 AI 的方法,整合术前以及术中变量,开发了高效的 PD 术后 POPF 风险预测模型[17]。该研究回顾性分析了三星医疗中心在 2007~2016 年时间内的 1769 例 PD 患者的 38 项临床变量,分别运用随机森林(RF)、神经网络(NN)以及结合递归特征消除(RFE)的神经网络(NN + RFE)算法构建预测模型。其中,成功筛选出 16 项与 POPF 发生密切相关的关键风险因素的 NN + RFE 模型表现最出色, AUC 达到 0.74。并且该研究团队还进一步开发了一个免费的网络预测平台,为临床医生筛选高危患者、制定个体化防治策略提供了可靠的支持。

##### 4.1. 早期 ML

Lin Z 等人[18]针对 PD 术后 CR-POPF 的术前精准预测,开展了基于增强 CT 的放射组学研究。该回顾性研究纳入了北京大学第一医院 2013 年 4 月至 2019 年 12 月期间 250 例接受 PD 的患者,这些患者在术前 30 天内完成了增强 CT。研究团队按 7:3 比例,将患者随机分为训练集(175 例)和验证集(75 例)。研究分别构建了纯放射组学模型和整合放射组学特征、人口统计学变量以及常规影像学特征的联合模型。联合模型最终纳入 18 项特征,训练集/验证集 AUC 分别提升至 0.871 和 0.869。与 **FRS** 相比,该联合模型在预测性能上优越,在校准曲线和决策曲线分析(DCA)中均显示出更好的拟合度和临床净获益。该单中心回顾性研究表明,该联合模型可实现术前较为准确的 CR-POPF 风险分层,为优化围术期管理、制定个体化干预策略提供重要参考价值。

Shen ZY 等人[19]基于上海交通大学医学院瑞金医院 2010~2021 年间连续 2421 例 PD 患者的回顾性数据,开发并比较了四种机器学习算法(CatBoost、lightGBM、XGBoost、Random Forest)。该研究整合 62 项围术期临床变量,采用 5-fold 交叉验证,结果显示 CatBoost 模型性能最优,训练集平均 AUC 达 0.81 (95%CI 0.80~0.82),验证集 AUC 为 0.83。SHAP 值分析表明,术后前 7 天平均引流液淀粉酶(mean DFA)、

末次 DFA 以及 DFA 变化趋势是最重要的预测因子,其次为引流量以及炎症指标。而在纯术前以及术中变量模型中,性能出现下降(CatBoost AUC 仅 0.64),提示术后动态监测对准确预测的决定性作用。与传统 FRS 等评分系统相比,该 ML 框架在纳入术后数据的条件下提升了预测效能,并可通过风险计算支持个体化引流管拔除决策。尽管为单中心回顾性研究且缺乏外部验证,该工作仍为机器学习在 PD 术后胰瘘动态风险评估以及精准引流管理中的应用提供了重要证据。

Shi Y 等人[20]基于中国四家三级医院(沈阳盛京医院、辽宁肿瘤医院、天津肿瘤医院、广东省人民医院) 2009~2019 年间连续 990 例 PD 患者的回顾性多中心数据,开发并验证了一个整合增强 CT 特征与传统 FRS 元素的改良瘘风险评分模型(CT-FRS)。该模型通过 LASSO 回归从 26 项术前以及术中因素中筛选出 5 项核心预测变量:残余胰腺体积(RPV)、胰腺截面积、CT 脂肪评分、CT 萎缩评分以及主胰管直径。在预测集和独立外部验证集中,CT-FRS 的 C-index 分别达 0.825 和 0.807,明显优于原 FRS (0.794 和 0.741,  $p = 0.04$  和  $0.05$ )。尤其在中危人群(FRS 3~6)中,预测效能提升最为明显。进一步分析显示,CT-FRS 预测概率与胰腺残端组织学特征高度相关:纤维化和腺泡丰富度负相关、脂肪变性正相关(均  $p < 0.001$ )。该多中心外部验证研究证实,CT-FRS 作为术前无创风险评估工具,可以改善中危患者的 CR-POPF 预测准确性。与早期以单中心或小样本为主的 ML 探索相比,2023 年后多项大样本、多中心研究开始采用更先进的集成学习与深度学习框架,进一步提升了纯临床变量模型的预测性能,并为术前无创预测提供了新路径。

## 4.2. 近期纯临床 ML 及深度学习

Verma 等人[21]基于美国国家外科质量改进计划数据库的大样本多中心数据,开发了一种新型机器学习模型,用于围术期预测 PD 后 CR-POPF。该研究采用先进的集成学习算法,纳入易获取的临床变量,构建预测模型,并在内部验证以及独立外部验证队列中与 mFRS 进行头对头比较。结果显示,该机器学习模型在判别效能上优于 mFRS,表现出更强的预测准确性。该工作突出了机器学习在处理复杂非线性关系和复杂临床数据方面的优势,为超越传统线性评分系统提供了有力证据。尽管仍需进一步前瞻性验证,但该模型在大样本外部验证中的稳健表现,使其成为近年来纯临床变量机器学习模型的代表性进展。

Lee 等人[22]基于韩国高容量中心回顾性队列,开发了深度学习模型用于预测 PD 后 CR-POPF。该研究分别构建了仅依赖术前变量的预测模型以及整合围术期数据的模型。结果显示,深度学习架构在总体 POPF 和 CR-POPF 预测中均表现出优异性能,其中围术期模型判别能力进一步提升。该工作突显了深度学习在整合多维度临床数据、实现精准术前风险分层方面的潜力,为超越既往风险评分系统提供了新方法。尽管为单中心回顾性设计,其仍标志着人工智能从机器学习向深度学习演进在胰腺外科预测领域的最新进展。在深度学习推动围术期预测效能提升的同时,放射组学作为术前影像高通量分析的核心技术,正逐步与临床风险评分深度融合,以弥补传统因素(如胰腺质地、主胰管直径)的主观性和局限性。近期代表性工作进一步证实了这一方向的潜力。

Choubey AP 等人[23]在多中心回顾性队列中,系统评估了放射组学特征对临床风险评分系统预测 PD 术后 CR-POPF 的价值。该研究构建了纯放射组学模型、纯临床模型以及二者融合的联合模型。结果显示纯放射组学模型的预测性能与现有临床风险评分相当,而联合模型表现出最强的判别能力和临床净获益。该工作表明,术前 CT 高通量放射组学特征可有效补充传统临床变量的不足,捕获更多的形态学与预测信息,能够提升整体预测准确性。其标志着放射组学与临床评分融合已成为优于单一模态预测的重要方向。

## 4.3. 算法特征选择与优缺点

在特征选择方面,不同的 AI 模型筛选使用的核心预测变量具有高度相似一致性。这也间接从另一个

方面印证了 POPF 发生的关键病理生理机制是多因素、动态过程，而非单纯的“吻合口机械性漏出”。其中反复出现的临床指标主要包括主胰管直径、胰腺实质质地、BMI、术前血清白蛋白水平、术中出血量以及术后早期引流液淀粉酶的动态变化等。包括胰腺质地柔软、胰管直径细、胰腺肿物高危病理类型等在内的胰腺固有因素是公认的 POPF 不良预后因素[24]。而术后胰痿与高 BMI 呈正相关[25]，主要因为高 BMI 往往导致胰腺脂肪浸润增加，从而使胰腺质地更“软”、更“嫩”。由日本研究团队完成的另一项单中心回顾性研究中，他们探讨了术前 C 反应蛋白 - 白蛋白比值(CAR)是否能预测术后 POPF 的风险，结果提示：POPF 组术前 CAR 显著高于非 POPF 组，单因素分析：CAR  $\geq 0.05$  与 POPF 显著相关且多因素 logistic 回归证实 CAR  $\geq 0.05$  是 POPF 的独立危险因素[26]。术中出血量和术后早期引流液淀粉酶分别从术中损伤程度和术后胰腺残端渗漏活性两个方面反映了 POPF 的核心病理生理过程。在整合放射组学的模型中，影像学特征如胰腺残余体积(RPV)、胰腺截面积、CT 脂肪评分、CT 萎缩评分以及胰腺段的形态纹理特征被反复确认为重要贡献者[27]。这些特征与胰腺纤维化减少、脂肪变性增加及腺泡丰富度等组织学变化高度相关。它们构成了 POPF 的核心生物学标记。

不同算法在处理小样本和不平衡数据时的表现各不相同。像 XGBoost、CatBoost、随机森林这类树模型，它们自身自带、无需额外编程实现特征分析功能，即模型训练完成后，这类模型能自动计算并判断“哪些特征对预测结果更重要”，不用研究者再单独搭建分析流程。他们通常在中等样本量下表现出色，且不易过拟合。集成方法能在 SHAP 分析中能清晰揭示非线性贡献，常优于单一神经网络。而神经网络和深度学习模型在大数据量时潜力更大，但对小样本易过拟合，需要通过 Lasso 正则化来缓解。总体而言，XGBoost 和 CatBoost 在多数回顾性研究中显示出最佳平衡的判别性能和泛化潜力[28]。

尽管上述放射组学融合模型在内部验证中已展现出超越单一模型的判别能力与临床净获益，值得注意的是，与早期 ML 工作类似，这些进展大多仍基于回顾性队列，且外部验证覆盖面有限。为克服这些局限以及方法学质量参差不齐等问题，未来有待开展大规模、多中心、前瞻性研究，纳入更多人群与手术方式，以实现对这些模型的严格外部验证、优化以及临床推广，最终为 PD 术后胰痿的精准风险分层与个性化管理提供更可靠的证据基础。

## 5. 现状与限制

近年来，尽管多项预测模型在内部验证中展现出较好的判别效能，但 POPF 预测的准确性仍受限于多种风险因素的复杂交互、手术变异性以及病理生理机制的缺失。目前，POPF 预测模型的稳健性和临床推广仍受到许多关键因素的制约，这些因素在国际范围内尚未形成充分共识，主要体现在以下方面：诊断阈值的设定缺乏统一标准、预测变量的评估普遍未采用盲法设计，易引入系统性偏倚以及统计建模方法种类繁多且缺乏标准化，导致模型间结果难以进行可靠比较。因此，现有模型虽然在特定队列中表现优异，但整体预测效能仍有较大改进空间，需要通过标准化报告、严谨外部验证以及前瞻性多中心研究进一步改善。一项针对 PD 后胰痿预测模型报告质量的系统综述[29]采用 TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis)声明，对纳入的 52 个多变量预测模型进行了系统评估。结果显示，这些模型的整体平均 TRIPOD adherence 率为 65%，虽然高于其他领域已发表预测模型，但仍远未达到透明报告标准。常见报告缺陷包括模型细节、缺失值处理、过拟合评估、外部验证策略以及临床适用性讨论的缺失。该综述进一步筛选出 13 个兼具较高 TRIPOD adherence 率与优异预测性能的模型，建议作为临床实践的优先参考。该研究表明现有 PD 后 POPF 预测模型在报告透明度上的不足，强调严格遵循 TRIPOD 指南可以提升模型的可重复性、外部验证潜力以及临床转化价值。

另一个常常被多数预测模型忽略的重要因素是个体外科医生的经验与技能水平对 POPF 发生的影响。一项针对高瘘风险 PD 术后患者结局的研究[30]证实，外科医生经验是改善高 FRS (Fistula Risk Score 7~10

分)亚组结果的关键因素。该回顾性分析显示,高经验外科医生(累积 PD 例数  $\geq 50\sim 100$  例)在高危患者中降低 CR-POPF 发生率、严重并发症比例、失血量以及手术时间与低经验组形成鲜明对比。尽管肿瘤学指标(如切缘阴性率、淋巴结清扫数)无明显差异,但高经验组通过精细吻合技术、选择性预防措施以及更好处理并发症的能力,有效缓解了软胰腺、细主胰管等高危因素带来的风险。该研究强调,即使在高容量中心,个体外科医生经验也独立于医院总量,对高风险 PD 结局具有决定性影响。该发现提示外科医生经验可作为“人为优化”环节,进一步提升预测模型指导下的临床决策能力。在 2021 年 6 月发表的由韩国首尔国立大学医院的 Sungho Kim 等作者完成的一项针对单个外科医生开展纯 LPD 学习曲线的回顾性研究提示:LPD 手术失败率在第 61~83 例的过渡期,失败率才开始下降;而在第 84~119 例的熟练期,失败率显著降低并稳定[31]。若预测模型系统性地忽略这些外科医生层面的特异性,则会导致模型在外部验证中表现不稳定。因此,将外科医生经验作为独立变量或分层因素纳入模型开发将有助于进一步提升 POPF 预测的准确性与实用价值。

此外,地理区域患者的异质性、不同疾病特征以及不同医院围手术期管理的地域差异,均可导致 POPF 发生率以及关键风险因素的明显差异。一项针对美国深南部人群的外部验证研究[32]对 original FRS、a-FRS、ua-FRS、m-FRS 这四个瘘风险评分模型在 PD 术后不同种族 CR-POPF 预测中的适用性进行了评估。该单中心回顾性队列聚焦种族多样化特征的深南部患者。结果显示,种族差异是明显影响预测效能:在黑人亚组中,模型判别力下降、风险低估倾向明显,可能源于胰腺质地和 BMI 等种族特异性因素。该研究强调,西方主导的 FRS 系列模型在种族多样化地区应用时存在潜在偏倚。该结果提示需通过种族分层验证、纳入种族变量或开发本土化模型来提升预测的公平性和准确性。且另一项基于韩国全国多中心队列的外部验证研究[33]评估了西方开发的术后胰瘘评分系统在东方人群中的适用性。该回顾性分析纳入九所医院的连续 PD 患者,结果显示西方模型(如 FRS 和其改良版本)在韩国队列中的判别效能降低,尤其在中高危风险分层中表现欠佳。西方主导的 POPF 预测模型难以直接用至东方患者队列,需要开发并外部验证东方特异性模型,以提升预测准确性和临床适用性。该发现与美国深南部种族差异验证结果相呼应,共同揭示了地域、种族等人群异质性对现有模型预测能力的限制。因此,未来应优先开发并验证地域/种族特异性预测模型,以更准确地分层个体 CR-POPF 风险,并指导精准、个性化的围术期干预。该方向与 TRIPOD 报告质量提升和外科医生经验纳入相辅相成,有希望最终改善预测模型与真实临床实践的差距,实现更高效的胰腺外科风险管理。

## 6. 展望与未来方向

尽管近年来从传统线性评分到机器学习、深度学习、放射组学融合的预测模型在内部验证中已明显提升预测能力,并在术前影像分析和动态监测等方面展现出潜力。但回顾性设计、外部验证缺乏、方法学质量差异、地域和种族偏倚以及忽略外科医生经验等相关局限,仍制约预测模型在临床中的实际使用。要解决这些缺陷,让预测工具从“看起来准”变成“临床真能用”,未来可以重点往几个方向努力:首先,开展大规模、多中心、前瞻性研究,纳入多样化人群,以严格外部验证现有模型,并评估其在真实临床环境下的预测性能与临床净获益。

其次,继续推进多模型整合,把术前临床信息、CT/MRI 放射组学特征、术后早期引流数据甚至病理特征结合起来,甚至推动复杂 AI 与深度学习模型自动与现有电子病历系统或手术规划软件结合,以实现从实时更新的病历资料中动态的评估相关风险,实现类似于医院危急值制度的临床预警提醒机制,将模型部署为实时风险计算工具。发展可解释 AI 技术(如 SHAP、LIME 扩展)来提升模型的可信度,这在临床医患沟通中具有重要实际价值。SHAP 可直观展示每个临床数据对个体预测的正/负贡献幅度,帮助外科医生向患者解释“为什么你是高风险”或“哪些因素可通过干预改善降低并发症风险”,增强医患信任、改善医疗环境。术

前精准分层与围术期实时风险管控的更新，以期实现从静态评分向动态、个性化决策工具的转变。

第三，针对低风险患者专门建模，支持更早、更安全地拔管，配合 ERAS 路径减少不必要的干预，并降低医疗资源浪费。

第四，提高研究质量和报告规范，严格遵循 TRIPOD-AI 预测模型指南，主动评估种族、地域偏倚，最好把外科医生的经验作为变量或分层因素加进去，这样才能更真实地反映临床异质性。

最终，通过上述努力，有希望构建普适或区域适应性预测平台，为胰腺外科医生提供可靠的决策支持工具，让胰腺外科医生在术前就能更准确地判断风险，在围术期更有针对性地干预，最终降低 CR-POPF 的严重并发症发生率，缩短住院时间，也让医疗资源用得更合理。

## 参考文献

- [1] Xu, H., Bretthauer, M., Fang, F., Ye, W., Yin, L. and Adami, H. (2024) Dramatic Improvements in Outcome Following Pancreatoduodenectomy for Pancreatic and Periapillary Cancers. *British Journal of Cancer*, **131**, 747-754. <https://doi.org/10.1038/s41416-024-02757-w>
- [2] Giuliani, T., Marchegiani, G., Di Gioia, A., Amadori, B., Perri, G., Salvia, R., *et al.* (2022) Patterns of Mortality after Pancreatoduodenectomy: A Root Cause, Day-to-Day Analysis. *Surgery*, **172**, 329-335. <https://doi.org/10.1016/j.surg.2022.01.005>
- [3] McMillan, M.T., Christein, J.D., Callery, M.P., Behrman, S.W., Drebin, J.A., Hollis, R.H., *et al.* (2016) Comparing the Burden of Pancreatic Fistulas after Pancreatoduodenectomy and Distal Pancreatectomy. *Surgery*, **159**, 1013-1022. <https://doi.org/10.1016/j.surg.2015.10.028>
- [4] Damen, J.A.A.G., Hooft, L., Schuit, E., Debray, T.P.A., Collins, G.S., Tzoulaki, I., *et al.* (2016) Prediction Models for Cardiovascular Disease Risk in the General Population: Systematic Review. *BMJ*, **353**, i2416. <https://doi.org/10.1136/bmj.i2416>
- [5] Cruz, J.A. and Wishart, D.S. (2006) Applications of Machine Learning in Cancer Prediction and Prognosis. *Cancer Informatics*, **2**, 59-77. <https://doi.org/10.1177/117693510600200030>
- [6] Sarker, I.H. (2021) Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, **2**, Article No. 160. <https://doi.org/10.1007/s42979-021-00592-x>
- [7] Kourou, K., Exarchos, T.P., Exarchos, K.P., Karamouzis, M.V. and Fotiadis, D.I. (2015) Machine Learning Applications in Cancer Prognosis and Prediction. *Computational and Structural Biotechnology Journal*, **13**, 8-17. <https://doi.org/10.1016/j.csbj.2014.11.005>
- [8] Komura, D. and Ishikawa, S. (2018) Machine Learning Methods for Histopathological Image Analysis. *Computational and Structural Biotechnology Journal*, **16**, 34-42. <https://doi.org/10.1016/j.csbj.2018.01.001>
- [9] Judd, R.M. (2020) Machine Learning in Medical Imaging: All Journeys Begin with a Single Step. *JACC: Cardiovascular Imaging*, **13**, 696-698. <https://doi.org/10.1016/j.jcmg.2019.08.028>
- [10] Christodoulou, E., Ma, J., Collins, G.S., Steyerberg, E.W., Verbakel, J.Y. and Van Calster, B. (2019) A Systematic Review Shows No Performance Benefit of Machine Learning over Logistic Regression for Clinical Prediction Models. *Journal of Clinical Epidemiology*, **110**, 12-22. <https://doi.org/10.1016/j.jclinepi.2019.02.004>
- [11] Bassi, C., Marchegiani, G., Dervenis, C., Sarr, M., Abu Hilal, M., Adham, M., *et al.* (2017) The 2016 Update of the International Study Group (ISGPS) Definition and Grading of Postoperative Pancreatic Fistula: 11 Years After. *Surgery*, **161**, 584-591. <https://doi.org/10.1016/j.surg.2016.11.014>
- [12] Gaujoux, S., Cortes, A., Couvelard, A., Noullet, S., Clavel, L., Rebours, V., *et al.* (2010) Fatty Pancreas and Increased Body Mass Index Are Risk Factors of Pancreatic Fistula after Pancreatoduodenectomy. *Surgery*, **148**, 15-23. <https://doi.org/10.1016/j.surg.2009.12.005>
- [13] Callery, M.P., Pratt, W.B., Kent, T.S., Chaikof, E.L. and Vollmer, C.M. (2013) A Prospectively Validated Clinical Risk Score Accurately Predicts Pancreatic Fistula after Pancreatoduodenectomy. *Journal of the American College of Surgeons*, **216**, 1-14. <https://doi.org/10.1016/j.jamcollsurg.2012.09.002>
- [14] Mungroop, T.H., van Rijssen, L.B., van Klaveren, D., Smits, F.J., van Woerden, V., Linnemann, R.J., *et al.* (2019) Alternative Fistula Risk Score for Pancreatoduodenectomy (a-FRS): Design and International External Validation. *Annals of Surgery*, **269**, 937-943. <https://doi.org/10.1097/sla.0000000000002620>
- [15] Mungroop, T.H., Klompaker, S., Wellner, U.F., Steyerberg, E.W., Coratti, A., D'Hondt, M., *et al.* (2019) Updated Alternative Fistula Risk Score (ua-FRS) to Include Minimally Invasive Pancreatoduodenectomy: Pan-European Validation. *Annals of Surgery*, **273**, 334-340. <https://doi.org/10.1097/sla.0000000000003234>

- [16] Pande, R., Halle-Smith, J.M., Phelan, L., Thorne, T., Panikkar, M., Hodson, J., *et al.* (2022) External Validation of Postoperative Pancreatic Fistula Prediction Scores in Pancreatoduodenectomy: A Systematic Review and Meta-Analysis. *HPB*, **24**, 287-298. <https://doi.org/10.1016/j.hpb.2021.10.006>
- [17] Han, I.W., Cho, K., Ryu, Y., Shin, S.H., Heo, J.S., Choi, D.W., *et al.* (2020) Risk Prediction Platform for Pancreatic Fistula after Pancreatoduodenectomy Using Artificial Intelligence. *World Journal of Gastroenterology*, **26**, 4453-4464. <https://doi.org/10.3748/wjg.v26.i30.4453>
- [18] Lin, Z., Tang, B., Cai, J., Wang, X., Li, C., Tian, X., *et al.* (2021) Preoperative Prediction of Clinically Relevant Postoperative Pancreatic Fistula after Pancreatoduodenectomy. *European Journal of Radiology*, **139**, Article 109693. <https://doi.org/10.1016/j.ejrad.2021.109693>
- [19] Shen, Z., Chen, H., Wang, W., Xu, W., Zhou, Y., Weng, Y., *et al.* (2022) Machine Learning Algorithms as Early Diagnostic Tools for Pancreatic Fistula Following Pancreatoduodenectomy and Guide Drain Removal: A Retrospective Cohort Study. *International Journal of Surgery*, **102**, Article 106638. <https://doi.org/10.1016/j.ijso.2022.106638>
- [20] Shi, Y., Gao, F., Qi, Y., Lu, H., Ai, F., Hou, Y., *et al.* (2020) Computed Tomography-Adjusted Fistula Risk Score for Predicting Clinically Relevant Postoperative Pancreatic Fistula after Pancreatoduodenectomy: Training and External Validation of Model Upgrade. *eBioMedicine*, **62**, Article 103096. <https://doi.org/10.1016/j.ebiom.2020.103096>
- [21] Verma, A., Balian, J., Hadaya, J., Premji, A., Shimizu, T., Donahue, T., *et al.* (2023) Machine Learning-Based Prediction of Postoperative Pancreatic Fistula Following Pancreatoduodenectomy. *Annals of Surgery*, **280**, 325-331. <https://doi.org/10.1097/sla.0000000000006123>
- [22] Lee, W., Park, H.J., Lee, H., Song, K.B., Hwang, D.W., Lee, J.H., *et al.* (2024) Deep Learning-Based Prediction of Post-Pancreatoduodenectomy Pancreatic Fistula. *Scientific Reports*, **14**, Article No. 5089. <https://doi.org/10.1038/s41598-024-51777-2>
- [23] Choubey, A.P., Magnin, J., Gagnière, J., Midya, A., Steinharter, J.A., Yamashita, R., *et al.* (2025) Postoperative Pancreatic Fistula after Pancreatoduodenectomy: Can Radi-Omics Improve Clinical Risk Scores? *Annals of Surgery*. Publish Ahead of Print. <https://doi.org/10.1097/sla.0000000000000691>
- [24] Schuh, F., Mihaljevic, A.L., Probst, P., *et al.* (2023) A Simple Classification of Pancreatic Duct Size and Texture Predicts Postoperative Pancreatic Fistula: A classification of the International Study Group of Pancreatic Surgery. *Annals of Surgery*, **277**, e597-e608.
- [25] Fang, C., Chen, Q., Yang, J., Xiang, F., Fang, Z. and Zhu, W. (2016) Body Mass Index and Stump Morphology Predict an Increased Incidence of Pancreatic Fistula after Pancreatoduodenectomy. *World Journal of Surgery*, **40**, 1467-1476. <https://doi.org/10.1007/s00268-016-3413-5>
- [26] Funamizu, N., Sogabe, K., Shine, M., Honjo, M., Sakamoto, A., Nishi, Y., *et al.* (2022) Association between the Preoperative C-Reactive Protein-to-Albumin Ratio and the Risk for Postoperative Pancreatic Fistula Following Distal Pancreatectomy for Pancreatic Cancer. *Nutrients*, **14**, Article 5277. <https://doi.org/10.3390/nu14245277>
- [27] Ingwersen, E.W., Rijssenbeek, P.M.W., Marquering, H.A., Kazemier, G. and Daams, F. (2024) Radiomics for the Prediction of a Postoperative Pancreatic Fistula Following a Pancreatoduodenectomy: A Systematic Review and Radiomic Score Quality Assessment. *Pancreatology*, **24**, 306-313. <https://doi.org/10.1016/j.pan.2023.12.007>
- [28] Schlanger, D., Graur, F., Popa, C., Moiş, E. and Al Hajjar, N. (2022) The Role of Artificial Intelligence in Pancreatic Surgery: A Systematic Review. *Updates in Surgery*, **74**, 417-429. <https://doi.org/10.1007/s13304-022-01255-z>
- [29] Alhulaili, Z.M., Linnemann, R.J., Dascau, L., Pleijhuis, R.G. and Klaase, J.M. (2023) A Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis Analysis to Evaluate the Quality of Reporting of Postoperative Pancreatic Fistula Prediction Models after Pancreatoduodenectomy: A Systematic Review. *Surgery*, **174**, 684-691. <https://doi.org/10.1016/j.surg.2023.04.058>
- [30] Casciani, F., Trudeau, M.T., Asbun, H.J., Ball, C.G., Bassi, C., Behrman, S.W., *et al.* (2021) Surgeon Experience Contributes to Improved Outcomes in Pancreatoduodenectomies at High Risk for Fistula Development. *Surgery*, **169**, 708-720. <https://doi.org/10.1016/j.surg.2020.11.022>
- [31] Kim, S., Yoon, Y., Han, H., Cho, J.Y., Choi, Y. and Lee, B. (2020) Evaluation of a Single Surgeon's Learning Curve of Laparoscopic Pancreatoduodenectomy: Risk-Adjusted Cumulative Summation Analysis. *Surgical Endoscopy*, **35**, 2870-2878. <https://doi.org/10.1007/s00464-020-07724-z>
- [32] Blunck, C.K., Vickers, S.M., Wang, T.N., Dudeja, V., Reddy, S. and Rose, J.B. (2022) External Validation of Four Pancreatic Fistula Risk Score Models in the Deep South US: Do Racial Disparities Affect Pancreatic Fistula Prediction? *The American Journal of Surgery*, **224**, 557-561. <https://doi.org/10.1016/j.amjsurg.2022.02.042>
- [33] Kang, J.S., Park, T., Han, Y., Lee, S., Kim, J.R., Kim, H., *et al.* (2019) Clinical Validation of Scoring Systems of Postoperative Pancreatic Fistula after Pancreatoduodenectomy: Applicability to Eastern Cohorts? *HepatoBiliary Surgery and Nutrition*, **8**, 211-218. <https://doi.org/10.21037/hbsn.2019.03.17>